UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/943,699	11/10/2010	Mike Kennewick	62MV-218255	4185

133759 7590 02/19/2020

Sheppard Mullin Richter & Hampton LLP
650 Town Center Drive, 10th Floor
Costa Mesa, CA 92626

EXAMINER	
HALE, TIM B	

ART UNIT	PAPER NUMBER
3685	

NOTIFICATION DATE	DELIVERY MODE
02/19/2020	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

SheppardMullin_Pair@firsttofile.com
dmipdocketing@sheppardmullin.com

UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

In re Patent of 9,502,025	:	
Issue date: November 22, 2016	:	
Application No. 12/943,699	:	NOTICE
Filed: November 10, 2010	:	
Attorney Docket No. 62MV-218255	:	

This is a notice regarding the request for acceptance of a fee deficiency submission under 37 CFR 1.28(c) filed August 13, 2019.

The Office no longer investigates or rejects original or reissue applications under 37 CFR 1.56. **1098 Off. Gaz. Pat. Office 502 (January 3, 1989)**. Therefore, nothing in this Notice is intended to imply that an investigation was done.

The fee deficiency submission under 37 CFR 1.28(c) is hereby **ACCEPTED**.

It is noted that the $4688.00 fee deficiency has been charged from the petitioner's deposit account.

This application is no longer entitled to small entity status. Accordingly, all future fees paid in this application must be paid at the undiscounted rate.

The file does not indicate a change of address has been submitted, although the address given on the petition differs from the address of record. If appropriate, a change of address should be filed in accordance with MPEP 601.03. A courtesy copy of this decision is being mailed to the address given on the petition; however, the Office will mail all future correspondence solely to the address of record.

Telephone inquiries concerning this decision should be directed to Paula Britton at 571-272-1556. All other inquiries concerning the examination or status of the application is accessible in the PAIR system at http://portal.uspto.gov/.

/Irvin Dingle/
Irvin Dingle
Lead Paralegal Specialist
Office of Petitions

Cc: Mark A. Patrick
2099 Pennsylvania Avenue, NW
Suite 100
Washington, DC 20006-6801

Transmittal Communication on Petition	Application/Control No. 12/943,699	Applicant(s)/Patent Under Reexamination Kennewick et al.	
	Deciding Official HALE, TIM B	Office of Petitions OPET	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address. --

(ADDITIONAL PARTY'S CORRESPONDENCE ADDRESS)

Mark A. Patrick
2099 Pennsylvania Avenue, NW
Suite 100
Washington, DC 20006-6801

Enclosed is a copy of the latest communication from the United States Patent and Trademark Office in the above-identified Application/Patent.



United States Patent and Trademark Office

Office of the Chief Financial Officer

Document Code:WFEE

User :C41722

Sale Accounting Date:02/07/2020

Sale Item Reference Number	Effective Date
12943699	08/13/2019

Document Number	Fee Code	Fee Code Description	Amount Paid	Payment Method
I202027B56210717	1461	1.28(C) SUBMISSIONS - APPLIC FILE FEE	$4,688.00	Deposit Account

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY.DOCKET NO./TITLE	REQUEST ID
12/943,699	11/10/2010	Mike Kennewick	62MV-218255	93670

Acknowledgement of Loss of Entitlement to Entity Status Discount

The entity status change request below filed through Private PAIR on 08/14/2019 has been accepted.

CERTIFICATIONS:

Change of Entity Status:

X Applicant changing to regular undiscounted fee status.

NOTE: Checking this box will be taken to be notification of loss of entitlement to small or micro entity status, as applicable.

This portion must be completed by the signatory or signatories making the entity status change in accordance with 37 CFR 1.4(d)(4).

Signature:	/Mark A. Patrick/
Name:	Mark A. Patrick
Registration Number:	72958

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

INVENTORS :	Mike KENNEWICK *et al.*	CONFIRMATION NO.:	4185
SERIAL NUMBER :	12/943,699	PATENT NO.:	9,502,025
FILING DATE :	November 10, 2010	ISSUE DATE:	November 22, 2016
FOR :	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE		

DEFICIENCY PAYMENT UNDER 37 C.F.R. § 1.28(c)

Commissioner for Patents
P. O. Box 1450
Alexandria, VA 22313-1450

Sir:

This paper is being submitted to address a good faith error in payment of small entity fees. In this application, small entity status was established in good faith, and fees as a small entity were paid in good faith. It was later discovered that such status as a small entity was established in error. As such, Applicant believes small entity status is no longer appropriate in the instant application. Applicant notes that the erroneous payments of small entity fees were made inadvertently, without deceptive intent.

Pursuant to 37 C.F.R. § 1.28(c)(1)-(2), the fees paid in error and an itemization of the deficiencies owed for each previous fee erroneously paid as a small entity is attached as **APPENDIX A**, including an itemization of the requirements of 37 C.F.R. § 1.28(c)(2)(ii).

Accordingly, pursuant to 37 C.F.R. § 1.28(c), Applicant requests that the Commissioner excuse the good faith error in payment of small entity fees, and charge the deficiency payment owed, which is calculated to be **$4,688.00,** to Deposit Account No. 504561 **(Ref. No. 62MV-218255).**

Respectfully submitted,

SHEPPARD MULLIN RICHTER & HAMPTON LLP

By /Mark A. Patrick/
Mark A. Patrick
Registration No.: 72,958

Customer No.
133759

2099 Pennsylvania Avenue, NW
Suite 100
Washington, D.C. 20006-6801
Telephone: (202) 747-2192

Dated: August 13, 2019

APPENDIX A

ITEMIZATION OF DEFICIENCY PAYMENTS

DEFICIENT PAYMENT	TYPE OF FEE ERRONEOUSLY PAID AS SMALL ENTITY 37 C.F.R. 1.28(c)(2)(ii)(A)	CURRENT FEE FOR NON-SMALL ENTITY 37 C.F.R. 1.28(c)(2)(ii)(A)	SMALL ENTITY FEE ACTUALLY PAID 37 C.F.R. 1.28(c)(2)(ii)(B)	DATE SMALL ENTITY FEE PAID 37 C.F.R. 1.28(c)(2)(ii)(B)	DEFICIENCY OWED 37 C.F.R. 1.28(c)(2)(ii)(C)
FIRST DEFICIENT PAYMENT	Basic Filing Fee – Utility – Fee Code 4011	$300.00	$82.00	Application filed November 10, 2010	$300.00 - $82.00 = **$218.00**
SECOND DEFICIENT PAYMENT	Utility Search Fee – Fee Code 2111	$660.00	$270.00	Application filed November 10, 2010	$660.00 - $270.00 = **$390.00**
THIRD DEFICIENT PAYMENT	Utility Examination Fee – Fee Code 2311	$760.00	$110.00	Application filed November 10, 2010	$760.00 - $110.00 = **$650.00**
FOURTH DEFICIENT PAYMENT	Surcharge – Late filing fee, search fee, examination fee, inventor's oath or declaration – Fee Code 2051	$160.00	$65.00	Response to Notice to File Missing Parts filed January 26, 2011	$160.00 - $65.00 = **$95.00**
FIFTH DEFICIENT PAYMENT	Extension for response within first month – Fee Code 2251	$200.00	$75.00	Response filed January 30, 2013	$200.00 - $75.00 = **$125.00**
SIXTH DEFICIENT PAYMENT	Notice of Appeal – Fee Code 2401	$800.00	$400.00	Notice of Appeal filed June 5, 2013	$800.00 - $400.00 = **$400.00**
SEVENTH DEFICIENT PAYMENT	Submission of an Information Disclosure Statement – Fee Code 2806	$240.00	$90.00	Information Disclosure Statement filed June 5, 2013	$240.00 - $90.00 = **$150.00**

EIGHTH DEFICIENT PAYMENT	Extension for response within first month – Fee Code 2251	$200.00	$100.00	Response filed September 5, 2013	$200.00 - $100.00 = **$100.00**
NINTH DEFICIENT PAYMENT	Request for Oral Hearing – Fee Code 2403	$1,300.00	$650.00	Request for Oral Hearing filed February 27, 2014	$1,300.00 - $650.00 = **$650.00**
TENTH DEFICIENT PAYMENT	Forwarding an appeal in an application or ex parte reexamination proceeding to the Board – Fee code 2413	$2,240.00	$1,000.00	Reply Brief filed February 27, 2014	$2,240.00 - $1,000.00 = **$1,240.00**
ELEVENTH DEFICIENT PAYMENT	Submission of an Information Disclosure Statement – Fee Code 2806	$240.00	$90.00	Information Disclosure Statement filed May 11, 2015	$240.00 - $90.00 = **$150.00**
TWELFTH DEFICIENT PAYMENT	Utility Issue Fee – Fee Code 2501	$1,000.00	$480.00	Issue Fee paid October 21, 2016	$1,000.00 - $480.00 = **$520.00**
TOTAL DEFICIENCY PAYMENT OWED **37 C.F.R. 1.28(c)(2)(ii)(D)**	$218.00 + $390.00 + $650.00 + $95.00 + $125.00 + $400.00 + $150.00 + $100.00 + $650.00 + $1,240.00 + $150.00 + $520.00 = **$4,688.00** Total Deficiency Owed				

Electronic Acknowledgement Receipt	
EFS ID:	36858781
Application Number:	12943699
International Application Number:	
Confirmation Number:	4185
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Customer Number:	133759
Filer:	Mark Andrew Patrick/Susan Trader
Filer Authorized By:	Mark Andrew Patrick
Attorney Docket Number:	62MV-218255
Receipt Date:	13-AUG-2019
Filing Date:	10-NOV-2010
Time Stamp:	13:47:21
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Petition for review by the Office of Petitions	62MV-218255_Deficiency_Payment.pdf	123238 079e6742ad1693e0960319a7639990062a043d15	no	4
Warnings:					

Information:	
Total Files Size (in bytes):	123238

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	PATENT NUMBER	GROUP ART UNIT	REQUEST ID
12/943,699	9502025	3685	30149

PAIR Correspondence Address/Fee Address Change

The following fields have been changed to Customer Number 00197 on 12/09/2016 via Private PAIR in view of the certification copied below that authorized the change.

- Maintenance Fee Address

The address for Customer Number 00197 is:
00197
CPA GLOBAL LIMITED
2318 Mill Road 12th Floor
ALEXANDRIA, VA 22314

I certify, in accordance with 37 CFR 1.4(d)(4) that I am:

An attorney or Agent of Record registered to practice before the Patent and Trademark Office who has been given power of attorney in this application

Signature:	/James G. Gatto/
Name:	James G. Gatto
Registration Number:	32694

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	ISSUE DATE	PATENT NO.	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/943,699	11/22/2016	9502025	45HV-218255	4185

133759 7590 11/02/2016

Sheppard Mullin Richter & Hampton LLP
12275 El Camino Real, Suite 200
San Diego, CA 92130

ISSUE NOTIFICATION

The projected patent number and issue date are specified above.

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)
(application filed on or after May 29, 2000)

The Patent Term Adjustment is 1345 day(s). Any patent to issue from the above-identified application will include an indication of the adjustment on the front page.

If a Continued Prosecution Application (CPA) was filed in the above-identified application, the filing date that determines Patent Term Adjustment is the filing date of the most recent CPA.

Applicant will be able to obtain more detailed information by accessing the Patent Application Information Retrieval (PAIR) WEB site (http://pair.uspto.gov).

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Application Assistance Unit (AAU) of the Office of Data Management (ODM) at (571)-272-4200.

APPLICANT(s) (Please see PAIR WEB site http://pair.uspto.gov for additional applicants):

Mike Kennewick, Bellevue, WA;
Lynn Elise Armstrong, Woodinville, WA;

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation, and commercialization of new technologies. The USA offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to encourage and facilitate business investment. To learn more about why the USA is the best country in the world to develop technology, manufacture products, and grow your business, visit SelectUSA.gov.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	*Unassigned*
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	Herewith (November 10, 2010)	*Unassigned*

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	1A	4,430,669	02/07/1984	Cheung	358	122	
	1B	5,155,743	10/13/1992	Jacobs	375	28	
	1C	5,274,560	12/28/1993	LaRue	364	444	
	1D	5,377,350	12/27/1994	Skinner	395	600	
	1E	5,386,556	01/31/1995	Hedin et al.	395	600	
	1F	5,424,947	06/13/1995	Nagao et al.	364	419.08	
	1G	5,471,318	11/28/1995	Ahuja et al.	358	400	
	1H	5,475,733	12/12/1995	Eisdorfer et al.	379	52	
	1I	5,488,652	01/30/1996	Bielby *et al.*	379	88	
	1J	5,499,289	03/12/1996	Bruno et al.	379	220	
	1K	5,500,920	03/19/1996	Kupiec	395	2.79	
	1L	5,517,560	05/14/1996	Greenspan	379	114	
	1M	5,533,108	07/02/1996	Harris et al.	379	201	
	1N	5,537,436	07/16/1996	Bottoms et al.	375	222	
	1O	5,539,744	07/23/1996	Chu et al.	370	60	
	1P	5,557,667	09/17/1996	Bruno et al.	379	201	
	1Q	5,563,937	10/08/1996	Bruno et al.	379	201	
	1R	5,590,039	12/31/1996	Ikeda et al.	395	759	
	1S	5,617,407	04/01/1997	Bareis	369	275.3	
	1T	5,633,922	05/27/1997	August et al.	379	220	
	1U	5,675,629	10/07/1997	Raffel et al.	379	58	
	1V	5,696,965	12/09/1997	Dedrick	395	610	
	1W	5,708,422	01/13/1998	Blonder et al.	340	825.34	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO
	1X	WO 99/46763	09/16/1999	International				
	1Y	WO 00/21232	04/13/2000	International				
	1Z	WO 00/46792	08/10/2000	International				
	1AA	WO 01/78065	10/18/2001	International				
	1BB	WO 2004/072954	~~08/26/2004~~	International 08-2004				

Change(s) applied to document, /D.D./ 9/27/2016

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

	1CC	REUTERS, "IBM to Enable Honda Drivers to Talk to Cars", Charles Schwab & Co., Inc., July 28, 2002, 1 page.
	1DD	LIN, Bor-shen, *et al.*, "A Distributed Architecture for Cooperative Spoken Dialogue Agents with Coherent Dialogue State and History", ASRU'99, 1999, 4 pages.

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

402465373v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

PART B - FEE(S) TRANSMITTAL

Complete and send this form, together with applicable fee(s), to: Mail **Mail Stop ISSUE FEE**
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
or Fax (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

133759 7590 07/22/2016
Sheppard Mullin Richter & Hampton LLP
12275 El Camino Real, Suite 200
San Diego, CA 92130

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

Certificate of Mailing or Transmission
I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being facsimile transmitted to the USPTO (571) 273-2885, on the date indicated below.

(Depositor's name)
(Signature)
(Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/943,699	11/10/2010	Mike Kennewick	45HV-218255	4185

TITLE OF INVENTION: SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE

APPLN. TYPE	ENTITY STATUS	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	SMALL	$480	$0	$0	$480	10/24/2016

EXAMINER	ART UNIT	CLASS-SUBCLASS
HALE, TIM B	3685	705-059000

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).
- ☐ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.
- ☐ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list
(1) The names of up to 3 registered patent attorneys or agents OR, alternatively,
(2) The name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.

1 Sheppard Mullin Richter & Hampton LLP
2
3

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE: VoiceBox Technologies Corporation
(B) RESIDENCE: (CITY and STATE OR COUNTRY): Bellevue, Washington

Please check the appropriate assignee category or categories (will not be printed on the patent) : ☐ Individual ■ Corporation or other private group entity ☐ Government

4a. The following fee(s) are submitted:
- ■ Issue Fee
- ☐ Publication Fee (No small entity discount permitted)
- ☐ Advance Order - # of Copies _____

4b. Payment of Fee(s): **(Please first reapply any previously paid issue fee shown above)**
- ☐ A check is enclosed.
- ☐ Payment by credit card. Form PTO-2038 is attached.
- ■ The director is hereby authorized to charge the required fee(s), any deficiency, or credits any overpayment, to Deposit Account Number 501067 (enclose an extra copy of this form).

5. **Change in Entity Status** (from status indicated above)
- ☐ Applicant certifying micro entity status. See 37 CFR 1.29 — NOTE: Absent a valid certification of Micro Entity Status (see forms PTO/SB/15A and 15B), issue fee payment in the micro entity amount will not be accepted at the risk of application abandonment.
- ☐ Applicant asserting small entity status. See 37 CFR 1.27 — NOTE: If the application was previously under micro entity status, checking this box will be taken to be a notification of loss of entitlement to micro entity status.
- ☐ Applicant changing to regular undiscounted fee status. — NOTE: Checking this box will be taken to be a notification of loss of entitlement to small or micro entity status, as applicable.

NOTE: This form must be signed in accordance with 37 CFR 1.31 and 1.33. See 37 CFR 1.4 for signature requirements and certifications.

Authorized Signature /Hean L. Koo/ Date October 21, 2016
Typed or printed name Hean L. Koo Registration No. 61,214

PTOL-85 Part B (10-13) Approved for use through 10/31/2013. OMB 0651-0033 U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Amazon/VB Assets
Exhibit 1003
Page 15

Electronic Patent Application Fee Transmittal

Application Number:	12943699
Filing Date:	10-Nov-2010
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Filer:	Hean Lee Koo/Susan Trader
Attorney Docket Number:	45HV-218255

Filed as Small Entity

Filing Fees for Utility under 35 USC 111(a)

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
UTILITY APPL ISSUE FEE	2501	1	480	480

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Extension-of-Time:				
Miscellaneous:				
		Total in USD ($)		**480**

Electronic Acknowledgement Receipt	
EFS ID:	27282199
Application Number:	12943699
International Application Number:	
Confirmation Number:	4185
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Customer Number:	133759
Filer:	Hean Lee Koo/Susan Trader
Filer Authorized By:	Hean Lee Koo
Attorney Docket Number:	45HV-218255
Receipt Date:	21-OCT-2016
Filing Date:	10-NOV-2010
Time Stamp:	10:35:07
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	DA
Payment was successfully received in RAM	$480
RAM confirmation Number	102116INTEFSW00007923501067
Deposit Account	
Authorized User	
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Issue Fee Payment (PTO-85B)	45HV-218255_IssueFee.pdf	88603 23a5d4e7cfe865cf6adcd876478cbcf73cbcd367	no	1
Warnings:					
Information:					
2	Fee Worksheet (SB06)	fee-info.pdf	30569 e7f2b5bef226f93ddc87b318a44ddbf745c910fd	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	119172		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/943,699	11/10/2010	Mike Kennewick	45HV-218255	4185

133759 7590 09/22/2016

Sheppard Mullin Richter & Hampton LLP
12275 El Camino Real, Suite 200
San Diego, CA 92130

EXAMINER
HALE, TIM B

ART UNIT	PAPER NUMBER
3685	

NOTIFICATION DATE	DELIVERY MODE
09/22/2016	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

dmipdocketing@sheppardmullin.com

UNITED STATES DEPARTMENT OF COMMERCE
U.S. Patent and Trademark Office
Address : COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450

APPLICATION NO./ CONTROL NO.	FILING DATE	FIRST NAMED INVENTOR / PATENT IN REEXAMINATION	ATTORNEY DOCKET NO.
12/943,699	10 November, 2010	KENNEWICK ET AL.	45HV-218255

Sheppard Mullin Richter & Hampton LLP
12275 El Camino Real, Suite 200
San Diego, CA 92130

EXAMINER	
TIM HALE	
ART UNIT	PAPER
3685	20160906

DATE MAILED:

Please find below and/or attached an Office communication concerning this application or proceeding.

Commissioner for Patents

Corrected IDS filed 6/5/2013

/JOHN HAYES/ Supervisory Patent Examiner, Art Unit 3685	/TIM HALE/ Examiner, Art Unit 3685

PTO-90C (Rev.04-03)

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	A	6,122,613	09/19/2000	Baker	704	235	
	B	6,356,869	03/12/2002	Chapados *et al.*	704	275	
	C	6,505,155	01/07/2003	Vanbuskirk *et al.*	704	246	
	D	6,754,627	06/22/2004	Woodward	704	235	
	E	7,197,460	03/27/2007	Gupta *et al.*	704	270.1	
	F	7,421,393	09/02/2008	Di Fabbrizio *et al.*	704	275	
	G	7,890,324	02/15/2011	Bangalore *et al.*	704	231	
	H	8,219,399	07/10/2012	Lutz *et al.*	704	270	
	I	8,255,224	08/28/2012	Singleton *et al.*	704	275	
	J	8,447,607	05/21/2013	Weider *et al.*	704	250	
	K	8,452,598	05/28/2013	Kennewick *et al.*	704	257	
	L	2002/0143532	10/03/2002	McLean *et al.*	704	235	
	M	2009/0076827	03/19/2009	Bulitta *et al.*	704	275	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED
/TIM B HALE/	09/13/2016

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

403997232v1

/T.B.H/

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.B.H/

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Inventor(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	A	8,200,485	06/12/2012	Lee et al.	704	231	
	B	8,983,839	03/17/2015	Kennewick *et al.*	704	254	
	C	9,015,049	04/21/2015	Baldwin *et al.*			
	D	2001/0039492	11/08/2001	Nemoto	704	231	
	E	2005/0080632	04/14/2005	Endo *et al.*	704	277	
	F	2008/0147396	06/19/2008	Wang	704	246	
	G	2008/0162471	07/03/2008	Bernard	707	5	
	H	2010/0094707	04/15/2010	Freer	705	14.54	
	I	2011/0022393	01/27/2011	Waller *et al.*	704	270	
	J	2015/0066627	03/05/2015	Freeman *et al.*	705	14.41	
	K	2015/0073910	03/12/2015	Kennewick *et al.*	705	14.58	
	L	2015/0095159	04/02/2015	Kennewick et al.	705	14.58	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED
/TIM B HALE/	06/29/2016

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

405243468v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.B.H/

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/943,699	11/10/2010	Mike Kennewick	45HV-218255	4185

133759 7590 08/29/2016

Sheppard Mullin Richter & Hampton LLP
12275 El Camino Real, Suite 200
San Diego, CA 92130

EXAMINER
HALE, TIM B

ART UNIT	PAPER NUMBER
3685	

NOTIFICATION DATE	DELIVERY MODE
08/29/2016	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

dmipdocketing@sheppardmullin.com

UNITED STATES DEPARTMENT OF COMMERCE
U.S. Patent and Trademark Office
Address : COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450

APPLICATION NO./ CONTROL NO.	FILING DATE	FIRST NAMED INVENTOR / PATENT IN REEXAMINATION	ATTORNEY DOCKET NO.
12/943,699	10 November, 2010	KENNEWICK ET AL.	45HV-218255

	EXAMINER	
Sheppard Mullin Richter & Hampton LLP 12275 El Camino Real, Suite 200 San Diego, CA 92130	TIM HALE	
	ART UNIT	PAPER
	3685	20160815

DATE MAILED:

Please find below and/or attached an Office communication concerning this application or proceeding.

Commissioner for Patents

Corrected IDS and provided updated claim index.

/JOHN HAYES/ Supervisory Patent Examiner, Art Unit 3685	/TIM HALE/ Examiner, Art Unit 3685

PTO-90C (Rev.04-03)

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	12943699	KENNEWICK ET AL.
	Examiner	**Art Unit**
	TIM HALE	3685

CPC						
Symbol					**Type**	**Version**
G10L		15	/	18	F	2013-01-01
G06Q		30	/	0601	I	2013-01-01
G06F		17	/	28	I	2013-01-01
			/			
			/			
			/			
			/			
			/			
			/			
			/			
			/			
			/			
			/			
			/			
			/			

CPC Combination Sets							
Symbol					**Type**	**Set**	**Ranking**	**Version**
			/					
			/					

/TIM HALE/ Examiner.Art Unit 3685	06/29/2016	**Total Claims Allowed:**	
(Assistant Examiner)	(Date)	19	
/JOHN HAYES/ Supervisory Patent Examiner.Art Unit 3685	06/30/2016	O.G. Print Claim(s)	O.G. Print Figure
(Primary Examiner)	(Date)	1	1

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	12943699	KENNEWICK ET AL.
	Examiner	**Art Unit**
	TIM HALE	3685

US ORIGINAL CLASSIFICATION		INTERNATIONAL CLASSIFICATION										
CLASS	SUBCLASS	CLAIMED					NON-CLAIMED					
705	59	G	1	0	L	15 / 00 (2013.01.01)						

CROSS REFERENCE(S)						
CLASS	SUBCLASS (ONE SUBCLASS PER BLOCK)					

/TIM HALE/ Examiner.Art Unit 3685	06/29/2016	**Total Claims Allowed:**	
(Assistant Examiner)	(Date)	19	
/JOHN HAYES/ Supervisory Patent Examiner.Art Unit 3685	06/30/2016	O.G. Print Claim(s)	O.G. Print Figure
(Primary Examiner)	(Date)	1	1

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	12943699	KENNEWICK ET AL.
	Examiner	**Art Unit**
	TIM HALE	3685

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original
1	1	11	17												
	2	12	18												
	3	13	19												
	4	14	20												
2	5	15	21												
3	6	16	22												
4	7		23												
5	8	17	24												
6	9	18	25												
7	10	19	26												
8	11														
	12														
	13														
	14														
9	15														
10	16														

/TIM HALE/ Examiner.Art Unit 3685	06/29/2016	**Total Claims Allowed:**	
(Assistant Examiner)	(Date)	19	
/JOHN HAYES/ Supervisory Patent Examiner.Art Unit 3685	06/30/2016	O.G. Print Claim(s)	O.G. Print Figure
(Primary Examiner)	(Date)	1	1

IN THE UNITED STATES PATENT & TRADEMARK OFFICE

In Re Patent Application Of:	Mike KENNEWICK *et al.*	**Attorney Docket No.:**	45HV-218255
Serial No.:	12/943,699	**Confirmation No.:**	4185
Filing Date:	November 10, 2010	**Examiner:**	Tim Hale
		Art Unit:	3685

For: SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE

SUBSTANCE OF EXAMINER INTERVIEW

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Applicant presents this substance of an Examiner Interview conducted via telephone on June 22, 2016 (hereinafter "First Interview"), between Examiner Hale and Applicant's representatives Hean Koo (Reg. No. 61,214) and Mark Patrick (Reg. No. 72,958). During the First Interview, the claimed invention was discussed along with the Patent Trial and Appeal Board decision received April 21, 2016. Examiner Hale has also provided an Applicant-Initiated Interview Summary dated June 27, 2016.

Applicant presents this substance of an Examiner Interview conducted via telephone on June 30, 2016 (hereinafter "Second Interview"), between Examiner Hale and Applicant's representative Mark Patrick (Reg. No. 72,958). During the Second Interview, allowable subject matter to be potentially entered as an Examiner Amendment was discussed in order to expedite prosecution of this application. Examiner Hale has also provided an Examiner-Initiated Interview Summary dated July 22, 2016.

Amazon/VB Assets
Exhibit 1003
Page 29

Respectfully submitted,

Date: **July 25, 2016**

SHEPPARD MULLIN RICHTER & HAMPTON LLP

By: / Mark A. Patrick /
Mark A. Patrick
Registration No. 72,958

Customer No. 133759

2099 Pennsylvania Avenue, NW, Suite 100
Washington, DC 20006-6801
Direct Dial: 202-747-2192
Fax: 202-747-3853

Electronic Acknowledgement Receipt	
EFS ID:	26447425
Application Number:	12943699
International Application Number:	
Confirmation Number:	4185
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Customer Number:	133759
Filer:	Mark Andrew Patrick/Masako Onda
Filer Authorized By:	Mark Andrew Patrick
Attorney Docket Number:	45HV-218255
Receipt Date:	25-JUL-2016
Filing Date:	10-NOV-2010
Time Stamp:	17:04:02
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Applicant summary of interview with examiner	45HV218255_IntvwSummary.pdf	25373 a8c51b08b0d6dbe2fd31e040df22a77b9b38f743	no	2

Warnings:

Information:	
Total Files Size (in bytes):	25373

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

NOTICE OF ALLOWANCE AND FEE(S) DUE

133759 7590 07/22/2016
Sheppard Mullin Richter & Hampton LLP
12275 El Camino Real, Suite 200
San Diego, CA 92130

EXAMINER	
HALE, TIM B	
ART UNIT	PAPER NUMBER
3685	

DATE MAILED: 07/22/2016

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/943,699	11/10/2010	Mike Kennewick	45HV-218255	4185

TITLE OF INVENTION: SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE

APPLN. TYPE	ENTITY STATUS	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	SMALL	$480	$0	$0	$480	10/24/2016

THE APPLICATION IDENTIFIED ABOVE HAS BEEN EXAMINED AND IS ALLOWED FOR ISSUANCE AS A PATENT. PROSECUTION ON THE MERITS IS CLOSED. THIS NOTICE OF ALLOWANCE IS NOT A GRANT OF PATENT RIGHTS. THIS APPLICATION IS SUBJECT TO WITHDRAWAL FROM ISSUE AT THE INITIATIVE OF THE OFFICE OR UPON PETITION BY THE APPLICANT. SEE 37 CFR 1.313 AND MPEP 1308.

THE ISSUE FEE AND PUBLICATION FEE (IF REQUIRED) MUST BE PAID WITHIN THREE MONTHS FROM THE MAILING DATE OF THIS NOTICE OR THIS APPLICATION SHALL BE REGARDED AS ABANDONED. THIS STATUTORY PERIOD CANNOT BE EXTENDED. SEE 35 U.S.C. 151. THE ISSUE FEE DUE INDICATED ABOVE DOES NOT REFLECT A CREDIT FOR ANY PREVIOUSLY PAID ISSUE FEE IN THIS APPLICATION. IF AN ISSUE FEE HAS PREVIOUSLY BEEN PAID IN THIS APPLICATION (AS SHOWN ABOVE), THE RETURN OF PART B OF THIS FORM WILL BE CONSIDERED A REQUEST TO REAPPLY THE PREVIOUSLY PAID ISSUE FEE TOWARD THE ISSUE FEE NOW DUE.

HOW TO REPLY TO THIS NOTICE:

I. Review the ENTITY STATUS shown above. If the ENTITY STATUS is shown as SMALL or MICRO, verify whether entitlement to that entity status still applies.

If the ENTITY STATUS is the same as shown above, pay the TOTAL FEE(S) DUE shown above.

If the ENTITY STATUS is changed from that shown above, on PART B - FEE(S) TRANSMITTAL, complete section number 5 titled "Change in Entity Status (from status indicated above)".

For purposes of this notice, small entity fees are 1/2 the amount of undiscounted fees, and micro entity fees are 1/2 the amount of small entity fees.

II. PART B - FEE(S) TRANSMITTAL, or its equivalent, must be completed and returned to the United States Patent and Trademark Office (USPTO) with your ISSUE FEE and PUBLICATION FEE (if required). If you are charging the fee(s) to your deposit account, section "4b" of Part B - Fee(s) Transmittal should be completed and an extra copy of the form should be submitted. If an equivalent of Part B is filed, a request to reapply a previously paid issue fee must be clearly made, and delays in processing may occur due to the difficulty in recognizing the paper as an equivalent of Part B.

III. All communications regarding this application must give the application number. Please direct all communications prior to issuance to Mail Stop ISSUE FEE unless advised to the contrary.

IMPORTANT REMINDER: Utility patents issuing on applications filed on or after Dec. 12, 1980 may require payment of maintenance fees. It is patentee's responsibility to ensure timely payment of maintenance fees when due.

Amazon/VB Assets
Exhibit 1003
Page 33

PART B - FEE(S) TRANSMITTAL

Complete and send this form, together with applicable fee(s), to: Mail **Mail Stop ISSUE FEE**
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
or Fax (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

133759 7590 07/22/2016
Sheppard Mullin Richter & Hampton LLP
12275 El Camino Real, Suite 200
San Diego, CA 92130

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

Certificate of Mailing or Transmission
I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being facsimile transmitted to the USPTO (571) 273-2885, on the date indicated below.

(Depositor's name)
(Signature)
(Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/943,699	11/10/2010	Mike Kennewick	45HV-218255	4185

TITLE OF INVENTION: SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE

APPLN. TYPE	ENTITY STATUS	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	SMALL	$480	$0	$0	$480	10/24/2016

EXAMINER	ART UNIT	CLASS-SUBCLASS
HALE, TIM B	3685	705-059000

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).
- ❑ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.
- ❑ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list
(1) The names of up to 3 registered patent attorneys or agents OR, alternatively,
(2) The name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.
1 _____
2 _____
3 _____

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE (B) RESIDENCE: (CITY and STATE OR COUNTRY)

Please check the appropriate assignee category or categories (will not be printed on the patent) : ❑ Individual ❑ Corporation or other private group entity ❑ Government

4a. The following fee(s) are submitted:
- ❑ Issue Fee
- ❑ Publication Fee (No small entity discount permitted)
- ❑ Advance Order - # of Copies _____

4b. Payment of Fee(s): **(Please first reapply any previously paid issue fee shown above)**
- ❑ A check is enclosed.
- ❑ Payment by credit card. Form PTO-2038 is attached.
- ❑ The director is hereby authorized to charge the required fee(s), any deficiency, or credits any overpayment, to Deposit Account Number _____ (enclose an extra copy of this form).

5. **Change in Entity Status** (from status indicated above)
- ❑ Applicant certifying micro entity status. See 37 CFR 1.29 — NOTE: Absent a valid certification of Micro Entity Status (see forms PTO/SB/15A and 15B), issue fee payment in the micro entity amount will not be accepted at the risk of application abandonment.
- ❑ Applicant asserting small entity status. See 37 CFR 1.27 — NOTE: If the application was previously under micro entity status, checking this box will be taken to be a notification of loss of entitlement to micro entity status.
- ❑ Applicant changing to regular undiscounted fee status. — NOTE: Checking this box will be taken to be a notification of loss of entitlement to small or micro entity status, as applicable.

NOTE: This form must be signed in accordance with 37 CFR 1.31 and 1.33. See 37 CFR 1.4 for signature requirements and certifications.

Authorized Signature _____ Date _____

Typed or printed name _____ Registration No. _____

PTOL-85 Part B (10-13) Approved for use through 10/31/2013. OMB 0651-0033 U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Amazon/VB Assets
Exhibit 1003
Page 34

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/943,699	11/10/2010	Mike Kennewick	45HV-218255	4185

133759 7590 07/22/2016

Sheppard Mullin Richter & Hampton LLP
12275 El Camino Real, Suite 200
San Diego, CA 92130

EXAMINER	
HALE, TIM B	

ART UNIT	PAPER NUMBER
3685	

DATE MAILED: 07/22/2016

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)
(Applications filed on or after May 29, 2000)

The Office has discontinued providing a Patent Term Adjustment (PTA) calculation with the Notice of Allowance.

Section 1(h)(2) of the AIA Technical Corrections Act amended 35 U.S.C. 154(b)(3)(B)(i) to eliminate the requirement that the Office provide a patent term adjustment determination with the notice of allowance. See Revisions to Patent Term Adjustment, 78 Fed. Reg. 19416, 19417 (Apr. 1, 2013). Therefore, the Office is no longer providing an initial patent term adjustment determination with the notice of allowance. The Office will continue to provide a patent term adjustment determination with the Issue Notification Letter that is mailed to applicant approximately three weeks prior to the issue date of the patent, and will include the patent term adjustment on the patent. Any request for reconsideration of the patent term adjustment determination (or reinstatement of patent term adjustment) should follow the process outlined in 37 CFR 1.705.

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Customer Service Center of the Office of Patent Publication at 1-(888)-786-0101 or (571)-272-4200.

PTOL-85 (Rev. 02/11)

Amazon/VB Assets
Exhibit 1003
Page 35

OMB Clearance and PRA Burden Statement for PTOL-85 Part B

The Paperwork Reduction Act (PRA) of 1995 requires Federal agencies to obtain Office of Management and Budget approval before requesting most types of information from the public. When OMB approves an agency request to collect information from the public, OMB (i) provides a valid OMB Control Number and expiration date for the agency to display on the instrument that will be used to collect the information and (ii) requires the agency to inform the public about the OMB Control Number's legal significance in accordance with 5 CFR 1320.5(b).

The information collected by PTOL-85 Part B is required by 37 CFR 1.311. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450. Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Notice of Allowability	Application No. 12/943,699	Applicant(s) KENNEWICK ET AL.	
	Examiner TIM HALE	Art Unit 3685	AIA (First Inventor to File) Status No

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☒ This communication is responsive to *6/27/2016*.
   ☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on_____.
2. ☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.
3. ☒ The allowed claim(s) is/are *1, 5-11, 15-22 and 24-26*. As a result of the allowed claim(s), you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see http://www.uspto.gov/patents/init_events/pph/index.jsp or send an inquiry to PPHfeedback@uspto.gov .
4. ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

   Certified copies:

   a) ☐ All b) ☐ Some *c) ☐ None of the:
   1. ☐ Certified copies of the priority documents have been received.
   2. ☐ Certified copies of the priority documents have been received in Application No. _____ .
   3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

   * Certified copies not received: _____.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

5. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
   ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____.

   Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).
6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1. ☐ Notice of References Cited (PTO-892)
2. ☒ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date 5/11/2015
3. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material
4. ☒ Interview Summary (PTO-413), Paper No./Mail Date *6/29/2016* .
5. ☒ Examiner's Amendment/Comment
6. ☒ Examiner's Statement of Reasons for Allowance
7. ☐ Other _____.

/TIM HALE/	/JOHN HAYES/
Examiner, Art Unit 3685	Supervisory Patent Examiner, Art Unit 3685

EXAMINER'S AMENDMENT

1. An examiner's amendment to the record appears below. Should the changes and/or additions be unacceptable to applicant, an amendment may be filed as provided by 37 CFR 1.312. To ensure consideration of such an amendment, it MUST be submitted no later than the payment of the issue fee.

2. Authorization for this examiner's amendment was given in a telephone interview with Mark Patrick on 6/29/2016.

3. The application had been amended as follow:

1. (Currently Amended) A system for providing a natural language content dedication service, comprising:

one or more processors; and

one or more non-transitory computer readable mediums storing executable instructions that when executed by the one or more processors cause the one or more processors to:

receive a first utterance that includes a natural language utterance;

determine, based on processing of the first utterance by a speech recognition engine, one or more words or phrases of the first utterance;

<u>provide the one or more words or phrases as an input to a conversational language processor;</u>

<u>interpret the first utterance, at the conversational language processor, based on the one or more words or phrases;</u>

identify, based on ~~a natural language processing of the one or more words or phrases of~~ the interpretation of the first utterance, content to dedicate to a recipient;

initiate, based on the identified content, a dedication to the recipient;

receive a second utterance to be associated with the dedication; [[and]]

determine, based on a processing of the second utterance by the speech recognition engine, one or more words or phrases of the second utterance;

provide the one or more words or phrases of the second utterance as textual annotations within metadata of the content; and

send information to enable the recipient to access the content and the second utterance.

2-4. (Canceled)

5. (Previously Presented) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to provide an image as one or more tags within metadata of the content.

6. (Previously Presented) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to annotate the content with the second utterance.

7. (Previously Presented) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to purchase the content from a content provider to enable the recipient to access the content.

8. (Currently Amended) The system of claim 7, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to:

identify, for the dedication, a purchase option that includes one or more of a buy-to-own purchase option, a pay-to-~~pay~~ play purchase option, or a paid subscription purchase option; and

bill a user associated with the first utterance based on the purchase option.

9. (Previously Presented) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to: send, to the recipient, an electronic link to access the content as the information to enable the recipient to access the content; and

send, to the recipient, the second utterance in an encoded audio format as the information to enable the recipient to access the second utterance.

10. (Original) The system of claim 1, wherein the content includes one or more of a song or a video.

11. (Currently Amended) A method for providing a natural language content dedication service, the method being implemented in a computer system having one or more processors programmed with computer program instructions that, when executed by the one or more processors, cause the computer system to perform the method, the method comprising:

receiving, by one or more processors, a first utterance that includes a natural language utterance;

determining, by the one or more processors, one or more words or phrases of the first utterance based on processing of the first utterance by a speech recognition engine;

providing, by the one or more processors, the one or more words or phrases as an input to a conversational language processor;

interpreting, by the one or more processors, the first utterance, at the conversational language processor, based on the one or more words or phrases;

identifying, by the one or more processors, content to dedicate to a recipient based on ~~a natural language processing of the one or more words or phrases of~~ the interpretation of the first utterance;

initiating, by the one or more processors, a dedication to the recipient based on the identified content;

receiving, by the one or more processors, a second utterance to be associated with the dedication; [[and]]

determining, by the one or more processors, one or more words or phrases of the second utterance based on a processing of the second utterance by the speech recognition engine;

providing, by the one or more processors, the one or more words or phrases of the second utterance as textual annotations within metadata of the content; and

sending, by the one or more processors, information to enable the recipient to access the content and the second utterance.

12-14. (Canceled)

15. (Previously Presented) The method of claim 11, further comprising providing an image as one or more tags within metadata of the content.

16. (Previously Presented) The method of claim 11, further comprising annotating the content with the second utterance.

17. (Previously Presented) The method of claim 11, further comprising purchasing the content from a content provider to enable the recipient to access the content.

18. (Currently Amended) The method of claim 17, further comprising:

identifying, for the dedication, a purchase option that includes one or more of a buy-to-own purchase option, a pay-to-~~pay~~ play purchase option, or a paid subscription purchase option; and

billing a user associated with the first utterance based on the purchase option.

19. (Previously Presented) The method of claim 11, further comprising:

sending, to the recipient, an electronic link to access the content as the information to enable the recipient to access the content; and

sending, to the recipient, the second utterance in an encoded audio format as the information to enable the recipient to access the second utterance.

20. (Original) The method of claim 11, wherein the content includes one or more of a song or a video.

21. (Previously Presented) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to:

determine that a user is interacting with a media item; and

identify, based on the natural language processing and the interaction of the user with the media item, the media item as the content to dedicate to the recipient.

22. (Currently Amended) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to:

determine a media item that [[is]] a user is playing;

determine that the user provided the first utterance during the playing of the media item by the user; and

identify, based on the natural language processing and the determination that the user provided the first utterance during the playing of the media item, the media item as the content to dedicate to the recipient.

23. (Canceled).

24. (Previously Presented) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to provide an electronic link associated with the content as one or more tags within metadata of the content.

25. (Previously Presented) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to provide a voice tag as one or more tags within metadata of the content.

26. (Previously Presented) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to provide the second utterance within the content.

Reasons for Allowance

4. The following is an examiner's statement of reasons for allowance:

The present invention is directed to processing natural language to customize the content for recipients of the dedications, and delivering the customized content to the recipients of the dedications. Content dedication is well known in the art. For example, US Publication 20020032752 to Gold (hereafter "Gold") discloses dedicating content to a recipient (See Fig. 2 [0036-0038]).

The closest prior art US Publication 20040193420 to Kennewick (hereafter "Kennewick") discloses: "A system for providing a natural language content dedication service, comprising: one or more processors and one or more non-transitory computer readable mediums storing executable instructions that when executed by the one or more processors cause the processors to: receive, a first utterance that includes a natural language utterance; and determine, based on the processing of the first utterance by a speech recognition engine, one or more words or phrases of the first utterance."

However, the combination of prior art does not provide for the two step process of natural language processing during a dedication as reflected in the claim language of determine, based on processing of the first utterance by a speech recognition engine, one or more words or phrases of the first utterance; provide the one or more words or phrases as an input to a conversational language processor; interpret the first utterance, at the conversational language processor, based on the one or more words or phrases;

determine, based on a processing of the second utterance by the speech recognition engine, one or more words or phrases of the second utterance; and provide the one or more words or phrases of the second utterance as textual annotations within metadata of the content."

Conclusion

5. Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Tim B Hale whose telephone number is (571)270-3323. The examiner can normally be reached on M-F.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John Hayes can be reached on (571)272-6708. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300. Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status

information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/JOHN HAYES/
Supervisory Patent Examiner, Art Unit 3685

Examiner-Initiated Interview Summary	**Application No.** 12/943,699	**Applicant(s)** KENNEWICK ET AL.	
	Examiner TIM HALE	**Art Unit** 3685	

All participants (applicant, applicant's representative, PTO personnel):

(1) *TIM HALE*. (3)_____.

(2) *Mark Patrick (Reg. No. 65,473)* . (4)_____.

Date of Interview: *30 June 2016*.

Type: ☒ Telephonic ☐ Video Conference
☐ Personal [copy given to: ☐ applicant ☐ applicant's representative]

Exhibit shown or demonstration conducted: ☐ Yes ☐ No.
If Yes, brief description: _____.

Issues Discussed ☒101 ☐112 ☐102 ☐103 ☐Others
(For each of the checked box(es) above, please describe below the issue and detailed description of the discussion)

Claim(s) discussed: *All*.

Identification of prior art discussed: _____.

Substance of Interview
(For each issue discussed, provide a detailed description and indicate if agreement was reached. Some topics may include: identification or clarification of a reference or a portion thereof, claim interpretation, proposed amendments, arguments of any applied references etc...)

Discussed possible amendments to claims to overcome 101 rejection. Agreement reached with regard to claim amendments that would place the claims 21-23 within the scope of eligible subject matter. Examiner given client authorization to proceed with allowance by Examiner Amendment.

Applicant recordation instructions: It is not necessary for applicant to provide a separate record of the substance of interview.

Examiner recordation instructions: Examiners must summarize the substance of any interview of record. A complete and proper recordation of the substance of an interview should include the items listed in MPEP 713.04 for complete and proper recordation including the identification of the general thrust of each argument or issue discussed, a general indication of any other pertinent matters discussed regarding patentability and the general results or outcome of the interview, to include an indication as to whether or not agreement was reached on the issues raised.

☐ Attachment

/TIM HALE/ Examiner, Art Unit 3685	

EAST Search History

EAST Search History (Interference)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
S42	5873	@ad<"20091110" and g10l15/$.cpc.	USPAT	AND	ON	2016/06/29 16:17
S43	395	@ad<"20091110" and (g10l15/18.cpc. or g06f17/28.cpc. or go6q30/0601.cpc.)	USPAT	AND	ON	2016/06/29 16:18
S44	4	@ad<"20091110" and (g10l15/18.cpc. or g06f17/28.cpc. or go6q30/0601.cpc.) and (utterance same metadata)	USPAT	AND	ON	2016/06/29 16:18
S45	0	@ad<"20091110" and (g10l15/18.cpc. or g06f17/28.cpc. or go6q30/0601.cpc.) and (utterance same metadata) same content	USPAT	AND	ON	2016/06/29 16:18
S46	1	@ad<"20091110" and (g10l15/18.cpc. or g06f17/28.cpc. or go6q30/0601.cpc.) and (utterance same metadata) same word	USPAT	AND	ON	2016/06/29 16:19
S47	0	@ad<"20091110" and (g10l15/18.cpc. or g06f17/28.cpc. or go6q30/0601.cpc.) and (utterance same metadata) same word same phrase	USPAT	AND	ON	2016/06/29 16:19

6/29/2016 4:20:47 PM
C:\Users\thale\Documents\EAST\Workspaces\12943699.wsp

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	12943699	KENNEWICK ET AL.
	Examiner	Art Unit
	TIM HALE	3685

CPC- SEARCHED		
Symbol	Date	Examiner

CPC COMBINATION SETS - SEARCHED		
Symbol	Date	Examiner

US CLASSIFICATION SEARCHED			
Class	Subclass	Date	Examiner

SEARCH NOTES		
Search Notes	Date	Examiner
East text search	9/12/2012	TBH
East Text Search including inventor and assignee(s)	6/29/2016	TBH

INTERFERENCE SEARCH			
US Class/ CPC Symbol	US Subclass / CPC Group	Date	Examiner
G10l 15/18	G06F 17/28; GO6Q 30/0601	6/29/2016	TBH

EAST Search History

EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	7839535	@ad<"20091110"	US-PGPUB; USPAT	AND	ON	2016/06/29 16:13
L2	3686	@ad<"20091110" and (kennewick.in. or armstrong.in.)	US-PGPUB; USPAT	AND	ON	2016/06/29 16:13
L3	13	@ad<"20091110" and (kennewick.in. or armstrong.in.) and (content same dedica$5)	US-PGPUB; USPAT	AND	ON	2016/06/29 16:14
L4	40	@ad<"20091110" and (kennewick.in. or armstrong.in.) and (natural same language)	US-PGPUB; USPAT	AND	ON	2016/06/29 16:14
L5	140	@ad<"20091110" and (kennewick.in. or armstrong.in.) and (language same process$5)	US-PGPUB; USPAT	AND	ON	2016/06/29 16:14
L6	4	@ad<"20091110" and (kennewick.in. or armstrong.in.) and (language same process$5) same (song or movie)	US-PGPUB; USPAT	AND	ON	2016/06/29 16:15
L7	34	@ad<"20091110" and (kennewick.in. or armstrong.in.) and (language same process$5) same (song or movie or content)	US-PGPUB; USPAT	AND	ON	2016/06/29 16:15
L8	2	@ad<"20091110" and (voicebox$5.as.) and (language same process$5) same (song or movie)	US-PGPUB; USPAT	AND	ON	2016/06/29 16:15
L9	2	@ad<"20091110" and (voicebox$.as.) and (language same process$5) same (song or movie)	US-PGPUB; USPAT	AND	ON	2016/06/29 16:16
L10	26	@ad<"20091110" and (voicebox$.as.) and (language same process$5)	US-PGPUB; USPAT	AND	ON	2016/06/29 16:16
L11	0	@ad<"20091110" and (voicebox$.as.) and (content same dedicat$5)	US-PGPUB; USPAT	AND	ON	2016/06/29 16:16
S1	2	"20040068443"	US-PGPUB; USPAT	AND	ON	2014/11/04 14:50
S2	1	"20040068433"	US-PGPUB; USPAT	AND	ON	2014/11/04 14:53
S3	2	nemovicher.in.	US-PGPUB; USPAT	AND	ON	2014/11/04 14:58
S4	2	"20060053201"	US-PGPUB; USPAT	AND	ON	2014/11/04 15:06
S5	928	@ad<"20050222" and license same ((e-	US-	AND	ON	2014/11/04

		mail) or (electronic adj2 mail) or (electroinc adj2 address))	PGPUB; USPAT			15:35
S6	78	@ad<"20050222" and (license adj2 plate) same ((e-mail) or (electronic adj2 mail) or (electroinc adj2 address))	US-PGPUB; USPAT	AND	ON	2014/11/04 15:36
S7	1	nemovicher.in. and register	US-PGPUB; USPAT	AND	ON	2014/11/04 17:03
S8	1	"20020007453" and regist$5	US-PGPUB; USPAT	AND	ON	2014/11/04 17:04
S9	1	"20020007453" and node	US-PGPUB; USPAT	AND	ON	2014/11/05 13:57
S10	1	"8719172".pn.	US-PGPUB; USPAT	AND	ON	2014/11/05 14:23
S11	0	@ad<"20050222" and (license adj2 server) and (moter adj2 vehicle)	US-PGPUB; USPAT	AND	ON	2014/11/05 14:43
S12	0	@ad<"20050222" and (license adj2 server) and (motr adj2 vehicle)	US-PGPUB; USPAT	AND	ON	2014/11/05 14:43
S13	8	@ad<"20050222" and (license adj2 server) and (motor adj2 vehicle)	US-PGPUB; USPAT	AND	ON	2014/11/05 14:43
S14	1	"20020107809" and signature	US-PGPUB; USPAT	AND	ON	2014/11/05 14:48
S15	4	2002/0026478 and sever	US-PGPUB; USPAT	AND	ON	2014/11/05 15:38
S16	108	2002/0026478 and database	US-PGPUB; USPAT	AND	ON	2014/11/05 15:39
S17	1	"20020026478" and database	US-PGPUB; USPAT	AND	ON	2014/11/05 15:39
S18	0	"20020026478" and sever	US-PGPUB; USPAT	AND	ON	2014/11/05 15:39
S19	1	"20020026478" and server	US-PGPUB; USPAT	AND	ON	2014/11/05 15:39
S20	1	"20020007453" and signature	US-PGPUB; USPAT	AND	ON	2014/11/05 18:10
S21	1	"20020007453"	US-PGPUB; USPAT	AND	ON	2015/08/12 13:33
S22	1	"20020007453" and address	US-PGPUB; USPAT	AND	ON	2015/08/12 13:33
S23	1	"20020007453" and address$5	US-PGPUB; USPAT	AND	ON	2015/08/12 13:34
S24	2	"20060053201"	US-	AND	ON	2015/08/12

Amazon/VB Assets
Exhibit 1003
Page 52

			PGPUB; USPAT			13:36
S25	145	@ad<"20050222" and ginter.in.	US-PGPUB; USPAT	AND	ON	2015/08/12 13:39
S26	43	@ad<"20050222" and ginter.in. and gateway	US-PGPUB; USPAT	AND	ON	2015/08/12 13:40
S27	1	"20040133793" and e-mail	US-PGPUB; USPAT	AND	ON	2015/08/12 13:48
S28	1	"20040133793" and email	US-PGPUB; USPAT	AND	ON	2015/08/12 13:49
S29	1	"20040133793" and mail	US-PGPUB; USPAT	AND	ON	2015/08/12 13:50
S30	1	"20040133793" and ISO	US-PGPUB; USPAT	AND	ON	2015/08/12 13:56

6/29/2016 4:17:01 PM
C:\Users\thale\Documents\EAST\Workspaces\12943699.wsp

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	12943699	KENNEWICK ET AL.
	Examiner	**Art Unit**
	TIM HALE	3685

CPC						
Symbol					**Type**	**Version**
G10L		15	/	18	F	2013-01-01
G06Q		30	/	0601	I	2013-01-01
G06F		17	/	28	I	2013-01-01
			/			
			/			
			/			
			/			
			/			
			/			
			/			
			/			
			/			
			/			
			/			

CPC Combination Sets							
Symbol					**Type**	**Set**	**Ranking**	**Version**
			/					
			/					

/TIM HALE/ Examiner.Art Unit 3685	06/29/2016	**Total Claims Allowed:**	
(Assistant Examiner)	(Date)	19	
/JOHN HAYES/ Supervisory Patent Examiner.Art Unit 3685	06/30/2016	O.G. Print Claim(s)	O.G. Print Figure
(Primary Examiner)	(Date)	1	1

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	12943699	KENNEWICK ET AL.
	Examiner	**Art Unit**
	TIM HALE	3685

US ORIGINAL CLASSIFICATION		INTERNATIONAL CLASSIFICATION										
CLASS	SUBCLASS	CLAIMED					NON-CLAIMED					
705	59	G	1	0	L	15 / 00 (2013.01.01)						
CROSS REFERENCE(S)												
CLASS	SUBCLASS (ONE SUBCLASS PER BLOCK)											

/TIM HALE/ Examiner.Art Unit 3685	06/29/2016	**Total Claims Allowed:**	
(Assistant Examiner)	(Date)	19	
/JOHN HAYES/ Supervisory Patent Examiner.Art Unit 3685	06/30/2016	O.G. Print Claim(s)	O.G. Print Figure
(Primary Examiner)	(Date)	1	1

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	12943699	KENNEWICK ET AL.
	Examiner	**Art Unit**
	TIM HALE	3685

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original
2	1	11	17												
	2	12	18												
	3	13	19												
	4	4	20												
2	5	15	21												
3	6	16	22												
4	7		23												
5	8	17	24												
6	9	18	25												
7	10	19	26												
8	11														
	12														
	13														
	14														
9	15														
10	16														

/TIM HALE/ Examiner.Art Unit 3685	06/29/2016	**Total Claims Allowed:**	
(Assistant Examiner)	(Date)	19	
/JOHN HAYES/ Supervisory Patent Examiner.Art Unit 3685	06/30/2016	O.G. Print Claim(s)	O.G. Print Figure
(Primary Examiner)	(Date)	1	1

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Inventor(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	A	8,200,485	06/12/2012	Lee et al.	704	231	
	B	8,983,839	03/17/2015	Kennewick *et al.*	704	254	
	C	9,015,049	04/21/2015	Baldwin *et al.*			
	D	2001/0039492	11/08/2001	Nemoto	704	231	
	E	2005/0080632	04/14/2005	Endo *et al.*	704	277	
	F	2008/0147396	06/19/2008	Wang	704	246	
	G	2008/0162471	07/03/2008	Bernard	707	5	
	H	2010/0094707	04/15/2010	Freer	705	14.54	
	I	2011/0022393	01/27/2011	Waller *et al.*	704	270	
	J	2015/0066627	03/05/2015	Freeman *et al.*	705	14.41	
	K	2015/0073910	03/12/2015	Kennewick *et al.*	705	14.58	
	L	2015/0095159	04/02/2015	Kennewick et al.	705	14.58	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED
/TIM B HALE/	06/29/2016

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

405243468v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.B.H/

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	12943699	KENNEWICK ET AL.
	Examiner	Art Unit
	TIM HALE	3685

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	09/18/2012	06/29/2016							
	1	✓	=							
	2	✓	-							
	3	✓	-							
	4	✓	-							
	5	✓	=							
	6	✓	=							
	7	✓	=							
	8	✓	=							
	9	✓	=							
	10	✓	=							
	11	✓	=							
	12	✓	-							
	13	✓	-							
	14	✓	-							
	15	✓	=							
	16	✓	=							
	17	✓	=							
	18	✓	=							
	19	✓	=							
	20	✓	=							
	21		=							
	22		=							
	23		-							
	24		=							
	25		=							
	26		=							

Amazon/VB Assets
Exhibit 1003
Page 58

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/943,699	11/10/2010	Mike Kennewick	45HV-218255	4185

133759 7590 06/27/2016

Sheppard Mullin Richter & Hampton LLP
12275 El Camino Real, Suite 200
San Diego, CA 92130

EXAMINER
HALE, TIM B

ART UNIT	PAPER NUMBER
3685	

NOTIFICATION DATE	DELIVERY MODE
06/27/2016	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

dmipdocketing@sheppardmullin.com

Applicant-Initiated Interview Summary	**Application No.** 12/943,699	**Applicant(s)** KENNEWICK ET AL.	
	Examiner TIM HALE	**Art Unit** 3685	

All participants (applicant, applicant's representative, PTO personnel):

(1) *TIM HALE*. (3)_____.

(2) *Ngai B. Zhang (Reg. No. 65,473)*. (4)_____.

Date of Interview: *22 June 2016*.

Type: ☒ Telephonic ☐ Video Conference
☐ Personal [copy given to: ☐ applicant ☐ applicant's representative]

Exhibit shown or demonstration conducted: ☐ Yes ☐ No.
If Yes, brief description: _____.

Issues Discussed ☒101 ☐112 ☐102 ☐103 ☐Others
(For each of the checked box(es) above, please describe below the issue and detailed description of the discussion)

Claim(s) discussed: *21-23*.

Identification of prior art discussed: _____.

Substance of Interview
(For each issue discussed, provide a detailed description and indicate if agreement was reached. Some topics may include: identification or clarification of a reference or a portion thereof, claim interpretation, proposed amendments, arguments of any applied references etc...)

Discussed board decission reversing rejection of claims 21-23 and 101 eligibliity of said claims. Although no specific agreement was reached, Examiner noted that a two step technical process whereby text is identified from an utterance and the text is then parsed to determine meaning was present in the claims..

Applicant recordation instructions: The formal written reply to the last Office action must include the substance of the interview. (See MPEP section 713.04). If a reply to the last Office action has already been filed, applicant is given a non-extendable period of the longer of one month or thirty days from this interview date, or the mailing date of this interview summary form, whichever is later, to file a statement of the substance of the interview

Examiner recordation instructions: Examiners must summarize the substance of any interview of record. A complete and proper recordation of the substance of an interview should include the items listed in MPEP 713.04 for complete and proper recordation including the identification of the general thrust of each argument or issue discussed, a general indication of any other pertinent matters discussed regarding patentability and the general results or outcome of the interview, to include an indication as to whether or not agreement was reached on the issues raised.

☐ Attachment

/TIM HALE/ Examiner, Art Unit 3685	/JOHN HAYES/ Supervisory Patent Examiner, Art Unit 3685

Amazon/VB Assets
Exhibit 1003
Page 60

Summary of Record of Interview Requirements

Manual of Patent Examining Procedure (MPEP), Section 713.04, Substance of Interview Must be Made of Record
A complete written statement as to the substance of any face-to-face, video conference, or telephone interview with regard to an application must be made of record in the application whether or not an agreement with the examiner was reached at the interview.

Title 37 Code of Federal Regulations (CFR) § 1.133 Interviews
Paragraph (b)

In every instance where reconsideration is requested in view of an interview with an examiner, a complete written statement of the reasons presented at the interview as warranting favorable action must be filed by the applicant. An interview does not remove the necessity for reply to Office action as specified in §§ 1.111, 1.135. (35 U.S.C. 132)

37 CFR §1.2 Business to be transacted in writing.
All business with the Patent or Trademark Office should be transacted in writing. The personal attendance of applicants or their attorneys or agents at the Patent and Trademark Office is unnecessary. The action of the Patent and Trademark Office will be based exclusively on the written record in the Office. No attention will be paid to any alleged oral promise, stipulation, or understanding in relation to which there is disagreement or doubt.

The action of the Patent and Trademark Office cannot be based exclusively on the written record in the Office if that record is itself incomplete through the failure to record the substance of interviews.

It is the responsibility of the applicant or the attorney or agent to make the substance of an interview of record in the application file, unless the examiner indicates he or she will do so. It is the examiner's responsibility to see that such a record is made and to correct material inaccuracies which bear directly on the question of patentability.

Examiners must complete an Interview Summary Form for each interview held where a matter of substance has been discussed during the interview by checking the appropriate boxes and filling in the blanks. Discussions regarding only procedural matters, directed solely to restriction requirements for which interview recordation is otherwise provided for in Section 812.01 of the Manual of Patent Examining Procedure, or pointing out typographical errors or unreadable script in Office actions or the like, are excluded from the interview recordation procedures below. Where the substance of an interview is completely recorded in an Examiners Amendment, no separate Interview Summary Record is required.

The Interview Summary Form shall be given an appropriate Paper No., placed in the right hand portion of the file, and listed on the "Contents" section of the file wrapper. In a personal interview, a duplicate of the Form is given to the applicant (or attorney or agent) at the conclusion of the interview. In the case of a telephone or video-conference interview, the copy is mailed to the applicant's correspondence address either with or prior to the next official communication. If additional correspondence from the examiner is not likely before an allowance or if other circumstances dictate, the Form should be mailed promptly after the interview rather than with the next official communication.

The Form provides for recordation of the following information:
- Application Number (Series Code and Serial Number)
- Name of applicant
- Name of examiner
- Date of interview
- Type of interview (telephonic, video-conference, or personal)
- Name of participant(s) (applicant, attorney or agent, examiner, other PTO personnel, etc.)
- An indication whether or not an exhibit was shown or a demonstration conducted
- An identification of the specific prior art discussed
- An indication whether an agreement was reached and if so, a description of the general nature of the agreement (may be by attachment of a copy of amendments or claims agreed as being allowable). Note: Agreement as to allowability is tentative and does not restrict further action by the examiner to the contrary.
- The signature of the examiner who conducted the interview (if Form is not an attachment to a signed Office action)

It is desirable that the examiner orally remind the applicant of his or her obligation to record the substance of the interview of each case. It should be noted, however, that the Interview Summary Form will not normally be considered a complete and proper recordation of the interview unless it includes, or is supplemented by the applicant or the examiner to include, all of the applicable items required below concerning the substance of the interview.

A complete and proper recordation of the substance of any interview should include at least the following applicable items:
1) A brief description of the nature of any exhibit shown or any demonstration conducted,
2) an identification of the claims discussed,
3) an identification of the specific prior art discussed,
4) an identification of the principal proposed amendments of a substantive nature discussed, unless these are already described on the Interview Summary Form completed by the Examiner,
5) a brief identification of the general thrust of the principal arguments presented to the examiner,
   (The identification of arguments need not be lengthy or elaborate. A verbatim or highly detailed description of the arguments is not required. The identification of the arguments is sufficient if the general nature or thrust of the principal arguments made to the examiner can be understood in the context of the application file. Of course, the applicant may desire to emphasize and fully describe those arguments which he or she feels were or might be persuasive to the examiner.)
6) a general indication of any other pertinent matters discussed, and
7) if appropriate, the general results or outcome of the interview unless already described in the Interview Summary Form completed by the examiner.

Examiners are expected to carefully review the applicant's record of the substance of an interview. If the record is not complete and accurate, the examiner will give the applicant an extendable one month time period to correct the record.

Examiner to Check for Accuracy

If the claims are allowable for other reasons of record, the examiner should send a letter setting forth the examiner's version of the statement attributed to him or her. If the record is complete and accurate, the examiner should place the indication, "Interview Record OK" on the paper recording the substance of the interview along with the date and the examiner's initials.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/943,699	11/10/2010	Mike Kennewick	45HV-218255	4185

133759 7590 04/27/2016

Sheppard Mullin Richter & Hampton LLP
12275 El Camino Real, Suite 200
San Diego, CA 92130

EXAMINER
HALE, TIM B

ART UNIT	PAPER NUMBER
3685	

NOTIFICATION DATE	DELIVERY MODE
04/27/2016	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

dmipdocketing@sheppardmullin.com

RECORD OF ORAL HEARING

U.S. PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

Ex parte MIKE KENNEWICK and LYNN ELISE ARMSTRONG

Appeal 2014-004731
Application 12/943,699
Technology Center 3600

Oral Hearing Held: April 5, 2016

Before MURRIEL CRAWFORD, HUBERT LORIN, NINA L. MEDLOCK, *Administrative Patent Judges*.

APPEARANCES:

ON BEHALF OF THE APPELLANT:

JAMES G. GATTO, ESQ.
of: Sheppard Mullin Richter & Hampton LLP
2099 Pennsylvania Avenue, NW
Suite 100
Washington, D.C. 20006-6801
(202) 747-1945
jgatto@sheppardmullin.com

The above-entitled matter came on for hearing on Tuesday, April 5, 2016, commencing at 1:00 p.m. at the U.S. Patent and Trademark Office, 600 Dulany Street, Alexandria, Virginia.

P-R-O-C-E-E-D-I-N-G-S

1:25 p.m.

MS. PRADOS: 2014-004731 Mr. Gatto.

JUDGE LORIN: Okay. Great, thank you very much. Good afternoon counsel.

MR. GATTO: Good afternoon.

JUDGE LORIN: Let me repeat the appeal number so we're absolutely sure about this. This is 2014-004731.

MR. GATTO: Yes.

JUDGE LORIN: All right, fantastic. Okay. We are familiar with the record. You have 20 minutes. You may proceed.

MR. GATTO: Great. Thank you. Good afternoon and may it please the Board. We all like music. We share music with friends. We like to experience things and music can enhance that. Wouldn't it be really nice to be able to, when you hear a song that reminds you of a friend or an event, to be able to just easily send that song to a friend with a dedication and some message?

Well that's exactly what the invention does. And it's different than the prior art which the examiner seems to characterize as simply using natural language processing to send a dedication of content. The claims actually are more specific than that and it covers -- the first part is doing an utterance.

When you're hearing a song, you're playing a song, or it's on your iPhone or whatever, saying send that song to Judge Lorin. And

the app will identify the words you spoke, that's the speech recognition. And the natural language processing will take it, parse it and determine the intent of what you want there, that you want to do a dedication.

And so what it does, if it hears that song, it knows from the context of what's playing on your device that you're referring to a particular song. And by identifying the person, you know who it's being sent to. And it takes that information and initiates a request to send the content to the recipient.

And you can do a second utterance as a follow-on to say this song reminded me of our time at the beach last summer. And that utterance, that second utterance, is kind of the message that goes along with the dedication.

The prior art doesn't disclose that combination. If we look quickly at the prior art, the Kennewick reference which is the common inventor of this invention, is NLP system in general. It discloses natural language processing. The examiner doesn't even allege that it discloses content dedication. It's kind of more background for the NLP technology.

What the examiner relies on primarily for the content dedication is the Gold reference. And Gold does disclose content dedication but it's a much different application, it functions differently, and it achieves a different result.

With Gold, you go to a website, you do a search, you find songs, you select a song, and you can then send it to a recipient. It's not a

system that enables you to, you know, identify, to send a song that's being played right now. You have to kind of go out of the context of what you're doing, go to a separate site, and do that.

Be that as it may, it doesn't disclose the functionality of being able to parse an utterance and identify from a single utterance the content and a recipient which is what's specifically claimed in Claims 1 and 11.

JUDGE MEDLOCK: Counsel, can I just interrupt you?

MR. GATTO: Yes.

JUDGE MEDLOCK: You are suggesting, are you, that claim 1 requires that the utterance actually is made during the playing of the media item that you want to dedicate. It seems to me that that actually only appeared -- maybe I'm missing it. It appears to me that it only appears in the dependent claim.

MR. GATTO: Well in Claim 1, there's a couple of parts that read together I think do disclose it. I agree it is more specifically disclosed in some of the dependent claims. But the claim 1 recites that there's a first utterance that you identify based on that content to dedicate to a recipient. And you initiate based on a content dedication to the recipient.

So collectively, from a single utterance you're getting the content to be dedicated and the recipient to whom to send it to. I agree that some of the dependent claims further elaborate on that. But I

believe in Claim 1, that combination of recitations requires that the first utterance have that combination of the content and the recipient.

JUDGE MEDLOCK: I guess that wasn't my question.

MR. GATTO: I'm sorry.

JUDGE MEDLOCK: My question was whether or not you are suggesting that Claim 1 requires that the first utterance be made during the playing of the song that you wish to dedicate.

MR. GATTO: Claim 1 does not require that, no.

JUDGE MEDLOCK: Okay.

MR. GATTO: It would cover it. It doesn't require it, I agree.

JUDGE MEDLOCK: Okay.

MR. GATTO: Claim 21 and 22, for example, do have specific recitations. For example, Claim 21 says that you can determine that a user is interacting with a media item and then identify based on the NLP that the media item is the content to be dedicated. Claim 22 has similar language.

JUDGE MEDLOCK: Right. I think I may have misunderstood your statement.

MR. GATTO: Okay. So you know, really Gold is a very different system. You have to kind of go out of the context, go do a search, find songs to go do. It's a different context than what's in the claims here.

And Brachman, similarly, which the examiner relies on is really not about content dedication. The portions the examiner relies on say

you can use natural language processing to say, I like these songs. It's a way of you finding -- it's more of a search tool as opposed to a dedication tool.

It's a way for you to say I like this music or this radio station. And it's a way for you to find content. It's not disclosed as providing content dedication or the other features of the claims that we talked about.

The other thing I'd like -- just to go back to Claims 21 and 22. In the examiner's answer, it really doesn't really address the substance of the argument in Claim 21. It refers rather to Claim 22. And in connection with that, the examiner states that Anderson, you know, adding Anderson to the mix of the other references would remedy the deficiencies of those prior references with respect to the features of 21 and 22.

Because Anderson, according the examiner, in Paragraphs 48 and 67 does disclose that you can determine the status of an application that's being played on a mobile device. But again, there's no suggestion of using that for purposes of identifying a song or content to display, to dedicate rather.

What Anderson discloses is you can use the status of a song that's playing, for example, just to identify the title and artist. And so you can display it to the user who is listening. So again, there's no content dedication aspect there.

And interestingly, the examiner, in combining Kennewick, Gold, Brachman, and Anderson in the examiner's answer at Pages 14 to 15, recognizes that by using the natural language processing of Kennewick as an input to the system of Gold would have been an improvement. The examiner is saying because it's an improvement it would be obvious. That seems to fly in the face of patent law.

He's recognizing that this system is different, it's better. And saying based on that it would be obvious. Again, that seems to be an unsupported legal conclusion there.

So for all these reasons, you know, the examiner at best has natural language processing which is not used for content dedication. A content dedication system which you have to go to a separate website, do a search, find songs, and then do whatever process might be available for content dedication.

In contrast to what was specifically claimed here which is a much simpler elegant system that enables you to dedicate a song as it's being played without leaving the context of what you're doing.

JUDGE MEDLOCK: See that was my question. You say dedicate a song as it's being played. But Claim 1 doesn't really require that, does it?

MR. GATTO: Again, at least Claims 21 and 22 require that.

JUDGE MEDLOCK: Sure.

MR. GATTO: I would acknowledge that.

JUDGE MEDLOCK: Yes.

MR. GATTO: But even then, even with respect to claim 1, the specifics of what is claimed, the fact that you have a first utterance that identifies content and recipient which is what the combination of elements that I referred to earlier. That combination is not disclosed in Gold either.

In Gold, again, you have to go to a separate, you have to go to a site, do a search, identify content, select it, and then presumably select the recipient as well. That's a multi-step process and it doesn't, not only does it not have an utterance but doesn't have an analogous step of being able to select both content and a recipient through a single action.

Nor does Gold disclose a second utterance that's associated with the dedication. So again, the second utterance being the message that you want to send. So I think for those reasons, even Claims 1 and 11 are patentable over the combination of Gold, Brachman, and Kennewick.

JUDGE MEDLOCK: Well Gold does disclose, right, that you can record a voice message and have it associated with the content?

MR. GATTO: Right. You can send a message with it but it's not a second utterance in the context of having a previous utterance that identifies the content and the recipient. Yes, I agree.

So you know, our premise, and we kind of kept the Brief very simple, is that these elements are just missing from the prior art. And as such, the obviousness rejection is improper and should be overturned.

JUDGE LORIN: I just have one question. When you say dedicate to, what does that mean?

MR. GATTO: Dedicate to? So in the context of the way it's described in the specification, dedication is sending content to someone. And the dedication is a message saying, this song makes me think of you or whatever the message you want to do with that dedication.

JUDGE LORIN: Is that the same as associating or corresponding?

MR. GATTO: Yes. The dedication gets associated with the song that's sent, yes. There's different ways disclosed in which that --

JUDGE LORIN: So it's just a relationship is established?

MR. GATTO: Yes.

JUDGE LORIN: Okay.

MR. GATTO: And I believe in some of the claims -- and there's different ways in which that relationship is established. In claim 1, for example, you send information to enable the recipient to access the content and the second utterance which is the dedication.

In Claim 6, you specifically annotate the content with the second utterance. So you actually include it along with the content. And in Claim 9, you send an audio file of the utterance to enable the recipient to access that second utterance. So there's different ways in which the association can occur.

JUDGE MEDLOCK: I have one other question. I believe in your Brief you say that the first utterance -- I think you may have mentioned here as well. You say the first utterance has to identify not only the content but also the recipient.

MR. GATTO: Yes.

JUDGE MEDLOCK: Will you explain that to me in detail?

MR. GATTO: Yes. And I think that was the answer I gave to the different question you asked earlier. In that in claim 1 there is an utterance and you identify based on that utterance the content to dedicate to a recipient. And you initiate, based on the identified content, a dedication to the recipient.

So that, our view is the combination of those claim elements make clear that the content and the recipient are determined from the first utterance.

JUDGE MEDLOCK: So basically the clause "identify content to dedicate to a recipient," you're saying means identifying not only the content but also the recipient?

MR. GATTO: Yes. In the context of the claim as a whole, yes. That and the next element which is initiate a dedication to the recipient, the identified recipient. And that combination of features that are recited would require that the utterance include both the content and the recipient.

JUDGE MEDLOCK: Okay.

JUDGE LORIN: Okay. Any further questions?

Amazon/VB Assets
Exhibit 1003
Page 72

JUDGE MEDLOCK: No.

JUDGE LORIN: We have no further questions. Thank you very much.

MR. GATTO: Great. Thank you very much for your time.

(Whereupon, the above-entitled matter went off the record at 1:37 p.m.)

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/943,699	11/10/2010	Mike Kennewick	45HV-218255	4185

133759 7590 04/21/2016

Sheppard Mullin Richter & Hampton LLP
12275 El Camino Real, Suite 200
San Diego, CA 92130

EXAMINER
HALE, TIM B

ART UNIT	PAPER NUMBER
3685	

NOTIFICATION DATE	DELIVERY MODE
04/21/2016	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

dmipdocketing@sheppardmullin.com

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

Ex parte MIKE KENNEWICK and LYNN ELISE ARMSTRONG

Appeal 2014-004731[1]
Application 12/943,699[2]
Technology Center 3600

Before MURRIEL E. CRAWFORD, HUBERT C. LORIN, and NINA L. MEDLOCK, *Administrative Patent Judges*.

MEDLOCK, *Administrative Patent Judge*.

DECISION ON APPEAL

STATEMENT OF THE CASE

Appellants appeal under 35 U.S.C. § 134(a) from the Examiner's final rejection of claims 1, 5–11, and 15–26. An oral hearing was held on April 5, 2016. We have jurisdiction under 35 U.S.C. § 6(b).

We AFFIRM-IN-PART.

---

[1] Our decision references Appellants' Appeal Brief ("App. Br.," filed September 5, 2013) and Reply Brief ("Reply Br.," filed February 27, 2014), and the Examiner's Answer ("Ans.," mailed December 27, 2013) and Final Office Action ("Final Act.," mailed March 5, 2013). The record includes a transcript of the oral hearing held April 5, 2016.

[2] Appellants identify VoiceBox Technologies Corporation as the real party in interest. App. Br. 2.

CLAIMED INVENTION

Appellants' claimed invention

> generally relates to providing a natural language content dedication service in a voice services environment, and in particular, to detecting multi-modal device interactions that include requests to dedicate content, identifying the content requested for dedication from natural language utterances included in the multi-modal device interactions, processing transactions for the content requested for dedication, processing natural language to customize the content for recipients of the dedications, and delivering the customized content to the recipients of the dedications.

(Spec. ¶ 2).

Claim 1, reproduced below, is illustrative of the subject matter on appeal:

> 1. A system for providing a natural language content dedication service, comprising:
>
> one or more processors; and
>
> one or more non-transitory computer readable mediums storing executable instructions that when executed by the one or more processors cause the one or more processors to:
>
> receive a first utterance that includes a natural language utterance;
>
> determine, based on processing of the first utterance by a speech recognition engine, one or more words or phrases of the first utterance;
>
> identify, based on a natural language processing of the one or more words or phrases of the first utterance, content to dedicate to a recipient;
>
> initiate, based on the identified content, a dedication to the recipient;
>
> receive a second utterance to be associated with the dedication; and
>
> send information to enable the recipient to access the content and the second utterance.

Amazon/VB Assets
Exhibit 1003
Page 76

REJECTIONS

Claims 1, 7, 9, 10, 11, 17, 19, 20, and 26 are rejected under 35 U.S.C. § 103(a) as unpatentable over Kennewick (US 2004/0193420 A1, pub. Sept. 30, 2004), Gold (US 2002/0032752 A1, pub. Mar. 14, 2002), and Brachman (US 6,704,576 B1, iss. Mar. 9, 2004).

Claims 6 and 16 are rejected under 35 U.S.C. § 103(a) as unpatentable over Kennewick, Gold, Brachman, and Mahajan (US 2008/0140385 A1, pub. June 12, 2008).

Claims 5, 15, and 23–25 are rejected under 35 U.S.C. § 103(a) as unpatentable over Kennewick, Gold, Brachman, and Caine (US 2010/0036967 A1, pub. Feb. 11, 2010).

Claims 8 and 18 are rejected under 35 U.S.C. § 103(a) as unpatentable over Kennewick, Gold, Brachman, and Ginter (US 5,892,900, iss. Apr. 6, 1999).

Claims 21 and 22 are rejected under 35 U.S.C. § 103(a) as unpatentable over Kennewick, Gold, Brachman, and Anderson (US 2009/0276700 A1, pub. Nov. 5, 2009).

ANALYSIS

Independent claims 1 and 11 and dependent claims 7, 9, 10, 17, 19, 20, and 26

Appellants argue claims 1, 7, 9, 10, 11, 17, 19, 20, and 26 as a group (App. Br. 6–8). We select claim 1 as representative. The remaining claims stand or fall with claim 1. *See* 37 C.F.R. §41.37(c)(1)(iv).

Appellants maintain that the Examiner erred in rejecting claim 1 under 35 U.S.C. § 103(a) because none of Kennewick, Gold, and Brachman, alone or in combination, discloses or suggests "identify[ing], based on a natural

Amazon/VB Assets
Exhibit 1003
Page 77

language processing of the one or more words or phrases of the first utterance, content to dedicate to a recipient," as recited in claim 1 (App. Br. 7–8).

By way of background, Kennewick is directed to a system for capturing and processing natural language utterances so as to allow users to ask natural language questions or give natural language commands in a wide range of domains, e.g., in a vehicular telematics system (*see, e.g.*, Kennewick ¶¶ 28–34, 41, Abstract). Kennewick, thus, discloses that for each user utterance, including a question or query or a set of questions or queries, the system may perform multiple steps, including, *inter alia*, (1) capturing the user's question or query through accurate speech recognition; (2) parsing and interpreting the question or query; (3) determining the domain of expertise required and context, invoking the proper resources, including agents; and (4) formulating one or more queries to one or more local and/or network data sources or sending appropriate commands to local or remote devices or the system itself (*id.* ¶ 28).

Gold is directed to a method and system for sending dedications of popular songs through a communications medium such as the Internet (*see, e.g.*, Gold ¶¶ 31, 32, 36, 43). Gold discloses that, in accordance with the method, a dedication sender, using a personal computer or another device for accessing a Web-based dedication provider, enters song search criteria, e.g., a song title, and, in response, receives, from the provider, a list of songs matching the search criteria (*id.* ¶ 31). The sender selects a song from the received list, and enters transaction information including the email address of the recipient (*id.*). In addition to the list of songs, the sender also receives a phone number to use to leave a dedication message in the form of

Amazon/VB Assets
Exhibit 1003
Page 78

a voice message, which when combined with the dedicated song will create a complete dedication; alternatively, a text message may be used as a dedication message (*id.*). An email is sent to the recipient's email address containing a URL link to a dedication Web page (*id.* at Abstract). Retrieval of the dedication Web page causes retrieval of a dedication file containing URLs pointing to the voice dedication and the dedicated song; these URLs are extracted from the dedication file and used by the recipient to play back the voice dedication and dedicated song (*id.*).

Brachman discloses a method and system that allows a user to request data, e.g., multimedia content, using natural language (*see* Brachman, col. 4, ll. 22–26). Brachman discloses that the request may be entered by typing in a natural language search; alternatively, a voice means may be used to enter the user's natural language request (*id.*).

The Examiner takes the position in the Final Office Action that neither Kennewick nor Gold teaches the argued limitation, i.e., "identify[ing], based on a natural language processing of the one or more words or phrases of the first utterance, content to dedicate to a recipient," as recited in claim 1, and the Examiner cites Brachman to cure the deficiency of Kennewick and Gold (Final Act. 4–5 (citing Brachman, col. 4, ll. 5–27; col. 8, ll. 53–67; col. 9, ll. 1–11; claim 7)). However, the Examiner asserts in the Answer that the rejection of claim 1 is proper based on two separate rationales, i.e., I (Kennewick in view of Gold) and II (Kennewick in view of Gold and Brachman) (Ans. 4–8). Describing rationale I, the Examiner explains that Kennewick discloses the base method, i.e., capturing and processing a natural language speech utterance (*see, e.g.*, *id.* at 4–5 (citing Kennewick ¶ 41)), and that Gold teaches that "the sender reviews the song

Amazon/VB Assets
Exhibit 1003
Page 79

list and selects a song to dedicate to a recipient" (*id.* at 5 (citing Gold ¶ 36)). From this, the Examiner concludes that it would have been obvious to a person of ordinary skill in the art to use Kennewick to process a natural language utterance and to then route that utterance as input to the Gold system (*id.* at 5–6).

Turning to rationale II, the Examiner notes that "Gold teaches: a sender reviewing a song list and selecting a song to dedicate to a recipient (See [0036] of Gold)" and that "Brachman teaches: a user may request data using a natural language search and a voice means may be used to enter the user's natural language request. (See col. 4[,] l[l.] 22–26 of Brachman)" (*id.* at 6). And the Examiner concludes that

> [s]ince each individual element and its function are shown in the prior art, albeit shown in separate references, the difference between the claimed subject matter and the prior art rests not on any individual element or function but in the very combination itself that is in the substitution of content dedicated to the recipient for broadly defined content. Thus, the simple substitution of one known element for another producing a predictable result renders the claim obvious.

Id. at 7.

Responding to the Examiner's assertions in their Reply Brief, Appellants again maintain that Kennewick, Gold, and Brachman, even if combined, fail to disclose or suggest "identify[ing], based on a natural language processing of the one or more words or phrases of the first utterance, content to dedicate to a recipient," as recited in claim 1 — which Appellants assert requires that "both the content to be dedicated and the recipient are identified based on a natural language processing of one or more words or phrases of a first utterance" (Reply Br. 3–4).

Addressing the combination of Kennewick and Gold (i.e., the Examiner's rationale I), Appellants note that Gold teaches that a user may take a number of actions to dedicate content to a recipient, i.e., retrieving a song list, selecting a song from the song list, and providing a recipient's email address to send the song to the recipient (*id.* at 5 (citing Gold ¶ 36)). And Appellants argue that because "Gold's dedication process requires a user to manually handle each of the foregoing actions," the combination of Kennewick and Gold fails to disclose or suggest that one or more words or phrases of a first utterance may be processed to identify content to be dedicated and a recipient of the content, as called for in claim 1 (*id.*).

Appellants' argument is not persuasive at least it is not commensurate with the language of the claim. Appellants assert that claim 1 requires that both the content to be dedicated and the recipient are identified based on natural language processing of one or more words or phrases of the first utterance (Reply Br. 3–4). But claim 1 merely recites that the one or more words or phrases of the first utterance are processed to identify "content to dedicate to a recipient," i.e., that the first utterance is processed to identify the content, e.g., the particular song, that will be the subject of the dedication.

The Examiner cites Kennewick as disclosing a system and method for capturing and processing natural language speech utterances, and finds that it would have been obvious to a person of ordinary skill in the art to use Kennewick to process a natural language utterance and to route the utterance as input, e.g., to identify content to be dedicated, to the Gold system. Appellants have presented no persuasive argument or technical reasoning to explain why one of ordinary skill in the art, using no more than ordinary

creativity, would not have combined the teachings of Gold and Kennewick, as the Examiner proposes.

Addressing the Examiner's assertion in the Answer that Gold and Brachman, in combination, disclose identifying content to be dedicated to a recipient based on natural language processing of words or phrases of a first utterance, Appellants assert that Brachman, at best, teaches natural language search queries as opposed to natural language processing that identifies a particular content and/or a recipient to whom the content is to be dedicated (Reply Br. 5). Appellants, thus, argue that although Brachman's natural language search may produce a set of content results, "the natural language search does not identify either (i) the content to dedicate or (ii) the recipient to whom the content is to be dedicated" because a user would need to manually look through the set of content results to identify a particular content item of interest (*id.* at 6 (emphasis omitted)).

Appellants' argument is not persuasive at least because the Examiner relies on the combination of Gold and Brachman as disclosing the argued limitation, and not on Brachman alone. *See In re Merck & Co.*, 800 F.2d 1091, 1097 (Fed. Cir. 1986) ("Non-obviousness cannot be established by attacking references individually where the rejection is based upon the teachings of a combination of references."). We also fail to see why, and Appellants do not explain why, Brachman's natural language search query could not be used to identify a single content item, rather than a set of content results, which single content item is the particular content item of interest.

Amazon/VB Assets
Exhibit 1003
Page 82

In view of the foregoing, we sustain the Examiner's rejection of claim 1 under 35 U.S.C. § 103(a). We also sustain the Examiner's rejection of claims 7, 9, 10, 11, 17, 19, 20, and 26, which fall with claim 1.

Dependent claims 6 and 16

Appellants argue claims 6 and 16 together (App. Br. 8–9). We select claim 6 as representative. Claim 16 stands or falls with claim 6. *See* 37 C.F.R. § 41.37(c)(1)(iv).

Claim 6 depends from claim 1, and recites that "the executable instructions, when executed by the one or more processors, cause the one or more processors to annotate the content with the second utterance."

We are not persuaded by Appellants' argument that the Examiner erred in rejecting claim 6 because "the relied-upon portion of Mahajan (i.e., ¶ 40)" fails to disclose or suggest the claimed annotation feature (App. Br. 9). Appellants acknowledge that the cited portion of Mahajan teaches that a written transcript of audio/visual content may be presented on a user interface as the audio/visual content is being played (*id.*). But Appellants argue that presentation of a transcription of content on a user interface as the content is being played is not "annotation of content (that is to be dedicated to a recipient) with an utterance . . . much less annotation of content with a second utterance that is to be associated with a dedication to a recipient" (*id.* (emphasis omitted); *see also* Reply Br. 7–8).

Appellants' argument is not persuasive at least because claim 6 is rejected as unpatentable over the combination of Kennewick, Gold, Brachman, and Mahajan, and not over any one of them alone. As described above, non-obviousness cannot be established by attacking references

Amazon/VB Assets
Exhibit 1003
Page 83

individually where, as here, the rejection is based on the teachings of a combination of references. *Merck*, 800 F.2d at 1097.

Appellants' argument also is not persuasive to the extent that Appellants argue that a written transcript associated with the content, as disclosed in Mahajan, is not an annotation. The term "annotate" is generally understood to mean "to supply (a text) with critical or explanatory notes; comment upon in notes." *See* RANDOM HOUSE KERNERMAN WEBSTER'S COLLEGE DICTIONARY, Random House, Inc. (2010), accessed at http://www.thefreedictionary.com/annotate (last visited on April 1, 2016). And we find no indication in the Specification that Appellants have given the term a meaning different from its ordinary and customary meaning.

We agree with the Examiner that, under a broadest reasonable interpretation standard, "annotate the content with the second utterance" can properly be construed as a written transcript associated with the content, as disclosed in Mahajan (Ans. 9). To the extent that Appellants argue that presenting a "written transcription of content with the same content is not annotation of content," as called for in claim 6 (Reply Br. 7), we note, as the Examiner observes, that Gold discloses combining a voice message with the dedicated song, and, thus, discloses "annotat[ing] the content with the second utterance" under a broadest reasonable interpretation standard (Ans. 9 (citing Gold ¶¶ 31, 32)).

In view of the foregoing, we sustain the Examiner's rejection under 35 U.S.C. § 103(a) of claim 6, and claim 16, which falls with claim 6.

Amazon/VB Assets
Exhibit 1003
Page 84

Dependent claims 5, 15, and 23–25

Appellants argue claims 5, 15, and 23–25 as a group (App. Br. 9–11). We find that these claims are substantially different in scope; therefore, we address the claims separately.

Claim 23

Claim 23 depends from claim 1, and recites that

> the executable instructions, when executed by the one or more processors, cause the one or more processors to: determine . . . one or more words or phrases of the second utterance; and provide the one or more words or phrases of the second utterance as textual annotations within metadata of the content.

We are persuaded by Appellants' argument that the Examiner erred in rejecting claim 23 under 35 U.S.C. § 103(a) because the combination of Kennewick, Gold, Brachman, and Caine fails to disclose or suggest "provid[ing] the one or more words or phrases of the second utterance as textual annotations within metadata of the content," as recited in claim 23 (App. Br. 9–11; *see also* Reply Br. 8–9).

The Examiner relies on the combination of Gold and Caine as disclosing the argued feature, and asserts in the Answer that Gold discloses adding a text message to the content (Ans.11 (citing Gold ¶ 31) and that Caine discloses changes to image information including tags, which are the part of the metadata (*id.* (citing Caine ¶¶ 79, 80)). However, we fail to see how, and the Examiner does not explain how, adding the text messages in Gold, which are not words or phrases of an utterance, to metadata of an image, as disclosed in Caine, discloses or suggests "provid[ing] the one or more words or phrases of the second utterance as textual annotations within metadata of the content," as recited in claim 23.

Amazon/VB Assets
Exhibit 1003
Page 85

Therefore, we do not sustain the Examiner's rejection of claim 23 under 35 U.S.C. § 103(a).

Claims 5, 15, 24, and 25

As described above, we find that claims 5, 15, 24, and 25 are substantially different in scope from claim 23. Because Appellants do not provide any argument in support of the patentability of claims 5, 15, 24, and 25, apart from the argument with respect to claim 23, we summarily sustain the Examiner's rejection of claims 5, 15, 24, and 25 under 35 U.S.C. § 103(a).

Dependent claims 8 and 18

Claims 8 and 18 ultimately depend from independent claims 1 and 11, respectively. Appellants do not present any argument in support of the patentability of claims 8 and 18 except to assert that the claims are allowable based on their dependence from claims 1 and 11 (App. Br. 11).

We are not persuaded for the reasons set forth above that the Examiner erred in rejecting claims 1 and 11 under 35 U.S.C. § 103(a). Therefore, we sustain the Examiner's rejection under 35 U.S.C. § 103(a) of claims 8 and 18.

Dependent claims 21 and 22

We are persuaded by Appellants' argument that the Examiner erred in rejecting claim 22 under 35 U.S.C. § 103(a) because the combination of Kennewick, Gold, Brachman, and Anderson fails to disclose or suggest identifying the content to be dedicated to a recipient based on a determination that a user provided a first utterance during the playing of the media item, i.e., "identify[ing], based . . . the determination that the user provided the first utterance during the playing of the media item, the media

item as the content to dedicate to the recipient," as recited in claim 22 (App. Br. 11–13; *see also* Reply Br. 10–11).

The Examiner cites Kennewick as disclosing natural language processing (Ans. 13–14 (citing Kennewick ¶ 41)), Gold as disclosing the selection of a song to dedicate to a recipient from a song list (*id.* (citing Gold ¶ 36)), and Anderson as disclosing that the context of a mobile terminal may be based on the applications currently being performed by the mobile terminal, e.g., whether a user is playing a games, interacting with a media item (*id.* (citing Anderson ¶¶ 48, 67)). The Examiner, thus, concludes that applying the technique of Gold, Brachman and Anderson to the teachings of Kennewick would have yielded predictable results because it would have been within the skill of those in the art to use Kennewick to process a natural language utterance and to route the utterance from Kennewick together with the context (music being played) of Anderson to the system of Gold (*id.* at 14–15).

The difficulty with the Examiner's rejection is that the Examiner does not explain why, and we fail to see why, a person of ordinary skill in the art would have been led by the cited references to identify the content to be dedicated to a recipient based on the context of the user's computer terminal. It is not enough to identify individual features of a claimed invention in the prior art; the Examiner must also provide some articulated reasoning with some rational underpinning to explain why a person of ordinary skill in the art would have combined the cited references to arrive at the claimed invention. *See KSR Int'l Co. v. Teleflex, Inc.*, 550 U.S. 398, 418 (2007).

Amazon/VB Assets
Exhibit 1003
Page 87

We are not persuaded that the Examiner has established a prima facie case of obviousness with respect to claim 22. Therefore, we do not sustain the Examiner's rejection of claim 22 under 35 U.S.C. § 103(a).

Claim 21 is substantially similar to claim 22 (i.e., claim 21 recites that the content to be dedicated is identified based on the natural language processing and the interaction of the user with the media item) and was rejected based on the same rationale applied with respect to claim 22. Therefore, we do not sustain the Examiner's rejection of claim 21 under 35 U.S.C. § 103(a) for substantially the same reasons set forth above with respect to claim 22.

DECISION

The Examiner's rejections of claims 1, 5–11, 15–20, and 24–26 under 35 U.S.C. § 103(a) are affirmed.

The Examiner's rejections of claims 21–23 under 35 U.S.C. § 103(a) are reversed.

No time period for taking any subsequent action in connection with this appeal may be extended under 37 C.F.R. § 1.136(a)(1)(iv).

AFFIRMED-IN-PART

Amazon/VB Assets
Exhibit 1003
Page 88

RECEIVED
CENTRAL FAX CENTER

MAR 0 9 2016

The United States Patent and Trademark Office
PATENT TRIAL AND APPEAL BOARD

SHEPPARD MULLIN RICHTER &
HAMPTON LLP
12275 EL CAMINO REAL
SUITE 200
SAN DIEGO, CA 92130

Appeal No:	2014-004,731
Appellant:	Mike Kennewick, et al.
Application No:	12/943,699
Hearing Room:	B
Hearing Docket:	B
Hearing Date:	Tuesday, April 05, 2016
Hearing Time:	01:00 PM
Location:	**Madison Building - East Wing** **600 Dulany Street, 9th Floor** **Alexandria, Virginia 22313-1450**

NOTICE OF HEARING - ALEXANDRIA, VIRGINIA
RESPONSE REQUIRED WITHIN 21 DAYS

Your attention is directed to 37 CFR § 41.47. The above identified appeal will be heard by the Patent Trial and Appeal Board on the date indicated. Hearings will commence at the time set, and as soon as the argument in one appeal is concluded, the succeeding appeal will be taken up. **The time allowed for argument is 20 minutes** unless additional time is requested and approved before the argument commences. **If the application involved in this appeal has been published, the hearing will be open to the public.**

CONFIRMATION OF ATTENDANCE OR WAIVER OF THE HEARING IS REQUIRED WITHIN 21 DAYS OF THE MAILING DATE OF THIS NOTICE. Failure to respond may subject Appellant(s) to waiver of the oral hearing. If Appellant is no longer interested in having an oral hearing, Appellant must still file a waiver of oral hearing with the Board. This allows the panel to promptly act on the appeal without waiting for the oral hearing date.

Confirmation or waiver of the hearing should be indicated by completing the form below and returning it to the Board. This form may be filed with the Board by any one of the following three alternative methods:

1. Via the USPTO Electronic Filing System (EFS) at http://www.uspto.gov/patents/process/file/efs/
2. ***PREFERRED***: Facsimile transmitted to: The USPTO Central fax number (official copy): **(571) 273-8300** ***and*** the PTAB Hearing fax number (courtesy copy): **(571) 273-9797.**
3. By mail at the PTAB mailing address: Patent Trial and Appeal Board
   United States Patent and Trademark Office
   P.O. BOX 1450
   Alexandria, Virginia 22313-1450

In all communications relating to this appeal, please identify the appeal by its number.

RECEIVED
CENTRAL FAX CENTER

MAR 0 9 2016

PLEASE ELECT ONE OF THE OPTIONS BELOW:

(●) I ELECT AN IN-PERSON HEARING AND CONFIRM - ATTENDANCE IN **ALEXANDRIA, VIRGINIA** *(EFS-Web selection: Confirmation of Hearing by Appellant)*

( ) I ELECT A TELEPHONIC HEARING - ATTENDANCE CONFIRMED *(EFS-Web selection: Confirmation of Hearing by Appellant)*

( ) I ELECT A VIDEO HEARING - ATTENDANCE CONFIRMED *(EFS-Web selection: Confirmation of Hearing by Appellant)*

( ) I ELECT TO WAIVE HEARING ATTENDANCE *(EFS-Web selection: Waiver of Hearing by Appellant)*

To aid the oral hearings staff in scheduling hearing rooms, please indicate the total number of participating and observing attendees if more than three are expected: _____
To aid the judges in determining whether any conflicts exist that may require a recusal, please list in the 'Comments' section the names of any additional person(s) who will be participating in the oral hearing. (Upon arrival at the **Madison East Building, 9th floor**, all persons presenting arguments must sign in at the Usher's desk.)

Comments/Special Requests:

James G. Gatto	32,694
Typed or Printed Name of Attorney/Agent/Appellant	Registration No.
James Gatto	March 9, 2016
Signature of Attorney/Agent/Appellant	Date

The 'Hearings' section of the PTAB webpage http://www.uspto.gov/patents-application-process/patent-trial-and-appeal-board/hearings provides additional information about oral hearings.

Please direct other inquiries to the PTAB Hearings Clerk at 571-272-9797.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/943,699	11/10/2010	Mike Kennewick	45HV-218255	4185

133759 7590 02/19/2016

Sheppard Mullin Richter & Hampton LLP
12275 El Camino Real, Suite 200
San Diego, CA 92130

EXAMINER	
HALE, TIM B	

ART UNIT	PAPER NUMBER
3685	

NOTIFICATION DATE	DELIVERY MODE
02/19/2016	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

dmipdocketing@sheppardmullin.com

The United States Patent and Trademark Office
PATENT TRIAL AND APPEAL BOARD

SHEPPARD MULLIN RICHTER &
HAMPTON LLP
12275 EL CAMINO REAL
SUITE 200
SAN DIEGO, CA 92130

Appeal No:	2014-004,731
Appellant:	Mike Kennewick, et al.
Application No:	12/943,699
Hearing Room:	B
Hearing Docket:	B
Hearing Date:	Tuesday, April 05, 2016
Hearing Time:	01:00 PM
Location:	**Madison Building - East Wing** **600 Dulany Street, 9th Floor** **Alexandria, Virginia 22313-1450**

NOTICE OF HEARING - ALEXANDRIA, VIRGINIA
RESPONSE REQUIRED WITHIN 21 DAYS

Your attention is directed to 37 CFR § 41.47. The above identified appeal will be heard by the Patent Trial and Appeal Board on the date indicated. Hearings will commence at the time set, and as soon as the argument in one appeal is concluded, the succeeding appeal will be taken up. **The time allowed for argument is 20 minutes** unless additional time is requested and approved before the argument commences. **If the application involved in this appeal has been published, the hearing will be open to the public.**

CONFIRMATION OF ATTENDANCE OR WAIVER OF THE HEARING IS REQUIRED WITHIN **21 DAYS** OF THE MAILING DATE OF THIS NOTICE. Failure to respond may subject Appellant(s) to waiver of the oral hearing. If Appellant is no longer interested in having an oral hearing, Appellant must still file a waiver of oral hearing with the Board. This allows the panel to promptly act on the appeal without waiting for the oral hearing date.

Confirmation or waiver of the hearing should be indicated by completing the form below and returning it to the Board. This form may be filed with the Board by any one of the following three alternative methods:

1. Via the USPTO Electronic Filing System (EFS) at http://www.uspto.gov/patents/process/file/efs/
2. ***PREFERRED***: Facsimile transmitted to: The USPTO Central fax number (official copy): **(571) 273-8300** ***and*** the PTAB Hearing fax number (courtesy copy): **(571) 273-9797.**
3. By mail at the PTAB mailing address: Patent Trial and Appeal Board
   United States Patent and Trademark Office
   P.O. BOX 1450
   Alexandria, Virginia 22313-1450

In all communications relating to this appeal, please identify the appeal by its number.

PLEASE ELECT ONE OF THE OPTIONS BELOW:

( ) I ELECT AN IN-PERSON HEARING AND CONFIRM - ATTENDANCE IN **ALEXANDRIA, VIRGINIA** *(EFS-Web selection: Confirmation of Hearing by Appellant)*

( ) I ELECT A TELEPHONIC HEARING - ATTENDANCE CONFIRMED *(EFS-Web selection: Confirmation of Hearing by Appellant)*

( ) I ELECT A VIDEO HEARING - ATTENDANCE CONFIRMED *(EFS-Web selection: Confirmation of Hearing by Appellant)*

( ) I ELECT TO WAIVE HEARING ATTENDANCE *(EFS-Web selection: Waiver of Hearing by Appellant)*

To aid the oral hearings staff in scheduling hearing rooms, please indicate the total number of participating and observing attendees if more than three are expected: _____

To aid the judges in determining whether any conflicts exist that may require a recusal, please list in the 'Comments' section the names of any additional person(s) who will be participating in the oral hearing. (Upon arrival at the **Madison East Building, 9th floor**, all persons presenting arguments must sign in at the Usher's desk.)

Comments/Special Requests:

_____

_____

_____ _____

Typed or Printed Name of Attorney/Agent/Appellant Registration No.

_____ _____

Signature of Attorney/Agent/Appellant Date

The 'Hearings' section of the PTAB webpage http://www.uspto.gov/patents-application-process/patent-trial-and-appeal-board/hearings provides additional information about oral hearings.

Please direct other inquiries to the PTAB Hearings Clerk at 571-272-9797.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
12/943,699	11/10/2010	Mike Kennewick	090611-0384422

CONFIRMATION NO. 4185

909
Pillsbury Winthrop Shaw Pittman, LLP (NV)
PO Box 10500
McLean, VA 22102

POWER OF ATTORNEY NOTICE

OC000000078724614

Date Mailed: 11/19/2015

NOTICE REGARDING CHANGE OF POWER OF ATTORNEY

This is in response to the Power of Attorney filed 11/09/2015.

- The Power of Attorney to you in this application has been revoked by the assignee who has intervened as provided by 37 CFR 3.71. Future correspondence will be mailed to the new address of record(37 CFR 1.33).

Questions about the contents of this notice and the requirements it sets forth should be directed to the Office of Data Management, Application Assistance Unit, at **(571) 272-4000** or **(571) 272-4200** or **1-888-786-0101**.

/qtran/

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
12/943,699	11/10/2010	Mike Kennewick	090611-0384422

CONFIRMATION NO. 4185

133759
Sheppard Mullin Richter & Hampton LLP
12275 El Camino Real, Suite 200
San Diego, CA 92130

POA ACCEPTANCE LETTER

OC000000078724649

Date Mailed: 11/19/2015

NOTICE OF ACCEPTANCE OF POWER OF ATTORNEY

This is in response to the Power of Attorney filed 11/09/2015.

The Power of Attorney in this application is accepted. Correspondence in this application will be mailed to the above address as provided by 37 CFR 1.33.

Questions about the contents of this notice and the requirements it sets forth should be directed to the Office of Data Management, Application Assistance Unit, at **(571) 272-4000** or **(571) 272-4200** or **1-888-786-0101**.

/qtran/

PTO/SB/80 (11-08)
Approved for use through 11/30/2011. OMB 0651-0035
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

POWER OF ATTORNEY TO PROSECUTE APPLICATIONS BEFORE THE USPTO

I hereby revoke all previous powers of attorney given in the application identified in the attached statement under 37 CFR 3.73(b).

I hereby appoint:

☑ Practitioners associated with the Customer Number: 133759

OR

☐ Practitioner(s) named below (if more than ten patent practitioners are to be named, then a customer number must be used):

Name	Registration Number	Name	Registration Number

as attorney(s) or agent(s) to represent the undersigned before the United States Patent and Trademark Office (USPTO) in connection with any and all patent applications assigned only to the undersigned according to the USPTO assignment records or assignment documents attached to this form in accordance with 37 CFR 3.73(b).

Please change the correspondence address for the application identified in the attached statement under 37 CFR 3.73(b) to:

☑ The address associated with Customer Number: 133759

OR

☐ Firm or Individual Name			
Address			
City		State	Zip
Country			
Telephone		Email	

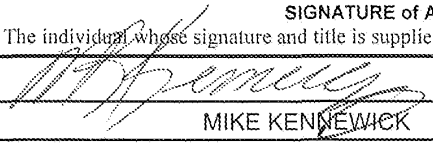
Assignee Name and Address:

VOICEBOX TECHNOLOGIES CORPORATION
11980 NE 24th Street, Suite 100
Bellevue, WA 98005

A copy of this form, together with a statement under 37 CFR 3.73(b) (Form PTO/SB/96 or equivalent) is required to be filed in each application in which this form is used. The statement under 37 CFR 3.73(b) may be completed by one of the practitioners appointed in this form if the appointed practitioner is authorized to act on behalf of the assignee, and must identify the application in which this Power of Attorney is to be filed.

SIGNATURE of Assignee of Record

The individual whose signature and title is supplied below is authorized to act on behalf of the assignee

Signature		Date
Name	MIKE KENNEWICK	Telephone
Title	CEO OF VOICEBOX TECHNOLOGIES CORPORATION	

This collection of information is required by 37 CFR 1.31, 1.32 and 1.33. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 3 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Electronic Acknowledgement Receipt	
EFS ID:	24023479
Application Number:	12943699
International Application Number:	
Confirmation Number:	4185
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Customer Number:	909
Filer:	James G. Gatto/Todd Yarborough
Filer Authorized By:	James G. Gatto
Attorney Docket Number:	090611-0384422
Receipt Date:	09-NOV-2015
Filing Date:	10-NOV-2010
Time Stamp:	14:01:23
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		VoiceBox_373b_POA_45HV_218255.pdf	736081 f68f703acf7fe794f40dc32dd8bba174a3723dda	yes	4

	Multipart Description/PDF files in .zip description		
	Document Description	**Start**	**End**
	Assignee showing of ownership per 37 CFR 3.73	1	3
	Power of Attorney	4	4
Warnings:			
Information:			
	Total Files Size (in bytes):	736081	

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/AIA/96 (08-12)
Approved for use through 01/31/2013. OMB 0651-0031
U.S. Patent and Trademark Office;U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of1995, no persons arerequired to respond to a collection of information unless it displays a valid OMB control number.

STATEMENT UNDER 37 CFR 3.73(c)

Applicant/Patent Owner: Mike KENNEWICK et al.

Application No./Patent No.: 12/943,699 Filed/Issue Date: November 10, 2010

Titled: SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE

Voicebox Technologies Corporation, a corporation

(Name of Assignee) (Type of Assignee, e.g., corporation, partnership, university, government agency, etc.)

states that, for the patent application/patent identified above, it is (choose **one** of options 1, 2, 3 or 4 below):

1. ☑ The assignee of the entire right, title, and interest.

2. ☐ An assignee of less than the entire right, title, and interest (check applicable box):

   ☐ The extent (by percentage) of its ownership interest is _____%. Additional Statement(s) by the owners holding the balance of the interest must be submitted to account for 100% of the ownership interest.

   ☐ There are unspecified percentages of ownership. The other parties, including inventors, who together own the entire right, title and interest are:

   Additional Statement(s) by the owner(s) holding the balance of the interest must be submitted to account for the entire right, title, and interest.

3. ☐ The assignee of an undivided interest in the entirety (a complete assignment from one of the joint inventors was made). The other parties, including inventors, who together own the entire right, title, and interest are:

   Additional Statement(s) by the owner(s) holding the balance of the interest must be submitted to account for the entire right, title, and interest.

4. ☐ The recipient, via a court proceeding or the like (*e.g.*, bankruptcy, probate), of an undivided interest in the entirety (a complete transfer of ownership interest was made). The certified document(s) showing the transfer is attached.

The interest identified in option 1, 2 or 3 above (not option 4) is evidenced by either (choose **one** of options A or B below):

A. ☐ An assignment from the inventor(s) of the patent application/patent identified above. The assignment was recorded in the United States Patent and Trademark Office at Reel _____, Frame _____, or for which a copy thereof is attached.

B. ☑ A chain of title from the inventor(s), of the patent application/patent identified above, to the current assignee as follows:

1. From: Mike KENNEWICK et al. To: Voicebox Technologies, Inc.

   The document was recorded in the United States Patent and Trademark Office at

   Reel 025813, Frame 0575, or for which a copy thereof is attached.

2. From: Voicebox Technologies, Inc. To: Voicebox Technologies Corporation

   The document was recorded in the United States Patent and Trademark Office at

   Reel 032621, Frame 0087, or for which a copy thereof is attached.

This collection of information is required by37 CFR3.73(b). The information is required toobtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentialityis governed by35 U.S.C. 122and 37 CFR1.11 and1.14. Thiscollection is estimated to take 12 minutes to complete, including gathering, preparing, and submittingthe completed application form to the USPTO.Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent tothe Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS.**SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

PTO/AIA/96 (08-12)
Approved for use through 01/31/2013. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

STATEMENT UNDER 37 CFR 3.73(c)

3. From: _____ To: _____

The document was recorded in the United States Patent and Trademark Office at

Reel _____, Frame _____, or for which a copy thereof is attached.

4. From: _____ To: _____

The document was recorded in the United States Patent and Trademark Office at

Reel _____, Frame _____, or for which a copy thereof is attached.

5. From: _____ To: _____

The document was recorded in the United States Patent and Trademark Office at

Reel _____, Frame _____, or for which a copy thereof is attached.

6. From: _____ To: _____

The document was recorded in the United States Patent and Trademark Office at

Reel _____, Frame _____, or for which a copy thereof is attached.

☐ Additional documents in the chain of title are listed on a supplemental sheet(s).

☐ As required by 37 CFR 3.73(c)(1)(i), the documentary evidence of the chain of title from the original owner to the assignee was, or concurrently is being, submitted for recordation pursuant to 37 CFR 3.11.

[NOTE: A separate copy (i.e., a true copy of the original assignment document(s)) must be submitted to Assignment Division in accordance with 37 CFR Part 3, to record the assignment in the records of the USPTO. See MPEP 302.08]

The undersigned (whose title is supplied below) is authorized to act on behalf of the assignee.

/James G. Gatto/	November 9, 2015
Signature	Date
James G. Gatto	32,694
Printed or Typed Name	Title or Registration Number

Privacy Act Statement

The **Privacy Act of 1974 (P.L. 93-579)** requires that yoube given certain informationin connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, pleasebe advised that: (1) the general authority forthe collection of thisinformation is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and(3) the principal purpose forwhich the information isused by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent applicationor patent. If you do not furnish the requested information,the U.S. Patent and Trademark Office may not be able to process and/or examineyour submission,which may result in termination of proceedings or abandonment of the applicationor expiration of the patent.

The informationprovided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the informationin order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an InternationalApplication filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (*i.e.*, GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, arecord may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from thissystem of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Inventor(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	A	8,200,485	06/12/2012	Lee et al.	704	231	
	B	8,983,839	03/17/2015	Kennewick *et al.*	704	254	
	C	9,015,049	04/21/2015	Baldwin *et al.*			
	D	2001/0039492	11/08/2001	Nemoto	704	231	
	E	2005/0080632	04/14/2005	Endo *et al.*	704	277	
	F	2008/0147396	06/19/2008	Wang	704	246	
	G	2008/0162471	07/03/2008	Bernard	707	5	
	H	2010/0094707	04/15/2010	Freer	705	14.54	
	I	2011/0022393	01/27/2011	Waller *et al.*	704	270	
	J	2015/0066627	03/05/2015	Freeman *et al.*	705	14.41	
	K	2015/0073910	03/12/2015	Kennewick *et al.*	705	14.58	
	L	2015/0095159	04/02/2015	Kennewick et al.	705	14.58	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

405243468v1

Electronic Patent Application Fee Transmittal

Application Number:	12943699
Filing Date:	10-Nov-2010
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Filer:	Ngai Bing Zhang/Susan Trader
Attorney Docket Number:	090611-0384422

Filed as Small Entity

Filing Fees for Utility under 35 USC 111(a)

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
Submission- Information Disclosure Stmt	2806	1	90	90
		Total in USD ($)		**90**

Electronic Acknowledgement Receipt	
EFS ID:	22307623
Application Number:	12943699
International Application Number:	
Confirmation Number:	4185
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Customer Number:	909
Filer:	Ngai Bing Zhang/Susan Trader
Filer Authorized By:	Ngai Bing Zhang
Attorney Docket Number:	090611-0384422
Receipt Date:	11-MAY-2015
Filing Date:	10-NOV-2010
Time Stamp:	14:51:07
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	$90
RAM confirmation Number	1209
Deposit Account	033975
Authorized User	TRADER, SUSAN L.
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Transmittal Letter	090611-0384422_SIDSTransmittal.pdf	523930 aaffcf954758815b9b14841b1a90da3404106743	no	3
Warnings:					
Information:					
2	Information Disclosure Statement (IDS) Form (SB08)	090611-0384422_PTO1449.pdf	542507 04483f7d375639ecf2c9c0033a01522a31a83d4b	no	1
Warnings:					
Information:					
This is not an USPTO supplied IDS fillable form					
3	Fee Worksheet (SB06)	fee-info.pdf	30919 e9c755e21fad26173858705a1db01c644082ff8f	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	1097356		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

INVENTORS :	Mike KENNEWICK *et al.*	CONFIRMATION NO.:	4185
SERIAL NUMBER :	12/943,699	EXAMINER:	Tim B. Hale
FILING DATE :	November 10, 2010	ART UNIT:	3685
FOR :	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE		

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT

Commissioner for Patents
P. O. Box 1450
Alexandria, VA 22313-1450

Sir:

Pursuant to the duty of disclosure under 37 C.F.R. §§ 1.56, 1.97 and 1.98, Applicants hereby make of record in the above-identified application the documents listed on the attached Form PTO-1449.

Applicants reserve the right to establish the patentability of the claimed invention over any of the information provided herewith, and/or to prove that this information may not be prior art, and/or to prove that this information may not be enabling for the teachings purportedly offered.

This Supplemental Information Disclosure Statement is being filed more than three months after the filing date of a national application other than a continued prosecution

Amazon/VB Assets
Exhibit 1003
Page 107

application under 37 C.F.R. § 1.53(d), or more than three months after the date of entry of the national stage as set forth in 37 C.F.R. § 1.491 in an international application, and after the mailing date of any of a Final Action under 37 C.F.R. § 1.113, a Notice of Allowance under 37 C.F.R. § 1.311, or an action that otherwise closes prosecution in the application, but on or before payment of the Issue Fee. 37 C.F.R. § 1.97(d).

Documents E and G-I were cited by the U.S. Patent Office in co-pending Application Serial No. 14/278,645, filed May 15, 2014, in an Office Action dated February 9, 2015. Documents A, D, and F were cited by the U.S. Patent Office in co-pending Application Serial No. 14/269,545, filed May 5, 2014, in an Office Action dated February 24, 2015. I hereby certify that no item of information contained in this Supplemental Information Disclosure Statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in this Supplemental Information Disclosure Statement was known to any individual designated in 37 C.F.R. § 1.56(c) more than three months prior to the filing of this Supplemental Information Disclosure Statement. 37 C.F.R. § 1.97(e)(2).

Please charge the $90.00 fee [37 C.F.R. § 1.17(p)] to Deposit Account No. 03-3975. In addition, please charge any fee deficiencies, or credit any overpayment, to Deposit Account No. 03-3975 as needed to ensure consideration of the Supplemental Information Disclosure Statement.

It is respectfully requested that the Examiner initial and return a copy of the enclosed PTO-1449.

Respectfully submitted,

PILLSBURY WINTHROP SHAW PITTMAN LLP

By /Ngai B. Zhang/
Ngai B. Zhang
Registration No.: 65,473

Customer No.
00909

P. O. Box 10500
McLean, Virginia 22102
Telephone: (703) 770-7900
Telefax: (703) 770-7901

Dated: May 11, 2015

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/943,699	11/10/2010	Mike Kennewick	090611-0384422	4185

909 7590 04/01/2014
Pillsbury Winthrop Shaw Pittman, LLP (NV)
PO Box 10500
McLean, VA 22102

EXAMINER	
HALE, TIM B	
ART UNIT	PAPER NUMBER
3685	

NOTIFICATION DATE	DELIVERY MODE
04/01/2014	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

docket_ip@pillsburylaw.com
heather.marthers@pillsburylaw.com

United States Patent and Trademark Office

Under Secretary of Commerce for Intellectual Property and
Director of the United States Patent and Trademark Office
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

PILLSBURY WINTHROP SHAW PITTMAN, LLP (NV)
PO BOX 10500
MCLEAN, VA 22102

Appeal No: 2014-004731
Application: 12/943,699
Appellant: Mike Kennewick et al.

Patent Trial and Appeal Board Docketing Notice

Application 12/943,699 was received from the Technology Center at the Board on March 03, 2014 and has been assigned Appeal No: 2014-004731.

In all future communications regarding this appeal, please include both the application number and the appeal number.

The mailing address for the Board is:

PATENT TRIAL and APPEAL BOARD
UNITED STATES PATENT AND TRADEMARK OFFICE
P.O. BOX 1450
ALEXANDRIA, VIRGINIA 22313-1450

Telephone inquiries can be made by calling 571-272-9797 and referencing the appeal number listed above.

By order of the Patent Trial and Appeal Board.

CM

PTO/SB/32 (07-09)
Approved for use through 01/31/2013. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

REQUEST FOR ORAL HEARING
BEFORE
THE PATENT TRIAL AND APPEAL BOARD

Docket Number (Optional)
090611-0384422

I hereby certify that this correspondence is being facsimile transmitted to the USPTO, EFS-Web transmitted to the USPTO, or deposited with the United States Postal Service as first class mail with sufficient postage in an envelope addressed to "Commissioner for Patents, P. O. Box 1450, Alexandria VA 22313-1450" [37 CFR 1.8(a)] on _____.

Signature _____

Typed or printed name _____

In re Application of: Mike KENNEWICK, et al.

Application Number: 12/943,699 | Filed: November 10, 2010

For: SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE

Art Unit: 3685 | Examiner: Tim B. Hale

Applicant hereby requests an oral hearing before the Patent Trial and Appeal Board in the appeal of the above-identified application.

The fee for this Request for Oral Hearing is (37 CFR 41.20(b)(3)) $ 1,300.00

☐ Applicant claims small entity status. See 37 CFR 1.27. Therefore, the fee shown above is reduced by half, and the resulting fee is: $ 650.00

☐ A check in the amount of the fee is enclosed.

☐ Payment by credit card. Form PTO-2038 is attached.

☐ The Director has already been authorized to charge fees in this application to a Deposit Account. I have enclosed a duplicate copy of this sheet.

☐ The Director is hereby authorized to charge any fees which may be required, or credit any overpayment to Deposit Account No. 033975 .

☐ Payment made via EFS-Web.

☐ A petition for an extension of time under 37 CFR 1.136(b) (PTO/SB/23) is enclosed. For extensions of time in reexamination proceedings, see 37 CFR 1.550.

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

I am the

☐ applicant.

☐ attorney or agent of record. Registration number 65,473

☐ attorney or agent acting under 37 CFR 1.34. Registration number if acting under 37 CFR 1.34. _____

/Ngai B. Zhang/
Signature

Ngai B. Zhang
Typed or printed name

February 27, 2014
Date

(703) 770-7532
Telephone number

NOTE: This form must be signed in accordance with 37 CFR 1.33. See 37 CFR 1.4 for signature requirements and certifications. Submit multiple forms if more than one signature is required, see below*.

☒ *Total of 1 forms are submitted.

This collection of information is required by 37 CFR 41.20(b)(3). The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11, 1.14 and 41.6. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

IN THE UNITED STATES PATENT & TRADEMARK OFFICE

IN RE PATENT APPLICATION OF:	Mike KENNEWICK et al.
SERIAL NO.:	12/943,699
FILING DATE:	November 10, 2010
ATTORNEY DOCKET NO.:	090611-0384422
CONFIRMATION NO.:	4185
EXAMINER:	Tim B. Hale
ART UNIT:	3685
FOR:	System and Method for Providing A Natural Language Content Dedication Service

APPELLANTS' REPLY BRIEF UNDER 37 C.F.R. § 41.41

MAIL STOP APPEAL BRIEF – PATENTS
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

This Reply Brief is being timely filed within two months of the Examiner's Answer (hereinafter "Answer") mailed **December 27, 2013**. This Brief responds to the new points raised by the Examiner.

I. STATUS OF CLAIMS

The status of the claims is as follows:

Pending:	Claims 1, 5-11, and 15-26 are pending.
Cancelled:	Claims 2-4 and 12-14 are canceled.
Rejected:	Claims 1, 5-11, and 15-26 stand rejected.
Allowed:	No claims have been allowed.
On Appeal:	Claims 1, 5-11, and 15-26 are appealed.

II. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

1. Claims 1, 7, 9, 10, 11, 17, 19, 20, and 26 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over U.S. Patent Application Pub. No. 2004/0193420 by Kennewick (hereinafter "Kennewick") in view of U.S. Patent Application Publication No. 2002/0032752 to Gold (hereinafter "Gold") and U.S. Patent No. 6,704,576 to Brachman (hereinafter "Brachman");
2. Dependent claims 6 and 16 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick, Gold, and Brachman, further in view of U.S. Patent Application Publication No. 2008/0140385 to Mahajan ("Mahajan");
3. Dependent claims 5, 15, and 23-25 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick, Gold, and Brachman, further in view of U.S. Patent Application Publication No. 2010/0036967 to Caine ("Caine");
4. Dependent claims 8 and 18 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick, Gold, and Brachman, further in view of U.S. Patent No. 5,892,900 to Ginter ("Ginter"); and
5. Dependent claims 21 and 22 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick, Gold, and Brachman, further in view of U.S. Patent Application Publication No. 2009/0276700 to Anderson ("Anderson").

III. RESPONSE TO EXAMINER'S ARGUMENTS

A. OVERVIEW

Appellants have the burden on appeal to the Board to demonstrate error in the Examiner's position. *See In re Kahn*, 441 F.3d 977, 985-86 (Fed. Cir. 2006) ("On appeal to the Board, an applicant can overcome a rejection [under § 103] by showing insufficient evidence of *prima facie* obviousness or by rebutting the *prima facie* case with evidence of secondary indicia of nonobviousness.") (quoting *In re Rouffet*, 149 F.3d 1350, 1355 (Fed. Cir. 1998)).

Amazon/VB Assets
Exhibit 1003
Page 114

In the Answer, just as in the 03/05/13 Final Office Action (hereinafter "Final Action"), the Examiner has failed to meet the requisite burden of establishing a *prima facie* case of obviousness for each of the alleged rejections.

In particular, even assuming arguendo that it were deemed legally proper to combine the relied-upon references of Kennewick, Gold, and Brachman for the rejection of each of independent claims 1 and 11 (which Appellants do not concede), the Examiner has failed to demonstrate that the three reference combination teaches or suggests all of the claim recitations of *at least* claims 1 and 11.

In addition, the Examiner has failed to demonstrate that various other combinations of the relied-upon references teach or suggest the further claim recitations of the dependent claims.

Based on the arguments presented herein, as well as in the Appeal Brief, it is clear that the Examiner has failed to meet the requisite burden of establishing a *prima facie* case of obviousness. The rejection of pending claims 1, 5-11, and 15-26 under 35 U.S.C. § 103(a) is therefore improper and must be reversed.

B. THE EXAMINER HAS FAILED TO ESTABLISH THAT THE ALLEGED COMBINATION OF KENNEWICK, GOLD, AND BRACHMAN TEACHES ALL OF THE CLAIM FEATURES OF AT LEAST INDEPENDENT CLAIMS 1 AND 11

Assuming arguendo that it were deemed legally proper to combine Kennewick, Gold, and Brachman in the manner alleged by the Examiner (which Appellants do not concede), the three references, even if combined, fail to disclose, teach, or suggest all of the claim recitations of *at least* independent claims 1 and 11. The Examiner has failed to demonstrate otherwise.

For example, independent claim 1 recites, *inter alia*, the features of:

> **identify, based on** a natural language processing of the one or more words or phrases of **the first utterance** [that includes a natural language utterance]**, content to dedicate to a recipient**; and
>
> initiate, based on the identified content, a dedication **to the recipient**.
>
> [Emphasis added].

Independent claim 11 includes similar recitations.

In particular, as claimed, both the content to be dedicated and the recipient are identified based on a natural language processing of one or more words or phrases of a first utterance.

As disclosed, for example, in paragraphs 105 and 106 of the Specification, an example of identifying content to dedicate to a recipient based on a natural language processing of words or phrases of a first utterance may involve: (1) processing an utterance "Buy that song and send it to Michael;" (2) determining that the utterance includes the words or phrases "Buy," "that song," "send it," and "to Michael;" and (3) utilizing the words or phrases as criteria to identify the content and the recipient for a dedication where "Buy" and "send it" may indicate a content dedication request, "that song" may indicate the content to dedicate, and "to Michael" may indicate the recipient to which the content is to be dedicated. The dedication may then be initiated for "Michael" (e.g., the recipient identified, along with the content to be dedicated, based on a natural language processing of one or more words or phrases of a first utterance).

1. THE EXAMINER'S NEW GROUND OF REJECTION IN HIS ANSWER FAILS TO OVERCOME THE DEFICIENCIES OF THE ALLEGED REJECTION IN THE FINAL ACTION

Perhaps in a telling sign of the weakness of the alleged rejection, the Examiner presents additional arguments in his Answer based on a new ground of rejection. Specifically, the Examiner now alleges that Kennewick and Gold teaches identification of content to be dedicated to a recipient based on a natural language processing of words or phrases of a first utterance.

However, to the contrary, the Examiner conceded in the Final Action that Kennewick and Gold fail to teach the claimed identification of content to be dedicated to a recipient. [Final Action, pgs. 4-5]. For example, on pages 4-5 of the Final Action, the Examiner states that:

> **Neither [Kennwick] nor Gold teach**, but Brachman teaches: "identify . . . , based on [] a natural language processing of the one or more words or phrases of the first utterance, content [to be dedicated to a recipient]."
>
> [Emphasis added, and original emphasis removed].

Kennewick teaches the use of natural language processing to formulate commands. [Kennewick, ¶ 41]. Gold teaches that a user may take a number of actions to dedicate content to a recipient, including retrieving a song list, selecting a song from the song list, and then providing a recipient's email address to send the song to the recipient. [Gold, ¶ 36]. Because Gold's dedication process requires a user to manually handle each of the foregoing actions, the Examiner recognized in the Final Action (based on his admission) that the combination of these two references fail to teach or suggest that one or more words or phrases of a first utterance may be processed to identify content to be dedicated and a recipient of the content, much less *at least* the foregoing claim recitation of independent claim 1 and 11.

For *at least* the foregoing reasons, the new ground of rejection of *at least* independent claims 1 and 11 under 35 U.S.C. § 103(a) is improper and should be reversed.

2. THE EXAMINER'S ALLEGED REJECTION IN THE FINAL ACTION IS IMPROPER AND SHOULD BE REVERSED

As explained in the Appeal Brief, on pages 7-8, the Examiner relies on a third reference (Brachman) in an effort to cure the admitted deficiencies of Kennewick and Gold. [*See also* Final Action, pgs. 4-5].

In the Answer, on page 7, the Examiner argues that Gold and Brachman in combination together teaches identification of content to be dedicated to a recipient based on a natural language processing of words or phrases of a first utterance. In particular, the Examiner states that:

> Gold teaches: a sending reviewing a song list and selecting a song to dedicate to a recipient . . .
>
> Brachman teaches: a user may request data using a natural language search and a voice means may be used to enter the use's natural language request.

Nevertheless, as explained in the Appeal Brief, Brachman (at best) teaches natural language search queries (as opposed to a natural language processing that identifies a particular content and/or a recipient to which the content is to be dedicated).

For example, in Brachman, "a user may request data using a natural language search (i.e., typing 'I really like Jazz and artists like the Beatles, but I do not like opera.'). Instead of the

web, a voice means may be used to enter the user's natural language request." [Brachman, col. 4, lines 23-26]. As such, while the natural language search in Brachman may produce a set of content results, a user would then need to manually look through the results to identify a particular content item of interest. Thus, the natural language search does not identify either (i) the content to dedicate or (ii) the recipient to which the content is to be dedicated.

Thus, Brachman does not teach or suggest the claimed features of identifying, based on a natural language processing of one or more words or phrases of a first utterance, content to dedicate to a recipient.

Brachman therefore fails to cure the admitted deficiencies of Kennewick and Gold. Thus, the rejection of *at least* independent claims 1 and 11 under 35 U.S.C. § 103(a) is improper, and should be reversed, for *at least* the foregoing reasons.

The dependent claims are allowable because they each depend from an allowable independent claim, as well as for the further features they recite.

C. DEPENDENT CLAIMS 6 AND 16 ARE PATENTABLE OVER THE ALLEGED COMBINATION OF KENNEWICK, GOLD, BRACHMAN, AND MAHAJAN

Dependent claims 6 and 16 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick, Gold, and Brachman, further in view of Mahajan. [Final Action, pg. 6]. For *at least* the reasons set forth above with respect to independent claims 1 and 11, the rejection of claims 6 and 16 under 35 U.S.C. § 103(a) is improper, and should be reversed, by virtue of their dependence on an allowable independent claim.

Moreover, dependent claims 6 and 16 are allowable for the further features they each recite. For example, claim 6 recites, *inter alia*, the feature of "annotate the content with the second utterance [that is to be associated with a dedication to a recipient]." [Emphasis added]. Claim 16 includes similar recitations.

As disclosed, for example, in paragraphs 96 and 97 of the Specification, an example of annotating content with a second utterance may involve: (1) capturing the second utterance (e.g., a dedication utterance "Dear Charlene, I was listening to this song and I thought of you.

Enjoy!") from a user that initiated a request to dedicate the content to a recipient based on a first utterance; (2) converting the second utterance into encoded audio; and (3) inserting the encoded audio of the second utterance into the content to be dedicated to the recipient.

1. MAHAJAN FAILS TO OVERCOME THE ADMITTED DEFICIENCIES OF KENNEWICK AND GOLD

In the Final Action, the Examiner relies on Mahajan in an effort to cure the admitted deficiencies of Kennewick and Gold. [Final Action, pg. 6].

In the Answer, on page 9, the Examiner argues that a presentation of a written transcription of audio/visual content with a presentation of the same audio/visual content (as taught by Mahajan) teaches the foregoing features of claims 6 and 16.[1]

However, as recited, the claimed second utterance (with which content to be dedicated is annotated) is not the same as the content. For example, independent claim 1 (from which claim 6 depends) recites, *inter alia*, the features of:

> **receive a first utterance** that includes a natural language utterance;
>
> determine, based on processing of the first utterance by a speech recognition engine, one or more words or phrases of the first utterance;
>
> **identify, based on** a natural language processing of the one or more words or phrases of **the first utterance, content to dedicate to a recipient**;
>
> initiate, based on the identified content, a dedication **to the recipient**; [and]
>
> **receive a second utterance** to be associated with the dedication.
>
> [Emphasis added].

In other words, as explained in the Appeal Brief, a presentation of a written transcription of content with the same content is not annotation of content (that is to be

---

[1] Contrary to the Examiner's own position in the Final Action, the Examiner appears to allege that Gold teaches the foregoing features of claims 6 and 16. Nevertheless, Gold simply teaches that a selected song and a voice message may be used to create a dedication. [Gold, ¶ 43; *see also* ¶¶ 31, 32]. The dedication may, for example, include a first URL to the selected song and a second URL to the voice message. [Gold, ¶ 43] The inclusion of the two URLs in a dedication does not, however, teach that annotation of content (to be dedicated to a recipient) with an

dedicated to a recipient) with the claimed second utterance (e.g., the utterance "Dear Charlene, I was listening to this song and I thought of you ...").

Mahajan therefore fails to cure the admitted deficiencies of Kennewick and Gold. Other references cited in the Final Action also do not overcome the foregoing deficiencies, and the Examiner has not established otherwise.

Thus, the rejection of dependent claims 6 and 16 under 35 U.S.C. § 103(a) is improper, and should be reversed, for *at least* the additional foregoing reasons.

D. DEPENDENT CLAIMS 5, 15, AND 23-25 ARE PATENTABLE OVER THE ALLEGED COMBINATION OF KENNEWICK, GOLD, BRACHMAN, AND CAINE

Dependent claims 5, 15, and 23-25 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick, Gold, and Brachman, further in view of Caine [Final Action, pg. 6]. For at least the reasons set forth above with respect to independent claims 1 and 11, the rejection of claims 5, 15, and 23-25 under 35 U.S.C. § 103(a) is improper, and should be reversed, by virtue of their dependence on an allowable independent claim.

Further, dependent claims 5, 15, and 23-25 are allowable for the further features they each recite. As one example, claim 23 recites, *inter alia*, the feature of:

> **provide the one or more words or phrases of the second utterance** [that is to be associated with a dedication to a recipient] **as textual annotations within metadata of the content** [that is to be dedicated to the recipient].
>
> [Emphasis added].

As disclosed, for example, in paragraphs 96 and 97 of the Specification, an example of annotating content with a second utterance may involve: (1) capturing the second utterance (e.g., a dedication utterance "Dear Charlene, I was listening to this song and I thought of you. Enjoy!") from a user that initiated a request to dedicate the content to a recipient based on a first utterance; (2) transcribing one or more words or phrases from the second utterance; and

---

utterance in the manner recited by claims 6 and 16.

(3) inserting the words or phrases of the second utterance within metadata of the content to be dedicated to the recipient.

In the Final Action, the Examiner relies on combined portions of both Gold and Caine in an effort to cure the admitted deficiencies of Kennewick. [Final Action, pg. 7].

In the Answer, on page 11, the Examiner asserts that Gold teaches adding a text message to content, and that Caine teaches adding tags as metadata to images.

Even assuming arguendo that these assertions regarding Gold and Caine are correct (which Appellants do not concede), the addition of a text message to content or tags as metadata to images do not teach that words or phrases of an utterance are provided as textual annotations within metadata of content (to be dedicated to a recipient). For example, among other issues, the text messages in Gold (that are alleged add to content) are not words or phrases of an utterance.

Gold and Caine therefore fails to cure the admitted deficiencies of Kennewick. Other references cited in the Final Action also do not overcome the foregoing deficiencies, and the Examiner has not established otherwise.

Thus, the rejection of dependent claims 5, 15, and 23-25 under 35 U.S.C. § 103(a) is improper, and should be reversed, for at least the additional foregoing reasons.

E. CLAIMS 8 AND 18 ARE PATENTABLE OVER THE ALLEGED COMBINATION OF KENNEWICK, GOLD, BRACHMAN, AND GINTER

Dependent claims 8 and 18 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick, Gold, and Brachman, further in view of Ginter. [Final Action, pg. 8]. For at least the reasons set forth above with respect to independent claims 1 and 11, the rejection of claims 8 and 18 under 35 U.S.C. § 103(a) is improper, and should be reversed, by virtue of their dependence on an allowable independent claim. In addition, the dependent claims are allowable for the further features they recite.

F. CLAIMS 21 AND 22 ARE PATENTABLE OVER THE ALLEGED COMBINATION OF KENNEWICK, GOLD, BRACHMAN, AND ANDERSON

Dependent claims 21 and 22 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick, Gold, and Brachman, further in view of Anderson. [Final Action, pgs. 8 and 9]. For at least the reasons set forth above with respect to independent claims 1 and 11, the rejection of claims 21 and 22 under 35 U.S.C. § 103(a) is improper, and should be reversed, by virtue of their dependence on an allowable independent claim.

In addition, dependent claims 21 and 22 are allowable for the further features they each recite. As an example, dependent claim 22 recites, *inter alia*, the features of:

> determine a media item that is a user is playing;
>
> **determine that the user provided the first utterance during the playing of the media item by the user; and**
>
> **identify, based on** the natural language processing and **the determination that the user provided the first utterance during the playing of the media item, the media item as the content to dedicate to the recipient.**
>
> [Emphasis added].

One example of identifying a media item as content to dedicate to a recipient based on (i) a natural language processing of one or more words or phrases of a first utterance and (ii) a determination that a user provided the first utterance during playing of the media item by the user is disclosed, for example, in paragraphs 105 and 106 of the Specification. As disclosed, an utterance "Buy that song and send it to Michael" may be processed and determined to include the words or phrases "Buy," "that song," "send it," and "to Michael." [Specification, ¶ 105]. A song playing on the user's radio device may be identified as the content to dedicate to the recipient based on the words/phrase "that song" and the determination that the user provided the utterance "Buy that song and send it to Michael" while the song was playing on the user's radio device. [Specification, ¶ 106].

In the Final Action, the Examiner relies on combined portions of both Gold and Anderson in an effort to cure the admitted deficiencies of Kennewick. [Final Action, pg. 9].

In the Answer, on page 14, the Examiner states that:

> Gold teaches the sending reviews a song list and selecting a song to dedicate to a recipient . . .
>
> Anderson teaches the context of a mobile terminal may be based on the applications currently being performed by the mobile terminal. For example, the context of a mobile terminal may be based on whether a user is . . . interacting with a media item (e.g., a photo, music track, video, etc.); context information may be received including information regarding the media content including the artist and song title that the mobile terminal is currently playing.

As explained in the Appeal Brief, a selection of a song (e.g., by clicking on the song in a song list) and a determination that a song is currently playing does not teach a determination that a user provided a first utterance during the playing of a media item by a user.

As such, Gold and Anderson also fail to teach identifying a media item as the content to dedicate to a recipient based on a natural language processing of words or phrases of a first utterance and based on the determination that a user provided the first utterance during the playing of the media item by the user.

Gold and Anderson therefore fails to cure the admitted deficiencies of Kennewick. Other references cited in the Final Action also do not overcome the foregoing deficiencies, and the Examiner has not established otherwise. Thus, the rejection of dependent claims 21 and 22 under 35 U.S.C. § 103(a) is improper, and should be reversed, for *at least* the additional foregoing reasons.

CONCLUSION

Appellants now appeal to this Honorable Board to promptly reverse all rejections of pending claims 1, 5-11, and 15-26, and issue a Decision in favor of Appellants. All of the claims are in condition for allowance.

Respectfully submitted,

Date: **February 27, 2014**

PILLSBURY WINTHROP SHAW PITTMAN LLP

By: /Ngai. B. Zhang/
Ngai B. Zhang
Registration No. 65,473

Customer No. 00909

P.O. Box 10500
McLean, Virginia 22102
Direct Dial: 703.770.7532
Main: 703.770.7900
Fax: 703.770.7901

Electronic Patent Application Fee Transmittal

Application Number:	12943699
Filing Date:	10-Nov-2010
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Filer:	Ngai Bing Zhang/Miriam Nelsen Awan
Attorney Docket Number:	090611-0384422

Filed as Small Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Request For Oral Hearing	2403	1	650	650
Appeal Forwarding Fee	2413	1	1000	1000
Post-Allowance-and-Post-Issuance:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Extension-of-Time:				
Miscellaneous:				
			Total in USD ($)	**1650**

Electronic Acknowledgement Receipt	
EFS ID:	18326212
Application Number:	12943699
International Application Number:	
Confirmation Number:	4185
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Customer Number:	909
Filer:	Ngai Bing Zhang/Miriam Nelsen Awan
Filer Authorized By:	Ngai Bing Zhang
Attorney Docket Number:	090611-0384422
Receipt Date:	27-FEB-2014
Filing Date:	10-NOV-2010
Time Stamp:	17:01:03
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	$1650
RAM confirmation Number	4299
Deposit Account	033975
Authorized User	
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: Charge any Additional Fees required under 37 C.F.R. Section 1.16 (National application filing, search, and examination fees) Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees)	

Charge any Additional Fees required under 37 C.F.R. Section 1.19 (Document supply fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.20 (Post Issuance fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.21 (Miscellaneous fees and charges)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		090611-0384422_RequestOralHearing_ReplyBrief.pdf	133039 4ac4330e35d665e4b5eac2807a397064c8baaafc	yes	13

Multipart Description/PDF files in .zip description		
Document Description	**Start**	**End**
Request for Oral Hearing	1	1
Reply Brief Filed	2	13

Warnings:					
Information:					
2	Fee Worksheet (SB06)	fee-info.pdf	32248 efc10430f4bb54eb2e1d59b7db09953018c41fe0	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	165287		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/943,699	11/10/2010	Mike Kennewick	090611-0384422	4185

909 7590 12/27/2013

Pillsbury Winthrop Shaw Pittman, LLP (NV)
PO Box 10500
McLean, VA 22102

EXAMINER	
HALE, TIM B	
ART UNIT	PAPER NUMBER
3685	

NOTIFICATION DATE	DELIVERY MODE
12/27/2013	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

docket_ip@pillsburylaw.com
heather.marthers@pillsburylaw.com

UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

BEFORE THE PATENT TRIAL AND APPEAL BOARD

Application Number: 12/943699
Filing Date: 10-Nov-2010
Appellant(s): Mike Kennewick

Ngai B. Zhang
Registration No. 65, 473
For Appellant

EXAMINER'S ANSWER

This is in response to the appeal brief filed 9/05/2013 appealing from the Office action mailed 03/05/2013.

(1) Grounds of Rejection to be Reviewed on Appeal

The examiner has no comment on the appellant's statement of the grounds of rejection to be reviewed on appeal. Every ground of rejection set forth in the Office action dated 03/05/2013 from which the appeal is taken (as modified by any advisory actions) is being maintained by the examiner.

The following grounds of rejection are applicable to the appealed claims:

Claims 1, 7, 9, 10, 11, 17, 19, 20 and 26 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent Publication 20040193420 to Kennewick in view of US Patent Publication 20020032752 to Gold and US Patent 6704576 to Brachman;

Claims 6 and 16 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent Publication 20040193420 to Kennewick in view of US Patent Publication 20020032752 to Gold US Patent 6704576 to Brachman and US Patent Publication 20080140385 to Mahajan;

Claims 5, 15, 23, 24, and 25 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent Publication 20040193420 to Kennewick in view of US Patent Publication 20020032752 to Gold, US Patent 6704576 to Brachman and US Patent Publication 20100036967 to Caine;

Claims 8 and 18 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent Publication 20040193420 to Kennewick in view of US Patent Publication 20020032752 to Gold US Patent 6704576 to Brachman and US Patent 5892900 to Ginter; and

Claims 21 and 22 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent Publication 20040193420 to Kennewick in view of US Patent Publication 20020032752 to Gold US Patent 6704576 to Brachman and US Publication 20090276700 to Anderson.

(2) Response to Arguments

A. Appellant's arguments that claims 1, 7, 9, 10, 11, 17, 19, 20, and 26 are patentable over the alleged *three-reference* combination of Kennewick, Gold, and Brachman.

Appellant argues that the rejection of claims 1, 7, 9, 10, 11, 17, 19, 20, and 26 under 35 U.S.C. § 103(a) as allegedly being unpatentable over Kennewick in view of Gold and Brachman is legally improper and should be reversed because the Examiner has failed to establish a *primafacie* case of obviousness.

Specifically, for claims 1 and 11, Appellant argues:

" . . . Brachman does not teach or suggest the claimed features of identifying based on a natural language processing of one or more words or phrases of a first utterance, content to dedicate to a recipient. Brachman therefore fails to cure the admitted deficiencies of Kennewick and Gold. The Examiner has failed to establish otherwise. "

Examiner respectfully disagrees based on two separate rationales (I and II).

I) Kennewick in view of Gold:

Kennewick teaches the base method as Kennewick teaches for each user command utterance, the system may execute a number of steps possibly

including: 1) capture the user's command through accurate speech recognition operating in a variety of real-world environments; 2) parse and interpret the command; 3) determine the domain for the command and context, invoking the proper resources, including agents as required; 4) gather required data including, device settings, and measurement data; 5) formulate device specific commands for the system or external devices; 6) route command to system, or external devices, including external devices connected to data networks; 7) receive and process results of command, including errors; and, 8) optionally, provide response to user indicating the success or failure of the command, and possibly including state information. (See [0041] of Kennewick)

Gold teaches the sender reviews the song list and selects a song to dedicate to a recipient. (See [0036] of Gold)

One of ordinary skill in the art would have recognized that applying the know technique of Gold would have yielded predictable results and resulted in an improved system. It would have been recognized that applying the technique of Gold to the teachings of Kennewick would have yielded predictable results because the level of ordinary skill in the art demonstrated by the references applied shows the ability to transfer data between the different systems (e.g., by routing the command from Kennewick to Gold as the song selection). Further,

using Kennewick to process a natural language utterance, and route the utterance as input to the system of Gold, would have been recognized by those of ordinary skill in the art as resulting in an improved system that would allow selection of content based on natural language.

II) Kennewick in view of Gold and Brachman

Gold and Brachman teach the identified claim language of "identify, based on a natural language processing of the one or more words or phrases of the first utterance [that includes a
natural language utterance], content to dedicate to a recipient."

Kennewick teaches:

"A system for providing a natural language content dedication service, comprising: [A method for providing a natural language content dedication service, comprising:]

one or more processors and one or more non-transitory computer readable mediums storing executable instructions that when executed by the one or more processors cause the processors to: [0012-0013][0014][0016][0095][0096][0097][0111]Figs. 1-3, 5-6

"receive[, by one or more processors,] a first utterance that includes a natural language utterance;[0028][0034][0038][0101]" and

"determine, [by the one or more processors] based on the processing of the first utterance by a speech recognition engine, one or more words or phrases of the first utterance;[0028-0034][0041][0123]

With regard to the identified language of "identify, based on a natural language processing of the one or more words or phrases of the first utterance [that includes a natural language utterance], content to dedicate to a recipient."

Gold teaches: a sender reviewing a song list and selecting a song to dedicate to a recipient (See [0036] of Gold)

Brachman teaches: a user may request data using a natural language search and a voice means may be used to enter the user's natural language request. (See col. 4 ln 22-26 of Brachman)

Since each individual element and its function are shown in the prior art, albeit shown in separate references, the difference between the claimed subject matter and the prior art rests not on any individual element or function but in the very combination itself that is in the substitution of content dedicated to the recipient for broadly defined content. Thus, the simple substitution of one known element for another producing a predictable result renders the claim obvious.

Claims 7, 9, 10, 17, 19, 20, and 26 depend from claims 1 and 11. Since claims 1 and 11 are non-allowable, the claims that depend from them are non-allowable.

B. Appellant's arguments that claims 6 and 16 are patentable over the alleged combination of Kennewick, Gold, Brachman, and Mahajan .

1. Appellant first argues that for at least the reasons set forth above in Section A with respect to independent claims 1 and 11, the rejection of claims 6 and 16 under 35 U.S.C. § 103(a) is improper, and should be reversed, by virtue of their dependence on an allowable independent claim.

Examiner respectfully disagrees.

Since claims 1 and 11 are non-allowable, the claims that depend from them are non-allowable.

2. Appellant also argues that dependent claims 6 and 16 are allowable for the further features they each recite. For example, claim 6 recites, *inter alia,* the feature of:

"**annotate the content with the second utterance [that is to be associated with a dedication to a recipient].**"

Claim 16 includes similar recitations.

Examiner respectfully disagrees.

Using the Broadest Reasonable Interpretation standard, "annotate the content with the second utterance" can be construed as a written transcript associated with the content as disclosed by Mahajan. See [0043] of Mahajan. Furthermore, adding a message to the content as taught by [0031-0032] of Gold reads on "annotate the content with the second utterance."

It would have been obvious to one of ordinary skill in the art to include in the combined system of Kennewick, Gold, and Brachman, the technique as taught by Mahajan since the claimed invention is merely a combination of old elements, and in the combination each element merely would have performed the same function as it did separately, and one of ordinary skill in the art would have recognized that the results of the combination were predictable.

C. Appellant's arguments that claims 5, 15, and 23-25 are patentable over the alleged combination of Kennewick, Gold, Brachman, and Caine.

1. Appellant first argues that for at least the reasons set forth above in Section A with respect to independent claims 1 and 11, the rejection of claims 5, 15, and 23-25 under 35 U.S.C. § 103(a) is improper, and should be reversed, by virtue of their dependence on an allowable independent claim.

Examiner respectfully disagrees. Since claims 1 and 11 are non-allowable, the claims that depend from them are non-allowable.

2. Appellant also argues that dependent claims 5, 15 and 23-25 are allowable for the further features they each recite. For example, claim 23 recites, *inter alia,* the feature of:

> "**providing one or more words or phrases of a second utterance [that is to be associated with a dedication to a recipient] as textual annotations within metadata of the content [that is to be dedicated to the recipient].“**

Examiner respectfully disagrees.

Appellant is reminded that the features recited in claims 5, 15, and 23-25 merely describe the metadata (“an image (5,15), textual annotations (23); and electronic link(24); voice tag(25) *[as one or more tags]* within the metadata”). Since the type of data in the

metadata is immaterial to the functioning of the claim as a whole, the limitations of claims 5, 15, and 23-25 are not accorded patentable weight.

Using the Broadest Reasonable Interpretation standard, "providing one or more words or phrases of a second utterance [that is to be associated with a dedication to a recipient] as textual annotations [that is to be dedicated to the recipient]" can be construed as, adding a text message to the content at taught by [0031] of Gold.

Caine, an analogous art of content storage, teaches changes to image information including tags (which are part of the metadata associated with the image) made by the user. See [0061][0079-0080] of Caine.

It would have been obvious to one of ordinary skill in the art to include in the combined system of Kennewick, Gold, and Brachman, the technique as taught by Caine since the claimed invention is merely a combination of old elements, and in the combination each element merely would have performed the same function as it did separately, and one of ordinary skill in the art would have recognized that the results of the combination were predictable.

D. Appellant's arguments that claims 8 and 18 are patentable over the alleged combination of Kennewick, Gold, Brachman, and Ginter.

1. Appellant argues that for at least the reasons set forth above in Section A with respect to independent claims 1 and 11, the rejection of claims 8 and 18 under 35 U.S.C. § 103(a) is improper, and should be reversed, by virtue of their dependence on an allowable independent claim.

Examiner respectfully disagrees. Since claims 1 and 11 are non-allowable, the claims that depend from them are non-allowable.

E. Appellant's arguments that claims 21 and 22 are patentable over the alleged combination of Kennewick, Gold, Brachman, and Anderson

1.Appellant argues that for at least the reasons set forth above in Section A with respect to independent claims 1 and 11, the rejection of claims 8 and 18 under 35 U.S.C. § 103(a) is improper, and should be reversed, by virtue of their dependence on an allowable independent claim.

Examiner respectfully disagrees. Since claims 1 and 11 are non-allowable, the claims that depend from them are non-allowable.

2. Appellant also argues that dependent claims 21 and 22 are allowable for the further features they each recite. For example, claim 22 recites, *inter alia,* the feature of

"**determine a media item that a user is playing;"**

"determine that the user provided the first utterance during the playing of the media item by the user;" and

"identify, based on the natural language processing and the determination that the user provided the first utterance during the playing of the media item, the media item as the content to dedicate to the recipient."

Kennewick teaches for each user command utterance, the system may execute a number of steps possibly including: 1) capture the user's command through accurate speech recognition operating in a variety of real-world environments; 2) parse and interpret the command; 3) determine the domain for the command and context, invoking the proper resources, including agents as required; 4) gather required data including, device settings, and measurement data; 5)) formulate device specific commands for the system or external devices; 6) route command to system, or external devices, including external devices connected to data networks; 7) receive and process

results of command, including errors; and, 8) optionally, provide response to user indicating the success or failure of the command, and possibly including state information.) See [0041] of Kennewick

Gold teaches the sender reviews the song list and selects a song to dedicate to a recipient. See [0036] of Gold.

Anderson teaches the context of a mobile terminal may be based on the applications currently being performed by the mobile terminal. For example, the context of a mobile terminal may be based on whether a user is playing a game, interacting with a media item (e.g., a photo, music track, video, etc.); context information may be received including information regarding the media content including the artist and song title that the mobile terminal is currently playing. See [0048][0067] of Anderson.

It would have been recognized that applying the technique of Gold, Brachman and Anderson to the teachings of Kennewick would have yielded predictable results because the level of ordinary skill in the art demonstrated by the references applied shows the ability to transfer data between the different systems (e.g., by routing the command from Kennewick and context (music being played) of Anderson to Gold). Further, using Kennewick to process a natural language utterance, and route the utterance as input to the system of Gold, would have been recognized by those of

ordinary skill in the art as resulting in an improved system that would allow selection of content based on natural language.

/TIM HALE/
Examiner, Art Unit 3685

/CALVIN L HEWITT II/
Supervisory Patent Examiner, Art Unit 3685

Conferees:

/CALVIN L HEWITT II/
Supervisory Patent Examiner, Art Unit 3685

/JAMES D NIGH/
Primary Examiner, Art Unit 3685

IN THE UNITED STATES PATENT & TRADEMARK OFFICE

IN RE PATENT APPLICATION OF:	Mike KENNEWICK *et al.*
SERIAL NO.:	12/943,699
FILING DATE:	November 10, 2010
ATTORNEY DOCKET NO.:	090611-0384422
CONFIRMATION NO.:	4185
EXAMINER:	Tim B. Hale
ART UNIT:	3685
FOR:	System and Method for Providing A Natural Language Content Dedication Service

APPELLANTS' BRIEF ON APPEAL UNDER 37 C.F.R. § 41.37

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA. 22313-1450

Dear Sir:

Further to the Notice of Appeal filed on **June 5, 2013**, Appellants respectfully submit Appellants' Brief on Appeal pursuant to 37 C.F.R. § 41.37.

The Director is authorized to charge any fees associated with the filing of this Appeal Brief, as well as any additional fees that may be due, or credit any overpayment of same to Deposit Account No. 033975 (**Ref. No. 090611-0384422**).

I. REAL PARTY IN INTEREST - *37 C.F.R. § 41.37(c)(1)(i)*

The real party in interest in this appeal is VoiceBox Technologies Corporation.

II. RELATED APPEALS AND INTERFERENCES - *37 C.F.R. § 41.37(c)(1)(ii)*

Appellants are aware of no related appeals or interferences.

III. STATUS OF CLAIMS - *37 C.F.R. § 41.37(c)(1)(iii)*

The status of the claims is as follows:

Pending: Claims 1, 5-11, and 15-26 are pending.

Cancelled: Claims 2-4 and 12-14 are canceled.

Rejected: Claims 1, 5-11, and 15-26 stand rejected.

Allowed: No claims have been allowed.

On Appeal: Claims 1, 5-11, and 15-26 are appealed.

Section VIII contains a correct listing of the pending claims.

IV. STATUS OF AMENDMENTS - *37 C.F.R. § 41.37(c)(1)(iv)*

No amendments have been filed subsequent to the mailing of the Final Office Action on March 5, 2013 (hereinafter "Final Action").

V. SUMMARY OF CLAIMED SUBJECT MATTER - *37 C.F.R. § 41.37(c)(1)(v)*

The following exemplary citations to the *as-filed* Specification and drawing figures demonstrating support for the **claim recitations** are not exclusive, as other examples of support for the claimed subject matter exist. As such, the following citations should not be viewed as limiting.

A. INDEPENDENT CLAIM 1

Independent claim 1 is directed to a system for providing a natural language content dedication service, comprising: one or more processors; and one or more non-transitory computer readable mediums storing executable instructions that when executed by the one or more processors cause the one or more processors to [Specification, *e.g.*, ¶¶ 10, 13, 91, 119; and FIG. 6]: receive a first utterance that includes a natural language utterance [Specification, *e.g.*, ¶¶ 10, 31, and 32; and FIG. 7]; determine, based on processing of the first utterance by a speech recognition engine, one or more words or phrases of the first utterance [Specification, *e.g.*, ¶¶ 36, 37, 94, 97, and 105; and FIG. 7]; identify, based on a natural language processing of the one or more words or phrases of the first utterance, content to dedicate to a recipient [Specification, *e.g.*, ¶¶ 95, 105, and 106; and FIG. 7]; initiate, based on the identified content, a dedication to the recipient [Specification, *e.g.*, ¶¶ 95, 105, and 106; and FIG. 7]; receive a second utterance to be associated with the dedication [Specification, *e.g.*, ¶¶ 96 and 106; and FIG. 7]; and send information to enable the recipient to access the content and the second utterance [Specification, *e.g.*, ¶¶ 21 and 118; and FIG. 8].

B. INDEPENDENT CLAIM 11

Independent claim 11 is directed to a method for providing a natural language content dedication service, comprising [Specification, *e.g.*, ¶¶ 10 and 12; and FIGS. 7 and 8]: receiving, by one or more processors, a first utterance that includes a natural language utterance[Specification, *e.g.*, ¶¶ 10, 31, and 32; and FIG. 7]; determining, by the one or more processors, one or more words or phrases of the first utterance based on processing of the first utterance by a speech recognition engine [Specification, *e.g.*, ¶¶

36, 37, 94, 97, and 105; and FIG. 7]; identifying, by the one or more processors, content to dedicate to a recipient based on a natural language processing of the one or more words or phrases of the first utterance [Specification, *e.g.*, ¶¶ 95, 105, and 106; and FIG. 7]; initiating, by the one or more processors, a dedication to the recipient based on the identified content [Specification, *e.g.*, ¶¶ 95, 105, and 106; and FIG. 7]; receiving, by the one or more processors, a second utterance to be associated with the dedication [Specification, *e.g.*, ¶¶ 96 and 106; and FIG. 7]; and sending, by the one or more processor, information to enable the recipient to access the content and the second utterance [Specification, *e.g.*, ¶¶ 21 and 118; and FIG. 8].

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL - *37 C.F.R. § 41.37(c)(1)(vi)*

1. Claims 1, 7, 9, 10, 11, 17, 19, 20, and 26 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over U.S. Patent Application Pub. No. 2004/0193420 by Kennewick (hereinafter "Kennewick") in view of U.S. Patent Application Publication No. 2002/0032752 to Gold (hereinafter "Gold") and U.S. Patent No. 6,704,576 to Brachman (hereinafter "Brachman");
2. Dependent claims 6 and 16 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick, Gold, and Brachman, further in view of U.S. Patent Application Publication No. 2008/0140385 to Mahajan ("Mahajan");
3. Dependent claims 5, 15, and 23-25 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick, Gold, and Brachman, further in view of U.S. Patent Application Publication No. 2010/0036967 to Caine ("Caine");
4. Dependent claims 8 and 18 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick, Gold, and Brachman, further in view of U.S. Patent No. 5,892,900 to Ginter ("Ginter"); and
5. Dependent claims 21 and 22 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick, Gold, and Brachman, further in view of U.S. Patent Application Publication No. 2009/0276700 to Anderson ("Anderson").

VII. ARGUMENTS - *37 C.F.R. § 41.37(c)(1)(vii)*

A. CLAIMS 1, 7, 9, 10, 11, 17, 19, 20, AND 26 ARE PATENTABLE OVER THE ALLEGED *THREE-REFERENCE* COMBINATION OF KENNEWICK, GOLD, AND BRACHMAN

The rejection of claims 1, 7, 9, 10, 11, 17, 19, 20, and 26 under 35 U.S.C. § 103(a) as allegedly being unpatentable over Kennewick in view of Gold and Brachman [Final Action, pg. 3] is legally improper and should be reversed because the Examiner has failed to establish a *prima facie* case of obviousness.

In rejecting claims under 35 U.S.C. § 103, it is incumbent upon the Examiner to establish a factual basis to support the legal conclusion of obviousness. *See In re Fine*, 837 F.2d 1071, 1073, 5 USPQ2d 1596, 1598 (Fed. Cir. 1988). In so doing, the Examiner must make the factual determinations set forth in *Graham v. John Deere Co.*, 383 U.S. 1, 17 (1966). Obviousness is a question of law based on the following underlying factual inquiries: "(a) the scope and content of the prior art; (b) the differences between the prior art and the claims at issue; (c) the level of ordinary skill in the art; and (d) objective evidence of nonobviousness." *Custom Accessories, Inc. v. Jeffrey-Allan Indus., Inc.*, 807 F.2d 955, 958, 1 U.S.P.Q.2d 1196, 1197 (Fed. Cir. 1986) (citing *Graham*, 383 U.S. at 17). "Against this background, the obviousness or nonobviousness of the subject matter is determined." *Graham*, 383 U.S. at 17.

To establish *prima facie* obviousness of a claimed invention, all the claim limitations must be taught or suggested by the prior art. *In re Royka*, 490 F.2d 981, 180 USPQ 580 (CCPA 1974). All words in a claim must be considered in judging the patentability of that claim against the prior art. *In re Wilson*, 424 F.2d 1382, 1385, 165 USPQ 494, 496 (CCPA 1970).

It is also well established that "the examiner bears the initial burden, on review of the prior art or an any other ground, of presenting a *prima facie* case of unpatentability. If that burden is met, the burden of coming forward with evidence or argument shifts to the applicant." *In re Oetiker*, 977 F.2d 1443, 1445, 24 U.S.P.Q.2d 1443, 1444 (Fed. Cir. 1992) (discussing *In re Piasecki*, 745 F.2d 1468, 1472, 223 USPQ 785, 788 (Fed. Cir. 1984)).

In the instant application, the Examiner has failed to establish a *prima facie* case of obviousness because, even assuming arguendo that it were deemed legally proper to combine Kennewick, Gold, and Brachman in the manner alleged by the Examiner (which Appellants do not concede), the alleged *three-reference* combination fails to teach or suggest all of the claim recitations of *at least* independent claims 1 and 11.

For example, independent claim 1 recites, *inter alia*, the feature of:

> **identify**, based on a natural language processing of the one or more words or phrases of the first utterance [that includes a natural language utterance], **content to dedicate to a recipient**.
>
> [Emphasis added].

Independent claim 11 includes similar recitations.

As disclosed, for example, in paragraphs 105 and 106 of the Specification, an example of identifying content to dedicate to a recipient based on a natural language processing of words or phrases of a first utterance may involve: (1) processing an utterance "Buy that song and send it to Michael;" (2) determining that the utterance includes the words or phrases "Buy," "that song," "send it," and "to Michael;" and (3) utilizing the words or phrases as criteria to identify the content and the recipient for a dedication where "Buy" and "send it" may indicate a content dedication request, "that song" may indicate the content to dedicate, and "to Michael" may indicate the recipient to which the content is to be dedicated.

In the Final Action, the Examiner concedes that the alleged combination of Kennewick and Gold fail to teach identification of content to dedicate to a recipient that is based on a natural language processing of one or more words or phrases of a first utterance. [Final Action, pgs. 4-5]. In an effort to cure the admitted deficiencies of Kennewick and Gold, the Examiner relies on a third reference – Brachman. [*Id.*]. However, the relied-upon portions of Brachman (i.e., col. 4, lines 5-27, col.8 lines 53-67, and col. 9, lines 1-11, and claim 7) fail to cure the admitted deficiencies of the alleged Kennewick/Gold combination.

In particular, at best, Brachman teaches natural language search queries. For example, in Brachman, "a user may request data using a natural language search (i.e., typing 'I really like Jazz and artists like the Beatles, but I do not like opera.'). Instead of

Amazon/VB Assets
Exhibit 1003
Page 151

the web, a voice means may be used to enter the user's natural language request." [Brachman, col. 4, lines 23-26]. As such, while the natural language search in Brachman may produce content results, the natural language search does not identify either (i) the content to dedicate or (ii) the recipient to which the content is to be dedicated.

Thus, Brachman does not teach or suggest the claimed features of identifying, based on a natural language processing of one or more words or phrases of a first utterance, content to dedicate to a recipient. Brachman therefore fails to cure the admitted deficiencies of Kennewick and Gold. The Examiner has failed to establish otherwise.

For at least this reason, the rejection of at least independent claims 1 and 11 under 35 U.S.C. § 103(a) is improper, and should be reversed. Dependent claims 7, 9, 10, 17, 19, 20, and 26 are allowable because they each ultimately depend from an allowable independent claim, as well as for the further features they recite.

B. CLAIMS 6 AND 16 ARE PATENTABLE OVER THE ALLEGED COMBINATION OF KENNEWICK, GOLD, BRACHMAN, AND MAHAJAN

Dependent claims 6 and 16 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick, Gold, and Brachman, further in view of Mahajan [Final Action, pg. 6]. For at least the reasons set forth above in Section A with respect to independent claims 1 and 11, the rejection of claims 6 and 16 under 35 U.S.C. § 103(a) is improper, and should be reversed, by virtue of their dependence on an allowable independent claim.

Moreover, dependent claims 6 and 16 are allowable for the further features they each recite. For example, claim 6 recites, *inter alia*, the feature of "annotate the content with the second utterance [that is to be associated with a dedication to a recipient]." [Emphasis added]. Claim 16 includes similar recitations.

As disclosed, for example, in paragraphs 96 and 97 of the Specification, an example of annotating content with a second utterance may involve: (1) capturing the second utterance (e.g., a dedication utterance "Dear Charlene, I was listening to this and I thought of you. Enjoy!") from a user that initiated a request to dedicate the content to a recipient based on a first utterance; (2) converting the second utterance

encoded audio; and (3) inserting the encoded audio of the second utterance into the content to be dedicated to the recipient.

In the Final Action, the Examiner concedes that the alleged combination of Kennewick and Gold fail to teach annotation of content (that is to be dedicated to a recipient) with a second utterance. [Final Action, pg. 6]. In an effort to cure the admitted deficiencies of Kennewick and Gold, the Examiner relies on Mahajan. [*Id.*]. However, the relied-upon portion of Mahajan (i.e., ¶ 40) also fails to teach the claimed annotation feature.

For example, the relied-upon portion of Mahajan teaches that a written transcript of audio/visual content may be presented on a user interface as the audio/visual content is being played. Presentation of a transcription of content on a user interface as the content is being played is not, however, annotation of content (that is to be dedicated to a recipient) with an utterance (e.g., the utterance "Dear Charlene, I was listening to this song and I thought of you …"), much less annotation of content with a second utterance that is to be associated with a dedication to a recipient.

Mahajan therefore does not teach or suggest the claimed features of annotating content (that is to be dedicated to a recipient) with a second utterance (that is to be associated with a dedication to the recipient). Thus, Mahajan fails to cure the admitted deficiencies of Kennewick and Gold.

For at least each of the foregoing reasons, the rejection of claims 6 and 16 under 35 U.S.C. § 103(a) is improper, and should be reversed.

C. CLAIMS 5, 15, AND 23-25 ARE PATENTABLE OVER THE ALLEGED COMBINATION OF KENNEWICK, GOLD, BRACHMAN, AND CAINE

Dependent claims 5, 15, and 23-25 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick, Gold, and Brachman, further in view of Caine [Final Action, pg. 6]. For at least the reasons set forth above in Section A with respect to independent claims 1 and 11, the rejection of claims 5, 15, and 23-25 under 35 U.S.C. § 103(a) is improper, and should be reversed, by virtue of their dependence on an allowable independent claim.

Amazon/VB Assets
Exhibit 1003
Page 153

Further, dependent claims 5, 15, and 23-25 are allowable for the further features they each recite. As one example, claim 23 recites, *inter alia*, the feature of:

> **provide the one or more words or phrases of the second utterance** [that is to be associated with a dedication to a recipient] **as textual annotations within metadata of the content** [that is to be dedicated to the recipient].
>
> [Emphasis added].

As disclosed, for example, in paragraphs 96 and 97 of the Specification, an example of annotating content with a second utterance may involve: (1) capturing the second utterance (e.g., a dedication utterance "Dear Charlene, I was listening to this song and I thought of you. Enjoy!") from a user that initiated a request to dedicate the content to a recipient based on a first utterance; (2) transcribing one or more words or phrases from the second utterance; and (3) inserting the words or phrases of the second utterance within metadata of the content to be dedicated to the recipient.

In the Final Action, the Examiner concedes that Kennewick fails to teach providing one or more words or phrases of a second utterance (that is to be associated with a dedication to a recipient) as textual annotations within metadata of content (that is to be dedicated to the recipient). [Final Action, pg. 6]. In an effort to cure the admitted deficiencies of Kennewick, the Examiner relies on combined portions of both Gold and Caine. [Final Action, pg. 7].

However, the relied-upon portion of Gold (i.e., ¶ 43) teaches that a script located on a web server may contain a URL for a voice message and a URL for a song, and the relied-upon portion of Caine (i.e., ¶¶ 79-80) teaches that an image may be tagged. The Examiner fails to demonstrate how the inclusion of URLs in a web server script, the tagging of an image, or the alleged combination of the URL inclusion and the image tagging teach or suggest the claimed features of providing one or more words or phrases of a second utterance (that is to be associated with a dedication to a recipient) as textual annotations within metadata of the content (that is to be dedicated to the recipient). For example, among other issues, the URLs in Gold are not words or phrases of an utterance. Thus, Gold and Caine fail to cure the admitted deficiencies of Kennewick.

For at least the foregoing reasons, the rejection of claims 5, 15, and 23-25 under 35 U.S.C. § 103(a) is improper, and should be reversed.

D. CLAIMS 8 AND 18 ARE PATENTABLE OVER THE ALLEGED COMBINATION OF KENNEWICK, GOLD, BRACHMAN, AND GINTER

Dependent claims 8 and 18 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick, Gold, and Brachman, further in view of Ginter. [Final Action, pg. 8]. For at least the reasons set forth above in Section A with respect to independent claims 1 and 11, the rejection of claims 8 and 18 under 35 U.S.C. § 103(a) is improper, and should be reversed, by virtue of their dependence on an allowable independent claim. In addition, the dependent claims are allowable for the further features they recite.

E. CLAIMS 21 AND 22 ARE PATENTABLE OVER THE ALLEGED COMBINATION OF KENNEWICK, GOLD, BRACHMAN, AND ANDERSON

Dependent claims 21 and 22 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick, Gold, and Brachman, further in view of Anderson [Final Action, pgs. 8 and 9]. For at least the reasons set forth above in Section A with respect to independent claims 1 and 11, the rejection of claims 21 and 22 under 35 U.S.C. § 103(a) is improper, and should be reversed, by virtue of their dependence on an allowable independent claim.

In addition, dependent claims 21 and 22 are allowable for the further features they each recite. As an example, dependent claim 22 recites, *inter alia*, the features of:

> determine a media item that is a user is playing;
>
> **determine that the user provided the first utterance during the playing of the media item by the user; and**
>
> **identify, based on** the natural language processing and **the determination that the user provided the first utterance during the playing of the media item, the media item as the content to dedicate to the recipient**.
>
> [Emphasis added].

One example of identifying a media item as content to dedicate to a recipient based on (i) a natural language processing of one or more words or phrases of a first utterance and (ii) a determination that a user provided the first utterance during playing of the media item by the user is disclosed, for example, in paragraphs 105 and 106 of the Specification. As disclosed, an utterance "Buy that song and send it to Michael" may be processed and determined to include the words or phrases "Buy," "that song," "send it," and "to Michael." [Specification, ¶ 105]. A song playing on the user's radio device may be identified as the content to dedicate to the recipient based on the words/phrase "that song" and the determination that the user provided the utterance "Buy that song and send it to Michael" while the song was playing on the user's radio device. [Specification, ¶ 106].

In the Final Action, the Examiner alleges that Gold teaches identification of a media item as content to dedicate to a recipient based on (i) a natural language processing of one or more words or phrases of a first utterance and (ii) a determination that a user provided the first utterance during playing of the media item by the user. [Final Action, pg. 9]. The Examiner also alleges that Anderson teaches a determination of a media item that a user is playing and a determination that the user provided the first utterance during the playing of the media item by the user. [*Id.*].

The relied-upon portion of Gold (i.e., ¶ 36) teaches user selection of a song from a song list, and the relied-upon portions of Anderson (i.e., ¶¶ 48 and 67) teach determination of a song that is currently playing. There is, however, no indication in either Gold or Anderson that a media item is identified as content to be dedicated to a recipient based on a determination that a user provided a first utterance during playing of the media item by the user, much less identification of the media item based on both: (i) a natural language processing of one or more words or phrases of a first utterance; and (ii) the determination that the user provided the first utterance during the playing of the media item.

Thus, Gold and Anderson do not teach or suggest the further features of at least claim 22. Moreover, Kennewick does not cure the deficiencies of Gold and Anderson.

For at least the foregoing reasons, the rejection of claims 21 and 22 under 35 U.S.C. § 103(a) is improper, and should be reversed.

CONCLUSION

For at least the foregoing reasons, reversal of the rejections of claims 1, 5-11 and 15-26 under 35 U.S.C. § 103 is earnestly sought.

Respectfully submitted,

Date: **September 5, 2013**

PILLSBURY WINTHROP SHAW PITTMAN LLP

/Ngai B. Zhang/

By: _____

Ngai B. Zhang
Registration No. 65,473

Customer No. 00909

P.O. Box 10500
McLean, Virginia 22102
Direct Dial: 703.770.7532
Main: 703.770.7900
Fax: 703.770.7901

VIII. CLAIMS APPENDIX - *37 C.F.R. § 41.37(c)(1)(viii)- **CLAIMS CURRENTLY PENDING ON APPEAL***

1. (**Previously Presented**) A system for providing a natural language content dedication service, comprising:

one or more processors; and

one or more non-transitory computer readable mediums storing executable instructions that when executed by the one or more processors cause the one or more processors to:

receive a first utterance that includes a natural language utterance;

determine, based on processing of the first utterance by a speech recognition engine, one or more words or phrases of the first utterance;

identify, based on a natural language processing of the one or more words or phrases of the first utterance, content to dedicate to a recipient;

initiate, based on the identified content, a dedication to the recipient;

receive a second utterance to be associated with the dedication; and

send information to enable the recipient to access the content and the second utterance.

2-4. (**Canceled**)

5. (**Previously Presented**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to provide an image as one or more tags within metadata of the content.

6. (**Previously Presented**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to annotate the content with the second utterance.

7. (**Previously Presented**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to purchase the content from a content provider to enable the recipient to access the content.

8. (**Previously Presented**) The system of claim 7, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to:

identify, for the dedication, a purchase option that includes one or more of a buy-to-own purchase option, a pay-to-pay purchase option, or a paid subscription purchase option; and

bill a user associated with the first utterance based on the purchase option.

9. (**Previously Presented**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to: send, to the recipient, an electronic link to access the content as the information to enable the recipient to access the content; and

send, to the recipient, the second utterance in an encoded audio format as the information to enable the recipient to access the second utterance.

10. (**Original**) The system of claim 1, wherein the content includes one or more of a song or a video.

11. (**Previously Presented**) A method for providing a natural language content dedication service, comprising:

receiving, by one or more processors, a first utterance that includes a natural language utterance; determining, by the one or more processors, one or more words or phrases of the first utterance based on processing of the first utterance by a speech recognition engine;

identifying, by the one or more processors, content to dedicate to a recipient based on a natural language processing of the one or more words or phrases of the first utterance;

initiating, by the one or more processors, a dedication to the recipient based on the identified content;

receiving, by the one or more processors, a second utterance to be associated with the dedication; and

sending, by the one or more processor, information to enable the recipient to access the content and the second utterance.

12-14. (**Canceled**)

15. (**Previously Presented**) The method of claim 11, further comprising providing an image as one or more tags within metadata of the content.

16. (**Previously Presented**) The method of claim 11, further comprising annotating the content with the second utterance.

17. (**Previously Presented**) The method of claim 11, further comprising purchasing the content from a content provider to enable the recipient to access the content.

18. (**Previously Presented**) The method of claim 17, further comprising:

identifying, for the dedication, a purchase option that includes one or more of a buy-to-own purchase option, a pay-to-pay purchase option, or a paid subscription purchase option; and

billing a user associated with the first utterance based on the purchase option.

19. (**Previously Presented**) The method of claim 11, further comprising: sending, to the recipient, an electronic link to access the content as the information to enable the recipient to access the content; and

sending, to the recipient, the second utterance in an encoded audio format as the information to enable the recipient to access the second utterance.

20. (**Original**) The method of claim 11, wherein the content includes one or more of a song or a video.

21. (**Previously Presented**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to:

determine that a user is interacting with a media item; and

identify, based on the natural language processing and the interaction of the user with the media item, the media item as the content to dedicate to the recipient.

22. (**Previously Presented**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to:

determine a media item that is a user is playing;

determine that the user provided the first utterance during the playing of the media item by the user; and

identify, based on the natural language processing and the determination that the user provided the first utterance during the playing of the media item, the media item as the content to dedicate to the recipient.

23. (**Previously Presented**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to:

determine, based on processing of the second utterance by the speech recognition engine, one or more words or phrases of the second utterance; and

provide the one or more words or phrases of the second utterance as textual annotations within metadata of the content.

24. (**Previously Presented**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to provide an electronic link associated with the content as one or more tags within metadata of the content.

25. (**Previously Presented**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to provide a voice tag as one or more tags within metadata of the content.

26. (**Previously Presented**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to provide the second utterance within the content.

IX. EVIDENCE APPENDIX - *37 C.F.R. § 41.37(c)(1)(ix)*

For this Appeal, Appellants do not rely on any evidence submitted pursuant to 37 C.F.R. §§ 1.130, 1.131, or 1.132, or other evidence entered by the Examiner.

Amazon/VB Assets
Exhibit 1003
Page 164

X. RELATED PROCEEDINGS INDEX - *37 C.F.R. § 41.37(c)(1)(x)*

NONE.

PTO/SB/22 (03-13)
Approved for use through 3/31/2013. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PETITION FOR EXTENSION OF TIME UNDER 37 CFR 1.136(a)

Docket Number (Optional): 090611-0384422

Application Number	Filed
12/943,699	November 10, 2010

For: SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE

Art Unit	Examiner
3685	Tim B. Hale

This is a request under the provisions of 37 CFR 1.136(a) to extend the period for filing a reply in the above-identified application.

The requested extension and fee are as follows (check time period desired and enter the appropriate fee below):

	Fee	Small Entity Fee	Micro Entity Fee	
☒ One month (37 CFR 1.17(a)(1))	$200	$100	$50	$ 100.00
☐ Two months (37 CFR 1.17(a)(2))	$600	$300	$150	$
☐ Three months (37 CFR 1.17(a)(3))	$1,400	$700	$350	$
☐ Four months (37 CFR 1.17(a)(4))	$2,200	$1,100	$550	$
☐ Five months (37 CFR 1.17(a)(5))	$3,000	$1,500	$750	$

☒ Applicant asserts small entity status. See 37 CFR 1.27.

☐ Applicant certifies micro entity status. See 37 CFR 1.29.
Form PTO/SB/15A or B or equivalent must either be enclosed or have been submitted previously.

☐ A check in the amount of the fee is enclosed.

☐ Payment by credit card. Form PTO-2038 is attached.

☒ The Director has already been authorized to charge fees in this application to a Deposit Account.

☒ The Director is hereby authorized to charge any fees which may be required, or credit any overpayment, to Deposit Account Number 033975.

☒ Payment made via EFS-Web.

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

I am the

☐ applicant/inventor.

☐ assignee of record of the entire interest. See 37 CFR 3.71. 37 CFR 3.73(b) statement is enclosed (Form PTO/SB/96).

☒ attorney or agent of record. Registration number 65,473.

☐ attorney or agent acting under 37 CFR 1.34. Registration number _____.

/Ngai B. Zhang/	September 5, 2013
Signature	Date
Ngai Zhang	(703) 770-7532
Typed or printed name	Telephone Number

NOTE: This form must be signed in accordance with 37 CFR 1.33. See 37 CFR 1.4 for signature requirements and certifications. Submit multiple forms if more than one signature is required, see below*.

☒ * Total of 1 forms are submitted.

This collection of information is required by 37 CFR 1.136(a). The information is required to obtain or retain a benefit by the public, which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 6 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

American LegalNet, Inc.
www.FormsWorkFlow.com

Electronic Patent Application Fee Transmittal

Application Number:	12943699
Filing Date:	10-Nov-2010
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Filer:	Ngai Bing Zhang/Miriam Nelsen Awan
Attorney Docket Number:	090611-0384422

Filed as Small Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				
Extension - 1 month with $0 paid	2251	1	100	100

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
			Total in USD ($)	**100**

Electronic Acknowledgement Receipt	
EFS ID:	16774494
Application Number:	12943699
International Application Number:	
Confirmation Number:	4185
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Customer Number:	909
Filer:	Ngai Bing Zhang/Miriam Nelsen Awan
Filer Authorized By:	Ngai Bing Zhang
Attorney Docket Number:	090611-0384422
Receipt Date:	05-SEP-2013
Filing Date:	10-NOV-2010
Time Stamp:	16:11:51
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	$100
RAM confirmation Number	3234
Deposit Account	033975
Authorized User	
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: Charge any Additional Fees required under 37 C.F.R. Section 1.16 (National application filing, search, and examination fees) Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees)	

Charge any Additional Fees required under 37 C.F.R. Section 1.19 (Document supply fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.20 (Post Issuance fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.21 (Miscellaneous fees and charges)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		090611-0384422_AppealBrief_EOT.pdf	745841 956d153f3178253a5c703a56eccd95b55cb34537	yes	22

	Multipart Description/PDF files in .zip description		
	Document Description	**Start**	**End**
	Appeal Brief Filed	1	21
	Extension of Time	22	22

Warnings:					
Information:					
2	Fee Worksheet (SB06)	fee-info.pdf	30702 42c274be721e48455adffcd39712d17379036030	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	776543		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	A	6,122,613	09/19/2000	Baker	704	235	
	B	6,356,869	03/12/2002	Chapados *et al.*	704	275	
	C	6,505,155	01/07/2003	Vanbuskirk *et al.*	704	246	
	D	6,754,627	06/22/2004	Woodward	704	235	
	E	7,197,460	03/27/2007	Gupta *et al.*	704	270.1	
	F	7,421,393	09/02/2008	Di Fabbrizio *et al.*	704	275	
	G	7,890,324	02/15/2011	Bangalore *et al.*	704	231	
	H	8,219,399	07/10/2012	Lutz *et al.*	704	270	
	I	8,255,224	08/28/2012	Singleton *et al.*	704	275	
	J	8,447,607	05/21/2013	Weider *et al.*	704	250	
	K	8,452,598	05/28/2013	Kennewick *et al.*	704	257	
	L	2002/0143532	10/03/2002	McLean *et al.*	704	235	
	M	2009/0076827	03/19/2009	Bulitta *et al.*	704	275	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

403997232v1

Electronic Patent Application Fee Transmittal

Application Number:	12943699
Filing Date:	10-Nov-2010
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Filer:	Ngai Bing Zhang/Miriam Nelsen Awan
Attorney Docket Number:	090611-0384422

Filed as Small Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Notice of Appeal	2401	1	400	400
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
Submission- Information Disclosure Stmt	2806	1	90	90
		Total in USD ($)		**490**

Electronic Acknowledgement Receipt	
EFS ID:	15959902
Application Number:	12943699
International Application Number:	
Confirmation Number:	4185
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Customer Number:	909
Filer:	Ngai Bing Zhang/Miriam Nelsen Awan
Filer Authorized By:	Ngai Bing Zhang
Attorney Docket Number:	090611-0384422
Receipt Date:	05-JUN-2013
Filing Date:	10-NOV-2010
Time Stamp:	17:07:26
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	$490
RAM confirmation Number	4140
Deposit Account	033975
Authorized User	
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: Charge any Additional Fees required under 37 C.F.R. Section 1.16 (National application filing, search, and examination fees) Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees)	

Charge any Additional Fees required under 37 C.F.R. Section 1.19 (Document supply fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.20 (Post Issuance fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.21 (Miscellaneous fees and charges)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		090611-0384422_NOA_IDS.pdf	61236 928accb4156c4e2daef91cc5ab258f72495b23fb	yes	5

Multipart Description/PDF files in .zip description		
Document Description	**Start**	**End**
Notice of Appeal Filed	1	1
Transmittal Letter	2	4
Information Disclosure Statement (IDS) Form (SB08)	5	5

Warnings:

Information:

2	Fee Worksheet (SB06)	fee-info.pdf	32442 5fa71a2a011511196b0184ab41e704a7a3c0d227	no	2

Warnings:

Information:

Total Files Size (in bytes):	93678

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/31 (09-12)
Approved for use through 01/31/2013. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

NOTICE OF APPEAL FROM THE EXAMINER TO THE BOARD OF PATENT APPEALS AND INTERFERENCES

Docket Number (Optional): 090611-0384422

I hereby certify that this correspondence is being facsimile transmitted to the USPTO EFS-Web transmitted to the USPTO, or or deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to "Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450" [37 CFR 1.8(a)]

on _____

Signature _____

Typed or printed name _____

In re Application of: MIKE KENNEWICK, et al.

Application Number	Filed
12/943,699	November 10, 2010

For SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE

Art Unit	Examiner
3685	Tim B. Hale

Applicant hereby **appeals** to the Board of Patent Appeals and Interferences from the last decision of the examiner.

The fee for this Notice of Appeal is (37 CFR 41.20 (b)(1)) $ 800.00

☒ Applicant claims small entity status. See 37 CFR 1.27. Therefore, the fee shown above is reduced by half, and the resulting fee is: $ 400.00

☐ A check in the amount of the fee is enclosed,

☐ Payment by credit card. Form PTO-2038 is attached.

☐ The Director has already been authorized to charge fees in this application to a Deposit Account.

☒ The Director is hereby authorized to charge any fees which may be required, or credit any overpayment to Deposit Account No. 033975.

☐ A petition for an extension of time under 37 CFR 1.136(a) (PTO/SB/22) is enclosed.

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

I am the

☐ applicant/inventor.

☐ assignee of record of the entire interest. See 37 CFR 3.71. Statement under 37 CFR 3.73(b) is enclosed. (Form PTO/SB/96)

☒ attorney or agent of record. Registration number 65,473

☐ attorney or agent acting under 37 CFR 1.34. Registration number if acting under 37 CFR 1.34. _____

/Ngai Zhang/
Signature

Ngai Zhang
Typed or printed name

(703) 770-7532
Telephone number

June 5, 2013
Date

NOTE: Signatures of all the inventors or assignees of record of the entire interest or their representative(s) are required. Submit multiple forms if more than one signature is required, see below*.

☒ *Total of 1 forms are submitted.

This collection of information is required by 37 CFR 41.31. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11, 1.14 and 41.6. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

American LegalNet, Inc.
www.FormsWorkFlow.com

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANTS :	Mike KENNEWICK *et al.*	CONFIRMATION No.:	4185
SERIAL NUMBER :	12/943,699	EXAMINER:	Tim B. Hale
FILING DATE :	November 10, 2010	ART UNIT:	3685
FOR :	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE		

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT

Commissioner for Patents
P. O. Box 1450
Alexandria, VA 22313-1450

Sir:

Pursuant to the duty of disclosure under 37 C.F.R. §§ 1.56, 1.97 and 1.98, Applicants hereby make of record in the above-identified application the documents listed on the attached Form PTO-1449.

Applicants reserve the right to establish the patentability of the claimed invention over any of the information provided herewith, and/or to prove that this information may not be prior art, and/or to prove that this information may not be enabling for the teachings purportedly offered.

This Supplemental Information Disclosure Statement is being filed more than three months after the filing date of a national application other than a continued prosecution application under 37 C.F.R. § 1.53(d), or more than three months after the date of entry of the national stage as set forth in 37 C.F.R. § 1.491 in an international application, and after the

mailing date of any of a Final Action under 37 C.F.R. § 1.113, a Notice of Allowance under 37 C.F.R. § 1.311, or an action that otherwise closes prosecution in the application, but on or before payment of the Issue Fee. 37 C.F.R. § 1.97(d).

Documents A, G, L, and M were cited by the U.S. Patent Office in co-pending Application Serial No. 12/703,032, filed February 9, 2010, in an Office Action dated April 22, 2013. Documents H and I were cited by the U.S. Patent Office in co-pending Application Serial No. 13/619,421, filed September 14, 2012, in an Office Action dated April 24, 2013. Documents B-F were cited by the U.S. Patent Office in co-pending Application Serial No. 13/022,370, filed February 7, 2011, in an Office Action dated May 1, 2013. I hereby certify that no item of information contained in this Supplemental Information Disclosure Statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to my knowledge after making reasonable inquiry, no item of information contained in this Supplemental Information Disclosure Statement was known to any individual designated in 37 C.F.R. § 1.56(c) more than three months prior to the filing of this Supplemental Information Disclosure Statement. 37 C.F.R. § 1.97(e)(2).

Please charge the $90.00 fee [37 C.F.R. § 1.17(p)] to Deposit Account No. 03-3975. In addition, please charge any fee deficiencies, or credit any overpayment, to Deposit Account No. 03-3975 as needed to ensure consideration of the Supplemental Information Disclosure Statement.

It is respectfully requested that the Examiner initial and return a copy of the enclosed PTO-1449.

Respectfully submitted,

PILLSBURY WINTHROP SHAW PITTMAN LLP

By /Ngai B. Zhang/
Ngai B. Zhang
Registration No.: 65,473

Customer No.
00909

P. O. Box 10500
McLean, Virginia 22102
Telephone: (703) 770-7900
Telefax: (703) 770-7901

Dated: June 5, 2013

Amazon/VB Assets
Exhibit 1003
Page 179

IN THE UNITED STATES PATENT & TRADEMARK OFFICE

IN RE PATENT APPLICATION OF:	Mike KENNEWICK *et al.*
SERIAL NO.:	12/943,699
FILING DATE:	November 10, 2010
ATTORNEY DOCKET NO.:	090611-0384422
CONFIRMATION NO.:	4185
EXAMINER:	Tim B. Hale
ART UNIT:	3685
FOR:	System and Method for Providing A Natural Language Content Dedication Service

EXAMINER INTERVIEW SUMMARY

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Examiner Tim B. Hale is thanked for the courtesies extended to Applicants' representative (the undersigned) during a telephonic interview conducted on May 7, 2013 (the "Interview").

In the summary of the substance of the Interview provided by Examiner Hale [*see* 05/07/13 Interview Summary], the Examiner recites that "possible amendments [were discussed] to overcome 103" during the Interview. Applicants wish to clarify, for the record, that possible claim amendments were discussed during the Interview *solely* in an effort to expedite prosecution. Applicants maintain, however, that the rejections under § 103 are legally improper and should be withdrawn and, as such, *no* claim amendments are required to overcome the § 103 rejections. Accordingly, because the Examiner failed to establish a *prima facie* case of obviousness in the last Office Action, and because no Agreement was reached during the Interview, Applicants are now pursuing an Appeal in the above-referenced application.

Nonetheless, if the Examiner believes, for any reason, that personal communication will expedite prosecution of this application, the Examiner is invited to telephone the undersigned at the number provided.

Respectfully submitted,

Date: **May 31, 2013**

PILLSBURY WINTHROP SHAW PITTMAN LLP

By: **/Ngai B. Zhang/**
Ngai B. Zhang
Registration No. 65,473

Customer No. 00909

P.O. Box 10500
McLean, Virginia 22102
Direct Dial: 703.770.7532
Main: 703.770.7900
Fax: 703.770.7901

Electronic Acknowledgement Receipt

EFS ID:	15921958
Application Number:	12943699
International Application Number:	
Confirmation Number:	4185
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Customer Number:	909
Filer:	Ngai Bing Zhang/Miriam Nelsen Awan
Filer Authorized By:	Ngai Bing Zhang
Attorney Docket Number:	090611-0384422
Receipt Date:	31-MAY-2013
Filing Date:	10-NOV-2010
Time Stamp:	16:44:27
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Applicant summary of interview with examiner	090611-0384422_SummaryOfExaminerInterview.pdf	513592 df34955154750a164e1cf6c809ea920c70a2c435	no	2
Warnings:					
Information:					

Total Files Size (in bytes):	513592

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/943,699	11/10/2010	Mike Kennewick	090611-0384422	4185

909 7590 05/21/2013
Pillsbury Winthrop Shaw Pittman, LLP (NV)
PO Box 10500
McLean, VA 22102

EXAMINER	
HALE, TIM B	
ART UNIT	PAPER NUMBER
3685	

NOTIFICATION DATE	DELIVERY MODE
05/21/2013	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

docket_ip@pillsburylaw.com
margaret.drosos@pillsburylaw.com

Applicant-Initiated Interview Summary	**Application No.** 12/943,699	**Applicant(s)** KENNEWICK ET AL.	
	Examiner TIM HALE	**Art Unit** 3685	

All participants (applicant, applicant's representative, PTO personnel):

(1) *TIM HALE*. (3)_____.

(2) *Ngai Zhang*. (4)_____.

Date of Interview: *07 May 2013*.

Type: ☒ Telephonic ☐ Video Conference
☐ Personal [copy given to: ☐ applicant ☐ applicant's representative]

Exhibit shown or demonstration conducted: ☐ Yes ☐ No.
If Yes, brief description: _____.

Issues Discussed ☐101 ☐112 ☐102 ☒103 ☐Others
(For each of the checked box(es) above, please describe below the issue and detailed description of the discussion)

Claim(s) discussed: *6,22,23 and 26*.

Identification of prior art discussed: *Gold*.

Substance of Interview
(For each issue discussed, provide a detailed description and indicate if agreement was reached. Some topics may include: identification or clarification of a reference or a portion thereof, claim interpretation, proposed amendments, arguments of any applied references etc...)

Discussed possible amendments to overcome 103.

Applicant recordation instructions: The formal written reply to the last Office action must include the substance of the interview. (See MPEP section 713.04). If a reply to the last Office action has already been filed, applicant is given a non-extendable period of the longer of one month or thirty days from this interview date, or the mailing date of this interview summary form, whichever is later, to file a statement of the substance of the interview

Examiner recordation instructions: Examiners must summarize the substance of any interview of record. A complete and proper recordation of the substance of an interview should include the items listed in MPEP 713.04 for complete and proper recordation including the identification of the general thrust of each argument or issue discussed, a general indication of any other pertinent matters discussed regarding patentability and the general results or outcome of the interview, to include an indication as to whether or not agreement was reached on the issues raised.

☐ Attachment

/TIM HALE/ Examiner, Art Unit 3685	

Summary of Record of Interview Requirements

Manual of Patent Examining Procedure (MPEP), Section 713.04, Substance of Interview Must be Made of Record
A complete written statement as to the substance of any face-to-face, video conference, or telephone interview with regard to an application must be made of record in the application whether or not an agreement with the examiner was reached at the interview.

Title 37 Code of Federal Regulations (CFR) § 1.133 Interviews
Paragraph (b)

In every instance where reconsideration is requested in view of an interview with an examiner, a complete written statement of the reasons presented at the interview as warranting favorable action must be filed by the applicant. An interview does not remove the necessity for reply to Office action as specified in §§ 1.111, 1.135. (35 U.S.C. 132)

37 CFR §1.2 Business to be transacted in writing.
All business with the Patent or Trademark Office should be transacted in writing. The personal attendance of applicants or their attorneys or agents at the Patent and Trademark Office is unnecessary. The action of the Patent and Trademark Office will be based exclusively on the written record in the Office. No attention will be paid to any alleged oral promise, stipulation, or understanding in relation to which there is disagreement or doubt.

The action of the Patent and Trademark Office cannot be based exclusively on the written record in the Office if that record is itself incomplete through the failure to record the substance of interviews.

It is the responsibility of the applicant or the attorney or agent to make the substance of an interview of record in the application file, unless the examiner indicates he or she will do so. It is the examiner's responsibility to see that such a record is made and to correct material inaccuracies which bear directly on the question of patentability.

Examiners must complete an Interview Summary Form for each interview held where a matter of substance has been discussed during the interview by checking the appropriate boxes and filling in the blanks. Discussions regarding only procedural matters, directed solely to restriction requirements for which interview recordation is otherwise provided for in Section 812.01 of the Manual of Patent Examining Procedure, or pointing out typographical errors or unreadable script in Office actions or the like, are excluded from the interview recordation procedures below. Where the substance of an interview is completely recorded in an Examiners Amendment, no separate Interview Summary Record is required.

The Interview Summary Form shall be given an appropriate Paper No., placed in the right hand portion of the file, and listed on the "Contents" section of the file wrapper. In a personal interview, a duplicate of the Form is given to the applicant (or attorney or agent) at the conclusion of the interview. In the case of a telephone or video-conference interview, the copy is mailed to the applicant's correspondence address either with or prior to the next official communication. If additional correspondence from the examiner is not likely before an allowance or if other circumstances dictate, the Form should be mailed promptly after the interview rather than with the next official communication.

The Form provides for recordation of the following information:
- Application Number (Series Code and Serial Number)
- Name of applicant
- Name of examiner
- Date of interview
- Type of interview (telephonic, video-conference, or personal)
- Name of participant(s) (applicant, attorney or agent, examiner, other PTO personnel, etc.)
- An indication whether or not an exhibit was shown or a demonstration conducted
- An identification of the specific prior art discussed
- An indication whether an agreement was reached and if so, a description of the general nature of the agreement (may be by attachment of a copy of amendments or claims agreed as being allowable). Note: Agreement as to allowability is tentative and does not restrict further action by the examiner to the contrary.
- The signature of the examiner who conducted the interview (if Form is not an attachment to a signed Office action)

It is desirable that the examiner orally remind the applicant of his or her obligation to record the substance of the interview of each case. It should be noted, however, that the Interview Summary Form will not normally be considered a complete and proper recordation of the interview unless it includes, or is supplemented by the applicant or the examiner to include, all of the applicable items required below concerning the substance of the interview.

A complete and proper recordation of the substance of any interview should include at least the following applicable items:

1) A brief description of the nature of any exhibit shown or any demonstration conducted,
2) an identification of the claims discussed,
3) an identification of the specific prior art discussed,
4) an identification of the principal proposed amendments of a substantive nature discussed, unless these are already described on the Interview Summary Form completed by the Examiner,
5) a brief identification of the general thrust of the principal arguments presented to the examiner,
   (The identification of arguments need not be lengthy or elaborate. A verbatim or highly detailed description of the arguments is not required. The identification of the arguments is sufficient if the general nature or thrust of the principal arguments made to the examiner can be understood in the context of the application file. Of course, the applicant may desire to emphasize and fully describe those arguments which he or she feels were or might be persuasive to the examiner.)
6) a general indication of any other pertinent matters discussed, and
7) if appropriate, the general results or outcome of the interview unless already described in the Interview Summary Form completed by the examiner.

Examiners are expected to carefully review the applicant's record of the substance of an interview. If the record is not complete and accurate, the examiner will give the applicant an extendable one month time period to correct the record.

Examiner to Check for Accuracy

If the claims are allowable for other reasons of record, the examiner should send a letter setting forth the examiner's version of the statement attributed to him or her. If the record is complete and accurate, the examiner should place the indication, "Interview Record OK" on the paper recording the substance of the interview along with the date and the examiner's initials.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/943,699	11/10/2010	Mike Kennewick	090611-0384422	4185

909 7590 03/05/2013

Pillsbury Winthrop Shaw Pittman, LLP (NV)
PO Box 10500
McLean, VA 22102

EXAMINER	
HALE, TIM B	
ART UNIT	PAPER NUMBER
3685	

NOTIFICATION DATE	DELIVERY MODE
03/05/2013	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

docket_ip@pillsburylaw.com
margaret.drosos@pillsburylaw.com

Office Action Summary	Application No.	Applicant(s)	
	12/943,699	KENNEWICK ET AL.	
	Examiner	Art Unit	
	TIM HALE	3685	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

4)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

5)☒ Claim(s) *1,5-11 and 15-26* is/are pending in the application.

5a) Of the above claim(s) _____ is/are withdrawn from consideration.

6)☐ Claim(s) _____ is/are allowed.

7)☒ Claim(s) *1,5-11 and 15-26* is/are rejected.

8)☐ Claim(s) _____ is/are objected to.

9)☐ Claim(s) _____ are subject to restriction and/or election requirement.

* If any claims have been determined allowable, you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see http://www.uspto.gov/patents/init_events/pph/index.jsp or send an inquiry to PPHfeedback@uspto.gov.

Application Papers

10)☐ The specification is objected to by the Examiner.

11)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date *1/30/2013*.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

4) ☐ Other: _____.

Status of Claims

1. Claims 1, 5-11, 15-26 have been presented.

2. Claims 1, 5-11, 15-26 are rejected.

Response to Arguments

3. *Claim Rejections - 35 USC § 102 and 103*

Applicant's arguments with respect to claim *1, 5-11, 15-26* have been considered but are moot because the arguments do not apply to any of the references being used in the current rejection.

Examiner's Comments

4. Per claims 8 and 18, the type of data being collected by the system is non-functional as it does not alter how the machine functions (E.g., a purchase option that includes one or more of a buy-to-own purchase option, a pay-to-pay purchase option, or a paid subscription purchase option;) It has been held that nonfunctional descriptive material cannot render nonobvious an invention that would have otherwise been obvious (In re Gulack, 217 USPQ 401 (Fed. Cir. 1983), In re Ngai, 70 USPQ2d (Fed. Cir. 2004), In re Lowry, 32 USPQ2d 1031 (Fed. Cir. 1994); MPEP 2106.01)

Per claims 1 and 21-26, the "to…" language (e.g. the one or more processors to:…) does not have weight.

Per claims 1 and 11, "to enable…" does not have weight as "to enable" is an intended use in claim 1 and not a positively recited step in claim 11.

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

6. Claims 1, 7, 9, 10, 11, 17, 19, 20 and 26 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent Publication 20040193420 to Kennewick (hereafter "Kennewick") in view of US Patent Publication 20020032752 to Gold (hereafter "Gold") and US Patent 6704576 to Brachman (hereafter "Brachman").

Per claims 1 and 11, Kennewick teaches

"A system for providing a natural language content dedication service, comprising: [A method for providing a natural language content dedication service, comprising:]

one or more processors and one or more non-transitory computer readable mediums storing executable instructions that when executed by the one or more processors cause the processors to: [0012-0013][0014][0016][0095][0096][0097][0111]Figs. 1-3, 5-6

"receive[, by one or more processors,] a first utterance that includes a natural language utterance;[0028][0034][0038][0101]" and

"determine, [by the one or more processors] based on the processing of the first utterance by a speech recognition engine, one or more words or phrases of the first utterance;[0028-0034][0041][0123]

Kennewick does not teach, but Gold teaches:

"initiate, [by the one or more processors] based on the identified content, a dedication to the recipient" Fig. 2 [0036-0038]

"to dedicate for the recipient" Fig. 2 [0036-0038]

"receive [by the one or more processors] a second utterance to be associated with the dedication; "[0038][0039];Abstract

"send [by the one or more processors] information to enable the recipient to access the content and the second utterance." [0042][0043]

It would have been obvious to combine Kennewich with Gold to provide a system for distribution of content based on a natural language input.

Neither Kennewich nor Gold teach, but Brachman teaches:

"identify, [by the one or more processors], based on the a natural language processing of the one or more words or phrases of the first utterance, content" col. 4 ln 5-27; col. 8 ln 53-67; col. 9 ln 1-11; claim 7

It would have been obvious to combine Kennewich with Gold and Brachman to provide a system for distribution of musical content based on a natural language input.

Per claims 7 and 17, Gold teaches:" purchase[ing], the content from a content provider to enable the recipient to access the content." [0009][0038-0043]]

Per claims 9 and 19, Gold teaches, "send, to the recipient, an electronic link to access the content as the information to enable the recipient to access the content;" [0042]; Abstract; Claim 8

"send, to the recipient, the second utterance in an encoded audio format as the information to enable the recipient to access the second utterance;"[0039]

Per claims 10 and 20, Gold teaches, "wherein the content includes one or more of a song or a video." [0007]

Per claim 26, Gold teaches

"cause the one or more processors to provide the second utterance within the content."[0040]

7. Claims 6 and 16 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent Publication 20040193420 to Kennewick (hereafter "Kennewick") in view of US Patent Publication 20020032752 to Gold (hereafter "Gold"), US Patent 6704576 to Brachman (hereafter "Brachman") and US Patent Publication 20080140385 to Mahajan (hereafter "Mahajan").

Per claims 6 and 16, Mahajan teaches: "annotate[ing] the content with the second utterance." [0040]

It would have been obvious to combine Kennewick, Gold, Brachman and Mahajan to produce a system for distributing A/V content based on speech recognition.

8. Claims 5, 15, 23, 24, and 25 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent Publication 20040193420 to Kennewick (hereafter "Kennewick") in view of US Patent Publication 20020032752 to Gold (hereafter "Gold"), US Patent 6704576 to Brachman (hereafter "Brachman"), and US Patent Publication 20100036967 to Caine (hereafter "Caine").

Per claims 5 and 15, neither Kennewick nor Gold teach, but Caine teaches: "provide an image as one or more tags within metadata of the content." [0079][0080]

It would have been obvious to combine Kennewick with Gold, Brachman, and Caine to produce a system for modifying visual content associated with digital content.

Per claim 23,

Kennewick teaches:

"determine, based on processing of the second utterance by the speech recognition engine, one or more words or phrases of the second utterance;" [0028-0034][0041][0123]

Gold teaches:

"provide the one or more words or phrases of the second utterance as textual annotations" [0043]

Caine teaches

"within metadata of the content." [0079][0080]

Per claim 24,

Gold teaches:

"cause the one or more processors to provide an electronic link associated with the content" [0037]

Caine teaches:

"as one or more tags within metadata of the content." [0079][0080]

Per claim 25,

Gold teaches:

"cause the one or more processors to provide a voice tag"[0039]

Caine teaches:

"as one or more tags within metadata of the content." [0079][0080]

9. Claims 8 and 18 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent Publication 20040193420 to Kennewick (hereafter "Kennewick") in view of US Patent Publication 20020032752 to Gold (hereafter "Gold"), US Patent 6704576 to Brachman (hereafter "Brachman") and US Patent 5892900 to Ginter (hereafter "Ginter").

Per claims 8 and 18, Kennewick teaches:

"identify[ing], for the dedication, a purchase option that includes one or more of a buy-to-own purchase option, a pay-to-pay purchase option, or a paid subscription purchase option;" [0019]

Neither Kennewick, Gold, nor Brachman teaches, but Ginter teaches:

"bill[ing] a user associated with the first utterance" col. 33 ln 1-35 col. 47 ln 20-55

It would have been obvious to combine Kennewick, Gold, Brachman and Ginter to produce a system for distributing digital content

10. Claims 21 and 22 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent Publication 20040193420 to Kennewick (hereafter "Kennewick") in view of US Patent Publication 20020032752 to Gold (hereafter "Gold"), US Patent 6704576 to

Brachman (hereafter "Brachman") and US Publication 20090276700 to Anderson (hereafter "Anderson").

Per claim 21,

Gold teaches:

"identify, based on the natural language processing and the interaction of the user with the media item, the media item as the content to dedicate to the recipient."[0036]

Gold does not teach, but Anderson teaches:

"determine that a user is interacting with a media item;" [0048][0067] and

It would have been obvious to combine Kennewick, Gold, Brachman and Anderson to produce a system for determining user status.

Per claim 22,

Gold teaches:

"identify, based on the natural language processing and the determination that the user provided the first utterance during the playing of the media item, the media item as the content to dedicate to the recipient".[0036]

Anderson teaches:

"determine a media item that is a user is playing;"[0048][0067]

"determine that the user provided the first utterance during the playing of the media item by the user;" [0048][0067]

Conclusion

11. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to TIM HALE whose telephone number is (571)270-3794. The examiner can normally be reached on Monday - Friday 9am - 5pm EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Calvin Hewitt II can be reached on (571)272-6709. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300. Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/CALVIN L HEWITT II/
Supervisory Patent Examiner, Art Unit 3685

Notice of References Cited	Application/Control No. 12/943,699	Applicant(s)/Patent Under Reexamination KENNEWICK ET AL.	
	Examiner TIM HALE	Art Unit 3685	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
*	A	US-2009/0276700	11-2009	Anderson et al.	715/700
*	B	US-6,704,576	03-2004	Brachman et al.	455/503
	C	US-			
	D	US-			
	E	US-			
	F	US-			
	G	US-			
	H	US-			
	I	US-			
	J	US-			
	K	US-			
	L	US-			
	M	US-			

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	
	V	
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	1A	5,027,406	06/25/1991	Roberts *et al.*	381	43	
	1B	5,208,748	05/04/1993	Flores *et al.*	364	419	
	1C	5,357,596	10/18/1994	Takebayashi *et al.*	395	2.84	
	1D	5,577,165	11/19/1996	Takebayashi et al.	395	2.84	
	1E	5,608,635	03/04/1997	Tamai	364	449.3	
	1F	5,740,256	04/14/1998	Castello Da Costa *et al.*	361	94.7	
	1G	5,748,974	05/05/1998	Johnson	395	759	
	1H	5,839,107	11/17/1998	Gupta *et al.*	704	270	
	1I	5,848,396	12/08/1998	Gerace	705	10	
	1J	5,855,000	12/29/1998	Waibel *et al.*	704	235	
	1K	5,897,613	04/27/1999	Chan	704	210	
	1L	5,960,397	09/28/1999	Rahim	704	244	
	1M	5,960,399	09/28/1999	Barclay *et al.*	704	270	
	1N	5,960,447	09/28/1999	Holt *et al.*	707	500	
	1O	5,995,928	11/30/1999	Nguyen *et al.*	704	251	
	1P	6,018,708	01/25/2000	Dahan *et al.*	704	244	
	1Q	6,035,267	03/07/2000	Watanabe *et al.*	704	1	
	1R	6,049,607	04/11/2000	Marash *et al.*	379	410	
	1S	6,067,513	05/23/2000	Ishimitsu	704	233	
	1T	6,076,059	06/13/2000	Glickman *et al.*	704	260	
	1U	6,085,186	07/04/2000	Christianson *et al.*	707	3	
	1V	6,108,631	08/22/2000	Ruhl	704	270	
	1W	6,167,377	12/26/2000	Gillick *et al.*	704	240	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO
	1X	1 320 043 A2	06/18/2003	Europe				
	1Y	1 646 037	04/12/2006	Europe				

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

	1Z	"Statement in Accordance with the Notice from the European Patent Office dated 1 October 2007 Concerning Business Methods (OJ EPO 11/2007, 592-593), XP002456252

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	2A	6,173,266	01/09/2001	Marx *et al.*	704	270	
	2B	6,188,982	02/13/2001	Chiang	704	256	
	2C	6,199,043	03/06/2001	Happ	704	272	
	2D	6,208,964	03/27/2001	Sabourin	704	244	
	2E	6,233,561	05/15/2001	Junqua *et al.*	704	277	
	2F	6,236,968	05/22/2001	Kanevsky *et al.*	704	275	
	2G	6,246,990	06/12/2001	Happ	704	275	
	2H	6,266,636	07/24/2001	Kosaka *et al.*	704	244	
	2I	6,269,336	07/31/2001	Ladd *et al.*	704	270	
	2J	6,275,231	08/14/2001	Obradovich	345	349	
	2K	6,278,968	08/21/2001	Franz *et al.*	704	3	
	2L	6,301,560	10/09/2001	Masters	704	251	
	2M	6,362,748	03/26/2002	Huang	340	901	
	2N	6,366,882	04/02/2002	Bijl *et al.*	704	235	
	2O	6,377,913	04/23/2002	Coffman *et al.*	704	8	
	2P	6,393,403	05/21/2002	Majaniemi	704	275	
	2Q	6,397,181	05/28/2002	Li *et al.*	704	256	
	2R	6,405,170	06/11/2002	Phillips *et al.*	704	270	
	2S	6,429,813	08/06/2002	Feigen	342	357.13	
	2T	6,434,529	08/13/2002	Walker *et al.*	704	275	
	2U	6,456,974	09/24/2002	Baker *et al.*	704	270.1	
	2V	6,510,417	01/21/2003	Woods *et al.*	704	275	
	2W	6,522,746	02/18/2003	Marchok *et al.*	379	406.03	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	3A	6,549,629	04/15/2003	Finn *et al.*	381	92	
	3B	6,560,590	05/06/2003	Shwe *et al.*	706	55	
	3C	6,571,279	05/27/2003	Herz *et al.*	709	217	
	3D	6,578,022	06/10/2003	Foulger *et al.*	706	45	
	3E	6,584,439	06/24/2003	Geilhufe *et al.*	704	270	
	3F	6,587,858	07/01/2003	Strazza	707	102	
	3G	6,591,239	07/08/2003	McCall *et al.*	704	275	
	3H	6,594,367	07/15/2003	Marash *et al.*	381	92	
	3I	6,601,026	07/29/2003	Appelt *et al.*	704	9	
	3J	6,606,598	08/12/2003	Holthouse *et al.*	704	275	
	3K	6,622,119	09/16/2003	Ramaswamy *et al.*	704	9	
	3L	6,631,351	10/07/2003	Ramachandran *et al.*	704	270	
	3M	6,633,846	10/14/2003	Bennett *et al.*	704	257	
	3N	6,681,206	01/20/2004	Gorin *et al.*	704	243	
	3O	6,701,294	03/02/2004	Ball *et al.*	704	257	
	3P	6,704,396	03/09/2004	Parolkar *et al.*	379	88.17	
	3Q	6,704,708	03/09/2004	Pickering	704	235	
	3R	6,745,161	06/01/2004	Arnold *et al.*	704	7	
	3S	6,751,612	06/15/2004	Schuetze *et al.*	707	4	
	3T	6,754,485	06/22/2004	Obradovich *et al.*	455	414.1	
	3U	6,757,544	06/29/2004	Rangarajan *et al.*	455	456.1	
	3V	6,832,230	12/14/2004	Zilliacus *et al.*	707	203	
	3W	6,850,603	02/01/2005	Eberle *et al.*	379	88.16	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	4A	6,868,380	03/15/2005	Kroeker	704	240	
	4B	6,877,001	04/05/2005	Wolf *et al.*	707	3	
	4C	6,910,003	06/21/2005	Arnold *et al.*	704	4	
	4D	6,912,498	06/28/2005	Stevens *et al.*	704	235	
	4E	6,934,756	08/23/2005	Maes	709	227	
	4F	6,954,755	10/11/2005	Reisman	707	10	
	4G	6,959,276	10/25/2005	Droppo *et al.*	704	226	
	4H	6,968,311	11/22/2005	Knockeart *et al.*	704	270	
	4I	6,975,993	12/13/2005	Keiller	704	275	
	4J	6,983,055	01/03/2006	Luo	381	313	
	4K	7,003,463	02/21/2006	Maes *et al.*	704	270.1	
	4L	7,024,364	04/04/2006	Guerra *et al.*	704	270	
	4M	7,027,586	04/11/2006	Bushey *et al.*	379	265.09	
	4N	7,035,415	04/25/2006	Belt *et al.*	381	92	
	4O	7,043,425	05/09/2006	Pao	704	211	
	4P	7,054,817	05/30/2006	Shao	704	270	
	4Q	7,058,890	06/06/2006	George *et al.*	715	728	
	4R	7,069,220	06/27/2006	Coffman *et al.*	704	275	
	4S	7,072,834	07/04/2006	Zhou	704	244	
	4T	7,107,210	09/12/2006	Deng *et al.*	704	226	
	4U	7,107,218	09/12/2006	Preston	704	270	
	4V	7,110,951	09/19/2006	Lemelson *et al.*	704	270	
	4W	7,130,390	10/31/2006	Abburi	379	88.17	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	5A	7,143,037	11/28/2006	Chestnut	704	251	
	5B	7,143,039	11/28/2006	Stifelman *et al.*	704	270	
	5C	7,165,028	01/16/2007	Gong	704	233	
	5D	7,184,957	02/27/2007	Brookes *et al.*	704	246	
	5E	7,190,770	03/13/2007	Ando *et al.*	379	88.01	
	5F	7,197,069	03/27/2007	Agazzi *et al.*	375	233	
	5G	7,206,418	04/17/2007	Yang *et al.*	381	92	
	5H	7,207,011	04/17/2007	Mulvey *et al.*	715	812	
	5I	7,236,923	06/26/2007	Gupta	704	9	
	5J	7,272,212	09/18/2007	Eberle *et al.*	379	88.17	
	5K	7,277,854	10/02/2007	Bennett *et al.*	704	257	
	5L	7,283,829	10/16/2007	Christenson *et al.*	455	461	
	5M	7,283,951	10/16/2007	Marchisio *et al.*	704	9	
	5N	7,299,186	11/20/2007	Kuzunuki *et al.*	704	270.1	
	5O	7,305,381	12/04/2007	Poppink *et al.*	1	1	
	5P	7,321,850	01/22/2008	Wakita	704	10	
	5Q	7,328,155	02/05/2008	Endo *et al.*	704	251	
	5R	7,337,116	02/26/2008	Charlesworth *et al.*	704	254	
	5S	7,366,285	04/29/2008	Parolkar *et al.*	379	88.17	
	5T	7,366,669	04/29/2008	Nishitani *et al.*	704	256	
	5U	7,406,421	07/29/2008	Odinak *et al.*	704	275	
	5V	7,415,414	08/19/2008	Azara *et al.*	704	270	
	5W	7,447,635	11/04/2008	Konopka *et al.*	704	275	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	6A	7,454,608	11/18/2008	Gopalakrishnan *et al.*	713	100	
	6B	7,472,020	12/30/2008	Brulle-Drews	701	211	
	6C	7,472,060	12/30/2008	Gorin *et al.*	704	240	
	6D	7,478,036	01/13/2009	Shen *et al.*	704	9	
	6E	7,493,259	02/17/2009	Jones *et al.*	704	257	
	6F	7,516,076	04/07/2009	Walker *et al.*	704	275	
	6G	7,529,675	05/05/2009	Maes	704	270.1	
	6H	7,536,297	05/19/2009	Byrd *et al.*	704	10	
	6I	7,536,374	05/19/2009	Au	706	55	
	6J	7,542,894	06/02/2009	Murata	704	9	
	6K	7,546,382	06/09/2009	Healey *et al.*	709	246	
	6L	7,558,730	07/07/2009	Davis *et al.*	704	235	
	6M	7,574,362	08/11/2009	Walker *et al.*	704	275	
	6N	7,577,244	08/18/2009	Taschereau	379	218.01	
	6O	7,606,708	10/20/2009	Hwang	704	257	
	6P	7,640,006	12/29/2009	Portman *et al.*	455	412.1	
	6Q	7,676,365	03/09/2010	Hwang *et al.*	704	240	
	6R	7,676,369	03/09/2010	Fujimoto *et al.*	704	270	
	6S	7,684,977	03/23/2010	Morikawa	704	211	
	6T	7,729,916	06/01/2010	Coffman *et al.*	704	270	
	6U	7,729,918	06/01/2010	Walker *et al.*	704	275	
	6V	7,729,920	06/01/2010	Chaar *et al.*	704	275	
	6W	7,748,021	06/29/2010	Obradovich	725	105	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	7A	7,788,084	08/31/2010	Brun *et al.*	704	7	
	7B	7,801,731	09/21/2010	Odinak *et al.*	704	275	
	7C	7,831,426	11/09/2010	Bennett	704	252	
	7D	7,831,433	11/09/2010	Belvin *et al.*	704	275	
	7E	7,873,519	01/18/2011	Bennett	704	257	
	7F	7,873,523	01/18/2011	Potter *et al.*	704	275	
	7G	7,881,936	02/01/2011	Longe *et al.*	704	257	
	7H	7,894,849	02/22/2011	Kass *et al.*	455	550.1	
	7I	7,902,969	03/08/2011	Obradovich	340	439	
	7J	7,917,367	03/29/2011	Di Cristo *et al.*	704	270.1	
	7K	7,920,682	04/05/2011	Byrne *et al.*	379	88.18	
	7L	7,949,529	05/24/2011	Weider *et al.*	704	270	
	7M	7,949,537	05/24/2011	Walker *et al.*	704	275	
	7N	7,953,732	05/31/2011	Frank *et al.*	707	724	
	7O	7,974,875	07/05/2011	Quilici *et al.*	705	14.4	
	7P	7,983,917	07/19/2011	Kennewick *et al.*	704	257	
	7Q	7,984,287	07/19/2011	Gopalakrishnan *et al.*	713	100	
	7R	8,015,006	09/06/2011	Kennewick *et al.*	704	236	
	7S	8,060,367	11/15/2011	Keaveney	704	247	
	7T	8,069,046	11/29/2011	Kennewick *et al.*	704	257	
	7U	8,073,681	12/06/2011	Baldwin *et al.*	704	9	
	7V	8,077,975	12/13/2011	Ma *et al.*	382	187	
	7W	8,082,153	12/20/2011	Coffman *et al.*	704	270	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	8A	8,086,463	12/27/2011	Ativanichayaphong *et al.*	704	275	
	8B	8,112,275	02/07/2012	Kennewick *et al.*	704	240	
	8C	8,140,327	03/20/2012	Kennewick *et al.*	704	226	
	8D	8,140,335	03/20/2012	Kennewick *et al.*	704	257	
	8E	8,145,489	03/27/2012	Freeman *et al.*	704	257	
	8F	8,150,694	04/03/2012	Kennewick *et al.*	704	257	
	8G	8,155,962	04/10/2012	Kennewick *et al.*	704	257	
	8H	8,170,867	05/01/2012	Germain	704	9	
	8I	8,195,468	06/05/2012	Weider *et al.*	704	275	
	8J	8,219,599	07/10/2012	Tunstall-Pedoe	707	955	
	8K	8,224,652	07/17/2012	Wang *et al.*	704	275	
	8L	8,326,627	12/04/2012	Kennewick *et al.*	704	257	
	8M	8,326,634	12/04/2012	Di Cristo *et al.*	704	270.1	
	8N	8,326,637	12/04/2012	Baldwin *et al.*	704	275	
	8O	8,332,224	12/11/2012	Di Cristo *et al.*	704	257	
	8P	2001/0049601	12/06/2001	Kroeker *et al.*	704	254	
	8Q	2001/0054087	12/20/2001	Flom *et al.*	709	218	
	8R	2002/0015500	02/07/2002	Belt *et al.*	381	66	
	8S	2002/0022927	02/21/2002	Lemelson *et al.*	701	301	
	8T	2002/0035501	03/21/2002	Handel *et al.*	705	10	
	8U	2002/0069071	06/06/2002	Knockeart et al.	704	275	
	8V	2002/0087326	07/04/2002	Lee *et al.*	704	270.1	
	8W	2002/0133402	09/19/2002	Faber *et al.*	705	14	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	9A	2002/0161646	10/31/2002	Gailey *et al.*	705	14	
	9B	2002/0173961	11/21/2002	Guerra	704	258	
	9C	2002/0184373	12/05/2002	Maes	709	228	
	9D	2003/0016835	01/23/2003	Elko *et al.*	381	92	
	9E	2003/0065427	04/03/2003	Funk *et al.*	701	1	
	9F	2003/0110037	06/12/2003	Walker *et al.*	704	257	
	9G	2003/0115062	06/19/2003	Walker *et al.*	704	258	
	9H	2003/0120493	06/26/2003	Gupta	704	270.1	
	9I	2003/0144846	07/31/2003	Denenberg *et al.*	704	277	
	9J	2003/0158731	08/21/2003	Falcon *et al.*	704	231	
	9K	2003/0161448	08/28/2003	Parolkar *et al.*	379	88.17	
	9L	2003/0204492	10/30/2003	Wolf *et al.*	707	3	
	9M	2003/0206640	11/06/2003	Malvar *et al.*	381	93	
	9N	2003/0212550	11/13/2003	Ubale	704	215	
	9O	2003/0212558	11/13/2003	Matula	704	260	
	9P	2003/0212562	11/13/2003	Patel *et al.*	704	275	
	9Q	2003/0225825	12/04/2003	Healey *et al.*	709	203	
	9R	2004/0098245	05/20/2004	Walker *et al.*	704	1	
	9S	2004/0117804	06/17/2004	Scahill *et al.*	719	320	
	9T	2004/0140989	07/22/2004	Papageorge	345	700	
	9U	2004/0167771	08/26/2004	Duan *et al.*	704	10	
	9V	2004/0205671	10/14/2004	Sukehiro *et al.*	715	532	
	9W	2005/0021470	01/27/2005	Martin *et al.*	705	51	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	10A	2005/0125232	06/09/2005	Gadd	704	270.1	
	10B	2005/0137877	06/23/2005	Oesterling *et al.*	704	275	
	10C	2005/0143994	06/30/2005	Mori *et al.*	704	235	
	10D	2005/0216254	09/29/2005	Gupta *et al.*	704	9	
	10E	2005/0234727	10/20/2005	Chiu	704	270.1	
	10F	2005/0283752	12/22/2005	Fruchter *et al.*	717	100	
	10G	2006/0041431	02/23/2006	Maes	704	270.1	
	10H	2006/0047509	03/02/2006	Ding *et al.*	704	231	
	10I	2006/0217133	09/28/2006	Christenson *et al.*	455	461	
	10J	2006/0285662	12/21/2006	Yin *et al.*	379	88.16	
	10K	2007/0033020	02/08/2007	(Kelleher) Francois *et al.*	704	226	
	10L	2007/0043868	02/22/2007	Kumar *et al.*	709	226	
	10M	2007/0073544	03/29/2007	Millett *et al.*	704	277	
	10N	2007/0078708	04/05/2007	Yu *et al.*	705	14	
	10O	2007/0078709	04/05/2007	Rajaram	705	14	
	10P	2007/0118357	05/24/2007	Kasravi *et al.*	704	10	
	10Q	2007/0135101	06/14/2007	Ramati *et al.*	455	414.1	
	10R	2007/0162296	07/12/2007	Altberg *et al.*	705	1	
	10S	2007/0198267	08/23/2007	Jones *et al.*	704	257	
	10T	2008/0034032	02/07/2008	Healey *et al.*	709	203	
	10U	2008/0065386	03/13/2008	Cross *et al.*	704	270	
	10V	2008/0109285	05/08/2008	Reuther *et al.*	705	7	
	10W	2008/0147410	06/19/2008	Odinak	704	270.1	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	11A	2009/0052635	02/26/2009	Jones *et al.*	379	88.02	
	11B	2009/0067599	03/12/2009	Agarwal *et al.*	379	114.21	
	11C	2009/0106029	04/23/2009	DeLine *et al.*	704	275	
	11D	2009/0271194	10/29/2009	Davis *et al.*	704	235	
	11E	2009/0273563	11/05/2009	Pryor	345	157	
	11F	2009/0313026	12/17/2009	Coffman *et al.*	704	275	
	11G	2010/0029261	02/04/2010	Mikkelsen *et al.*	455	419	
	11H	2010/0185512	07/22/2010	Borger *et al.*	705	14.49	
	11I	2010/0312566	12/09/2010	Odinak *et al.*	704	275	
	11J	2011/0112827	05/12/2011	Kennewick *et al.*	704	9	
	11K	2011/0131036	06/02/2011	Di Cristo *et al.*	704	9	
	11L	2011/0131045	06/02/2011	Cristo *et al.*	704	249	
	11M	2011/0231182	09/22/2011	Weider *et al.*	704	9	
	11N	2011/0231188	09/22/2011	Kennewick *et al.*	704	236	
	11O	2012/0022857	01/26/2012	Baldwin *et al.*	704	9	
	11P	2012/0101809	04/26/2012	Kennewick *et al.*	704	9	
	11Q	2012/0101810	04/26/2012	Kennewick *et al.*	704	9	
	11R	2012/0109753	05/03/2012	Kennewick *et al.*	705	14.58	
	11S	2012/0150636	06/14/2012	Freeman *et al.*	705	14.49	
	11T	2012/0278073	11/01/2012	Weider et al.	704	235	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED
/Tim Hale/	02/25/2013

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	1A	5,027,406	06/25/1991	Roberts *et al.*	381	43	
	1B	5,208,748	05/04/1993	Flores *et al.*	364	419	
	1C	5,357,596	10/18/1994	Takebayashi *et al.*	395	2.84	
	1D	5,577,165	11/19/1996	Takebayashi et al.	395	2.84	
	1E	5,608,635	03/04/1997	Tamai	364	449.3	
	1F	5,740,256	04/14/1998	Castello Da Costa *et al.*	361	94.7	
	1G	5,748,974	05/05/1998	Johnson	395	759	
	1H	5,839,107	11/17/1998	Gupta *et al.*	704	270	
	1I	5,848,396	12/08/1998	Gerace	705	10	
	1J	5,855,000	12/29/1998	Waibel *et al.*	704	235	
	1K	5,897,613	04/27/1999	Chan	704	210	
	1L	5,960,397	09/28/1999	Rahim	704	244	
	1M	5,960,399	09/28/1999	Barclay *et al.*	704	270	
	1N	5,960,447	09/28/1999	Holt *et al.*	707	500	
	1O	5,995,928	11/30/1999	Nguyen *et al.*	704	251	
	1P	6,018,708	01/25/2000	Dahan *et al.*	704	244	
	1Q	6,035,267	03/07/2000	Watanabe *et al.*	704	1	
	1R	6,049,607	04/11/2000	Marash *et al.*	379	410	
	1S	6,067,513	05/23/2000	Ishimitsu	704	233	
	1T	6,076,059	06/13/2000	Glickman *et al.*	704	260	
	1U	6,085,186	07/04/2000	Christianson *et al.*	707	3	
	1V	6,108,631	08/22/2000	Ruhl	704	270	
	1W	6,167,377	12/26/2000	Gillick *et al.*	704	240	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO
	1X	1 320 043 A2	06/18/2003	Europe				
	1Y	1 646 037	04/12/2006	Europe				

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

	1Z	"Statement in Accordance with the Notice from the European Patent Office dated 1 October 2007 Concerning Business Methods (OJ EPO 11/2007, 592-593), XP002456252

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	2A	6,173,266	01/09/2001	Marx *et al.*	704	270	
	2B	6,188,982	02/13/2001	Chiang	704	256	
	2C	6,199,043	03/06/2001	Happ	704	272	
	2D	6,208,964	03/27/2001	Sabourin	704	244	
	2E	6,233,561	05/15/2001	Junqua *et al.*	704	277	
	2F	6,236,968	05/22/2001	Kanevsky *et al.*	704	275	
	2G	6,246,990	06/12/2001	Happ	704	275	
	2H	6,266,636	07/24/2001	Kosaka *et al.*	704	244	
	2I	6,269,336	07/31/2001	Ladd *et al.*	704	270	
	2J	6,275,231	08/14/2001	Obradovich	345	349	
	2K	6,278,968	08/21/2001	Franz *et al.*	704	3	
	2L	6,301,560	10/09/2001	Masters	704	251	
	2M	6,362,748	03/26/2002	Huang	340	901	
	2N	6,366,882	04/02/2002	Bijl *et al.*	704	235	
	2O	6,377,913	04/23/2002	Coffman *et al.*	704	8	
	2P	6,393,403	05/21/2002	Majaniemi	704	275	
	2Q	6,397,181	05/28/2002	Li *et al.*	704	256	
	2R	6,405,170	06/11/2002	Phillips *et al.*	704	270	
	2S	6,429,813	08/06/2002	Feigen	342	357.13	
	2T	6,434,529	08/13/2002	Walker *et al.*	704	275	
	2U	6,456,974	09/24/2002	Baker *et al.*	704	270.1	
	2V	6,510,417	01/21/2003	Woods *et al.*	704	275	
	2W	6,522,746	02/18/2003	Marchok *et al.*	379	406.03	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	3A	6,549,629	04/15/2003	Finn *et al.*	381	92	
	3B	6,560,590	05/06/2003	Shwe *et al.*	706	55	
	3C	6,571,279	05/27/2003	Herz *et al.*	709	217	
	3D	6,578,022	06/10/2003	Foulger *et al.*	706	45	
	3E	6,584,439	06/24/2003	Geilhufe *et al.*	704	270	
	3F	6,587,858	07/01/2003	Strazza	707	102	
	3G	6,591,239	07/08/2003	McCall *et al.*	704	275	
	3H	6,594,367	07/15/2003	Marash *et al.*	381	92	
	3I	6,601,026	07/29/2003	Appelt *et al.*	704	9	
	3J	6,606,598	08/12/2003	Holthouse *et al.*	704	275	
	3K	6,622,119	09/16/2003	Ramaswamy *et al.*	704	9	
	3L	6,631,351	10/07/2003	Ramachandran *et al.*	704	270	
	3M	6,633,846	10/14/2003	Bennett *et al.*	704	257	
	3N	6,681,206	01/20/2004	Gorin *et al.*	704	243	
	3O	6,701,294	03/02/2004	Ball *et al.*	704	257	
	3P	6,704,396	03/09/2004	Parolkar *et al.*	379	88.17	
	3Q	6,704,708	03/09/2004	Pickering	704	235	
	3R	6,745,161	06/01/2004	Arnold *et al.*	704	7	
	3S	6,751,612	06/15/2004	Schuetze *et al.*	707	4	
	3T	6,754,485	06/22/2004	Obradovich *et al.*	455	414.1	
	3U	6,757,544	06/29/2004	Rangarajan *et al.*	455	456.1	
	3V	6,832,230	12/14/2004	Zilliacus *et al.*	707	203	
	3W	6,850,603	02/01/2005	Eberle *et al.*	379	88.16	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	4A	6,868,380	03/15/2005	Kroeker	704	240	
	4B	6,877,001	04/05/2005	Wolf *et al.*	707	3	
	4C	6,910,003	06/21/2005	Arnold *et al.*	704	4	
	4D	6,912,498	06/28/2005	Stevens *et al.*	704	235	
	4E	6,934,756	08/23/2005	Maes	709	227	
	4F	6,954,755	10/11/2005	Reisman	707	10	
	4G	6,959,276	10/25/2005	Droppo *et al.*	704	226	
	4H	6,968,311	11/22/2005	Knockeart *et al.*	704	270	
	4I	6,975,993	12/13/2005	Keiller	704	275	
	4J	6,983,055	01/03/2006	Luo	381	313	
	4K	7,003,463	02/21/2006	Maes *et al.*	704	270.1	
	4L	7,024,364	04/04/2006	Guerra *et al.*	704	270	
	4M	7,027,586	04/11/2006	Bushey *et al.*	379	265.09	
	4N	7,035,415	04/25/2006	Belt *et al.*	381	92	
	4O	7,043,425	05/09/2006	Pao	704	211	
	4P	7,054,817	05/30/2006	Shao	704	270	
	4Q	7,058,890	06/06/2006	George *et al.*	715	728	
	4R	7,069,220	06/27/2006	Coffman *et al.*	704	275	
	4S	7,072,834	07/04/2006	Zhou	704	244	
	4T	7,107,210	09/12/2006	Deng *et al.*	704	226	
	4U	7,107,218	09/12/2006	Preston	704	270	
	4V	7,110,951	09/19/2006	Lemelson *et al.*	704	270	
	4W	7,130,390	10/31/2006	Abburi	379	88.17	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	5A	7,143,037	11/28/2006	Chestnut	704	251	
	5B	7,143,039	11/28/2006	Stifelman *et al.*	704	270	
	5C	7,165,028	01/16/2007	Gong	704	233	
	5D	7,184,957	02/27/2007	Brookes *et al.*	704	246	
	5E	7,190,770	03/13/2007	Ando *et al.*	379	88.01	
	5F	7,197,069	03/27/2007	Agazzi *et al.*	375	233	
	5G	7,206,418	04/17/2007	Yang *et al.*	381	92	
	5H	7,207,011	04/17/2007	Mulvey *et al.*	715	812	
	5I	7,236,923	06/26/2007	Gupta	704	9	
	5J	7,272,212	09/18/2007	Eberle *et al.*	379	88.17	
	5K	7,277,854	10/02/2007	Bennett *et al.*	704	257	
	5L	7,283,829	10/16/2007	Christenson *et al.*	455	461	
	5M	7,283,951	10/16/2007	Marchisio *et al.*	704	9	
	5N	7,299,186	11/20/2007	Kuzunuki *et al.*	704	270.1	
	5O	7,305,381	12/04/2007	Poppink *et al.*	1	1	
	5P	7,321,850	01/22/2008	Wakita	704	10	
	5Q	7,328,155	02/05/2008	Endo *et al.*	704	251	
	5R	7,337,116	02/26/2008	Charlesworth *et al.*	704	254	
	5S	7,366,285	04/29/2008	Parolkar *et al.*	379	88.17	
	5T	7,366,669	04/29/2008	Nishitani *et al.*	704	256	
	5U	7,406,421	07/29/2008	Odinak *et al.*	704	275	
	5V	7,415,414	08/19/2008	Azara *et al.*	704	270	
	5W	7,447,635	11/04/2008	Konopka *et al.*	704	275	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	6A	7,454,608	11/18/2008	Gopalakrishnan *et al.*	713	100	
	6B	7,472,020	12/30/2008	Brulle-Drews	701	211	
	6C	7,472,060	12/30/2008	Gorin *et al.*	704	240	
	6D	7,478,036	01/13/2009	Shen *et al.*	704	9	
	6E	7,493,259	02/17/2009	Jones *et al.*	704	257	
	6F	7,516,076	04/07/2009	Walker *et al.*	704	275	
	6G	7,529,675	05/05/2009	Maes	704	270.1	
	6H	7,536,297	05/19/2009	Byrd *et al.*	704	10	
	6I	7,536,374	05/19/2009	Au	706	55	
	6J	7,542,894	06/02/2009	Murata	704	9	
	6K	7,546,382	06/09/2009	Healey *et al.*	709	246	
	6L	7,558,730	07/07/2009	Davis *et al.*	704	235	
	6M	7,574,362	08/11/2009	Walker *et al.*	704	275	
	6N	7,577,244	08/18/2009	Taschereau	379	218.01	
	6O	7,606,708	10/20/2009	Hwang	704	257	
	6P	7,640,006	12/29/2009	Portman *et al.*	455	412.1	
	6Q	7,676,365	03/09/2010	Hwang *et al.*	704	240	
	6R	7,676,369	03/09/2010	Fujimoto *et al.*	704	270	
	6S	7,684,977	03/23/2010	Morikawa	704	211	
	6T	7,729,916	06/01/2010	Coffman *et al.*	704	270	
	6U	7,729,918	06/01/2010	Walker *et al.*	704	275	
	6V	7,729,920	06/01/2010	Chaar *et al.*	704	275	
	6W	7,748,021	06/29/2010	Obradovich	725	105	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	7A	7,788,084	08/31/2010	Brun *et al.*	704	7	
	7B	7,801,731	09/21/2010	Odinak *et al.*	704	275	
	7C	7,831,426	11/09/2010	Bennett	704	252	
	7D	7,831,433	11/09/2010	Belvin *et al.*	704	275	
	7E	7,873,519	01/18/2011	Bennett	704	257	
	7F	7,873,523	01/18/2011	Potter *et al.*	704	275	
	7G	7,881,936	02/01/2011	Longe *et al.*	704	257	
	7H	7,894,849	02/22/2011	Kass *et al.*	455	550.1	
	7I	7,902,969	03/08/2011	Obradovich	340	439	
	7J	7,917,367	03/29/2011	Di Cristo *et al.*	704	270.1	
	7K	7,920,682	04/05/2011	Byrne *et al.*	379	88.18	
	7L	7,949,529	05/24/2011	Weider *et al.*	704	270	
	7M	7,949,537	05/24/2011	Walker *et al.*	704	275	
	7N	7,953,732	05/31/2011	Frank *et al.*	707	724	
	7O	7,974,875	07/05/2011	Quilici *et al.*	705	14.4	
	7P	7,983,917	07/19/2011	Kennewick *et al.*	704	257	
	7Q	7,984,287	07/19/2011	Gopalakrishnan *et al.*	713	100	
	7R	8,015,006	09/06/2011	Kennewick *et al.*	704	236	
	7S	8,060,367	11/15/2011	Keaveney	704	247	
	7T	8,069,046	11/29/2011	Kennewick *et al.*	704	257	
	7U	8,073,681	12/06/2011	Baldwin *et al.*	704	9	
	7V	8,077,975	12/13/2011	Ma *et al.*	382	187	
	7W	8,082,153	12/20/2011	Coffman *et al.*	704	270	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	8A	8,086,463	12/27/2011	Ativanichayaphong *et al.*	704	275	
	8B	8,112,275	02/07/2012	Kennewick *et al.*	704	240	
	8C	8,140,327	03/20/2012	Kennewick *et al.*	704	226	
	8D	8,140,335	03/20/2012	Kennewick *et al.*	704	257	
	8E	8,145,489	03/27/2012	Freeman *et al.*	704	257	
	8F	8,150,694	04/03/2012	Kennewick *et al.*	704	257	
	8G	8,155,962	04/10/2012	Kennewick *et al.*	704	257	
	8H	8,170,867	05/01/2012	Germain	704	9	
	8I	8,195,468	06/05/2012	Weider *et al.*	704	275	
	8J	8,219,599	07/10/2012	Tunstall-Pedoe	707	955	
	8K	8,224,652	07/17/2012	Wang *et al.*	704	275	
	8L	8,326,627	12/04/2012	Kennewick *et al.*	704	257	
	8M	8,326,634	12/04/2012	Di Cristo *et al.*	704	270.1	
	8N	8,326,637	12/04/2012	Baldwin *et al.*	704	275	
	8O	8,332,224	12/11/2012	Di Cristo *et al.*	704	257	
	8P	2001/0049601	12/06/2001	Kroeker *et al.*	704	254	
	8Q	2001/0054087	12/20/2001	Flom *et al.*	709	218	
	8R	2002/0015500	02/07/2002	Belt *et al.*	381	66	
	8S	2002/0022927	02/21/2002	Lemelson *et al.*	701	301	
	8T	2002/0035501	03/21/2002	Handel *et al.*	705	10	
	8U	2002/0069071	06/06/2002	Knockeart et al.	704	275	
	8V	2002/0087326	07/04/2002	Lee *et al.*	704	270.1	
	8W	2002/0133402	09/19/2002	Faber *et al.*	705	14	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	9A	2002/0161646	10/31/2002	Gailey *et al.*	705	14	
	9B	2002/0173961	11/21/2002	Guerra	704	258	
	9C	2002/0184373	12/05/2002	Maes	709	228	
	9D	2003/0016835	01/23/2003	Elko *et al.*	381	92	
	9E	2003/0065427	04/03/2003	Funk *et al.*	701	1	
	9F	2003/0110037	06/12/2003	Walker *et al.*	704	257	
	9G	2003/0115062	06/19/2003	Walker *et al.*	704	258	
	9H	2003/0120493	06/26/2003	Gupta	704	270.1	
	9I	2003/0144846	07/31/2003	Denenberg *et al.*	704	277	
	9J	2003/0158731	08/21/2003	Falcon *et al.*	704	231	
	9K	2003/0161448	08/28/2003	Parolkar *et al.*	379	88.17	
	9L	2003/0204492	10/30/2003	Wolf *et al.*	707	3	
	9M	2003/0206640	11/06/2003	Malvar *et al.*	381	93	
	9N	2003/0212550	11/13/2003	Ubale	704	215	
	9O	2003/0212558	11/13/2003	Matula	704	260	
	9P	2003/0212562	11/13/2003	Patel *et al.*	704	275	
	9Q	2003/0225825	12/04/2003	Healey *et al.*	709	203	
	9R	2004/0098245	05/20/2004	Walker *et al.*	704	1	
	9S	2004/0117804	06/17/2004	Scahill *et al.*	719	320	
	9T	2004/0140989	07/22/2004	Papageorge	345	700	
	9U	2004/0167771	08/26/2004	Duan *et al.*	704	10	
	9V	2004/0205671	10/14/2004	Sukehiro *et al.*	715	532	
	9W	2005/0021470	01/27/2005	Martin *et al.*	705	51	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	10A	2005/0125232	06/09/2005	Gadd	704	270.1	
	10B	2005/0137877	06/23/2005	Oesterling *et al.*	704	275	
	10C	2005/0143994	06/30/2005	Mori *et al.*	704	235	
	10D	2005/0216254	09/29/2005	Gupta *et al.*	704	9	
	10E	2005/0234727	10/20/2005	Chiu	704	270.1	
	10F	2005/0283752	12/22/2005	Fruchter *et al.*	717	100	
	10G	2006/0041431	02/23/2006	Maes	704	270.1	
	10H	2006/0047509	03/02/2006	Ding *et al.*	704	231	
	10I	2006/0217133	09/28/2006	Christenson *et al.*	455	461	
	10J	2006/0285662	12/21/2006	Yin *et al.*	379	88.16	
	10K	2007/0033020	02/08/2007	(Kelleher) Francois *et al.*	704	226	
	10L	2007/0043868	02/22/2007	Kumar *et al.*	709	226	
	10M	2007/0073544	03/29/2007	Millett *et al.*	704	277	
	10N	2007/0078708	04/05/2007	Yu *et al.*	705	14	
	10O	2007/0078709	04/05/2007	Rajaram	705	14	
	10P	2007/0118357	05/24/2007	Kasravi *et al.*	704	10	
	10Q	2007/0135101	06/14/2007	Ramati *et al.*	455	414.1	
	10R	2007/0162296	07/12/2007	Altberg *et al.*	705	1	
	10S	2007/0198267	08/23/2007	Jones *et al.*	704	257	
	10T	2008/0034032	02/07/2008	Healey *et al.*	709	203	
	10U	2008/0065386	03/13/2008	Cross *et al.*	704	270	
	10V	2008/0109285	05/08/2008	Reuther *et al.*	705	7	
	10W	2008/0147410	06/19/2008	Odinak	704	270.1	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	11A	2009/0052635	02/26/2009	Jones *et al.*	379	88.02	
	11B	2009/0067599	03/12/2009	Agarwal *et al.*	379	114.21	
	11C	2009/0106029	04/23/2009	DeLine *et al.*	704	275	
	11D	2009/0271194	10/29/2009	Davis *et al.*	704	235	
	11E	2009/0273563	11/05/2009	Pryor	345	157	
	11F	2009/0313026	12/17/2009	Coffman *et al.*	704	275	
	11G	2010/0029261	02/04/2010	Mikkelsen *et al.*	455	419	
	11H	2010/0185512	07/22/2010	Borger *et al.*	705	14.49	
	11I	2010/0312566	12/09/2010	Odinak *et al.*	704	275	
	11J	2011/0112827	05/12/2011	Kennewick *et al.*	704	9	
	11K	2011/0131036	06/02/2011	Di Cristo *et al.*	704	9	
	11L	2011/0131045	06/02/2011	Cristo *et al.*	704	249	
	11M	2011/0231182	09/22/2011	Weider *et al.*	704	9	
	11N	2011/0231188	09/22/2011	Kennewick *et al.*	704	236	
	11O	2012/0022857	01/26/2012	Baldwin *et al.*	704	9	
	11P	2012/0101809	04/26/2012	Kennewick *et al.*	704	9	
	11Q	2012/0101810	04/26/2012	Kennewick *et al.*	704	9	
	11R	2012/0109753	05/03/2012	Kennewick *et al.*	705	14.58	
	11S	2012/0150636	06/14/2012	Freeman *et al.*	705	14.49	
	11T	2012/0278073	11/01/2012	Weider et al.	704	235	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED
/Tim Hale/	02/25/2013

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

Electronic Patent Application Fee Transmittal

Application Number:	12943699
Filing Date:	10-Nov-2010
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Filer:	Ngai Bing Zhang/Miriam Nelsen Awan
Attorney Docket Number:	090611-0384422

Filed as Small Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				
Extension - 1 month with $0 paid	2251	1	75	75

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
Submission- Information Disclosure Stmt	1806	1	180	180
		Total in USD ($)		**255**

Electronic Acknowledgement Receipt	
EFS ID:	14831382
Application Number:	12943699
International Application Number:	
Confirmation Number:	4185
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Customer Number:	909
Filer:	Ngai Bing Zhang/Miriam Nelsen Awan
Filer Authorized By:	Ngai Bing Zhang
Attorney Docket Number:	090611-0384422
Receipt Date:	30-JAN-2013
Filing Date:	10-NOV-2010
Time Stamp:	14:50:01
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	$255
RAM confirmation Number	1274
Deposit Account	033975
Authorized User	
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: Charge any Additional Fees required under 37 C.F.R. Section 1.21 (Miscellaneous fees and charges)	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Fee Worksheet (SB06)	fee-info.pdf	32589 8ff15b672d39b806f3ddd5763eb6e1d3c780c7f8	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	32589		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

IN THE UNITED STATES PATENT & TRADEMARK OFFICE

IN RE PATENT APPLICATION OF:	Mike KENNEWICK *et al.*
SERIAL NO.:	12/943,699
FILING DATE:	November 10, 2010
ATTORNEY DOCKET NO.:	090611-0384422
CONFIRMATION NO.:	4185
EXAMINER:	Tim B. Hale
ART UNIT:	3685
FOR:	System and Method for Providing A Natural Language Content Dedication Service

RESPONSE TO NON-FINAL OFFICE ACTION mailed 01 OCTOBER 2012

MAIL STOP AMENDMENT
Commissioner for Patents
P.O. Box 1450
Alexandria, VA. 22313-1450

Dear Sir:

In response to the Non-Final Office Action mailed **October 1, 2012**, please amend the above- identified application as follows:

Amendments to the Claims begin on page **2** of this paper.

Remarks begin on page **9** of this paper.

It is not believed that extensions of time or fees for net addition of claims are required beyond those that may otherwise be provided for in documents accompanying this paper. However, if additional extensions of time are necessary to prevent abandonment of this application, then such extensions of time are hereby petitioned for under 37 C.F.R. § 1.136(a), and any fees required therefore (including fees for net addition of claims) are hereby authorized to be charged to Deposit Account No. 033975 (**Ref. No. 090611-0384422**).

This listing of claims will replace all prior versions and listings of claims in the Application.

LISTING OF CLAIMS:

1. (**Currently Amended**) A system for providing a natural language content dedication service, comprising:

~~an electronic device configured to~~ one or more processors; and

one or more non-transitory computer readable mediums storing executable instructions that when executed by the one or more processors cause the one or more processors to:

receive a first utterance that includes a natural language utterance; ~~a multi-modal interaction that includes a request to dedicate content, wherein the multi-modal interaction includes a natural language utterance relating to the request to dedicate the content; and~~

~~capture an utterance dedicating the content for a recipient; and~~

~~a content dedication system in communication with the electronic device through a message interface, wherein the content dedication system is configured to:~~

determine, based on processing of the first utterance by a speech recognition engine, one or more words or phrases of the first utterance;

identify, based on a natural language processing of the one or more words or phrases of the first utterance, content to dedicate to a recipient;

initiate, based on the identified content, a dedication to the recipient;

receive a second utterance to be associated with the dedication; and

~~identify the content to dedicate for the recipient based on an interpretation of the natural language utterance included in the multi-modal interaction;~~

~~process a transaction for the identified content to dedicate for the recipient; and~~

send ~~a content dedication message to the recipient, wherein the content dedication message includes~~ information to enable ~~for~~ the recipient to access the content and the second utterance. ~~, and wherein the content dedication message further includes~~

~~information for the recipient to access the utterance dedicating the content for the recipient.~~

2-4. (**Canceled**)

5. (**Currently Amended**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to ~~wherein the tags dedicating the content for the recipient include one or more of an image or a picture to insert within album art metadata for the content~~ provide an image as one or more tags within metadata of the content.

6. (**Currently Amended**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to ~~the content dedication system is further configured to~~ annotate the content with the second utterance. ~~utterance dedicating the content for the recipient.~~

7. (**Currently Amended**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to ~~the content dedication system is further configured to~~ purchase the ~~identified~~ content from a content provider ~~to process the transaction for the identified content~~ to enable the recipient to access the content.

8. (**Currently Amended**) The system of claim 7, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to: ~~the content dedication system is further configured to:~~

identify, for the dedication, a purchase option ~~for the request to dedicate the content, wherein the purchase option~~ that includes one or more of a buy-to-own purchase option, a pay-to-pay purchase option, or a paid subscription purchase option; and

bill a user associated with the first utterance ~~of the electronic device for the natural~~

~~content dedication service~~ based on ~~identified~~ the purchase option.

9. (**Currently Amended**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to: ~~the content dedication message sent to the recipient includes an electronic link for the recipient to access the content and the utterance dedicating the content for the recipient.~~

send, to the recipient, an electronic link to access the content as the information to enable the recipient to access the content; and

send, to the recipient, the second utterance in an encoded audio format as the information to enable the recipient to access the second utterance.

10. (**Original**) The system of claim 1, wherein the content includes one or more of a song or a video.

11. (**Currently Amended**) A method for providing a natural language content dedication service, comprising:

receiving, by one or more processors, a first utterance that includes a natural language utterance; ~~at an electronic device, a multi-modal interaction that includes a request to dedicate content, wherein the multi-modal interaction includes a natural language utterance relating to the request to dedicate the content;~~

~~capturing an utterance dedicating the content for a recipient;~~

determining, by the one or more processors, one or more words or phrases of the first utterance based on processing of the first utterance by a speech recognition engine;

identifying, by the one or more processors, content to dedicate to a recipient based on a natural language processing of the one or more words or phrases of the first utterance;

initiating, by the one or more processors, a dedication to the recipient based on the identified content;

receiving, by the one or more processors, a second utterance to be associated with the dedication; and

~~identifying the content to dedicate for the recipient based on an interpretation of the natural language utterance included in the multi-modal interaction;~~

~~processing at a content dedication system, a transaction for the identified content to dedicate for the recipient; and~~

sending, by the one or more processor, ~~a content dedication message from the content dedication system to the recipient, wherein the content dedication message includes~~ information to enable ~~for~~ the recipient to access the content and the second utterance. ~~, and wherein the content dedication message further includes information for the recipient to access the utterance dedicating the content for the recipient.~~

12-14. (**Canceled**)

15. (**Currently Amended**) The method of claim 11, further comprising providing an image as one or more tags within metadata of the content. ~~wherein the tags dedicating the~~

~~content for the recipient include one or more of an image or a picture to insert within art metadata for the content.~~

16. (**Currently Amended**) The method of claim 11, further comprising annotating the content with the second utterance. ~~utterance dedicating the content for the recipient.~~

17. (**Currently Amended**) The method of claim 11, further comprising purchasing~~, by the content dedication system,~~ the ~~identified~~ content from a content provider ~~to process the transaction for the identified content~~ to enable the recipient to access the content.

18. (**Currently Amended**) The method of claim 17, further comprising:

identifying, ~~at the content dedication system,~~ for the dedication, a purchase option ~~for the request to dedicate the content, wherein the purchase option~~ that includes one or more of a buy-to-own purchase option, a pay-to-pay purchase option, or a paid subscription purchase option; and

billing~~, by the content dedication system,~~ a user associated with the first utterance ~~of the electronic device for the natural content dedication service~~ based on ~~identified~~ the purchase option.

19. (**Currently Amended**) The method of claim 11, further comprising: ~~wherein the content dedication message sent to the recipient includes an electronic link for the recipient to access the content and the utterance dedicating the content for the recipient.~~

sending, to the recipient, an electronic link to access the content as the information to enable the recipient to access the content; and

sending, to the recipient, the second utterance in an encoded audio format as the information to enable the recipient to access the second utterance.

20. (**Original**) The method of claim 11, wherein the content includes one or more of a song or a video.

21. (**New**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to:

determine that a user is interacting with a media item; and

identify, based on the natural language processing and the interaction of the user with the media item, the media item as the content to dedicate to the recipient.

22. (**New**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to:

determine a media item that is a user is playing;

determine that the user provided the first utterance during the playing of the media item by the user; and

identify, based on the natural language processing and the determination that the user provided the first utterance during the playing of the media item, the media item as the content to dedicate to the recipient.

23. (**New**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to:

determine, based on processing of the second utterance by the speech recognition engine, one or more words or phrases of the second utterance; and

provide the one or more words or phrases of the second utterance as textual annotations within metadata of the content.

24. (**New**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to provide an electronic link associated with the content as one or more tags within metadata of the content.

25. (**New**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to provide a voice tag as

one or more tags within metadata of the content.

26. (**New**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to provide the second utterance within the content.

REMARKS

In response to the Non-Final Office Action mailed **October 1, 2012** (hereinafter "Office Action"), claims 1, 5-9, 11, and 15-19 have been amended, and claims 2-4 and 12-14 have been canceled without prejudice or disclaimer. Claims 21-26 have been newly added. Therefore, claims 1, 5, 6-11, and 15-26 are pending. Support for the instant amendments is provided throughout the as-filed Specification. Thus, no new matter has been added. In view of the foregoing amendments and following comments, allowance of all the claims pending in the application is respectfully requested.

INFORMATION DISCLOSURE STATEMENT

Applicants are submitting herewith a Supplemental Information Disclosure Statement and respectfully request that the Examiner consider the cited references and provide a signed copy of the Form PTO-1449 for this submission with the next Office Action.

EXAMINER INTERVIEW

Applicants thank the Examiner for the courtesies extended to Applicants' representative (Ngai Zhang) during the telephonic interviews conducted on January 15, 2013 and on January 23, 2013. During the interviews, the subject matter of the application was explained and outstanding issues were discussed. Without acquiescing in the propriety of the Examiner's preliminary comments regarding alleged functional-language issues or the propriety of the alleged 35 U.S.C. §§ 101 and 112 rejections, it is noted that the Examiner indicated that amending the claims as presented herein would overcome the alleged issues in the preliminary comments and the alleged 35 U.S.C. §§ 101 and 112 rejections.

REJECTION UNDER 35 U.S.C. § 102

Claims 1-10 stand rejected under 35 U.S.C. § 102(b) as allegedly being anticipated by U.S. Patent Application Publication No. 2004/0193420 to Kennewick *et al.* ("Kennewick") [Office Action, pg. 7]. In particular, with respect to the alleged rejection, the Office Action alleges that a substantial portion of the recitations of independent claim 1 and all of the recitations of dependent claims 2-10 are directed to functional language. Based on this contention, the Office Action argues that it is not necessary that Kennewick teaches such claim recitations to reject the claims under 35 U.S.C. § 102(b). [Office Action, pg. 8]. Such a contention is misguided. There is no prohibition on functional claiming.

Nonetheless, *solely* in an effort to expedite prosecution, and in *no way* acquiescing in the propriety of the alleged rejection, claims 1, 5-9, 11, and 15-19 have been amended as suggested by the Examiner during the interviews. As indicated by the Examiner during the interviews, the amendment overcomes the alleged functional-language issues. In addition, because the Office Action fails to address each and every feature in independent claim 1, and the claims depending therefrom, in the alleged rejection, the alleged rejection is improper and should be withdrawn.

Moreover, claims 2-4 have been canceled without prejudice or disclaimer. As such, the alleged rejection with respect to claims 2-4 is now moot and should be withdrawn.

REJECTIONS UNDER 35 U.S.C. § 103

A. Claims 1, 2, 7, 9, 10, 11, 12, 17, 19, and 20

Claims 1, 2, 7, 9, 10, 11, 12, 17, 19, and 20 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over Kennewick in view of U.S. Patent Application Publication No. 2002/0032752 to Gold *et al*. ("Gold") [Office Action, pg. 8, ¶11].

Applicants disagree with the propriety of the alleged rejection. However, *solely* in an effort to expedite prosecution, and in *no way* acquiescing in the propriety of the alleged rejection, independent claim 1 has been amended to clarify aspects of the invention. In particular, claim 1 recites, *inter alia*, the features of:

> determine, based on a processing of the first utterance by a speech recognition engine, one or more words or phrases of the first utterance [that includes a natural language utterance];
>
> identify, based on a natural language processing of the one or more words or phrases of the first utterance, content to dedicate to a recipient;
>
> initiate, based on the identified content, a dedication to the recipient . . . . [Emphasis added].

Independent claim 11 includes similar recitations.

Even assuming arguendo that it were deemed legally proper to combine Kennewick and Gold in the manner alleged by the Examiner (which Applicants do not concede), the alleged combination still fails to teach or suggest *at least* the foregoing claim recitation.

In the Office Action, the Examiner relies on Kennewick to teach the foregoing claim recitation. [Office Action, pg. 9]. Applicants note, however, that while Kennewick discloses natural language processing, Kennewick does not describe *at least* all of the features of "identify, based on a natural language processing of the one or more words or phrases of the first utterance, content to dedicate to a recipient; [and] initiate, based on the identified content, a dedication to the recipient." Gold, which provides for a website that enables users to select a song (e.g., by clicking an electronic link) and a recipient to whom the song is to be sent (e.g., by typing in the recipient's email address), also does not provide for such features.

Accordingly, for at least the foregoing reasons, Kennewick and Gold, either alone or in combination, fail to teach or suggest each and every feature recited in independent claims 1 and 11. The alleged rejection under 35 U.S.C. § 103 is therefore improper and should be withdrawn. Dependent claims 7, 9, 10, 17, 19, and 20 are also allowable because they each ultimately depend from an allowable claim, as well as for the further features they recite.

Moreover, claims 2 and 12 have been canceled without prejudice or disclaimer. As such, the alleged rejection with respect to claims 2 and 12 is now moot and should be withdrawn.

B. Claims 3, 4, 6, 13, 14 and 16

Claims 3, 4, 6, 13, 14 and 16 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick and Gold, further in view of U.S. Patent Application Publication No. 2008/0140385 to Mahajan *et al*. ("Mahajan") [Office Action, pg. 12, ¶12].

Applicants disagree with the propriety of the alleged rejection. Nonetheless, based on the foregoing reasons in Section A, dependent claims 6 and 16 are allowable because they reach ultimately depend from an allowable claim. Moreover, dependent claims 6 and 16 are allowable for the further features they recite. For example, dependent claim 6 recites the feature of:

> annotate the content [that is identified as content to be dedicated to the recipient based on a processing of the first utterance] with the second utterance [that is to be associated with the content]. [Emphasis added]

Dependent claim 16 includes similar recitations.

Even assuming arguendo that it were deemed legally proper to combine Kennewick, Gold, and Mahajan in the manner alleged by the Examiner (which Applicants do not concede), the alleged combination still fails to teach or suggest *at least* the foregoing claim recitation.

In the Office Action, the Examiner acknowledges that neither Kennewick nor Gold, either alone or in combination, describe such annotations of content (that is identified as content to be dedicated to a recipient based on a processing of a first utterance) with a second utterance that is to be associated with the content. Instead, the Examiner relies on Mahajan to provide for such features. [Office Action, pg. 12]. However, Mahajan (at best) describes recognizing textual words of A/V content that is currently playing and displaying the recognized textual words as the A/V content is played. Thus, Mahajan also does not provide for the recited features of claims 6 and 16. Accordingly, the alleged rejection under 35 U.S.C. § 103 is therefore improper and should be withdrawn.

Moreover, claims 3, 4, 13, and 14 have been canceled without prejudice or disclaimer. As such, the alleged rejection with respect to claims 3, 4, 13, and 14 is now moot and should be withdrawn.

C. Claims 5, 8, 15, and 18

Claims 5 and 15 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick and Gold, further in view of U.S. Patent Application Publication No. 2010/0036967 to Caine *et al*. ("Caine") [Office Action, pg. 13, ¶13]. Claims 8 and 18 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick and Gold, further in view of U.S. Patent No. 5,892,900 to Golden *et al*. ("Golden") [Office Action, pg. 14, ¶14].

Applicants disagree with the propriety of the alleged rejections. Nonetheless, based on the foregoing reasons in Section A, dependent claims 5, 8, 15, and 18 are allowable because they reach ultimately depend from an allowable claim, as well as for the further features they recite. Accordingly, the alleged rejections under 35 U.S.C. § 103 are therefore improper and should be withdrawn.

NEW CLAIMS

New claims 21-26 are allowable *at least* by virtue of their dependence on independent claim 1, as well as for the further features they recite.

CONCLUSION

Having addressed each of the foregoing rejections, it is respectfully submitted that a full and complete response has been made to the outstanding Office Action and, as such, the application is in condition for allowance. Notice to that effect is respectfully requested.

If the Examiner believes, for any reason, that personal communication will expedite prosecution of this application, the Examiner is invited to telephone the undersigned at the number provided.

Respectfully submitted,

Date: **January 30, 2013**

PILLSBURY WINTHROP SHAW PITTMAN LLP

By: /Ngai Zhang/
Ngai B. Zhang
Registration No. 65,473

Customer No. 00909

P.O. Box 10500
McLean, Virginia 22102
Direct Dial: 703.770.7532
Main: 703.770.7900
Fax: 703.770.7901

PTO/SB/22 (10-12)
Approved for use through 1/31/2013. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARMENT OF COMMERCE
Under the paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless if displays a valid OMB control number.

PETITION FOR EXTENSION OF TIME UNDER 37 CFR 1.136(a)

Docket Number (Optional): 090611-0384422

Application Number	Filed
12/943,699	November 10, 2010

For: SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE

Art Unit	Examiner
3685	Tim B. Hale

This is a request under the provisions of 37 CFR 1.136(a) to extend the period for filing a reply in the above identified application.

The requested extension and fee are as follows (check time period desired and enter the appropriate fee below):

		Fee	Small Entity Fee	
☒	One month (37 CFR 1.17(a)(1))	$150	$75	$ 75.00
☐	Two months (37 CFR 1.17(a)(2))	$570	$285	$
☐	Three months (37 CFR 1.17(a)(3))	$1,290	$645	$
☐	Four months (37 CFR 1.17(a)(4))	$2,010	$1,005	$
☐	Five months (37 CFR 1.17(a)(5))	$2,730	$1,365	$

☒ Applicant claims small entity status. See 37 CFR 1.27.

☐ A check in the amount of the fee is enclosed.

☐ Payment by credit card. Form PTO-2038 is attached.

☐ The Director has already been authorized to charge fees in this application to a Deposit Account.

☒ The Director is hereby authorized to charge any fees which may be required, or credit any overpayment, to Deposit Account Number 033975.

☒ Payment made via EFS-Web.

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

I am the

☐ applicant/inventor.

☐ assignee of record of the entire interest. See 37 CFR 3.71. 37 CFR 3.73(b) statement is enclosed (Form PTO/SB/96).

☒ attorney or agent of record. Registration Number 65,473.

☐ attorney or agent acting under 37 CFR 1.34. Registration number _____.

/Ngai Zhang/	January 30, 2013
Signature	Date
Ngai Zhang	(703) 770-7532
Typed or printed name	Telephone Number

NOTE: This form must be signed in accordance with 37 CFR 1.33. See 37 CFR 1.4 for signature requirements and certifications. Submit multiple forms if more than one signature is required, see below*.

☒ *Total of 1 forms are submitted.

This collection of information is required by 37 CFR 1.136(a). The information is required to obtain or retain a benefit by the public, which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 6 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Mail Stop PCT, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

American LegalNet, Inc.
www.FormsWorkFlow.com

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANTS :	Mike KENNEWICK *et al.*	CONFIRMATION NO.:	4185
SERIAL NUMBER :	12/943,699	EXAMINER:	Tim B. Hale
FILING DATE :	November 10, 2010	ART UNIT:	3685
FOR :	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE		

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT

Commissioner for Patents
P. O. Box 1450
Alexandria, VA 22313-1450

Sir:

Pursuant to the duty of disclosure under 37 C.F.R. §§ 1.56, 1.97 and 1.98, Applicants hereby make of record in the above-identified application the documents listed on the attached Form PTO-1449.

Applicants have listed a publication date on the attached PTO-1449 based on information presently available to the undersigned. However, the listed publication date should not be construed as an admission that the information was actually published on the date indicated.

Applicants reserve the right to establish the patentability of the claimed invention over any of the information provided herewith, and/or to prove that this information may not be prior art, and/or to prove that this information may not be enabling for the teachings purportedly offered.

This Supplemental Information Disclosure Statement is being filed more than three months after the filing date of a national application other than a continued prosecution application under 37 C.F.R. § 1.53(d), or more than three months after the date of entry of the national stage as set forth in 37 C.F.R. § 1.491 in an international application, and after the mailing date of the first Office Action on the merits, but before the mailing date of any of a Final Action under 37 C.F.R. § 1.113, a Notice of Allowance under 37 C.F.R. § 1.311, or an action that otherwise closes prosecution in the application. 37 C.F.R. § 1.97(c).

Please charge the $180.00 fee [37 C.F.R. § 1.17(p)] to Deposit Account No. 03-3975. In addition, please charge any fee deficiencies, or credit any overpayment, to Deposit Account No. 03-3975 as needed to ensure consideration of the Supplemental Information Disclosure Statement.

Document 1Z was cited in a search report by a foreign patent office in a counterpart foreign application. Submission of an English language version of the search report that indicates the degree of relevance found by the foreign patent office is provided.

It is respectfully requested that the Examiner initial and return a copy of the enclosed PTO-1449.

Respectfully submitted,

PILLSBURY WINTHROP SHAW PITTMAN LLP

By /Ngai Zhang/
Ngai B. Zhang
Registration No.: 65,473

Customer No.
00909

P. O. Box 10500
McLean, Virginia 22102
Telephone: (703) 770-7900
Telefax: (703) 770-7901

Dated: January 30, 2013

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	1A	5,027,406	06/25/1991	Roberts *et al.*	381	43	
	1B	5,208,748	05/04/1993	Flores *et al.*	364	419	
	1C	5,357,596	10/18/1994	Takebayashi *et al.*	395	2.84	
	1D	5,577,165	11/19/1996	Takebayashi et al.	395	2.84	
	1E	5,608,635	03/04/1997	Tamai	364	449.3	
	1F	5,740,256	04/14/1998	Castello Da Costa *et al.*	361	94.7	
	1G	5,748,974	05/05/1998	Johnson	395	759	
	1H	5,839,107	11/17/1998	Gupta *et al.*	704	270	
	1I	5,848,396	12/08/1998	Gerace	705	10	
	1J	5,855,000	12/29/1998	Waibel *et al.*	704	235	
	1K	5,897,613	04/27/1999	Chan	704	210	
	1L	5,960,397	09/28/1999	Rahim	704	244	
	1M	5,960,399	09/28/1999	Barclay *et al.*	704	270	
	1N	5,960,447	09/28/1999	Holt *et al.*	707	500	
	1O	5,995,928	11/30/1999	Nguyen *et al.*	704	251	
	1P	6,018,708	01/25/2000	Dahan *et al.*	704	244	
	1Q	6,035,267	03/07/2000	Watanabe *et al.*	704	1	
	1R	6,049,607	04/11/2000	Marash *et al.*	379	410	
	1S	6,067,513	05/23/2000	Ishimitsu	704	233	
	1T	6,076,059	06/13/2000	Glickman *et al.*	704	260	
	1U	6,085,186	07/04/2000	Christianson *et al.*	707	3	
	1V	6,108,631	08/22/2000	Ruhl	704	270	
	1W	6,167,377	12/26/2000	Gillick *et al.*	704	240	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO
	1X	1 320 043 A2	06/18/2003	Europe				
	1Y	1 646 037	04/12/2006	Europe				

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

	1Z	"Statement in Accordance with the Notice from the European Patent Office dated 1 October 2007 Concerning Business Methods (OJ EPO 11/2007, 592-593), XP002456252

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	2A	6,173,266	01/09/2001	Marx *et al.*	704	270	
	2B	6,188,982	02/13/2001	Chiang	704	256	
	2C	6,199,043	03/06/2001	Happ	704	272	
	2D	6,208,964	03/27/2001	Sabourin	704	244	
	2E	6,233,561	05/15/2001	Junqua *et al.*	704	277	
	2F	6,236,968	05/22/2001	Kanevsky *et al.*	704	275	
	2G	6,246,990	06/12/2001	Happ	704	275	
	2H	6,266,636	07/24/2001	Kosaka *et al.*	704	244	
	2I	6,269,336	07/31/2001	Ladd *et al.*	704	270	
	2J	6,275,231	08/14/2001	Obradovich	345	349	
	2K	6,278,968	08/21/2001	Franz *et al.*	704	3	
	2L	6,301,560	10/09/2001	Masters	704	251	
	2M	6,362,748	03/26/2002	Huang	340	901	
	2N	6,366,882	04/02/2002	Bijl *et al.*	704	235	
	2O	6,377,913	04/23/2002	Coffman *et al.*	704	8	
	2P	6,393,403	05/21/2002	Majaniemi	704	275	
	2Q	6,397,181	05/28/2002	Li *et al.*	704	256	
	2R	6,405,170	06/11/2002	Phillips *et al.*	704	270	
	2S	6,429,813	08/06/2002	Feigen	342	357.13	
	2T	6,434,529	08/13/2002	Walker *et al.*	704	275	
	2U	6,456,974	09/24/2002	Baker *et al.*	704	270.1	
	2V	6,510,417	01/21/2003	Woods *et al.*	704	275	
	2W	6,522,746	02/18/2003	Marchok *et al.*	379	406.03	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	3A	6,549,629	04/15/2003	Finn *et al.*	381	92	
	3B	6,560,590	05/06/2003	Shwe *et al.*	706	55	
	3C	6,571,279	05/27/2003	Herz *et al.*	709	217	
	3D	6,578,022	06/10/2003	Foulger *et al.*	706	45	
	3E	6,584,439	06/24/2003	Geilhufe *et al.*	704	270	
	3F	6,587,858	07/01/2003	Strazza	707	102	
	3G	6,591,239	07/08/2003	McCall *et al.*	704	275	
	3H	6,594,367	07/15/2003	Marash *et al.*	381	92	
	3I	6,601,026	07/29/2003	Appelt *et al.*	704	9	
	3J	6,606,598	08/12/2003	Holthouse *et al.*	704	275	
	3K	6,622,119	09/16/2003	Ramaswamy *et al.*	704	9	
	3L	6,631,351	10/07/2003	Ramachandran *et al.*	704	270	
	3M	6,633,846	10/14/2003	Bennett *et al.*	704	257	
	3N	6,681,206	01/20/2004	Gorin *et al.*	704	243	
	3O	6,701,294	03/02/2004	Ball *et al.*	704	257	
	3P	6,704,396	03/09/2004	Parolkar *et al.*	379	88.17	
	3Q	6,704,708	03/09/2004	Pickering	704	235	
	3R	6,745,161	06/01/2004	Arnold *et al.*	704	7	
	3S	6,751,612	06/15/2004	Schuetze *et al.*	707	4	
	3T	6,754,485	06/22/2004	Obradovich *et al.*	455	414.1	
	3U	6,757,544	06/29/2004	Rangarajan *et al.*	455	456.1	
	3V	6,832,230	12/14/2004	Zilliacus *et al.*	707	203	
	3W	6,850,603	02/01/2005	Eberle *et al.*	379	88.16	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	4A	6,868,380	03/15/2005	Kroeker	704	240	
	4B	6,877,001	04/05/2005	Wolf *et al.*	707	3	
	4C	6,910,003	06/21/2005	Arnold *et al.*	704	4	
	4D	6,912,498	06/28/2005	Stevens *et al.*	704	235	
	4E	6,934,756	08/23/2005	Maes	709	227	
	4F	6,954,755	10/11/2005	Reisman	707	10	
	4G	6,959,276	10/25/2005	Droppo *et al.*	704	226	
	4H	6,968,311	11/22/2005	Knockeart *et al.*	704	270	
	4I	6,975,993	12/13/2005	Keiller	704	275	
	4J	6,983,055	01/03/2006	Luo	381	313	
	4K	7,003,463	02/21/2006	Maes *et al.*	704	270.1	
	4L	7,024,364	04/04/2006	Guerra *et al.*	704	270	
	4M	7,027,586	04/11/2006	Bushey *et al.*	379	265.09	
	4N	7,035,415	04/25/2006	Belt *et al.*	381	92	
	4O	7,043,425	05/09/2006	Pao	704	211	
	4P	7,054,817	05/30/2006	Shao	704	270	
	4Q	7,058,890	06/06/2006	George *et al.*	715	728	
	4R	7,069,220	06/27/2006	Coffman *et al.*	704	275	
	4S	7,072,834	07/04/2006	Zhou	704	244	
	4T	7,107,210	09/12/2006	Deng *et al.*	704	226	
	4U	7,107,218	09/12/2006	Preston	704	270	
	4V	7,110,951	09/19/2006	Lemelson *et al.*	704	270	
	4W	7,130,390	10/31/2006	Abburi	379	88.17	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	5A	7,143,037	11/28/2006	Chestnut	704	251	
	5B	7,143,039	11/28/2006	Stifelman *et al.*	704	270	
	5C	7,165,028	01/16/2007	Gong	704	233	
	5D	7,184,957	02/27/2007	Brookes *et al.*	704	246	
	5E	7,190,770	03/13/2007	Ando *et al.*	379	88.01	
	5F	7,197,069	03/27/2007	Agazzi *et al.*	375	233	
	5G	7,206,418	04/17/2007	Yang *et al.*	381	92	
	5H	7,207,011	04/17/2007	Mulvey *et al.*	715	812	
	5I	7,236,923	06/26/2007	Gupta	704	9	
	5J	7,272,212	09/18/2007	Eberle *et al.*	379	88.17	
	5K	7,277,854	10/02/2007	Bennett *et al.*	704	257	
	5L	7,283,829	10/16/2007	Christenson *et al.*	455	461	
	5M	7,283,951	10/16/2007	Marchisio *et al.*	704	9	
	5N	7,299,186	11/20/2007	Kuzunuki *et al.*	704	270.1	
	5O	7,305,381	12/04/2007	Poppink *et al.*	1	1	
	5P	7,321,850	01/22/2008	Wakita	704	10	
	5Q	7,328,155	02/05/2008	Endo *et al.*	704	251	
	5R	7,337,116	02/26/2008	Charlesworth *et al.*	704	254	
	5S	7,366,285	04/29/2008	Parolkar *et al.*	379	88.17	
	5T	7,366,669	04/29/2008	Nishitani *et al.*	704	256	
	5U	7,406,421	07/29/2008	Odinak *et al.*	704	275	
	5V	7,415,414	08/19/2008	Azara *et al.*	704	270	
	5W	7,447,635	11/04/2008	Konopka *et al.*	704	275	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	6A	7,454,608	11/18/2008	Gopalakrishnan *et al.*	713	100	
	6B	7,472,020	12/30/2008	Brulle-Drews	701	211	
	6C	7,472,060	12/30/2008	Gorin *et al.*	704	240	
	6D	7,478,036	01/13/2009	Shen *et al.*	704	9	
	6E	7,493,259	02/17/2009	Jones *et al.*	704	257	
	6F	7,516,076	04/07/2009	Walker *et al.*	704	275	
	6G	7,529,675	05/05/2009	Maes	704	270.1	
	6H	7,536,297	05/19/2009	Byrd *et al.*	704	10	
	6I	7,536,374	05/19/2009	Au	706	55	
	6J	7,542,894	06/02/2009	Murata	704	9	
	6K	7,546,382	06/09/2009	Healey *et al.*	709	246	
	6L	7,558,730	07/07/2009	Davis *et al.*	704	235	
	6M	7,574,362	08/11/2009	Walker *et al.*	704	275	
	6N	7,577,244	08/18/2009	Taschereau	379	218.01	
	6O	7,606,708	10/20/2009	Hwang	704	257	
	6P	7,640,006	12/29/2009	Portman *et al.*	455	412.1	
	6Q	7,676,365	03/09/2010	Hwang *et al.*	704	240	
	6R	7,676,369	03/09/2010	Fujimoto *et al.*	704	270	
	6S	7,684,977	03/23/2010	Morikawa	704	211	
	6T	7,729,916	06/01/2010	Coffman *et al.*	704	270	
	6U	7,729,918	06/01/2010	Walker *et al.*	704	275	
	6V	7,729,920	06/01/2010	Chaar *et al.*	704	275	
	6W	7,748,021	06/29/2010	Obradovich	725	105	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	7A	7,788,084	08/31/2010	Brun *et al.*	704	7	
	7B	7,801,731	09/21/2010	Odinak *et al.*	704	275	
	7C	7,831,426	11/09/2010	Bennett	704	252	
	7D	7,831,433	11/09/2010	Belvin *et al.*	704	275	
	7E	7,873,519	01/18/2011	Bennett	704	257	
	7F	7,873,523	01/18/2011	Potter *et al.*	704	275	
	7G	7,881,936	02/01/2011	Longe *et al.*	704	257	
	7H	7,894,849	02/22/2011	Kass *et al.*	455	550.1	
	7I	7,902,969	03/08/2011	Obradovich	340	439	
	7J	7,917,367	03/29/2011	Di Cristo *et al.*	704	270.1	
	7K	7,920,682	04/05/2011	Byrne *et al.*	379	88.18	
	7L	7,949,529	05/24/2011	Weider *et al.*	704	270	
	7M	7,949,537	05/24/2011	Walker *et al.*	704	275	
	7N	7,953,732	05/31/2011	Frank *et al.*	707	724	
	7O	7,974,875	07/05/2011	Quilici *et al.*	705	14.4	
	7P	7,983,917	07/19/2011	Kennewick *et al.*	704	257	
	7Q	7,984,287	07/19/2011	Gopalakrishnan *et al.*	713	100	
	7R	8,015,006	09/06/2011	Kennewick *et al.*	704	236	
	7S	8,060,367	11/15/2011	Keaveney	704	247	
	7T	8,069,046	11/29/2011	Kennewick *et al.*	704	257	
	7U	8,073,681	12/06/2011	Baldwin *et al.*	704	9	
	7V	8,077,975	12/13/2011	Ma *et al.*	382	187	
	7W	8,082,153	12/20/2011	Coffman *et al.*	704	270	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	8A	8,086,463	12/27/2011	Ativanichayaphong *et al.*	704	275	
	8B	8,112,275	02/07/2012	Kennewick *et al.*	704	240	
	8C	8,140,327	03/20/2012	Kennewick *et al.*	704	226	
	8D	8,140,335	03/20/2012	Kennewick *et al.*	704	257	
	8E	8,145,489	03/27/2012	Freeman *et al.*	704	257	
	8F	8,150,694	04/03/2012	Kennewick *et al.*	704	257	
	8G	8,155,962	04/10/2012	Kennewick *et al.*	704	257	
	8H	8,170,867	05/01/2012	Germain	704	9	
	8I	8,195,468	06/05/2012	Weider *et al.*	704	275	
	8J	8,219,599	07/10/2012	Tunstall-Pedoe	707	955	
	8K	8,224,652	07/17/2012	Wang *et al.*	704	275	
	8L	8,326,627	12/04/2012	Kennewick *et al.*	704	257	
	8M	8,326,634	12/04/2012	Di Cristo *et al.*	704	270.1	
	8N	8,326,637	12/04/2012	Baldwin *et al.*	704	275	
	8O	8,332,224	12/11/2012	Di Cristo *et al.*	704	257	
	8P	2001/0049601	12/06/2001	Kroeker *et al.*	704	254	
	8Q	2001/0054087	12/20/2001	Flom *et al.*	709	218	
	8R	2002/0015500	02/07/2002	Belt *et al.*	381	66	
	8S	2002/0022927	02/21/2002	Lemelson *et al.*	701	301	
	8T	2002/0035501	03/21/2002	Handel *et al.*	705	10	
	8U	2002/0069071	06/06/2002	Knockeart et al.	704	275	
	8V	2002/0087326	07/04/2002	Lee *et al.*	704	270.1	
	8W	2002/0133402	09/19/2002	Faber *et al.*	705	14	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	9A	2002/0161646	10/31/2002	Gailey *et al.*	705	14	
	9B	2002/0173961	11/21/2002	Guerra	704	258	
	9C	2002/0184373	12/05/2002	Maes	709	228	
	9D	2003/0016835	01/23/2003	Elko *et al.*	381	92	
	9E	2003/0065427	04/03/2003	Funk *et al.*	701	1	
	9F	2003/0110037	06/12/2003	Walker *et al.*	704	257	
	9G	2003/0115062	06/19/2003	Walker *et al.*	704	258	
	9H	2003/0120493	06/26/2003	Gupta	704	270.1	
	9I	2003/0144846	07/31/2003	Denenberg *et al.*	704	277	
	9J	2003/0158731	08/21/2003	Falcon *et al.*	704	231	
	9K	2003/0161448	08/28/2003	Parolkar *et al.*	379	88.17	
	9L	2003/0204492	10/30/2003	Wolf *et al.*	707	3	
	9M	2003/0206640	11/06/2003	Malvar *et al.*	381	93	
	9N	2003/0212550	11/13/2003	Ubale	704	215	
	9O	2003/0212558	11/13/2003	Matula	704	260	
	9P	2003/0212562	11/13/2003	Patel *et al.*	704	275	
	9Q	2003/0225825	12/04/2003	Healey *et al.*	709	203	
	9R	2004/0098245	05/20/2004	Walker *et al.*	704	1	
	9S	2004/0117804	06/17/2004	Scahill *et al.*	719	320	
	9T	2004/0140989	07/22/2004	Papageorge	345	700	
	9U	2004/0167771	08/26/2004	Duan *et al.*	704	10	
	9V	2004/0205671	10/14/2004	Sukehiro *et al.*	715	532	
	9W	2005/0021470	01/27/2005	Martin *et al.*	705	51	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	10A	2005/0125232	06/09/2005	Gadd	704	270.1	
	10B	2005/0137877	06/23/2005	Oesterling *et al.*	704	275	
	10C	2005/0143994	06/30/2005	Mori *et al.*	704	235	
	10D	2005/0216254	09/29/2005	Gupta *et al.*	704	9	
	10E	2005/0234727	10/20/2005	Chiu	704	270.1	
	10F	2005/0283752	12/22/2005	Fruchter *et al.*	717	100	
	10G	2006/0041431	02/23/2006	Maes	704	270.1	
	10H	2006/0047509	03/02/2006	Ding *et al.*	704	231	
	10I	2006/0217133	09/28/2006	Christenson *et al.*	455	461	
	10J	2006/0285662	12/21/2006	Yin *et al.*	379	88.16	
	10K	2007/0033020	02/08/2007	(Kelleher) Francois *et al.*	704	226	
	10L	2007/0043868	02/22/2007	Kumar *et al.*	709	226	
	10M	2007/0073544	03/29/2007	Millett *et al.*	704	277	
	10N	2007/0078708	04/05/2007	Yu *et al.*	705	14	
	10O	2007/0078709	04/05/2007	Rajaram	705	14	
	10P	2007/0118357	05/24/2007	Kasravi *et al.*	704	10	
	10Q	2007/0135101	06/14/2007	Ramati *et al.*	455	414.1	
	10R	2007/0162296	07/12/2007	Altberg *et al.*	705	1	
	10S	2007/0198267	08/23/2007	Jones *et al.*	704	257	
	10T	2008/0034032	02/07/2008	Healey *et al.*	709	203	
	10U	2008/0065386	03/13/2008	Cross *et al.*	704	270	
	10V	2008/0109285	05/08/2008	Reuther *et al.*	705	7	
	10W	2008/0147410	06/19/2008	Odinak	704	270.1	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	11A	2009/0052635	02/26/2009	Jones *et al.*	379	88.02	
	11B	2009/0067599	03/12/2009	Agarwal *et al.*	379	114.21	
	11C	2009/0106029	04/23/2009	DeLine *et al.*	704	275	
	11D	2009/0271194	10/29/2009	Davis *et al.*	704	235	
	11E	2009/0273563	11/05/2009	Pryor	345	157	
	11F	2009/0313026	12/17/2009	Coffman *et al.*	704	275	
	11G	2010/0029261	02/04/2010	Mikkelsen *et al.*	455	419	
	11H	2010/0185512	07/22/2010	Borger *et al.*	705	14.49	
	11I	2010/0312566	12/09/2010	Odinak *et al.*	704	275	
	11J	2011/0112827	05/12/2011	Kennewick *et al.*	704	9	
	11K	2011/0131036	06/02/2011	Di Cristo *et al.*	704	9	
	11L	2011/0131045	06/02/2011	Cristo *et al.*	704	249	
	11M	2011/0231182	09/22/2011	Weider *et al.*	704	9	
	11N	2011/0231188	09/22/2011	Kennewick *et al.*	704	236	
	11O	2012/0022857	01/26/2012	Baldwin *et al.*	704	9	
	11P	2012/0101809	04/26/2012	Kennewick *et al.*	704	9	
	11Q	2012/0101810	04/26/2012	Kennewick *et al.*	704	9	
	11R	2012/0109753	05/03/2012	Kennewick *et al.*	705	14.58	
	11S	2012/0150636	06/14/2012	Freeman *et al.*	705	14.49	
	11T	2012/0278073	11/01/2012	Weider et al.	704	235	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 320 043 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
18.06.2003 Bulletin 2003/25

(51) Int Cl.[7]: **G06F 17/30**

(21) Application number: **02258126.8**

(22) Date of filing: **26.11.2002**

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR IE IT LI LU MC NL PT SE SK TR
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **13.12.2001 GB 0129788**

(71) Applicant: **Hewlett-Packard Company**
Palo Alto, CA 94304 (US)

(72) Inventor: **Belrose, Guillaume**
Bristol BS2 8HT (GB)

(74) Representative: **Squibbs, Robert Francis**
Hewlett-Packard Limited,
IP Section,
Building 3
Filton Road
Stoke Gifford, Bristol BS 34 8QZ (GB)

(54) Multi-modal picture

(57) A system for presenting a multi-modal picture includes picture presentation equipment (37) for displaying an image (38) of the picture and for enabling a user to interact with the picture by selecting specific picture features and asking a specific query relating to the feature. A voice browser system (3) controlled according to dialog scripts (46) associated with the picture, determines an appropriate response having regard to the spoken user query and the picture feature being pointed to. Each picture can have multiple narrators associated with it and the can choose which narrator is currently active. Picture authoring apparatus is also provided.

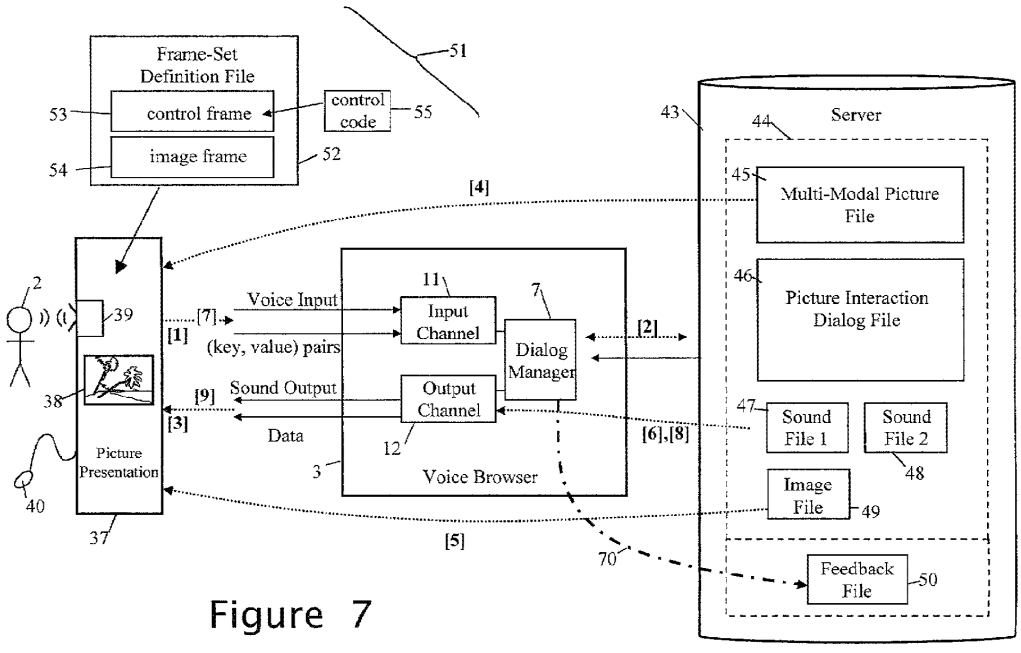


Figure 7

EP 1 320 043 A2

Amazon/VB Assets
Exhibit 1003
Page 255

Description

Field of the Invention

[0001] The present invention relates to multi-modal pictures with which a use can interact by spoken dialog exchanges.

Background of the Invention

[0002] In recent years there has been an explosion in the number of services available over the World Wide Web on the public internet (generally referred to as the "web"), the web being composed of a myriad of pages linked together by hyperlinks and delivered by servers on request using the HTTP protocol. Each page comprises content marked up with tags to enable the receiving application (typically a GUI browser) to render the page content in the manner intended by the page author; the markup language used for standard web pages is HTML (HyperText Markup Language).

[0003] However, today far more people have access to a telephone than have access to a computer with an Internet connection. Sales of cellphones are outstripping PC sales so that many people have already or soon will have a phone within reach where ever they go. As a result, there is increasing interest in being able to access web-based services from phones. 'Voice Browsers' offer the promise of allowing everyone to access web-based services from any phone, making it practical to access the Web any time and any where, whether at home, on the move, or at work.

[0004] Voice browsers allow people to access the Web using speech synthesis, pre-recorded audio, and speech recognition. Figure 1 of the accompanying drawings illustrates the general role played by a voice browser. As can be seen, a voice browser is interposed between a user 2 and a voice page server 4. This server 4 holds voice service pages (text pages) that are marked-up with tags of a voice-related markup language (or languages). When a page is requested by the user 2, it is interpreted at a top level (dialog level) by a dialog manager 7 of the voice browser 3 and output intended for the user is passed in text form to a Text-To-Speech (TTS) converter 6 which provides appropriate voice output to the user. User voice input is converted to text by speech recognition module 5 of the voice browser 3 and the dialog manager 7 determines what action is to be taken according to the received input and the directions in the original page. The voice input / output interface can be supplemented by keypads and small displays.

[0005] In general terms, therefore, a voice browser can be considered as a largely software device which interprets a voice markup language and generate a dialog with voice output, and possibly other output modalities, and / or voice input, and possibly other modalities (this definition derives from a working draft, dated September 2000, of the Voice browser Working Group of the World Wide Web Consortium).

[0006] Voice browsers may also be used together with graphical displays, keyboards, and pointing devices (e.g. a mouse) in order to produce a rich "multimodal voice browser". Voice interfaces and the keyboard, pointing device and display maybe used as alternate interfaces to the same service or could be seen as being used together to give a rich interface using all these modes combined.

[0007] Some examples of devices that allow multimodal interactions could be multimedia PC, or a communication appliance incorporating a display, keyboard, microphone and speaker/headset, an in car Voice Browser might have display and speech interfaces that could work together, or a Kiosk.

[0008] Some services may use all the modes together to provide an enhanced user experience, for example, a user could touch a street map displayed on a touch sensitive display and say "Tell me how I get here?". Some services might offer alternate interfaces allowing the user flexibility when doing different activities. For example while driving speech could be used to access services, but a passenger might used the keyboard.

[0009] Figure 2 of the accompanying drawings shows in greater detail the components of an example voice browser for handling voice pages 15 marked up with tags related to four different voice markup languages, namely:

- tags of a dialog markup language that serves to specify voice dialog behaviour;
- tags of a multimodal markup language that extends the dialog markup language to support other input modes (keyboard, mouse, etc.) and output modes (large and small screens);
- tags of a speech grammar markup language that serve to specify the grammar of user input; and
- tags of a speech synthesis markup language that serve to specify voice characteristics, types of sentences, word emphasis, etc.

[0010] When a page 15 is loaded into the voice browser, dialog manager 7 determines from the dialog tags and multimodal tags what actions are to be taken (the dialog manager being programmed to understand both the dialog and multimodal languages 19). These actions may include auxiliary functions 18 (available at any time during page processing) accessible through APIs and including such things as database lookups, user identity and validation, telephone call control etc. When speech output to the user is called for, the semantics of the output is passed, with any

Amazon/VB Assets
Exhibit 1003
Page 256

associated speech synthesis tags, to output channel 12 where a language generator 23 produces the final text to be rendered into speech by text-to-speech converter 6 and output to speaker 17. In the simplest case, the text to be rendered into speech is fully specified in the voice page 15 and the language generator 23 is not required for generating the final output text; however, in more complex cases, only semantic elements are passed, embedded in tags of a natural language semantics markup language (not depicted in Figure 2) that is understood by the language generator. The TTS converter 6 takes account of the speech synthesis tags when effecting text to speech conversion for which purpose it is cognisant of the speech synthesis markup language 25.

[0011] User voice input is received by microphone 16 and supplied to an input channel of the voice browser. Speech recogniser 5 generates text which is fed to a language understanding module 21 to produce semantics of the input for passing to the dialog manager 7. The speech recogniser 5 and language understanding module 21 work according to specific lexicon and grammar markup language 22 and, of course, take account of any grammar tags related to the current input that appear in page 15. The semantic output to the dialog manager 7 may simply be a permitted input word or may be more complex and include embedded tags of a natural language semantics markup language. The dialog manager 7 determines what action to take next (including, for example, fetching another page) based on the received user input and the dialog tags in the current page 15.

[0012] Any multimodal tags in the voice page 15 are used to control and interpret multimodal input/output. Such input/output is enabled by an appropriate recogniser 27 in the input channel 11 and an appropriate output constructor 28 in the output channel 12.

[0013] Whatever its precise form, the voice browser can be located at any point between the user and the voice page server. Figures 3 to 5 illustrate three possibilities in the case where the voice browser functionality is kept all together; many other possibilities exist when the functional components of the voice browser are separated and located in different logical/physical locations.

[0014] In Figure 3, the voice browser 3 is depicted as incorporated into an end-user system 8 (such as a PC or mobile entity) associated with user 2. In this case, the voice page server 4 is connected to the voice browser 3 by any suitable data-capable bearer service extending across one or more networks 9 that serve to provide connectivity between server 4 and end-user system 8. The data-capable bearer service is only required to carry text-based pages and therefore does not require a high bandwidth.

[0015] Figure 4 shows the voice browser 3 as co-located with the voice page server 4. In this case, voice input/output is passed across a voice network 9 between the end-user system 8 and the voice browser 3 at the voice page server site. The fact that the voice service is embodied as voice pages interpreted by a voice browser is not apparent to the user or network and the service could be implemented in other ways without the user or network being aware.

[0016] In Figure 5, the voice browser 3 is located in the network infrastructure between the end-user system 8 and the voice page server 4, voice input and output passing between the end-user system and voice browser over one network leg, and voice-page text data passing between the voice page server 4 and voice browser 3 over another network leg. This arrangement has certain advantages; in particular, by locating expensive resources (speech recognition, TTS converter) in the network, they can be used for many different users with user profiles being used to customise the voice-browser service provided to each user.

[0017] It is known to enhance pictures by providing associated speech annotations. It is an object of the present invention to provide further enhancements to such pictures.

Summary of the Invention

[0018] According to one aspect of the present invention, there is provided a system for presenting information concerning a picture to a user, the system comprising:

- a data store for holding responses, specific to said picture, in respect of specific user queries concerning particular picture features;
- a manually-operable feature-selection arrangement for enabling a user to select a feature in a displayed image of the picture, and for providing an output indication regarding what said particular feature, if any, the user has thereby selected;
- a voice dialog input-output subsystem including a speech recogniser for interpreting queries from a user;
- a control arrangement responsive to a user selecting a said particular feature and asking a specific query regarding that feature, to output the corresponding stored response.

[0019] Advantageously, each response is associated with a specified narrator and the user can opt to receive only the response of a selected narrator for either a current query or for all queries.

[0020] Preferably, provision is made for providing feedback to an author of the picture or its associated responses in the event that a user makes a query about a picture feature that is not a said particular picture feature or, although

Amazon/VB Assets
Exhibit 1003
Page 257

a query about such a feature, is not a said specific query for which a response is available.

[0021] According to another aspect of the present invention, there is provided a multi-modal picture specified by data held on at least one data carrier, this data comprising:

- picture image data for displaying a picture image;
- response data indicative of voice responses intended to be given to specific user queries concerning particular picture features of the picture;
- first control data for enabling a determination to be made as to which said particular feature in the picture image, if any, a user is selecting when using a selection arrangement to indicate a feature in the displayed image; and
- second control data for determining, on the basis of a spoken user query and on which said particular picture feature is selected by the user using the selection arrangement, which said response is to be used to reply to the user query.

[0022] According to a further aspect of the present invention, there is provided a multi-modal picture comprising a hard-copy picture image, and data held on at least one data carrier, this data comprising:

- response data indicative of voice responses intended to be given to specific user queries concerning particular picture features;
- first control data for enabling a determination to be made as to which said particular feature in the picture image, if any, a user is selecting when using a selection arrangement to indicate a feature of the image; and
- second control data for determining, on the basis of a spoken user query and on which said particular picture feature is selected by the user using the selection arrangement, which said response is to be used to reply to the user query.

[0023] The present invention also encompasses the overall process of generating a multi-modal picture and using it to convey information about particular picture features to a user. Thus, according to a still further aspect of the present invention, there is provided a method of conveying information about particular features in a picture, the method comprising the steps of:

(a) creating the following specifically-associated data:

- picture image data for displaying a picture image;
- response data indicative of voice responses intended to be given to specific user queries concerning particular picture features of the picture;
- first control data for enabling a determination to be made as to which said particular feature in the picture image, if any, a user is selecting when using a selection arrangement to indicate a feature in the displayed image; and
- second control data for determining, on the basis of a spoken user query and on which said particular picture feature is selected by the user, which said response is to be used to reply to the user query;

(b) using the image data to display an image of the picture;
(c) having a user use a manually-operated selection arrangement to select a feature in the displayed image and using the first control data to determine which said particular feature in the picture image, if any, the user is selecting;
(d) receiving and interpreting a spoken query from the user to determine if a said specific query is being asked; and
(e) using the second control data to determine, on the basis of the said particular feature determined as being selected in step (c) and the said specific query determined as being asked in step (d), which said response is to be used to reply and thereupon using the response data to output the corresponding voice response.

[0024] Similarly, according to a yet further aspect of the present invention, there is provided a method of conveying information about particular features in a hard-copy picture, the method comprising the steps of:

(a) creating the following specifically-associated data:

- response data indicative of voice responses intended to be given to specific user queries concerning particular picture features in said picture;
- first control data for enabling a determination to be made as to which said particular feature in the picture, if any, a user is selecting when using a selection arrangement to indicate a feature of the picture; and
- second control data for determining, on the basis of a spoken user query and on which said particular picture

Amazon/VB Assets
Exhibit 1003
Page 258

feature is selected by the user, which said response is to be used to reply to the user query;

(b) making the picture and data available to a user;
(c) having the user use a manually-operated selection arrangement to select a feature in the picture and using the first control data to determine which said particular feature in the picture, if any, the user is selecting;
(d) receiving and interpreting a spoken query from the user to determine if a said specific query is being asked; and
(e) using the second control data to determine, on the basis of the said particular feature determined as being selected in step (c) and the said specific query determined as being asked in step (d), which said response is to be used to reply and thereupon using the response data to output the corresponding voice response.

[0025] According to another aspect of the present invention, there is provided apparatus for authoring a multi-modal picture, comprising:

- a first tool for defining image hotspots associated with particular picture-image features;
- a second tool with speech recognition capability, for recording user responses input by voice, to user-specified queries each associated with a particular said picture-image feature; and
- means for automatically generating control data for determining, on the basis of a spoken user query and on which said particular picture feature is selected by a user, which said response is to be used to reply to the user query.

[0026] With respect to the above-mentioned feature regarding the provision of feedback on user queries for which responses are not available, a further independent aspect of the present invention provides a method of collecting user-interest information regarding a picture, the method comprising the steps of:

(a) presenting a picture image to a user, particular picture features in the image each having associated information that is presented to the user upon the user requesting such information;
(b) recognising a user information request, in respect of a picture feature, made by the user at least selecting that picture feature using a feature-selection arrangement; and
(c) in response to recognition of a user information request in step (b) concerning a picture feature that is other than a said particular picture feature, outputting an identifier of the concerned picture feature independently of the presented image.

[0027] This aspect of the invention may be implemented in the form of a system for collecting user-interest information regarding a picture, the system comprising:

- a data store for holding information concerning particular picture features;
- a feature-selection arrangement for enabling a user to select a feature in a displayed view of the picture, and for providing an output indication regarding what said particular feature, if any, the user has thereby selected;
- an information-output subsystem for outputting information about picture features to the user;
- a control arrangement for recognising a user information request, in respect of a picture feature, made by the user at least selecting that picture feature using the feature-selection arrangement, the control arrangement including:

    - first output means arranged to respond to a user making an information request concerning a said particular feature, by outputting the corresponding stored information via the information-output subsystem; and
    - second output means arranged to respond to a user making an information request concerning a picture feature that is other than a said particular picture feature, by outputting an identifier of the concerned picture feature for informing an author of the picture or its associated information regarding the unsatisfied information request.

Brief Description of the Drawings

[0028] A method and apparatus embodying the invention will now be described, by way of nonlimiting example, with reference to the accompanying diagrammatic drawings, in which:

. **Figure 1** is a diagram illustrating the role of a voice browser;
. **Figure 2** is a diagram showing the functional elements of a voice browser and their relationship to different types of voice markup tags;
. **Figure 3** is a diagram showing a voice service implemented with voice browser functionality located in an end-user system;

Amazon/VB Assets
Exhibit 1003
Page 259

. **Figure 4** is a diagram showing a voice service implemented with voice browser functionality co-located with a voice page server;
. **Figure 5** is a diagram showing a voice service implemented with voice browser functionality located in a network between the end-user system and voice page server; and
. **Figure 6** shows an example picture image of a multi-modal picture embodying the invention;
. **Figure 7** is a diagram of a system for presenting a multi-modal picture to a user;
. **Figure 8** is a diagram showing constituent dialog blocks of a picture interaction dialog file of the Figure 6 multi-modal picture;
. **Figure 9** is a diagram of apparatus for authoring a multi-modal picture;
. **Figure 10** shows, in respect of the Figure 6 image, a user seeking information from two picture features for which there is no associated information; and
. **Figure 11** shows the Figure 6 picture image enhanced with upper and lower information bars.

Best Mode of Carrying Out the Invention

[0029] In the following description, voice dialog interaction with a user is described based on a voice page server serving a dialog page with embedded voice markup tags to a multi-modal voice browser. Unless otherwise indicated, the foregoing description of voice browsers, and their possible locations and access methods is to be taken as applying also to the described embodiments ofthe invention. Furthermore, although voice-browser based forms of voice dialog services are preferred, the present invention in its widest conception, is not limited to these forms of voice dialog service system and other suitable systems will be apparent to persons skilled in the art.

Multi-Modal Picture

[0030] Figure 6 depicts a multi-modal picture comprising a displayed picture image 30 (here shown as being of a holiday island taken, for example, whilst the author of the picture was on holiday) with which a recipient (referred to below as the "user") can interact using multiple modalities and, in particular, by spoken dialogues and the use of a pointing arrangement, such as a cursor controlled by a mouse, stylus or keyboard keys, a touch-screen detection arrangement, etc. Thus, the user can query the picture using speech input, the query being either a general query about the picture or a specific query about a particular feature (item or area) of the picture indicated by use of the pointing arrangement (for example, the user uses a mouse to move a cursor over a picture feature and then clicks a mouse button). Of course, only certain picture features will have been set up to be queried and in the Figure 6 example, there are three such features, namely a coconut tree 31, a first person 32, and a second person 33; the dashed boxes around these features represent the "hotspot" image areas set up to encompass the features, the dashed lines generally not being visible (though displaying a hotspot boundary can be used as one way of indicating the hotspot to the user - a more typical way would be, in the case of the pointing arrangement being a mouse-controlled cursor, to change the cursor image as it moved into and out of a hotspot).

[0031] Processing functionality associated with the multi-modal picture is arranged to recognize one or more one or more general queries and, for each picture feature set up for selection, one or more specific queries. For each query set to be recognized, there is a corresponding response which is output to the user when the query is recognized, this output generally being in spoken form, generated using either speech synthesis or pre-recorded audio content. Thus, for example:

General Queries - available general queries can include picture location, the date of the picture was taken, and a description of the general subject of the picture. For instance:

User: "Describe the picture"
System: "This is a picture of the XYZ beach"
User: "What is the date?"
System: "The picture was taken November last year"
User: "What is the location?"
System: "XYZ beach is in Martinique, an island that is part of the French West Indies."

(In this and other dialog examples given in the present specification, the "System" is the functionality used to present the multi-modal picture to the user and to provide for multi-modal interaction).

Specific Queries - typically, the same specific queries are available for all the selectable picture features; example specific queries are: "What it is?", "Any story?", etc. For instance:

Amazon/VB Assets
Exhibit 1003
Page 260

User: "What is it"? "
System: "This is my cousin John".
User: "Any story?"
System: "John goes fishing quite often in the Caribbean Sea. One day, he just escaped a white shark attack."

[0032] The responses from the system can, of course, be more than just a simple single turn response and may involve a multi-turn structured dialog. Provision may also be made for different levels of content, for instance an initial level of public information and a lower level of private information only accessible under password control (or by some other security check procedure). Alternatively, there may be both a brief and a much fuller response .

[0033] A further possibility is to have responses from more than one narrator. Thus, a user can first obtain a response from one person associated with the picture and then a response from a different person also associated with the picture. For example, the user receives a picture from his/her parents. In the picture, the user's mother appears to be talking to someone, but the user's father looks bored. The user interaction might therefore proceed as follows:

User: "Mother, who is she?" [User uses pointing arrangement to indicate a person in the picture].
System: [In pre-recorded voice of user's mother] "This is my colleague from work. We had such a good time."
User: "Father, can you please describe the picture?"
System: [In pre-recorded voice of user's father] "We went to this rather boring evening with your mother's colleagues. They ended up talking about clothes all the time."

[0034] From the foregoing general description of a multi-modal picture, it will be appreciated that the term "multi-modal picture" covers more than just the picture image and includes also the associated dialog and the behaviour of the picture in response to user input.

Example Implementation

[0035] Figure 7 depicts an example implementation of a system for presenting a multi-modal picture.

[0036] The multi-modal picture is specified by a set 44 of related files held on a server 43, these files specifying both the image associated with the multi-modal picture and also the picture interaction dialogue associated with the picture. At least the dialogue and image files of set 44 are specifically associated with each other- that is, the association of these files is prespecified (in advance of user interaction) with the contents of the dialogue file being specifically tailored to the picture represented by the image file of the same set 44.

[0037] User 2 has local picture-presentation equipment 37; in the present case, this equipment is standard processing platform, such as a PC or portable computer, arranged to run a graphical web browser such as Microsoft's Internet Explorer. Other implementations of the picture presentation equipment 37 can alternatively be provided such as a mobile phone with a graphical display and suitable programs.

[0038] The picture-presentation equipment 37 displays the picture image of the multi-modal picture on display 38 and also provides for the input and output audio data via audio functionality 39, this audio input and output being transferred to and from a voice browser 3 (here shown as located separately from the picture picture-presentation equipment - for example, a network-based voice browser - but alternatively, integrated with the picture-presentation equipment 37). The picture-presentation equipment also has an associated pointing arrangement 40 depicted in Figure 7 as a mouse for controlling a cursor movable over the displayed picture image; other forms of pointing arrangements for selecting features of the picture image are, of course, also possible such as the use of touch pads, tracker balls or joysticks for moving an image cursor; touch-sensitive displays and other arrangements (such as a matrix of infra-red beams immediately overlying the display) for enabling a user to use a finger or stylus to point directly at a feature of the displayed image; etc..

[0039] The voice browser 3 comprises input channel 11, dialog manager 7 and output channel 12. The voice browser provides interactive dialog with the user, via the picture-presentation equipment 37, in respect of the currently-presented multi-modal picture, this dialog being in accordance with picture interaction dialog data retrieved from server 43.

[0040] The Figure 7 voice browser 3, as well as providing for voice input and output exchanges with the picture-presentation equipment 37, is also arranged to receive input data from the equipment in the form of key,value pairs indicating, for example, the selection of a particular picture feature by the user using the pointing arrangement 40 of the picture-presentation equipment. This data input is used by the dialog manager 7 in determining the course of the dialog with the user. The voice browser 3 can also provide data back to the picture-presentation equipment 37.

[0041] The picture-presentation equipment 37, voice browser 3 and server 3 inter-communicate via any suitable communications infrastructure or direct links.

[0042] Considering in more detail the files involved in the presentation of a multi-modal picture, these comprise, in addition to the set 44 of files that specify a multi-modal picture, a set of generic files 51. The multi-modal picture files

Amazon/VB Assets
Exhibit 1003
Page 261

44 comprise:

- the picture file 45, this file including a source reference for the picture image file 49 to be displayed and map data defining the image hotspots;
- a picture interaction dialog file 46 containing dialog scripts;
- one or more sound files 47, 48 (such as " .wav" files) containing audio data;
- the image file 49 (such as a ".jpg" or ".gif' file) containing the image data to be displayed.

[0043] The generic files 51 can be stored locally in the picture-presentation equipment 37 or retrieved from a remote location such as the server 43. The generic files comprise:

- a frame-set definition file 52 defining two frames 53, 54 into which page files can be independently loaded; one frame 53 is used to hold a file 55 containing control code (the contents of this frame not being visible), and the other frame 54 being used to hold the picture file 45 for the multi-modal picture to be presented.
- the control code file 55 to be loaded into frame 53, the control code being in the form of a number of scripts the main purpose of which is to provide key,value pairs to the voice browser according to events detected by the browser software run by the picture-presentation equipment 37 - in particular, clicking on an image hotspot as defined in file 45 is arranged to trigger a corresponding script in the control code file 55 whereby to cause a corresponding key,value pair to be passed to the voice browser 3 to inform it that a particular picture feature (corresponding to the activated hotspot) has been selected by the user.

[0044] To use the generic files to present a particular multi-modal picture, it is necessary to provide a reference to the multi-modal picture. This reference can, for example, be manually input by the user into an initial form displayed in the image frame 54 and then used by a script to request the loading of a file into frame 54. Alternatively, the multi-modal picture reference can be included as data in a query string attached to the URL of the frame-set definition file 52 (this URL and query string being, for example, provided to the user by the author of the multi-modal picture); in this case, in response to a request for the frame-set definition file 52, server-side code could, for example, extract the data from the query string and place it in the file source reference in the definition line for frame 54 in the frame-set definition file before that file is returned to the user.

[0045] In the present example, the multi-modal picture reference used to initiate presentation of the multi-modal picture is a reference to the picture interaction dialog file 46 to be loaded into the voice browser, rather than a reference to the picture file 45 that is to be loaded into the image frame 54. Thus, the multi-modal picture reference, however obtained, is passed in a key,value pair to the voice browser 3; voice browser 3 thereupon retrieves the picture interaction dialog file 46 to the dialog manager 7 of the voice browser. The file 46 includes a reference to the picture image file 45 to be loaded into the frame 54 and this reference is returned to the picture-presentation equipment 37 where it is used to retrieve the picture file 45. Where the picture-presentation equipment 37 comprises a standard web browser, one way of achieving the above is to have the dialog file reference sent (either as a source reference in the frame-set definition file or by a script in, for example, file 55) to the voice browser in a request for a file to load into frame 54, the voice browser sending back the picture file reference as a redirection.

[0046] It will be appreciated that the multi-modal picture reference passed into the generic files could have been that of the picture file 45, the latter then being retrieved into frame 54 and including an "onLoad" event script for passing to the voice browser a reference to the interaction dialog file.

[0047] The various steps involved in presenting a multi-modal picture according to the Figure 7 implementation are summarized below with reference to the reference numerals in square brackets in Figure 7:

[1] the generic files 51 have been loaded into the picture-presentation equipment 37 and a multi-modal picture reference has been provided; as a result, a key,value pair including a reference to the picture interaction dialog file 46 is passed to the voice browser 3;

[2] voice browser retrieves the dialog file 46 from the server 43 and dialog manger 7 uses the file to control further interaction with the user;

[3] the first action taken by the dialog manager 7 under the control of dialog file 46 is to pass to the equipment the reference for the picture file 45 along with a voice greeting to the user;

[4] the picture file 45 (including a reference to the image file 49) is retrieved from server 43 and loaded into the image frame 54;

[5] the image file 49 is retrieved from the server 43 and displayed on display 38;

[6] meanwhile, the dialog manager 7 causes a sound (in sound file 47) to be played to the user to indicate that the picture is ready to receive user input (this sound can simply be an appropriate background sound such as, for the Figure 6 picture image, the sound of the sea);

Amazon/VB Assets
Exhibit 1003
Page 262

[7] the user queries the picture by voice input (and possibly also by pointing to a particular area of the picture, this being indicated by a corresponding key,value pair sent to the voice browser along with the user voice input);
[8] dialog manager 7 acknowledges the receipt of the user query by causing an acknowledgement sound (in sound file 48) to be played back to the user;
[9] the dialog manager 7, having determined the appropriate response to the user query, outputs this response.

Steps [8] and [9] are repeated as many times as required by the user. In due course the user asks to exit and the dialog is terminated by the dialog manager.

[0048] Figure 8 illustrates the contents of the picture interaction dialog file 46. This file contains a number of dialog blocks 60 to 73 that contain dialog elements and/or control structures relating to dialog progression. Thus, dialog block 60 provides the initial greeting and causes the picture file reference to be passed to the equipment 37 (in step [3] above). Block 61 defines the query grammar and represents a waiting state for the dialog pending the receipt of a query from the user.

[0049] Block 62 carries out an analysis of a recognized query to determine whether it is an exit request (if so, an exit dialog block 63 is entered), a generic request, or a specific request; generic and specific requests are further analyzed to determine the nature of the query (that is, what "action" - type of information - is being requested). For a general query, the available actions are, in the present example, "date", "description", and "location; for a specific query, the action types are, in the present example, "what" and "story". Depending on the outcome of the action analysis, the dialog manager proceeds to one of blocks 64 - 66 (for a general query) or one of blocks 67 and 68 (for a specific query). The analysis carried out by dialog block 62 is on the basis of voice input only.

[0050] If the query was a general one, then block 64 is used to answer a date query, block 65 is used to respond to a description query, and block 66 is used to respond to a location query.

[0051] If the query was a specific "what" query, block 67 determines the identity of the picture feature (object) involved using the key,value pair provided to the voice browser; depending on the object identity, the dialog manager proceeds either to a "what" dialog block 70 for a coconut tree or to a "what" dialog block 71 for a French tourist. Similarly if the query was a specific "story" query, block 68 determines the identity of the picture feature (object) involved using the key,value pair provided to the voice browser; depending on the object identity, the dialog manager proceeds either to a "story" dialog block 72 for a coconut tree or to a "story" dialog block 73 for a French tourist.

[0052] After a response is provided by any one of the dialog blocks 64-66 or 70-73, the dialog manager returns to dialog block 61.

[0053] An example dialog interaction based on the Figure 6 picture and the Figure 8 dialog structure is given below; in this example, there is one narrator called Roger.

Example Interaction:

[0054]

System: "This picture was sent to you by Guillaume. You can talk to Roger to ask for information about it."
User: "Roger, Describe the picture"
System: "This is a picture of the XYZ beach."
User: "What is the location?"
System: "This picture was taken in Martinique, an island located in the French West Indies."
User: "What is it" [and clicks on feature 31]
System: "This is a coconut tree. This kind of trees is very common in the Caribbean."
User: "Any story?"
System: "When I was little, I used to get coconuts from my garden. There is a sweet liquid inside that is very tasty. We use coconut to make ice creams and all sorts of nice desserts."
User: "What is it?" [and clicks on 32]
System: "I guess this is a French tourist"
User: "Any story?"
System: "Tourism is one of the main source of revenue in Martinique. A lot of people come from France when it is winter over there."
User: "Exit"
System: "Thanks for using this digital picture frame."

[0055] The Appendix to the present description includes a detailed script example of the Figure 8 dialog interaction as well as the HTML source of a picture file 45.

Amazon/VB Assets
Exhibit 1003
Page 263

Multi-modal picture authoring tool.

[0056] Figure 9 illustrates apparatus for enabling an author 80 to author a multi-modal picture. The apparatus comprises a computing platform 81 providing a graphical user interface and running a tool, such as Microsoft's FrontPage product, for authoring mark-up language pages and for creating image hotspot maps. The apparatus further comprises a speech interface system 82 (here shown as a voice browser though other forms of speech interface system can be used). The speech interface system 82 permits the author 80 to interact with the apparatus by voice and is set up to recognize command words such as "Record".

[0057] Different people, known as narrators, can author different aspects of the same picture. The apparatus keeps a record of narrators known to it.

[0058] The apparatus is arranged to interact with one or more narrators to build up, in memory, the set of files 44 that specify a multi-modal picture, it being assumed that the picture image file 47 around which the multi-modal picture is to be built has already been loaded into memory (for example, from a digital camera) and is displayed via the graphical user interface (GUI) of platform 81. The process of building the required files is controlled by a top-level authoring program 90 that has three main steps 91-93 as will be more fully explained below.

Identifying the narrator - (step 91)

[0059] The first step of the authoring program is to identify the current narrator. The narrator speaks his/her name into the speech interface system 82; if the name is known to the apparatus, the system replies with a recognition greeting. However, if the narrator's name is not already known to the apparatus, the system asks the narrator to create a new profile (basically, input his/her name and any other required information), using appropriate data collection screens displayed on the graphical user interface of the computing platform 81.

[0060] Example: the apparatus knows the following names: "Lawrence" and "Marianne".

Apparatus: "What is your name?"
Narrator: "Steve".
Apparatus: "Sorry, I do not know this name, please write it down."
Narrator: [inputs "Steve" via a data collection screen of the GUI].

[0061] The authoring program uses the narrator's name to customize a greeting dialog block of a template picture interaction dialog file 46.

Adding general information about the picture - (step 92)

[0062] After identification, the narrator can input general information concerning the picture image such as the date, the location or the description, via a spoken dialogue. In the following example, the command words recognized by the speech interface system 82 are shown in bold whilst the nature of information being recorded (corresponding to the query "action" type of Figure 8) is indicated by underlining. The words indicating the nature of the information are either pre-designated to the system (effectively limiting the classification of information to be input) or else the system can be arranged to analyze narrator "Record" commands to determine the nature of the information to be recorded.

Example:

[0063]

Narrator: **"Record** description"
Apparatus: [Plays a beep].
Narrator: "This is a picture of me and John fishing in the Caribbean sea". [The apparatus records this input, either directly as sound data or as a text data after the input has been subject to speech recognition by the system 82]
Narrator: **"Write** date".
Apparatus: [displays date capture screen on GUI].
Narrator: [inputs date information via GUI].
Narrator: **"Record** story"
Apparatus: [Plays a beep.]
Narrator: "This day, John was attacked by a white shark." [The apparatus records this input]

Amazon/VB Assets
Exhibit 1003
Page 264

[0064] The authoring program uses the input from the narrator to create corresponding dialog blocks, similar to those described above with reference to Figure 8, in dialog file 46.

Adding specific information - (step 93)

[0065] The narrator can also input information concerning a specific feature of the picture image. To do this, the narrator indicates the picture feature of interest by using the GUI to draw a "hotspot" boundary around the feature. The apparatus responds by asking the narrator to input a label for the feature via an entry screen displayed on the GUI.. The authoring program uses the input from the narrator to create the multi-media picture file 45 holding an image hotspot map with appropriate links to the control code scripts.

[0066] The narrator can then enter further information using the speech interface system or the GUI. The narrator can record or write multiple descriptions or stories for a single area of the picture, for example, to give different level of details.

Narrator: **"Record** description".
Apparatus: [Plays beep.]
Narrator: "It is a whale" [The apparatus records this input]
Narrator: **"Record** story"
Apparatus: [Plays beep.]
Narrator: "We saw this whale on the way to Dominica." [The apparatus records this input]
Narrator: **"Record next"** (indicates that further details for the story are to be recorded).
Narrator: "We were crossing the Guadeloupe channel when we saw it." [The apparatus records this input]

[0067] Again, the authoring program uses the input from the narrator to create corresponding dialog blocks; thus, for the above example, where a "whale" hotspot has been designated by the user, the authoring program generates a set of dialogs blocks: 'whaleDescription', 'whaleStoryl', 'whaleStory2', etc.

[0068] After the first narrator has finished inputting information, other narrators can enter information in the same manner.

User Feedback

[0069] It will be appreciated that the authoring of a multi-modal picture, and in particular the adding of the dialog data, can be quite involved. As a result, it is quite likely that the author will not always include information that a recipient may be interested in.

[0070] It is therefore useful to be able to monitor a user's interaction with a multi modal picture to see if the user tries to access missing information. For example, a user receives the Figure 6 multi modal picture from a friend. Two objects 35,36 (a building and an island) in the picture image intrigues the user (see Figure 9) who therefore clicks on each picture feature concerned and asks for more information ("What is it?"). Unfortunately, there are no hotspots associated with either picture feature and therefore in each case the voice browser comes back with the response "Sorry, there is no information about this item."

[0071] However, the coordinates corresponding to each picture feature the user clicked are known to the browser used to display the picture image and the control script (for example, in file 55) can be used to pass the coordinates as key,value pairs to the voice browser 3. At the same time, a user verbal query is also passed to the voice browser. The voice browser first determines whether the query is a general one and if it is, the voice browser ignores the received coordinate values; however, if the voice browser determines that the query is a specific one, then it determines from the key,value pairs received that the user has indicated a picture feature for which there is no corresponding information available. In this case, the voice browser logs the received coordinate values and the associated "action" type (in the example given above, "what" or "story") in a feedback file 50 that forms part of the set 44 of related files associated with the multimodal picture (see arrow 70 in Figure 7). For example, upon the user clicking on the island feature 36 in Figure 9 and asking "what is this?", the data (action='what'; coordx =400;coordy=300) is logged to file 50. Such logging functionality is, for example, provided by a further dialog block of Figure 8. The logged coordinates provide, together with an indication of the picture concerned, a picture-feature identifier that identifies the picture feature about which information has been requested by the user.

[0072] Alternatively or additionally to logging the "desired-information" feedback data in file 50, the author of the multi-modal picture can be sent a message (for example, an e-mail message) explaining the query from the user such as "John wants a description of this object.". This message includes a picture feature identifier that identifies the picture feature concerned. The picture feature identifier can take the form of explicit data items indicative of the picture concerned and the coordinates of the feature in the picture or may more directly indicate the feature by including either

Amazon/VB Assets
Exhibit 1003
Page 265

image data representing the relevant portion of the picture image or image data showing the picture image with a marking indicating the position in the image of the feature concerned (both such forms of picture-feature indication can also be included the feedback file 50 additionally or alternatively to the feature coordinates). The picture-feature indication need not be sent to the file 50 (or included in a message to the author) at the time the system detects that the user has asked for information about a non-hotspot picture feature; instead, the indication and related query can be temporarily stored and output, together with other accumulated similar indications and queries, either at a specified time or event (such as termination of viewing of the picture by the user) or upon request from the picture author.

[0073] The author can then decide whether or not to add in further information by adding an additional hotspot and additional dialogs (for example, using the above-described authoring apparatus).

[0074] Of course, the same general feedback process can be used where although a selected picture feature is associated with an existing hotspot, there is no existing query "action" corresponding to a user's query. Furthermore, a similar feedback process can be used where user queries are input by means other than speech (such as, for example, via a keyboard or by a hand-writing or stroke recognition system) and, indeed, where there are no explicit user queries such as when the selecting of picture feature is taken as a request for information about that feature.

Variants

[0075] Many variants are, of course, possible to the arrangements described above. For example, Figure 11 illustrates a variant form of the multi-modal picture in which the picture image 30 is accompanied by upper and lower information bars 100 and 101 respectively. The upper information bar 100 indicates the narrators associated with the picture whilst the lower information bar 101 indicates what types of general and specific queries are available for use. These information bars assists the user in appreciating what queries can be put and to whom. To further assist the user, speaking (or clicking on) a narrator's name is preferably arranged to indicate what hotspots are associated with that narrator and where these hotspots are located in the image - thus, in Figure 11, narrator "Vivian" has been selected and hotspot 32 is indicated on the picture image by a dashed hotspot boundary line. Query types used by "Vivian" can also be indicated by highlighting these types (in the Figure 11 example, "Vivian" has used general query type 'description' and specific query types 'what' and 'story'). Preferably, once a narrator has been selected, that narrator remains selected until a different (or all) narrators is subsequently selected, whereby with a narrator selected, only the responses of that narrator will be used in responding to the users' queries. However, it is also possible to arrange for selection of a narrator to be effective only for a single query. In this latter case, a convenient way of providing for user selection of a narrator to answer a query about a particular feature is for a list of narrators of responses about that feature to be displayed whenever the user points to that feature.

[0076] Instead of identifying narrators by name as shown in Figure 11, other forms of identifier can be used such as an image of each narrator.

[0077] As regards the generic control code, this can be provided in the form of a Java applet or any other suitable form and is not limited to the use of client-side scripts of the form described above. Furthermore, rather than the frame-set definition file 52 and control code file 55 being generic in form, they can be made specific to each multi-modal picture.

[0078] As already indicated, picture invocation can be initiated in a variety of ways. As a further example, picture invocation can be arranged to be effected by the creator sending the user a reference to the picture interaction dialog file, the user sending this reference to the voice browser to enable the latter to retrieve the dialog file; the dialog file is arranged to cause the return to the user of the frame-set definition file (or a reference to it) with the latter already including the correct source reference to the picture file as well as to the control code file.

[0079] Persons skilled in the art will appreciate that there are very many ways of implementing multi-modal pictures and the supporting functionality ranging from the multiple file approach described above to having just a single data file containing all the necessary data and arranged to be executed by specifically adapted software. It will also be appreciated that whilst in the described embodiments the course of the system-user interaction is controlled by the control logic embedded in the picture interaction dialog file and interpreted by the dialog manager 7 of the voice browser, it is possible to provide this functionality separate from the voice browser and the response scripts (thus, in general terms, the dialog block 62, 67 and 68 would be used by a separate control arrangement for determining, on the basis of the user voice and pointing-arrangement inputs, which of multiple stored responses is the appropriate one to play back to the user). Furthermore, it is not necessary to explicitly identify a selected feature from the coordinates output by the pointing arrangement as a separate step to choosing an appropriate response to a particular user query; thus, a picture feature could simply be specified in the control logic of the dialog interaction file by its coordinate values (of range of values) whereby this control logic tests coordinate values output by the pointing arrangement against coordinate values of particular features in the course of selecting a response to a user query.

[0080] With respect to the selection of a picture feature of interest to a user, in the described embodiments this has been carried out by means of a user-operated pointing arrangement whereby data input about the picture feature of interest is generated through a manual operation. In addition to the described arrangements for manually effecting

Amazon/VB Assets
Exhibit 1003
Page 266

feature selection, it is possible to use an arrangement in which the coordinates of the feature of interest are manually input using a keypad (a coordinate system being displayed as part of the image presented to the user, or being presented around the boundary of the display area). Another possible feature-selection arrangement is one based on specifying, via keypad input, a particular area of the display where either the display is divided into labelled areas or there is a direct mapping between keypad keys and display areas. It is also possible to label each feature of interest in the image with a reference and have the user effect feature selection by keypad input of the appropriate reference.

[0081] Further as regards the determination of which picture feature is being selected by a user, it should be noted that determining the picture feature of interest from the image coordinates or image area identified by whatever selection arrangement is being used, can be done in ways other than that described above in which image coordinates generated by the feature-selection arrangement are mapped to picture features using predetermined mapping data as described above. For example, the image can have data encoded into it that labels particular picture features, the pointing arrangement being arranged to read this label data when over the corresponding picture feature. Technology for embedding auxiliary data in picture image data in a way that does not degrade the image for the average human viewer is already known in the art.

[0082] The picture image can be a hard-copy image carrying markings (such as infra-red ink markings) intended to be read by a sensor of a suitable pointing arrangement whereby to determine what particular picture feature is being pointed to; the markings can be a pattern of markings enabling the pointing arrangement to determine the position of its sensor on the image (in which case, an image map can be used to translate the coordinates into picture features) or the markings can be feature labels appropriately located on the image. Even without special markings added to the image, the image can still be a hard-copy image provided the image is located in a reference position relative to which the pointing arrangement can determine the position of a pointing element (so that an image map can be used to translate pointing-element position into picture features). Other manually-operated selection arrangements, such as those based on explicit coordinate input via a keypad, can also be used.

[0083] Where image map data is used to translate image coordinates to picture features, the image map data can be held separately from the picture file and only accessed when needed; this facilitates updating of the image map data. A reference (e.g. URL) to the image map data can be included in the picture file or, where the image is a hard-copy image, in markings carried by the image.

[0084] Preferably, the described embodiments are applied to pictures of scenes and places in the real world such as a tourist might take with a camera. However, it is also possible to apply the embodiments to topographic pictures that are primarily intended to convey map-type information.

Amazon/VB Assets
Exhibit 1003
Page 267

APPENDIX

This Appendix forms an integral part of the description.

Example of Multi-Modal Picture file.

The following HTML file was generated using Microsoft FrontPage and is the Multi-Modal Picture file 45 for the picture depicted in Figure 6; the HTML code pulls in image file "beachPicture.jpg" and defines three "hotspots" which when clicked result in a call to a function (method) of the control code file 55 with the passing of a parameter value indicating which hotspot has been clicked on.

File Name: MMPicture1.htm

```
<html>
<head>
<meta http-equiv="Content-Type" content="text/html; charset=windows-1252">
<meta name="GENERATOR" content="Microsoft FrontPage 4.0">
<meta name="ProgId" content="FrontPage.Editor.Document">
<title>Frame Page</title>
</head>
<body>
<p><map name="FPMap0">
<area href="http://localhost/mmpic1/dpfBlank.htm"
onclick="parent.parent.scriptFrame.pushParams('object=coconutTree')"
target="invisible" coords="128, 84, 253, 236" shape="rect">
<area href="http://localhost/mmpic1/dpfBlank.htm"
onclick="parent.parent.scriptFrame.pushParams('object=frenchTourist')"
target="invisible" shape="rect" coords="204, 266, 219, 282">
<area href="http://localhost/mmpic1/dpfBlank.htm"
onclick="parent.parent.scriptFrame.pushParams('object=frenchTourist')"
target="invisible" shape="rect" coords="309, 257, 324, 269"></map>
<img border="0" src="beachPicture.jpg" align="center" usemap="#FPMap0"
width="569" height="375"></p>
</body>
</html>
```


Amazon/VB Assets
Exhibit 1003
Page 268

Example of Picture Interaction Dialogue file

[0085] The following is an example implementation of the Figure 8 picture interaction dialog file 46. In this example, there is one narrator called Roger. The user can say things like:"Roger, what is the date?".

```
<talkml>
    <!--variables declaration-->
    <create name="object"          expr="'undefined'"/>
    <create name="action"          expr="'what'"/>
    <create name="type"        expr="'generic'"/>
    <!-- Specify here the url of the display webserver-->
    <display-server url="localhost"/>

    <dialog name="welcome">
        <say> Welcome to the digital picture frame. </say>
        <!-- Display the Multi Modal picture of beach. -->
        <display-show url="MMPicture1.htm"/>
        <say>
            This picture was sent to you by Guillaume. You can talk to Roger to
            ask for information about it.
        </say>
        <goto next="getUserRequest"/>
    </dialog>

    <dialog name="getUserRequest">
        <play src="sound1.wav"/>
        <display-listen wait_after="2000"/>
        <speech wait_after="2000">
            sentence =  narrator question | question | exit {action='exit'};
            question =  genericQuestion {type='generic'} | specificQuestion
                        {type='specific'};
```


Amazon/VB Assets
Exhibit 1003
Page 269

```
        narrator = roger {narrator='roger'};
        roger    = "roger";
        genericQuestion = date {action='date'}| location {action='location'}|
                          description {action='description'};
        date     = "what is the date";
        location = "what is the location";
        description = "describe the picture";
        specificQuestion= what {action='what'} | story {action='story'};
        what     = "what is it";
        story    = "any story";
        exit     = "exit";
    </speech>

    <recognised>
        <play src="sound2.wav"/>
        <goto next="filterRequest"/>
    </recognised>

    <unrecognised>
        <play src="sound2.wav"/>
        <say> I am sorry, I have some trouble understanding you.</say>
        <retry/>
    </unrecognised>

</dialog>


<dialog name="filterRequest">
    <if cond="action='exit'">
        <goto next="exit"/>
    </if>
```


Amazon/VB Assets
Exhibit 1003
Page 270

```
        <if cond="type='generic'">
                <if cond="action='description'"><goto next="description"/></if>
                <if cond="action='date'"><goto next="date"/></if>
                <if cond="action='location'"><goto next="location"/></if>
        </if>

        <if cond="type='specific'">
                <if cond="action='what'">    <goto next="what"/> </if>
                <if cond="action='story'">    <goto next="story"/></if>
        </if>

</dialog>


<dialog name="description">
        <say>
                This is a picture of the beach.
        </say>
        <goto next="getUserRequest"/>
</dialog>


<dialog name="date">
        <say>
                Guillaume took this picture during his last trip home to
                Martinique last November
        </say>
        <goto next="getUserRequest"/>
</dialog>


<dialog name="location">
```


Amazon/VB Assets
Exhibit 1003
Page 271

```
        <say>
            This picture was taken in Martinique, an island located in the
            French West Indies.
        </say>
        <goto next="getUserRequest"/>
</dialog>

<dialog name="what">
        <if cond="object='coconutTree'"><goto next="coconutTreeWhat"/></if>
        <if cond="object='frenchTourist'"><goto next="frenchTouristWhat"/></if>
        <say>
            Apparently, you did not click on any parts of the picture. Try
            again.
        </say>
        <goto next="getUserRequest"/>
</dialog>

<dialog name="story">
        <if cond="object='coconutTree'"><goto next="coconutTreeStory"/></if>
        <if cond="object='frenchTourist'"><goto next="frenchTouristStory"/></if>
        <say>
            Apparently, you did not click on any parts of the picture. Try
            again.
        </say>
        <goto next="getUserRequest"/>
</dialog>

<dialog name="frenchTouristWhat">
        <say>
            I guess this is a French tourist.
        </say>
```


Amazon/VB Assets
Exhibit 1003
Page 272

```
        <goto next="getUserRequest"/>
    </dialog>


    <dialog name="coconutTreeWhat">
        <say>
            This is a coconut tree. This kind of trees is very common in the
            Caribbean.
        </say>
        <goto next="getUserRequest"/>
    </dialog>


    <dialog name="frenchTouristStory">
        <say>
            Tourism is one of the main source of revenue in Martinique. A lot
            of people come from France when it is winter over there.
        </say>
        <goto next="getUserRequest"/>
    </dialog>


    <dialog name="coconutTreeStory">
        <say>
            When I was little,I used to get coconuts from my garden. There is
            a sweet liquid inside that is very tasty. We use coconut to make
            ice creams and all sorts of nice desserts.
        </say>
        <goto next="getUserRequest"/>
    </dialog>
```


Amazon/VB Assets
Exhibit 1003
Page 273

```
<dialog name="exit">
        <say> Thanks for using this digital picture frame. </say>
</dialog>

</talkml>
```

Claims

1. A system for presenting information concerning a picture to a user, the system comprising:

   - a data store for holding responses, specific to said picture, in respect of specific user queries concerning particular picture features;
   - a manually-operable feature-selection arrangement for enabling a user to select a feature in a displayed image of the picture, and for providing an output indication regarding what said particular feature, if any, the user has thereby selected;
   - a voice dialog input-output subsystem including a speech recogniser for interpreting queries from a user;
   - a control arrangement responsive to a user selecting a said particular feature and asking a specific query regarding that feature, to output the corresponding stored response.

2. A system according to claim 1, wherein image-map data is associated with the picture image for mapping image coordinates to said particular features, the selection arrangement being arranged to use the image-map data to determine what picture feature is selected by the user.

3. A system according to claim 1, wherein said image includes label data positioned in the region of a said particular picture feature to indicate the identity of that feature, the selection arrangement being arranged to read the label data to determine what picture feature is selected by the user.

4. A system according to any one of claims 1 to 3, further comprising a display subsystem for displaying said image provided to it in the form of digital image data.

5. A system according to any one of claims 1 to 3, wherein the picture image is a hard-copy image.

6. A system according to any one of the preceding claims, wherein the data store is arranged also to hold responses concerning general queries that are not associated with any particular picture feature, the control arrangement being arranged to respond to the user voice input of a general query by returning the appropriate response.

7. A system according to any one of the preceding claims, wherein the voice interface subsystem and control arrangement are arranged to cooperate in recognising multiple different queries in respect of a particular picture feature selected by the user.

8. A system according to claim 1, wherein each of at least some of the responses is associated with a specified narrator, the system being arranged to permit a user to receive only the response of a user-selected narrator in respect of at least one query.

9. A system according to claim 8, wherein user selection of a narrator is arranged to be effected by voice input, the control arrangement being arranged to respond to the selection of a particular specified narrator by user voice input, by using only the said responses associated with that narrator in providing a response to a user query.

10. A system according to claim 9, further comprising means for displaying, along with said picture image, identifiers of narrators associated with the picture.

11. A system according to claim 9, further comprising a display subsystem for displaying said image provided to it in

Amazon/VB Assets
Exhibit 1003
Page 274

the form of digital image data and identifiers of narrators associated with the picture; the display subsystem being arranged to respond to selection of a particular specified narrator by user voice input by indicating on the displayed image the said particular features for which responses are available concerning that narrator.

12. A system according to claim 8, further comprising a display subsystem for displaying said image provided to it in the form of digital image data, and identifiers of narrators associated with the picture; user selection of a narrator being arranged to be effected by the user using said selection arrangement to select a displayed narrator identifier, and the control arrangement being arranged to respond to the selection of a particular narrator by using only the said responses associated with that narrator in providing a response to a user query.

13. A system according to claim 12, wherein the display subsystem is arranged to respond to selection of a particular specified narrator to indicate on the displayed image the said particular features for which responses are available concerning the currently-selected narrator.

14. A system according to any one of the preceding claims, wherein said selection arrangement is a pointing arrangement usable by the user to point to a feature of interest in a displayed view of the picture.

15. A system according to any one of the preceding claims, further comprising a feedback arrangement arranged to respond to a user query concerning a picture feature that is other than a said particular picture feature, by outputting at least an identifier of the concerned picture feature for informing an author of the picture or its associated responses regarding the unsatisfied query.

16. A system according to any one of the preceding claims, further comprising a feedback arrangement arranged to respond to a user query that although concerning a said particular picture feature is not a said specific query, by outputting an indication of the query and an identifier of the concerned picture feature for informing an author of the picture or its associated responses regarding the unsatisfied query.

17. A system according to claim 15, wherein the feedback arrangement is further arranged to output an indication of the user query together with said picture feature identifier.

18. A system according to any one of claims 15 to 17, wherein the picture feature identifier arranged to be output by the feedback arrangement comprises an indication of the picture concerned and image coordinates of the picture feature.

19. A system according to any one of claims 15 to 17, further comprising means for converting digital picture data into said displayed image of the picture, the picture feature identifier arranged to be output by the feedback arrangement comprising digital data representing a corresponding portion of the image.

20. A system according to any one of claims 15 to 17, further comprising means for converting digital picture data into said displayed image of the picture, the picture feature identifier arranged to be output by the feedback arrangement comprising digital data representing the image with a marking indicating the position of the feature in the image.

21. A system according to any one of claims 15 to 20, wherein the feedback arrangement is operative to output said picture feature identifier to a log file.

22. A system according to any one of claims 15 to 20, wherein the feedback arrangement is operative to output said picture feature identifier in the form of an electronic message addressed to a person associated with the generation of the picture or its associated responses.

23. A multi-modal picture specified by data held on at least one data carrier, this data comprising:

- picture image data for displaying a picture image;
- response data indicative of voice responses intended to be given to specific user queries concerning particular picture features of the picture;
- first control data for enabling a determination to be made as to which said particular feature in the picture image, if any, a user is selecting when using a selection arrangement to indicate a feature in the displayed image; and
- second control data for determining, on the basis of a spoken user query and on which said particular picture

Amazon/VB Assets
Exhibit 1003
Page 275

feature is selected by the user using the selection arrangement, which said response is to be used to reply to the user query.

24. A multi-modal picture according to claim 23, wherein the first control data comprises image-map data mapping image coordinates to said particular features.

25. A multi-modal picture according to claim 23, wherein the first control data comprises label data arranged to be positioned in the displayed image in the region of each said particular picture feature to indicate the identity of that feature.

26. A multi-modal picture according to claim 23, wherein at least some of the said responses are associated with a narrator identified in the response data, the second control data enabling the determination of which said response is to be used to reply to the user query to be restricted to those responses associated with a said narrator that has been selected by the user.

27. A multi-modal picture according to any one of claims 23 to 26, wherein the picture is of a non-topographic real-world scene.

28. A multi-modal picture comprising a hard-copy picture image, and data held on at least one data carrier, this data comprising:

- response data indicative of voice responses intended to be given to specific user queries concerning particular picture features;
- first control data for enabling a determination to be made as to which said particular feature in the picture image, if any, a user is selecting when using a selection arrangement to indicate a feature of the image; and
- second control data for determining, on the basis of a spoken user query and on which said particular picture feature is selected by the user using the selection arrangement, which said response is to be used to reply to the user query.

29. A multi-modal picture according to claim 28, wherein the first control data comprises image-map data mapping image coordinates to said particular features.

30. A multi-modal picture according to claim 28, wherein the first control data comprises label data arranged to be positioned in or on the image in the region of each said particular picture feature to indicate the identity of that feature.

31. A multi-modal picture according to claim 28, wherein at least some of the said responses are associated with a narrator identified in the response data, the second control data enabling the determination of which said response is to be used to reply to the user query to be restricted to those responses associated with a said narrator that has been selected by the user.

32. A multi-modal picture according to any one of claims 28 to 31, wherein the picture is of a non-topographic real-world scene.

33. Apparatus for authoring a multi-modal picture, comprising:

- a first tool for defining image hotspots associated with particular picture-image features;
- a second tool with speech recognition capability, for recording user responses input by voice, to user-specified queries each associated with a particular said picture-image feature; and
- means for automatically generating control data for determining, on the basis of a spoken user query and on which said particular picture feature is selected by a user, which said response is to be used to reply to the user query.

Amazon/VB Assets
Exhibit 1003
Page 276

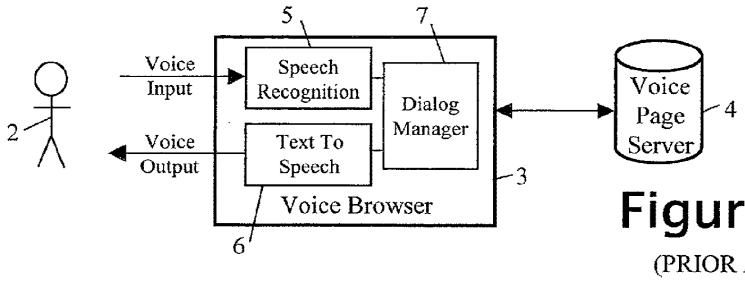


Figure 1
(PRIOR ART)

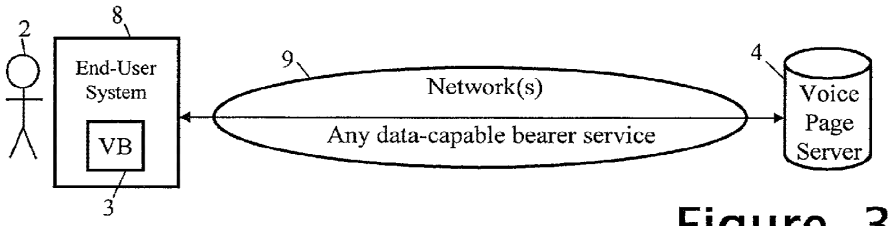


Figure 3
(PRIOR ART)

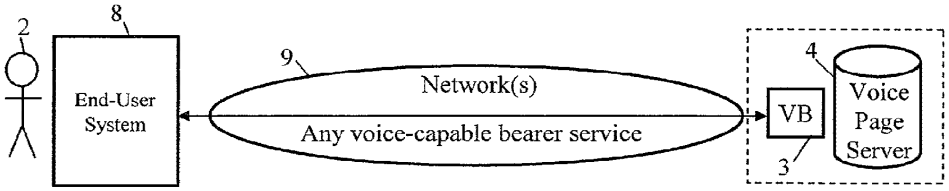


Figure 4
(PRIOR ART)

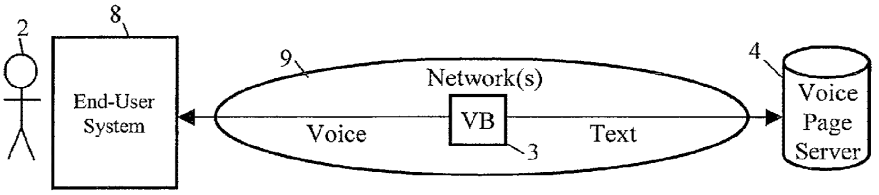


Figure 5
(PRIOR ART)

Amazon/VB Assets
Exhibit 1003
Page 277

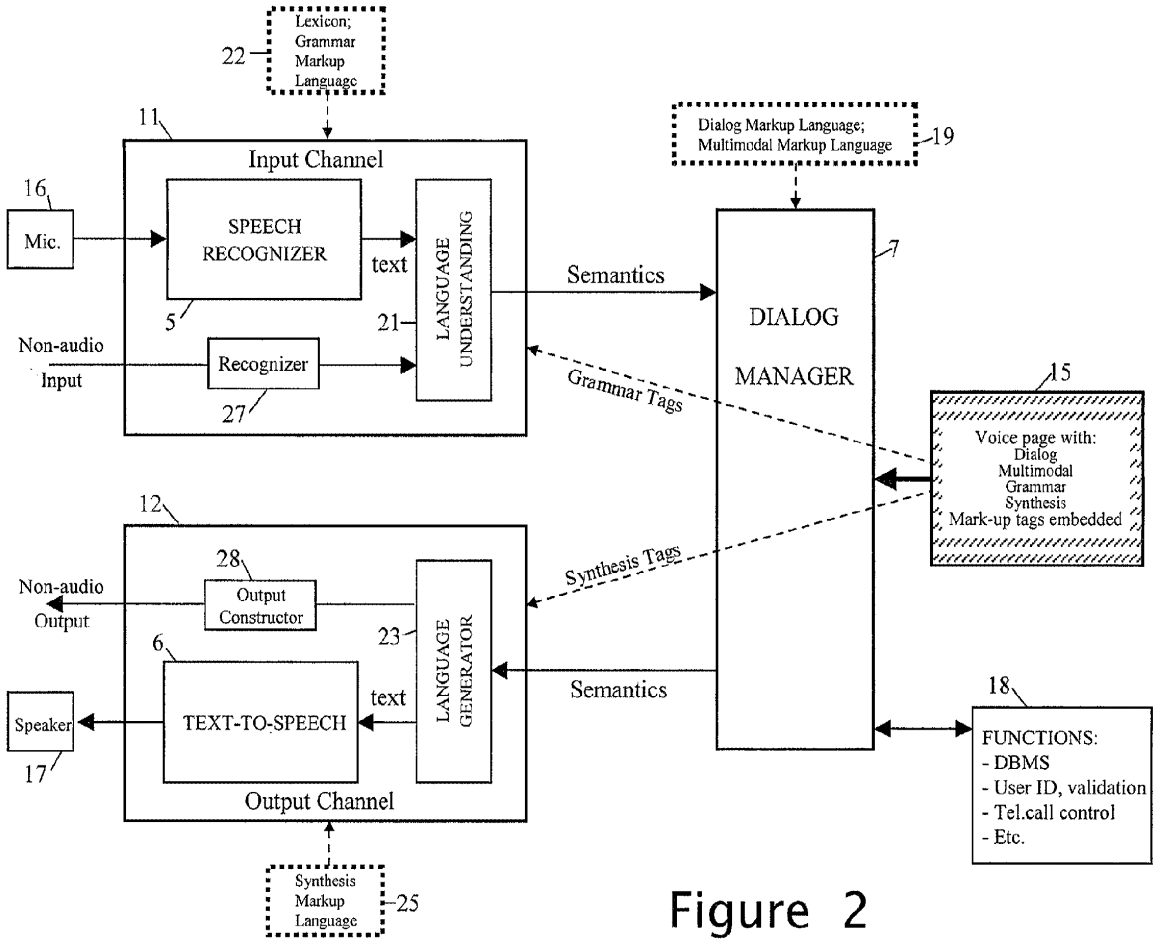


Figure 2
(PRIOR ART)

Amazon/VB Assets
Exhibit 1003
Page 278

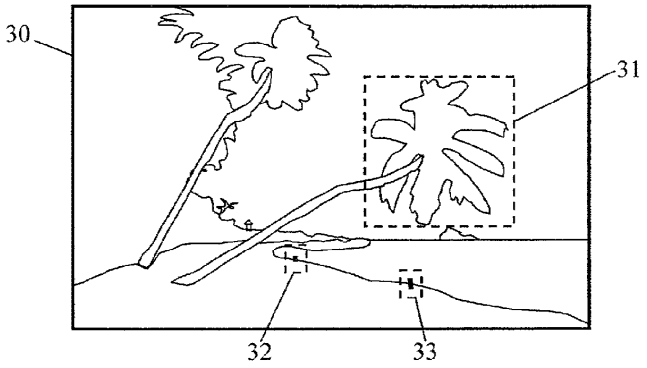


Figure 6

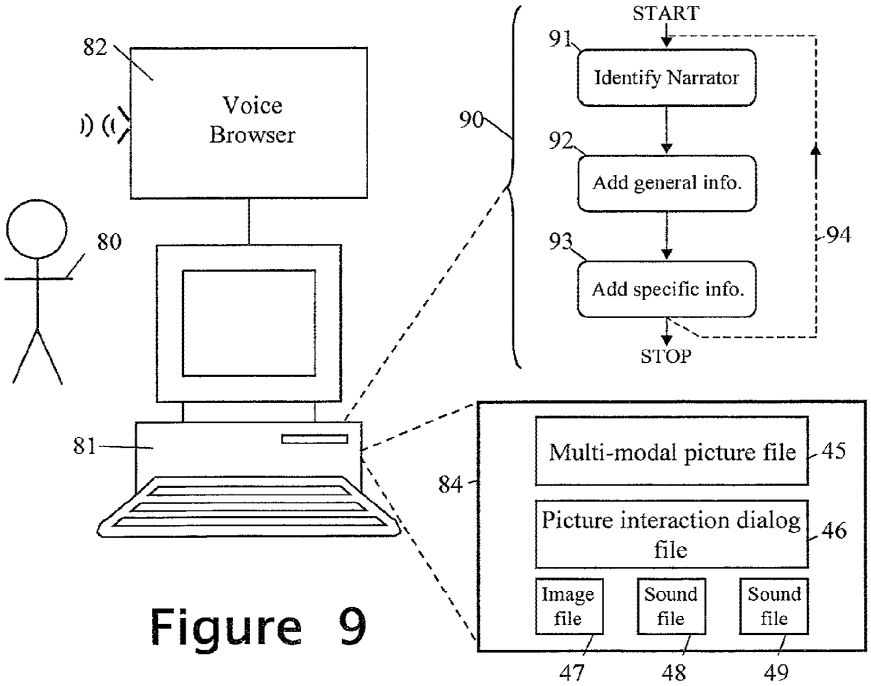


Figure 9

Amazon/VB Assets
Exhibit 1003
Page 279

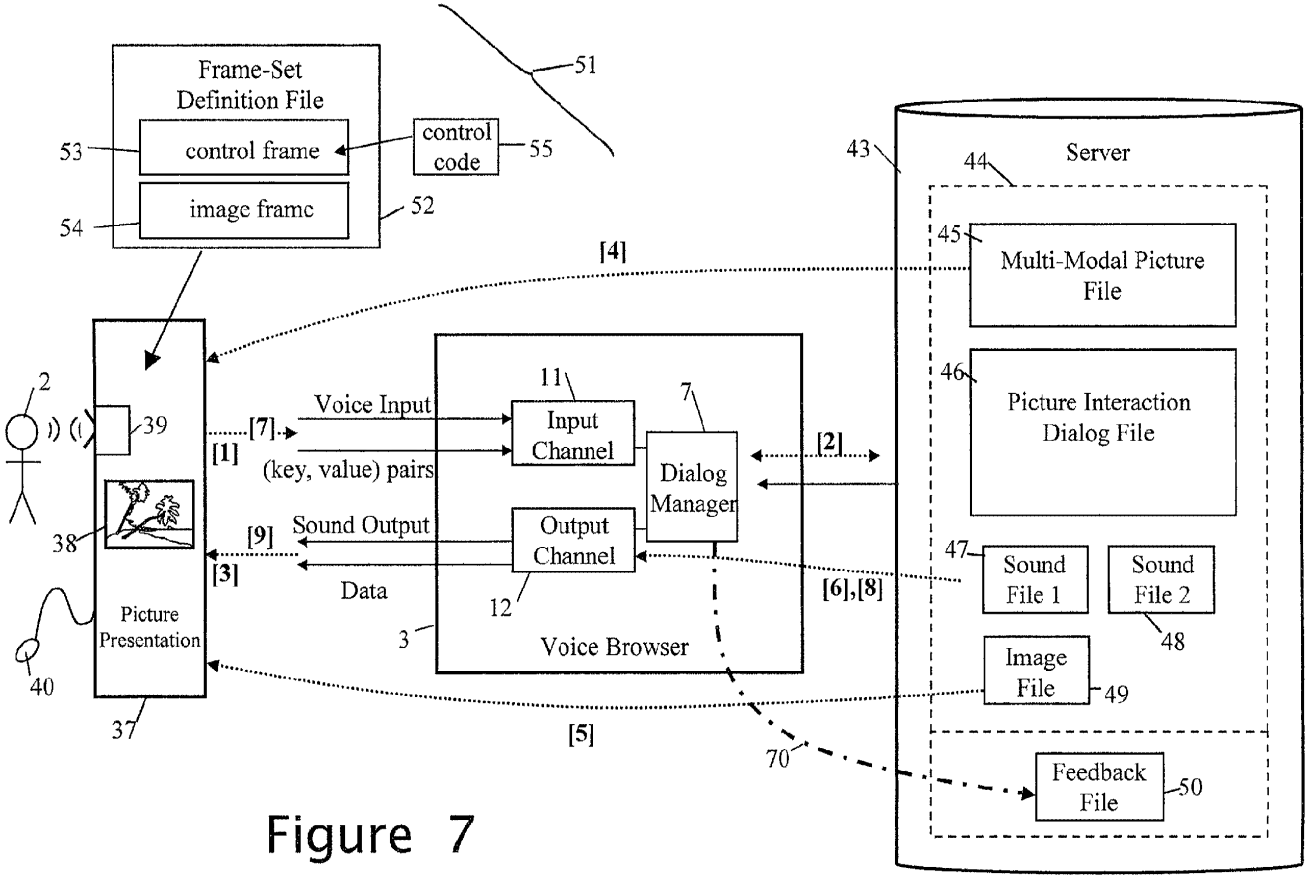
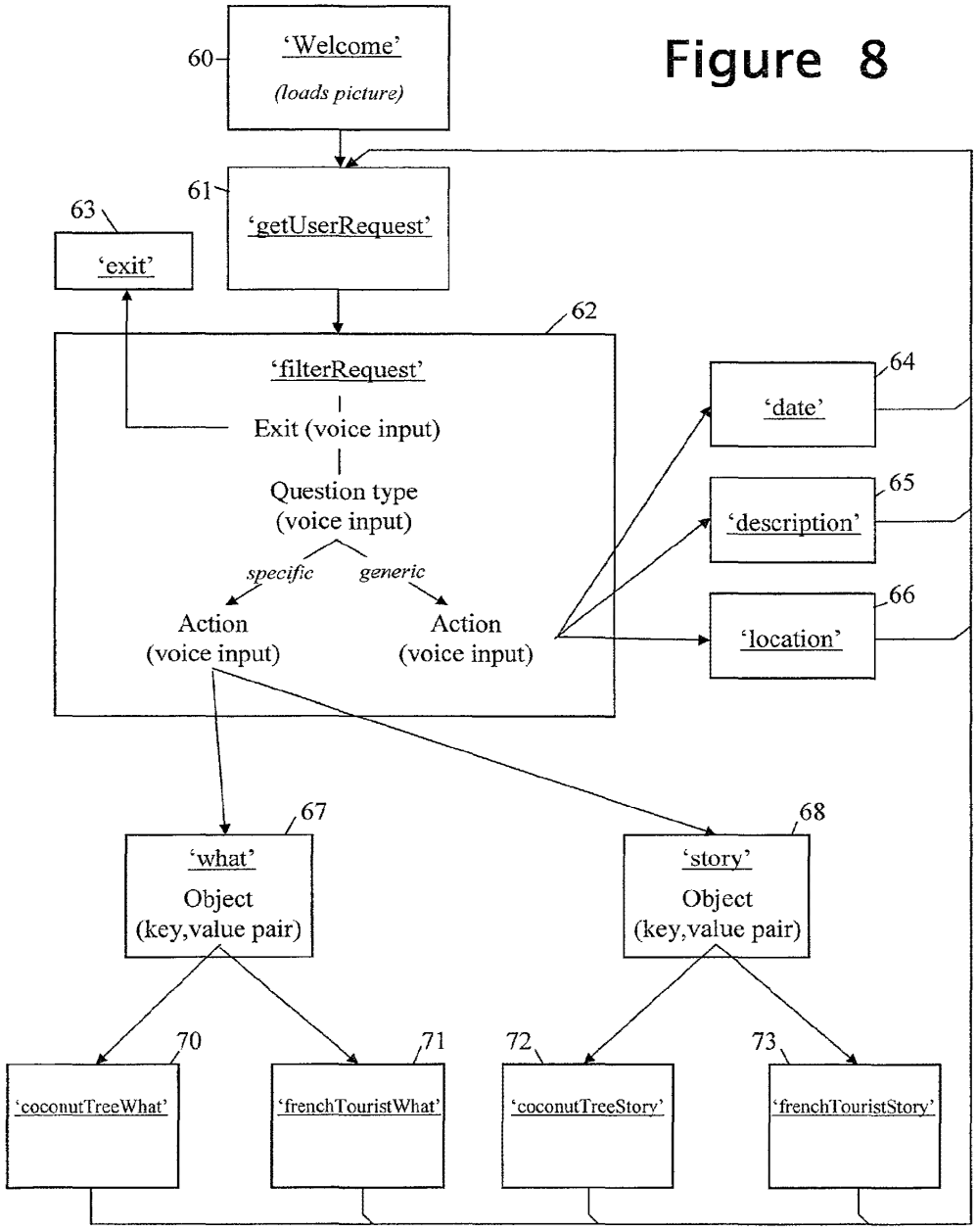


Figure 7

EP 1 320 043 A2

Amazon/VB Assets
Exhibit 1003
Page 280

Figure 8



Amazon/VB Assets
Exhibit 1003
Page 281

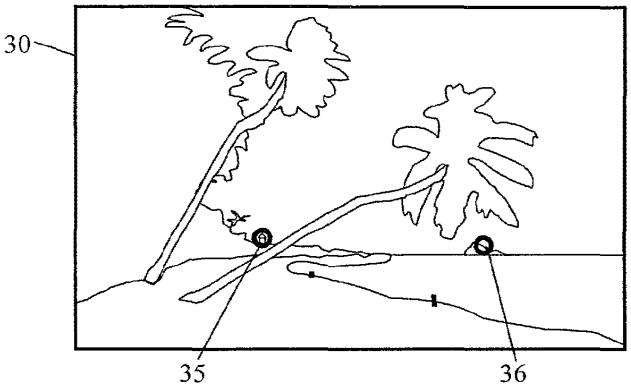


Figure 10

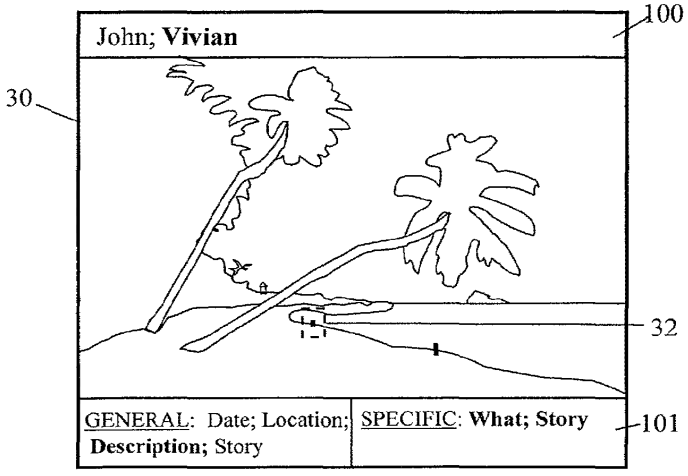


Figure 11

Amazon/VB Assets
Exhibit 1003
Page 282

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 646 037 B1

(12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention of the grant of the patent:
26.03.2008 Bulletin 2008/13

(51) Int Cl.:
G10L 15/26 *(2006.01)*

(21) Application number: **05291989.1**

(22) Date of filing: **23.09.2005**

(54) Method and apparatus for enhancing speech recognition accuracy by using geographic data to filter a set of words

Verfahren und Vorrichtung für erhöhte Spracherkennungspräzision unter Verwendung geografischer Daten zur Filterung einer Wortgruppe

Procédé et appareil pour améliorer la précision de la reconnaissance vocale en utilisant des données géographiques pour filtrer un lot de mots

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU LV MC NL PL PT RO SE SI SK TR

(30) Priority: **29.09.2004 US 953671**

(43) Date of publication of application:
12.04.2006 Bulletin 2006/15

(73) Proprietor: **FRANCE TELECOM**
75015 Paris (FR)

(72) Inventors:
- **Waters, Keith**
  West Newton
  MA 02465 (US)
- **Simon, Benoît Benjamin**
  Boston
  MA 02114 (US)

(56) References cited:
EP-A- 0 661 688
EP-A- 1 233 407
US-B1- 6 230 132

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Printed by Jouve, 75001 PARIS (FR)

EP 1 646 037 B1

Description

FIELD OF THE INVENTION

[0001] The invention relates generally to speech recognition systems and methods and, more particularly, to systems and methods for enhancing speech recognition accuracy by using geographic data to reduce the size of a set of words that must be checked to effect speech recognition.

BACKGROUND OF THE INVENTION

[0002] While visiting an unfamiliar city, a mobile handset user may wish to locate a business, landmark, street, or other place of interest. It would be quite advantageous if the user could place a voice call to a remote server at a designated telephone number, speaking his or her request for information into the handset. The handset would execute a speech recognition procedure, transforming the spoken request into a server query, and receiving the requested information from the remote server in the form of an electronic signal for producing audible speech. For example, a user in Manhattan may wish to locate the closest Neiman Marcus store. She dials a designated telephone number and speaks a request such as "where is the closest Neiman Marcus?" The handset uses speech recognition to convert the request into an electronic query for receipt by the remote server. Based upon the current location of the handset, the remote server accesses a database to retrieve the address and phone number of the nearest Neiman Marcus. The address and phone number are sent from the server to the user's handset in the form of an electronic signal, whereupon the handset converts the electronic signal into audible speech.

[0003] Unfortunately, there are significant hurdles to overcome if one wishes to use existing speech recognition algorithms to carry out the foregoing procedures. More specifically, it is a challenge to develop speech recognition algorithms, which provide a desired level of accuracy in certain operational environments, such as that of mobile telephonic devices. In general, the processing power and memory capabilities of mobile telephonic devices are limited by size and cost constraints, thereby placing limits on the extent to which the complexity of existing algorithms can be increased. Moreover, the existence of ambient noise severely degrades the performance of many speech recognition algorithms. In view of these shortcomings, prior art techniques have been developed to improve the performance of speech recognition algorithms in applications where processing power may be limited. These prior art techniques are directed toward enhancing a recognition model by utilizing improved acoustic models. The following definitions are presented herein for purposes of understanding and clarity:

[0004] *Lexical Model:* A lexical model breaks down individual words of a syntactic model (defined hereinafter) into a sequence of phonemes (also defined hereinafter).

[0005] *Acoustic Model:* An acoustic model breaks down the sequence of phonemes determined by the lexical model into a Markov model of cepstrum coefficients, also termed vector sequences. Cepstrum coefficients are generated by taking the Fourier Transform of the power spectrum of a speech signal.

[0006] *Recognition Model:* A recognition model is a set of rules and operations that are used by a recognition engine to recognize individual words. The recognition model includes a fully deployed speech recognition system which encompasses the acoustic model, the lexical model and the syntactic model as follows: recognition model = (acoustic model) x (lexical model) x (syntactic model).

[0007] U.S. Patent Application Publication No. US2003236099 discloses a technique for improving the accuracy of a speech recognition algorithm by selecting an appropriate acoustic model in accordance with the location of a mobile telephonic device ("mobile handset", or just "handset" for the sake of convenience). As indicated above, an acoustic model includes a set of cepstrum coefficient statistics for each phoneme in a language. Several different acoustic models can be precompiled, each representing a unique acoustical environment, such that a speech recognition system can select an acoustic model appropriate for a given set of environmental conditions. For example, each acoustic model is prepared by individually recording several hundred speakers in a controlled, model-specific laboratory environment that has known acoustical properties illustratively specified in terms of noise, echo, reverberation, and frequency response. Each model-specific laboratory environment is intended to represent a real-world environment, such as a warehouse, an office, a shopping mall, an airport, or the interior of an automobile. Hence, acoustic models incorporate properties of specific environments as well as characteristics of the human voice. The laboratory environment may include artificially generated sounds to simulate, for example, traffic, conversations, computer keyboard clicks, and forklift trucks.

[0008] Pursuant to this prior art technique, position information is retrieved from a location-based service (LBS), as defined hereinafter, based upon the current location of the mobile handset. The retrieved position information is then employed to select the acoustical properties of the environment surrounding the handset. An acoustic model that approximates the acoustical properties of the type of environment for the current location of the handset is selected from a set of previously prepared acoustic models. The selected acoustic model is then utilized to improve the accuracy of the speech recognition algorithm.

[0009] The technique disclosed in U.S. Patent Application Publication No. US2003236099 presents many practical shortcomings. Existing location-based services, using the Global Positioning System (GPS), will not

Amazon/VB Assets
Exhibit 1003
Page 284

determine a user's position to a requisite accuracy of, say, one to two meters, nor will such services determine a user's altitude with sufficient accuracy for many speech recognition applications. The impact of insufficient position accuracy is that, if a user is situated near the entrance of a building, it is not possible to accurately determine whether the user is, in fact, inside the building. The impact of insufficient altitude accuracy is that it is not possible to determine with certainty whether a user is within the building or standing on the roof. An acoustic model that is appropriate for an interior building lobby is unlikely to be totally appropriate for use in an outdoor setting. At the very least, selection of an inappropriate acoustic model based upon an inaccurate determination of a user's actual position will degrade the accuracy of many speech recognition algorithms.

[0010] Another shortcoming of the technique disclosed in U.S. Patent Application Publication No. US2003236099 is the need for creating an acoustic model. First of all, it is not feasible to create a unique acoustic model for every type of location. Secondly, even if an acoustic model is developed for a certain type of location, an actual location may not possess the acoustical characteristics embodied in the model. For example, due to variations in construction materials, furnishings, and room configurations, an acoustic model that attempts to represent a typical conference room, for example, may not adequately represent a specific conference room in a specific office building. Accordingly, what is needed is an improved technique for improving the accuracy of speech recognition algorithms.

[0011] EP1233407 recites a navigation system with speech recognition capacities for a moving vehicle. Whenever the distance between the vehicle and known active locations becomes greater then a given threshold, a pre-existing word list, corresponding to the active location, is acquired by the navigation system for subsequent use with the speech recognition engine.

[0012] US6230132 recites another navigation system with speech recognition functions. A lexicon is generated at periodic intervals with a limited number of words so as to facilitate the speech recognition functions. Furthermore, when a user needs to use the navigation system, he/she has to make a vocal request with the name of a location so that a lexicon can be generated.

[0013] EP0661688 recites a single device navigation apparatus, on board a car, with speed recognition functions. Pre-existing libraries are used and updated at regular intervals to facilitate the speech recognition engine.

SUMMARY OF THE INVENTION

[0014] The present invention is defined by the appended independent claims.

[0015] Other aspects and features of the present invention will become apparent from the following detailed description considered in conjunction with the accompanying drawings. It is to be understood, however, that the drawings are designed solely for purposes of illustration and not as a definition of the limits of the invention, for which reference should be made to the appended claims. It should be further understood that, unless otherwise indicated, the drawings are merely intended to conceptually illustrate the structures and procedures described herein.

BRIEF DESCRIPTION OF THE DRAWINGS

[0016] In the drawings:

[0017] FIG. 1 is a diagrammatic representation of a geographic area of interest.

[0018] FIG. 2 is a data structure diagram setting forth a portion of an illustrative element database.

[0019] FIG. 3 is a data structure diagram setting forth an illustrative subset of words selected from the element database of FIG. 2.

[0020] FIG. 4 is a generalized information flow diagram showing a set of fundamental operations for the present invention.

[0021] FIG. 5 is a detailed block diagram setting forth information flow for a speech recognition system constructed in accordance with a preferred embodiment of the present invention.

DETAILED DESCRIPTION OF THE DRAWINGS

[0022] The following terms used herein are defined as follows:

[0023] *Geographic Area of Interest:* Illustratively defined in terms of a zip code, group of zip codes, cell identification (Cell-ID) information, caller ID (Call-ID) information, the boundaries of a state, county, city, town, place, or other governmental or municipal subdivision, standard USGS (United States Geological Survey) grids, locations within a fixed range of latitude and longitude, or locations within a specified geometric boundary. The geographic area of interest may, but need not, be set automatically based upon the present location of a handset. The geographic area of interest may, but need not, be designated by a user based upon manual, verbal, or user-specified input. The geographic area of interest can be a coarse geographic area of interest or a refined geographic area of interest, with the latter being a portion of the former.

[0024] *E911:* As mandated by the FCC (Federal Communications Commission), a standard for a system by which wireless carriers are able to pinpoint the location of a mobile device in an emergency.

[0025] Element: Name of at least one of a street, a business, a merchant, an individual's name, a place, a point of interest, a transportation facility, an individual household, an activity, and a landmark.

[0026] *GPS*: Global Positioning System -- Enables geographic position determination at a land-based portable device equipped to receive signals from three or more orbiting GPS satellites.

[0027] *LBS:* Location-Based Service -- Any service

Amazon/VB Assets
Exhibit 1003
Page 285

that obtains information concerning the current geographical location of a mobile device user. Illustrative examples of location-based services include maps displayed in an automobile that track a motorists' position using GPS, or automatic localization services provided by dialing 911 on an E911-compliant mobile telephone network.

[0028] *Mobile Device:* Any portable device enabled for wireless communication, such as a wireless telephone, a personal digital assistant, or a laptop computer equipped with a wireless modem.

[0029] *Phoneme:* A phoneme represents a unique human uttered sound. Individual words, as pronounced by a human speaker, consist of a sequence of phonemes.

[0030] *Syntactic Model:* A set of rules and operations that specify allowed word sequences (phrases) of one or more words for subsequent use in a speech recognition procedure. Allowed word sequences are set forth in a list of static allowed phrases. In addition, the phrases can be optionally weighted by occurrence probabilities. Otherwise, word sequences can be in the form of stochastic or probabilistic lists, such as a Markov model of word sequences. Syntactic models are usually designed to be specific to a particular speech engine.

[0031] *Word:* A unit of language that native speakers can identify, and which communicates a meaning. Words are the building blocks from which phrases and sentences are constructed. Illustratively, for elements, "The Empire State Building" is the name of an element that contains four words. A street with the name "125th Street" includes the six words "one hundred and twenty fifth street". The street name "K Street" has the words "K" (which is a word for purposes of this invention) and "street".

[0032] The invention presumes that a very large element database is available for some sizeable, designated geographic area, such as a city, state, county, or country. One example of a known element database is TIGER (Topologically Integrated Geographic Encoding and Referencing system), available from the United States Census Bureau. TIGER compounds localized street names, rivers and railroads into a series of large regional databases.

[0033] The element database contains a set of words for elements at locations that are within the designated geographic area. In accordance with an aspect of the invention, enhanced speech recognition accuracy is provided by using geographic data to automatically select a subset of the words in the element database for use with a speech recognition procedure. The subset of words is selected by selecting only those words in the element database that are situated within a geographic area of interest. The geographic area of interest, either coarse or refined (depending on the circumstances, as explained below) is illustratively defined in terms of a zip code, group of zip codes, cell identification (Cell-ID) information, caller ID (Call-ID) information, the boundaries of a state, county, city, town, place, or other governmental or municipal subdivision, standard USGS (United States Geological Survey) grids, locations within a fixed range of latitude and longitude, or locations within a specified geometric boundary.

[0034] By selecting a subset of words, the total number of words used in the speech recognition procedure is substantially reduced. For example, a large element database containing street and business names used throughout all five boroughs of New York City can be reduced to a substantially shorter list of street and business names covering only the Grand Central Station area by specifying zip code 10017. Such a reduction in the total number of potential street names to be recognized, significantly improves speech recognition accuracy.

[0035] Pursuant to a further aspect of the invention, position information is obtained from an LBS based upon the current location of a mobile handset, illustratively using any of GPS coordinates, zip codes, and cellular base station triangulation. The obtained position information can be utilized to automatically define a coarse geographic area of interest. Let us say, as an example, that zip codes are used for this purpose. Thus, once the current location of the handset is determined, the zip code which includes that location automatically becomes the coarse geographic area of interest. Typically, the system can determine which localization method to use based on a knowledge of call configuration. For example, when a call is initiated, the Cell-ID can be used to determine the approximate location of the device. Alternatively, the application itself can choose which methods to employ, or the application can prompt the user to select which method to use. An illustrative prompt might include "*Do you want to use Cell-ID, or GPS to determine you location?*"

[0036] In accordance with a further aspect of the invention, an interface enables a user to refine the coarse geographic area of interest. For example, the user can select a specified geometrical shape within the coarse geographic area centered about the current location of the handset. This can be accomplished using key-based entries. Also, a graphical user interface can be used to draw a geometrical shape within the coarse geographic area. Additionally or alternatively, a Cell-ID can be specified within the geographic area zip code, thus defining a refined geographic area of interest for which the subset of words is to be selected. Combining automatic definition of a coarse geographic area of interest with additional, user-selectable geometric refinement is advantageous because it further reduces the total number of words that must be used in the speech recognition procedure. For example, a user-selectable rectangular region of 300 x 300 feet may contain only a few dozen elements that need to be recognized.

[0037] FIG. 1 is a diagrammatic representation of a geographic region of interest contained within a large area 100, such as a city. The current geographic location of a mobile device is determined to be point 401. In this illustrative example, the coarse geographic area of inter-

Amazon/VB Assets
Exhibit 1003
Page 286

est is defined by the zip code. Thus, the coarse geographic area of interest is defined by Zip code region 470 within which point 401 is located. A mobile device user has selected a specific radius 407 around point 401 for defining the refined geographic area of interest. Elements contained (in whole or in part) within this coarse geographic area include Street A 420, Street B 422, and a first building 406. Elements contained within this coarse geographic area of interest but which are not within this refined geographic area include Street C 424, second building 404, and third building 402.

[0038] An element database (FIG. 2, 500) includes a list of element names for at least one of streets, businesses, merchants, places, points of interest, transportation facilities, individual households, activities, or landmarks in large area 100 (FIG. 1). This database can be part of the LBS database. Each element name is associated with one or more corresponding parameters for identifying or associating the element with a refined geographic area of interest, and a parameter for identifying or associating the element with a coarse geographic area of interest. A set of phonemes corresponding to the element name can be computed in real time by recognition model compiler 213, using the outputted element name (501) in conjunction with a phonetizer. A phonetizer is a mechanism that provides phonemic transcriptions for each of a plurality of element names. Phonemic transcriptions comprise a phoneme, or a sequences of two or more phonemes, which indicate the manner in which an element name may be spoken. Illustratively, the element name "Madison" has a corresponding phonemic transcription "mäd è sun".

[0039] Element names are stored in an element name field 501. Each name in element name field 501 is associated with a corresponding parameter field (x,y) for the refined geographic area of interest 503, and a corresponding parameter field for the coarse geographic area of interest 505. These element names may or may not include elements that are within whatever refined geographic area of interest is selected by the user during operation of the invention (in this example, defined by radius 407 around point 401 of FIG. 1).

[0040] Refined geographic parameters are stored in the parameter field for the refined geographic area of interest 503 (FIG. 2). A refined geographic parameter sets forth a characteristic of an element in terms of how the refined geographic area of interest is defined. For example, if the refined geographic area is defined by the user as a shape (e.g., a circle) around the current location of the handset, the refined geographic parameters comprise (x,y) coordinates. In the illustrative example of FIG. 2, the refined geographic parameters are (x,y) coordinates determined with respect to an arbitrarily assigned reference point. When a first set and a second set of (x, y) coordinates are present in the parameter field for the refined geographic area of interest 503, this represents a line segment between the first set of (x,y) coordinates and the second set of (x,y) coordinates. Such a segment is used, for example, to define a street that extends from the first set of (x,y) coordinates to the second set of (x, y) coordinates.

[0041] Coarse geographic parameters are stored in the parameter field for the coarse geographic area of interest 505. A coarse geographic parameter sets forth a characteristic of an element in terms of how the coarse geographic area of interest is defined. For example, if the coarse geographic area of interest is defined in terms of one or more zip codes, the coarse geographic parameters in the parameter field for the coarse geographic area of interest 505 comprise one or more zip codes. For example, if Street A 420 (FIG. 1) crosses into two different zip code regions, then the parameter field for the coarse geographic area of interest 505 includes two zip codes, such as 10017 and 10036. However, other types of coarse geographic parameters could be employed to relate the element to a zip code system, such as standard USGS (United States Geological System) grids, regions within a fixed range of latitude and longitude, state boundaries, county boundaries, city boundaries, and/or boundaries of urbanized areas. For example, a lookup table could be used to determine the foregoing relationship.

[0042] Returning now to FIG. 1, and in accordance with one feature of the invention, only words corresponding to elements within the geographic area of interest are extracted from element database 500 (FIG. 2) to form word subset list 600 (FIG. 3). If the user defines the refined geographic area of interest as shown in FIG. 1, the limited set of words does not include Street C 424, second building 404, and third building 402. Use of this limited set of words enhances the accuracy of a subsequently performed speech recognition procedure because words pertaining to each and every element of large area 100 of FIG. 1 need not be considered.

[0043] FIG. 3 is a data structure diagram setting forth a portion of an illustrative word subset list 600. The names of one or more elements within the geographic area of interest are stored in an element name field 601. Optionally, word subset list 600 may include additional information, such as the probability that a particular element will be incorporated into a user's query based upon the distance of that element from the current position of the mobile device, so as to provide an enhanced level of performance.

[0044] FIG. 4 is a generalized information flow diagram describing a set of fundamental operations for the present invention. Geographic data acquisition mechanism 101 (the details of which are shown in FIG. 5) is equipped to collect geographic data corresponding to a mobile device's current geographic position and, based thereon, to determine and output a geographic area of interest. Geographic data includes, but is not limited to, any of: global positioning system (GPS) coordinates in latitude and longitude, and also altitude if needed, [x,y, z] coordinates, city name, place name, zip code area, geographic area, type of area (urban, suburban or coun-

Amazon/VB Assets
Exhibit 1003
Page 287

try), one or more cell identification (Cell-ID) parameters from a cellular network, or a caller identification (Call-ID) parameter from a terrestrial telephone network.

[0045] Word filtering mechanism 102 is equipped to receive a geographic area of interest from geographic data acquisition mechanism 101 and, based thereupon, select a subset of words from element database 500 (FIGs. 2 and 4). If a coarse geographic area of interest is being used, then the subset of words is selected by comparing coarse geographic parameters in the parameter field for the coarse geographic area of interest 505 corresponding to the elements in element name field 501 with coarse geographic parameters used to characterize the geographic area of interest. More specifically, if the coarse geographic area of interest is zip code 10017, then all the element names in element name field 501 will be selected which are associated with zip code 10017 in the parameter field for the coarse geographic area of interest 505. A similar operation is performed if a refined geographic area of interest is used. Of course, the parameter field for the refined geographic area of interest 503 will be used instead of the parameter field for the coarse geographic area of interest 505. The refined geographic parameters are combined with the coarse geographic parameters to enhance the word filtering mechanism 102 by using a two-step filtering process. A first step uses the parameter field for the coarse geographic area of interest 505, and a second step uses the parameter field for the refined geographic area of interest 503.

[0046] Word filtering mechanism 102 places the selected subset of words into a word subset list 600 (FIGs. 3 and 4). Speech recognition mechanism 103 (FIG. 4) is equipped to receive word subset list 600 (FIGS. 3 and 4) from word filtering mechanism 102 (FIG. 4) and, based thereupon, perform a speech recognition procedure. Illustratively, speech recognition mechanism 103 may be equipped to implement a speech recognition procedure by utilizing a (syntactic model) x (lexical model) x (acoustic model) paradigm in conjunction with Markov models. Alternatively or additionally, speech recognition mechanism 103 may utilize a neural network. More generally, the invention is applicable to any speech recognition mechanism that accepts and processes words or groups of words referred to as "grammar". Accordingly, word filtering mechanism 102 includes a grammar formatting mechanism that receives the word subset list 600 and formats the list into a grammar format usable by recognition model compiler 213.

[0047] Recognition model compiler 213 may or may not be required, depending upon the specific design characteristics of speech recognition mechanism 103. Recognition model compiler 213 parses (or analyzes) all of the entries in a formatted grammar, one after the other, and then, in one or more successive stages or "passes", constructs an output code readable by speech recognition mechanism 103. Since the speech recognition procedure is based upon a geographically refined subset of words, the overall accuracy of this procedure is enhanced relative to prior art designs. Speech recognition mechanism 103 can be implemented by a server or by a mobile handset.

[0048] A speech query from a mobile device user is received by an acoustic input mechanism 104 at the mobile device. Acoustic input mechanism 104 may be implemented using a microphone or other transducer equipped to transform acoustical vibrations into electronic signals. Acoustic input mechanism 104 feeds an electronic representation of user speech to speech recognition mechanism 103. Speech recognition mechanism 103 attempts to recognize one or more words in the electronic representation of user speech that are included in word subset list 600 retrieved from element database 500. If one or more words are recognized, speech recognition mechanism 103 sends an electronic message to an output mechanism 105 at the mobile device. The electronic message includes any recognized words in the speech query received from the mobile device user that were also included in word subset list 600. Output mechanism 105 converts the electronic message into audible and/or visual form for perception by the mobile device user. Illustratively, output mechanism 105 may be implemented using a loudspeaker coupled to audio amplifier circuitry and voice synthesizer circuitry for producing an audio message including one or more recognized words. Alternatively or additionally, output mechanism 105 may be implemented using a visual display device for displaying alphanumeric characters including one or more recognized words.

[0049] FIG. 5 is a detailed block diagram setting forth information flow for a speech recognition system constructed in accordance with an embodiment of the present invention. FIG. 5 shows details of geographic data acquisition mechanism 101. The depicted operations are performed either by a remote server 240 or by mobile device 250. Information flow commences upon receipt of a "start" signal at input mechanism 201 of mobile device 250. In practice, this "start" signal can be a request issued from a multimodal web page, an HTML page, or just a phone call. Alternatively, a user may place a call on a mobile handset to a designated telephone number to access a zip code form displayed on the mobile handset. The user fills in the zip code form using DTMF keypresses on the handset and validates the entered zip code by pressing DTMF digit "1" as a "start" signal. Basically, the "start" signal triggers a request for information.

[0050] Next, at mechanism 202, a decision is made as to whether the geographic area of interest will be determined locally (i.e., at mobile device 250), or at a remote location (i.e., at remote server 240). Depending upon the specifics of a given system application, this decision can be made by issuing a query to mobile device 250 in order to ascertain the configuration and capabilities of mobile device 250.

[0051] In cases where the geographic area of interest is determined locally, any one of two approaches may be employed. Pursuant to a first approach, mobile device

Amazon/VB Assets
Exhibit 1003
Page 288

250 may include an optional GPS receiver 207, whereupon acquired GPS data are forwarded to a geographic-area-of-interest defining mechanism 204 (server side, see FIG. 5). Mechanism 204 uses the GPS data corresponding to the current location of mobile device 250 to obtain coarse geographic parameters for characterizing the coarse geographic area of interest. According to a second approach, zip code data, for example, can be entered by the user at input mechanism 201 and forwarded to geographic-area-of-interest defining mechanism 204. Illustratively, zip code data are entered into input mechanism 201 by sequentially pressing keys on a DTMF (dual-tone, multifrequency) keypad. Mechanism 204 uses the entered zip code data as a coarse geographic parameter for characterizing the geographic area of interest.

[0052] In cases where the geographic area of interest is determined remotely, mechanism 202 forwards the request received from input mechanism 201 to a location-based service 203 at server 240. Alternatively or additionally, GPS data from GPS receiver 207 are conveyed to location-based service 203, if needed. In response to the received request, location-based service 203 retrieves, for example, zip code information or cell identification (Cell ID) data for mobile device 250. Cell-IDs data can be retrieved in a mobile communications network using the Automatic Number Identification (ANI) retrieved from the mobile device 250. ANI is a service that provides the receiver of a telephone call with the number of the calling phone. The method of providing this information is determined by the service provider. Cell ID data specifies the geographic location of the wireless base station currently in communication with mobile device 250 by identifying at least one of the telephone number and the Internet Protocol (IP) address assigned to mobile device 250. For example, mobile device 250 may send to location-based service 203 Cell ID data in the form of an automated number identification (ANI) parameter. Optionally or alternatively, location-based service 203 determines the geographic location of mobile device 250 by performing a triangulation procedure among a plurality of wireless base stations. The Cell ID data, Caller ID data, and/or triangulation data retrieved by location-based service 203 (ultimately through the telecommunication provider) are forwarded to geographic-area-of-interest defining mechanism 204.

[0053] Optionally, in cases where the geographic area of interest is determined remotely, LBS service 203 utilizes at least one of Caller ID data and Cell ID data to access a zip code database 243. Based upon the Caller ID data or the Cell ID data, a zip code is retrieved from zip code database 243 and then sent to geographic-area-of-interest defining mechanism 204. Although the illustrative embodiment of FIG. 5 shows zip code database 243 and location-based service 203 as separate elements, those skilled in the relevant art will easily recognize that zip code database 243 could, but need not, be incorporated into location-based service 203. Likewise, although element database 500 and location-based service 203 are shown as separate elements, those skilled in the relevant art will easily recognize that element database 500 could, but need not, be incorporated into location-based service 203.

[0054] Geographic-area-of-interest defining mechanism 204 processes incoming geographic data, illustratively comprising at least one of incoming GPS data, zip code data, Caller ID data, or Cell ID data, to generate one or more geographic parameters defining the geographic area of interest for mobile device 250. For example, as discussed above, if the incoming data are based upon zip code data, then the output of mechanism 204 is a zip code. This parameter, e.g., a zip code, is used to specify a geographic area of interest. The defined geographic area of interest may combine a coarse geographic area of interest with a refined geographic area of interest, as is illustrated in the following example. One or more zipcode(s) may be deduced from GPS coordinates, thus defining a coarse geographic area of interest. The GPS coordinates themselves, associated with an area of uncertainty, define a refined geographic area of interest. Such a defined geographic area of interest may be used to implement the two-step filtering process described above.

[0055] Optionally, as indicated by the dotted lines of FIG. 5, a user at mobile device 250 may wish to further refine the geographic area of interest defined by geographic-area-of-interest defining mechanism 204. If so, the geographic area of interest defined by geographic-area-of-interest defining mechanism 204 is referred to as a coarse geographic area of interest, and the area of interest as refined by the user is referred to as a refined geographic area of interest. The refined geographic area of interest comprises a subset of the coarse geographic area of interest. Input mechanism 201 accepts user input setting forth a geometric specification of a refined geographic area of interest. User input is illustratively received over connection line 275 in the form of keypad entries, graphical objects, sketches, textual descriptions, or vocal input. Examples of geometric specifications include squares, rectangles, octagons, triangles, circles, and freehand sketches. Geometry specification refinement mechanism 205 accepts user input from input mechanism 201, and refines the set of geographic parameters generated by geographic-area-of-interest defining mechanism 204.

[0056] The one or more geographic parameters generated by geographic-area-of-interest defining mechanism 204, as optionally refined by geometry specification refinement mechanism 205, are received by word filtering mechanism 102. Word filtering mechanism 102 is enabled to carry out its functions when two conditions are met. Firstly, mobile device 250 should be executing a speech-enabled application where the user has issued a request for information. This requirement is met upon generation of a request from input mechanism 201. Secondly, the geographic area of interest should be defined.

Amazon/VB Assets
Exhibit 1003
Page 289

This requirement is met once geographic-area-of-interest defining mechanism 204 generates the above-described geographic parameters.

[0057] When enabled, word filtering mechanism 102 uses the one or more geographic parameters received from mechanism 204 or 205 to determine whether or not each of a plurality of elements in element database 500 is situated within the geographic area of interest. The element names stored in element name field 501 (FIG. 2) may include elements that are within the geographic area of interest, as well as elements that do not lie within the geographic area of interest. Accordingly, word filtering mechanism 102 (FIG. 5) compares the one or more received parameters defining the geographic area of interest against one or more parameters stored in at least one of parameter field for the refined geographic area of interest 503 or parameter field for the coarse geographic area of interest 505 (FIG. 2), to determine whether or not each of a plurality of elements in element database 500 (FIGs. 2 and 5) is situated within the geographic area of interest. For example, Street A is characterized in the parameter field for the refined geographic area of interest 503 by a first and a second set of (x,y) coordinates. Accordingly, Street A is conceptualized as including a straight line segment connecting the first and second sets of (x,y) coordinates. If any portion of this straight line segment intersects the geographic area of interest, then Street A is included in word subset list 600 (FIGs. 3 and 5).

[0058] Word filtering mechanism 102 appends to word subset list 600 only those element names within the geographic area of interest. Optionally, if the user has made any refinements to the coarse geographic area of interest, only elements that are within the refined geographic area of interest are appended to word subset list 600.

[0059] The word subset list 600 of FIGs. 3 and 5 represents an illustrative output format for word filtering mechanism 102. Alternatively or additionally, the output of word filtering mechanism 102 may be conceptualized as a computer-readable stream of words representing one or more elements in a geographic area of interest. This stream, illustratively in text format, is compatible with at least one of a speech recognition engine or a voice browser embedding speech recognition mechanism 103 (FIG. 5). Optionally, word filtering mechanism 102 may reformat at least one of word subset list 600 or the computer-readable stream of words into a format such as an XML document, to provide a subset of words compatible with speech recognition mechanism 103.

[0060] Speech recognition mechanism 103 may be implemented at server 240 or at mobile device 250. In cases where speech recognition is performed at mobile device 250, the output from word filtering mechanism 102 (illustratively in the form of a stream of words corresponding to grammar subset list 600, FIG. 3) is sent from server 240 to mobile device 250. Word subset list 600 is then compiled by the recognition model compiler 213 at mobile device 250 to provide a speech recognition model for use by a speech recognition mechanism 103 responsive to the acoustic input mechanism 104. Speech recognition mechanism 103 then forwards recognized speech to the output mechanism 105 for generating an electronic output signal representing the recognized speech. Such an arrangement is advantageous because it reduces the burden on wireless communication resources. Transmissions between the server and the mobile device are reduced because only word subset list 600 needs to be communicated to the mobile device. The speech recognition steps can be performed at the mobile device with reference to word subset list 600, thereby enhancing the speed and efficiency of the speech recognition process.

[0061] If speech recognition is performed at server 240, then recognition model compiler 213 and speech recognition mechanism 103 are implemented at server 240 instead of at mobile device 250, wherein the output of speech recognition mechanism 103 is sent from server 240 to output mechanism 105 at mobile device 250.

[0062] The operational principles of the configurations shown in FIGs. 4 and 5 are best illustrated with reference to several illustrative scenarios. These scenarios clarify the location resolution process that will govern the filtering process performed by word filtering mechanism 102 in accordance with location data to be resolved. The scenarios consider use of any of several different types of inputs by which a location context will be determined.

[0063] Scenario One: Geographic area of interest to be determined by entered zip code and/or city. Zip code and/or city are entered into mobile device 250 by the user. Possible clients used by mobile device 250 to accept the entered zip code and/or city include a voice browser and/or a multimodal client. In the case of a phone call, the client is not the phone itself, but the voice browser (from the server point of view). Location data are provided in the form of a city and/or a zip code. The geographic area of interest corresponding to this location data is determined in the form of a city or zip code. In this case, location resolution is performed locally. The zip code and/or city are entered into mobile device 250 using DTMF keypresses and/or voice. An HTTP request is issued in response to the entered information after the information is received at input mechanism 201. In this case, word subset list 600 is generated by selecting a set of words and/or grammar pertaining to the entered city and/or zip code. The following steps are performed:

[0064] 1. User connects to geographic data acquisition mechanism 101 using mobile device 250. If mobile device 250 represents a multimodal client, an HTTP request is made by the client to a document server of word filtering mechanism 102 (FIGs. 4 and 5). If mobile device 250 represents a wireless telephone, an http request is made by the voice browser ("speaking" with the telephone) to the document server of word filtering mechanism 102.

[0065] 2. A first page is loaded into mobile device 250 or loaded onto the voice browser. The page is prompting the user to enter a zip code. The zip code by either be entered by DTMF (touch tones) or by voice. The grammar used to recognize the zip code is static.

Amazon/VB Assets
Exhibit 1003
Page 290

[0066] 3. The user enters the zip code into input mechanism 201. A second HTTP request is made by the client, with the zip code value as a parameter. As location-based service 203 is not employed in this scenario to resolve the location, location resolution is "local". The geographic area of interest is evaluated at geographic area of interest defining mechanism 204, and no further refinement of the geographical area of interest takes place.

[0067] 4. A second page is loaded into mobile device 250 (the "reduced" page). The document server of word filtering mechanism 102 computes reduced grammars associated with the page in the form of word subset list 600.

[0068] 5. The user can vocally interact with the page over mobile device 250, as the page is optimized for its location

[0069] Note: The geometry specification refinement mechanism 205 is not used in this scenario. If this mechanism were to be used, a "remote" location resolution must be performed so as to enable sending a map (for example) back to the multimodal client of mobile device 250.

[0070] Scenario Two: Geographic area of interest to be determined automatically using Caller ID or Cell ID.

[0071] Possible clients used by mobile device 250 for Caller Id or Cell ID include a voice browser and/or a multimodal client. Location data are provided in the form of a caller ID or cell ID parameter. The geographic area of interest is determined as a zip code or as an area ID that identifies a geographic area. Location resolution is performed remotely, not locally. Input mechanism 201 is not employed for purposes of location resolution. Mobile device 250 is programmed to issue an HTTP request while a first page is retrieved from a document server of word filtering mechanism 102. In this scenario, words and/or grammar from element database 500 are filtered to prepare word subset list 600 based upon ANI or Cell-ID information. Input mechanism 201 is completely transparent, in the sense that the request that trims the word filtering process appears at the first stage of the interaction, and there is no need of a data collecting stage, as in the previous scenario. The HTTP request can be issued by a voice browser (that provides ANI info) or a multimodal client at mobile device 250.

[0072] The following operational sequence is performed:

[0073] 1. User connects to geographic data acquisition mechanism 101 using mobile device 250. If the user is using a multimodal client at mobile device 250, an HTTP request is made by the client to the document server of word filtering mechanism 102. If mobile device 250 is a wireless telephone, an http request is made by the voice browser ("speaking" with the telephone) to the document server of word filtering mechanism 102. The ANI and/or the Cell ID are passed to the document server as part of the HTTP request. Location-based service 203 determines the geographic area of interest, in this case following the remote path indicated in FIG. 5 which commences at mechanism 202. Location-based service also determines geographic position of the user, an area of uncertainty with respect to the user's position, and one or more associated zip codes.

[0074] 2. A page is downloaded to mobile device 250 or loaded onto the voice browser. The geographic context is used by word filtering mechanism 102 to compute a word subset list 600 associated with the page in accordance with the previously determined geographic area of interest.

[0075] 3. The user can vocally interact with the page using mobile device 250. The page is optimized for the user's geographic area of interest.

[0076] Scenario Three: Geographic area of interest to be determined automatically using GPS data. Possible clients used by mobile device 250 for GPS include a multimodal client. Location data are provided in the form of GPS coordinates. The geographic area of interest is determined as a zip code, as (x,y) or (x,y,z) coordinates, optionally incorporating an area of uncertainty. Location resolution is performed locally, not remotely. Input mechanism 201 is not employed to determine geographic area of interest. An HTTP request is issued by mobile device 250 during the first page fetch. In this case, word filtering mechanism 102 is preparing a word subset list 600 by filtering element database 500 using GPS data. Input mechanism 201 is transparent, in the sense that the HTTP request that results in word filtering appears at the first stage of the interaction. Unlike the first scenario, there is no need for a data collecting stage. The following steps are performed:

[0077] 1. User connects to geographic data acquisition mechanism 101 using mobile device 250. If mobile device 250 utilizes a multimodal client, an HTTP request is made by the client to a document server of word filtering mechanism 102. The GPS information is sent as part of the http request. Location-based service 203 determines the geographic area of interest by adding coarse localization to the GPS information, and the coarse localization is sent to geographic area of interest defining mechanism 204.

[0078] 2. A page is downloaded to mobile device 250: The geographic area of interest in the form of the coarse and/or refined localization received at geographic area of interest defining mechanism 204 is used by word filtering mechanism 102 to compute a word subset list 600 and to place the word subset list into a page that is sent to mobile device 250, using the two-step filtering process described above.

[0079] 3. The user can vocally interact with the page, optimized for the geographic area of interest. Note:

[0080] 1. In cases where geometry specification refinement mechanism 205 is employed, an intermediate page is downloaded to mobile device 250, asking to the user to precise its area of interest (before step 2):

[0081] 1.a - An intermediate (i.e., "localization") page is downloaded to mobile device 250.

[0082] 1.b - The user inputs an area of interest into

Amazon/VB Assets
Exhibit 1003
Page 291

mobile device 250.

[0083] 1.c - A refined geographic area of interest is sent by mobile device 250 and incorporated into the geographic area of interest defined by geographic area of interest defining mechanism 204.

[0084] In contrast to prior art techniques, the present invention does not employ acoustic models to improve speech recognition accuracy. Instead, the invention reduces the total number of words that need to be considered by a speech recognition algorithm. The number of words is reduced by excluding words that do not relate to a geographic area of interest. In this manner, the reduced processing burden makes it possible to handle speech recognition at the handset. Moreover, the accuracy of the speech recognition algorithm is improved.

[0085] To summarize, the advantages of the invention include more accurate speech recognition due to utilization of a database that has been reduced in size (i.e., word subset list 600). The burden on communication resources is decreased by sending the reduced database to the handset, whereupon the handset performs speech recognition using only the reduced size database instead of a much larger database. The decreased burden on communication resources results in faster, more efficient speech recognition.

[0086] Thus, while there have been shown and described novel features of the invention as applied to embodiments thereof, it will be understood that various omissions and substitutions and changes in the form and details of the devices illustrated, and in their operation, may be made by those skilled in the art without departing from the scope of the invention, which is solely defined by the appended claims. For example, although an embodiment of the invention was described in terms of a user refining an automatically determined geographic area of interest, it is also possible for a user to augment an automatically determined coarse geographic area of interest.

[0087] In view of the foregoing, it is expressly intended that all combinations of those elements and/or method steps which perform substantially the same function in substantially the same way to achieve the same results may be incorporated in any other disclosed or described or suggested form or embodiment as a general matter of design choice in as much as they fall within the scope of the claims appended hereto.

Claims

1. A method for use with a mobile handset equipped to communicate with a remote server over a wireless communications network, the remote server equipped to access an element database (500) including words that describe elements at each of a plurality of geographically defined locations, the method comprising the steps of:

   determining (202) geographic data for the mobile handset following a request; and
   using (102) the determined geographic data to automatically select by the remote server a subset (600) of words for use with a speech recognition procedure;

   wherein the subset of words is selected from the element database, thereby reducing the number of words used in the speech recognition procedure.

2. The method of claim 1 wherein the determined geographic data includes at least one of global positioning system (GPS) position data, cell identity (Cell-ID) data, caller identification (Caller ID) data, place name data, or zip code data.

3. The method of claim 1 wherein the geographic data are derived from a signal received from the mobile handset, said signal being related to a current position of said mobile handset or to a manual input entered into said mobile handset.

4. The method of any preceding claim wherein the elements include at least one of street names, businesses, merchants, points of interest, transportation facilities, individual households, activities, or landmarks.

5. The method of any preceding claim wherein the mobile handset receives the subset of words from the remote server and executes the speech recognition procedure based upon the received subset of words, thereby reducing use of the communications network and the remote server.

6. A speech recognition system comprising:

   a mobile handset equipped to communicate with a remote server over a wireless network,
   a data acquisition mechanism (101) for determining geographic data for the mobile handset following a request input; and
   a selection mechanism (102) for using the determined geographic data to automatically select a subset of words (106) for use with a speech recognition procedure;

   wherein the subset of words is selected by the remote server from an element database (500) said remote server has access to, said element database including words that describe elements at each of a plurality of geographically defined locations, thereby reducing the number of words used in the speech recognition procedure.

7. The speech recognition system of claim 6 wherein the geographic data are derived from a signal re-

Amazon/VB Assets
Exhibit 1003
Page 292

ceived from the mobile handset, said signal being related to a current position of said mobile handset or to a manual input entered into said mobile handset.

8. The speech recognition system of claim 6 wherein the determined geographic data includes at least one of global positioning system (GPS) position data, cell identity (Cell-ID) data, caller identification (Caller ID) data, place name data, or zip code data.

9. The speech recognition system of any of claims 6 to 8 wherein the elements include at least one of street names, businesses, merchants, points of interest, transportation facilities, individual households, activities, or landmarks.

10. A speech recognition system for associating received speech with words retrieved from an element database, the system comprising:

means for determining a geographic area of interest following a request wherein speech is to be received, said means further comprising:

(i) a portable location determining mechanism associated with a mobile handset for generating an indication signal indicative of current geographic location, and
(ii) a server, in communication with the portable location determining mechanism, programmed to determine a geographic area of interest from the indication signal;

means for selecting a subset of words from the element database based upon the geographic area of interest; said means comprising the server programmed to extract from the element database only words that are associated with the geographic area of interest; and,
means for associating the received speech with words selected from the subset of words, said means being implemented on the mobile handset.

11. The speech recognition system of claim 10 further comprising:

means for accepting a signal from at least one graphical user interface for selecting a portion of the determined geographic area of interest to thereby specify a further limited geographic area of interest; and
means for further reducing the subset of words according to the further limited geographic area of interest to generate a further subset of words, such that the means for associating received speech with words only selects words from the further subset of words.

12. The speech recognition system of claim 11 wherein the means for accepting a signal includes:

an electronic display for displaying a map of the determined geographic area of interest; and
a processing mechanism for combining the accepted signal with the map of the determined geographic area of interest so as to cause a display of the further limited geographic area of interest on the electronic display.

13. A mobile handset (250) comprising speech recognition means for association of received speech with words received from a remote database, the mobile handset comprising:

means for receiving a request;
means for providing a signal representing a geographic area of interest following said request;
means for transmitting said signal to a remote server and for receiving from the remote server a subset of words from the remote database based upon the geographic area of interest; and
means for associating the received speech with words selected from the subset of words.

14. The mobile handset of claim 13 wherein the means for providing a signal comprises means for determining a geographic area of interest wherein speech is to be received.

15. The mobile handset of claim 13 or 14 further comprising:

means (201) for accepting at least one input used to determine a refined geographic area of interest in the geographic area of interest;
means (205) for reducing the subset of words according to the refined geographic area of interest;

and wherein:

the means for associating the received speech is arranged for association of speech only with the reduced subset of words.

Patentansprüche

1. Verfahren zur Verwendung mit einem mobilen Handapparat, der dafür ausgestattet ist, über ein drahtloses Kommunikationsnetz mit einem abgesetzten Server zu kommunizieren, wobei der abgesetzte Server dafür ausgestattet ist, auf eine Elementdatenbank (500) zuzugreifen, die Wörter enthält, die

Amazon/VB Assets
Exhibit 1003
Page 293

Elemente an jedem von mehreren geographisch definierten Orten beschreiben, wobei das Verfahren die folgenden Schritte umfaßt:

Bestimmen (202) geographischer Daten für den mobilen Handapparat nach einer Anforderung und Verwenden (102) der bestimmten geographischen Daten zum automatischen Auswählen einer Teilmenge (600) von Wörtern zur Verwendung mit einer Spracherkennungsprozedur durch den abgesetzten Server;

wobei die Teilmenge von Wörtern aus der Elementdatenbank ausgewählt wird, wodurch die Anzahl der in der Spracherkennungsprozedur verwendeten Wörter reduziert wird.

2. Verfahren nach Anspruch 1, wobei die bestimmten geographischen Daten Positionsdaten des globalen Navigationssystems (GPS) und/oder Daten der Zellenidentität (Cell-ID) und/oder Daten der Anruferidentifikation (Caller ID) und/oder Städtenamendaten und/oder Postleitzahldaten umfassen.

3. Verfahren nach Anspruch 1, wobei die geographischen Daten aus einem aus dem mobilen Handapparat empfangenen Signal abgeleitet werden, wobei das Signal eine aktuelle Position des mobilen Handapparats oder eine in den mobilen Handapparat eingegebene manuelle Eingabe betrifft.

4. Verfahren nach einem der vorhergehenden Ansprüche, wobei die Elemente Straßennamen und/oder Unternehmen und/oder Händler und/oder interessierende Punkte und/oder Transporteinrichtungen und/oder individuelle Haushalte und/oder Aktivitäten und/oder Sehenswürdigkeiten umfassen.

5. Verfahren nach einem der vorhergehenden Ansprüche, wobei der mobile Handapparat die Teilmenge von Wörtern von dem abgesetzten Server empfängt und auf der Basis der empfangenen Teilmenge von Wörtern die Spracherkennungsprozedur ausführt und **dadurch** die Benutzung des Kommunikationsnetzes und des abgesetzten Servers reduziert.

6. Spracherkennungssystem, umfassend:

einen mobilen Handapparat, der dafür ausgestattet ist, über ein drahtloses Netz mit einem abgesetzten Server zu kommunizieren,
einen Datenbeschaffungsmechanismus (101) zum Bestimmen von geographischen Daten für den mobilen Handapparat nach einer Anforderungseingabe und
einen Auswahlmechanismus (102) zum Verwenden der bestimmten geographischen Daten zum automatischen Auswählen einer Teilmenge von Wörtern (106) zur Verwendung mit einer Spracherkennungsprozedur;

wobei die Teilmenge von Wörtern durch den abgesetzten Server aus einer Elementdatenbank (500) ausgewählt wird, zu der der abgesetzte Server Zugang hat, wobei die Elementdatenbank Wörter enthält, die Elemente an jedem von mehreren geographisch definierten Orten beschreiben, wodurch die Anzahl der bei der Spracherkennungsprozedur verwendeten Wörter reduziert wird.

7. Spracherkennungssystem nach Anspruch 6, wobei die geographischen Daten aus einem aus dem mobilen Handapparat empfangenen Signal abgeleitet werden, wobei das Signal eine aktuelle Position des mobilen Handapparats oder eine in den mobilen Handapparat eingegebene manuelle Eingabe betrifft.

8. Spracherkennungssystem nach Anspruch 6, wobei die bestimmten geographischen Daten Positionsdaten des globalen Navigationssystems (GPS) und/oder Daten der Zellenidentität (Cell-ID) und/oder Daten der Anruferidentifikation (Caller ID) und/oder Städtenamendaten und/oder Postleitzahldaten umfassen.

9. Spracherkennungssystem nach einem der Ansprüche 6 bis 8, wobei die Elemente Straßennamen und/oder Unternehmen und/oder Händler und/oder interessierende Punkte und/oder Transporteinrichtungen und/oder individuelle Haushalte und/oder Aktivitäten und/oder Sehenswürdigkeiten umfassen.

10. Spracherkennungssystem zum Assoziieren von empfangener Sprache mit aus einer Elementdatenbank abgerufenen Wörtern, wobei das System folgendes umfaßt:

Mittel zum Bestimmen eines interessierenden geographischen Gebiets nach einer Aufforderung, wobei Sprache zu empfangen ist, wobei die Mittel ferner folgendes umfassen:

i) einen mit einem mobilen Handapparat assoziierten tragbaren Ortsbestimmungsmechanismus zum Erzeugen eines Anzeigesignals, das einen aktuellen geographischen Ort anzeigt und
ii) einen mit dem tragbaren Ortsbestimmungsmechanismus kommunizierenden Server, der dafür programmiert ist, aus dem Anzeigesignal ein interessierendes geographisches Gebiet zu bestimmen;

Mittel zum Auswählen einer Teilmenge von Wörtern aus der Elementdatenbank auf der Ba-

Amazon/VB Assets
Exhibit 1003
Page 294

sis des interessierenden geographischen Gebiets; wobei die Mittel den Server umfassen, der dafür programmiert ist, aus der Elementdatenbank nur Wörter zu extrahieren, die mit dem interessierenden geographischen Gebiet assoziiert sind und
Mittel zum Assoziieren der empfangenen Sprache mit aus der Teilmenge von Wörtern ausgewählten Wörtern, wobei die Mittel auf dem mobilen Handapparat implementiert sind.

11. Spracherkennungssystem nach Anspruch 10, ferner umfassend:

Mittel zum Annehmen eines Signals von mindestens einer graphischen Benutzeroberfläche zum Auswählen eines Teils des bestimmten interessierenden geographischen Gebiets, um **dadurch** ein weiter eingeschränktes interessierendes geographisches Gebiet zu spezifizieren und
Mittel zum weiteren Reduzieren der Teilmenge von Wörtern gemäß dem weiter eingeschränkten interessierenden geographischen Gebiet, um eine weitere Teilmenge von Wörtern zu erzeugen, dergestalt, daß die Mittel zum Assoziieren empfangener Sprache mit Wörtern nur Wörter aus der weiteren Teilmenge von Wörtern auswählen.

12. Spracherkennungssystem nach Anspruch 11, wobei die Mittel zum Annehmen eines Signals folgendes umfassen:

ein elektronisches Display zum Anzeigen einer Karte des bestimmten interessierenden geographischen Gebiets und
einen Verarbeitungsmechanismus zum Kombinieren des angenommenen Signals mit der Karte des bestimmten interessierenden geographischen Gebiets, um so eine Anzeige des weiter eingeschränkten interessierenden geographischen Gebiets auf dem elektronischen Display zu bewirken.

13. Mobiler Handapparat (150) mit Spracherkennungsmitteln zum Assoziieren empfangener Sprache mit aus einer abgesetzten Datenbank empfangenen Wörtern, wobei der mobile Handapparat folgendes umfaßt:

Mittel zum Empfangen einer Anforderung;
Mittel zum Bereitstellen eines Signals, das ein interessierendes geographisches Gebiet repräsentiert, nach der Anforderung;
Mittel zum Senden des Signals zu einem abgesetzten Server und zum Empfangen einer Teilmenge von Wörtern aus der abgesetzten Datenbank von dem abgesetzten Server auf der Basis des interessierenden geographischen Gebiets und
Mittel zum Assoziieren der empfangenen Sprache mit aus der Teilmenge von Wörtern ausgewählten Wörtern.

14. Mobiler Handapparat nach Anspruch 13, wobei die Mittel zum Bereitstellen eines Signals Mittel zum Bestimmen eines interessierenden geographischen Gebiets umfaßt, worin Sprache empfangen werden soll.

15. Mobiler Handapparat nach Anspruch 13 oder 14, ferner umfassen:

Mittel (210) zum Annehmen mindestens einer Eingabe, mit der ein verfeinertes interessierendes geographisches Gebiet in dem interessierenden geographischen Gebiet bestimmt wird;
Mittel (205) zum Reduzieren der Teilmenge von Wörtern gemäß dem verfeinerten interessierenden geographischen Gebiet und wobei

die Mittel zum Assoziieren der empfangenen Sprache dafür ausgelegt sind, Sprache nur mit der reduzierten Teilmenge von Wörtern zu assoziieren.

Revendications

1. Procédé pour une utilisation avec un combiné mobile équipé pour communiquer avec un serveur distant sur un réseau de communication sans fil, le serveur distant étant équipé pour accéder à une base de données d'éléments (500) comprenant des mots qui décrivent des éléments pour chacun d'une pluralité d'emplacements définis géographiquement, le procédé comprenant les étapes de :

détermination (202) de données géographiques pour le combiné mobile suite à une demande ; et utilisation (102) des données géographiques déterminées pour sélectionner automatiquement par le serveur distant un sous-ensemble (600) de mots pour une utilisation avec une procédure de reconnaissance vocale ;

dans lequel le sous-ensemble de mots est sélectionné parmi la base de données d'éléments, réduisant de ce fait le nombre de mots utilisés dans la procédure de reconnaissance vocale.

2. Procédé selon la revendication 1 dans lequel les données géographiques déterminées comprennent au moins une parmi des données de position du système de positionnement mondial (GPS), des données d'identité de cellule (identifiant de cellule), des

Amazon/VB Assets
Exhibit 1003
Page 295

données d'identification d'appelant (identifiant d'appelant), des données de nom de lieu ou des données de code postal.

3. Procédé selon la revendication 1 dans lequel les données géographiques sont déduites d'un signal reçu du combiné mobile, ledit signal étant lié à une position actuelle dudit combiné mobile ou à une entrée manuelle entrée dans ledit combiné mobile.

4. Procédé selon une revendication précédente quelconque dans lequel les éléments comprennent au moins un parmi des noms de rues, des entreprises, des commerçants, des centres d'intérêt, des installations de transport, des maisons individuelles, des activités ou des points de repère.

5. Procédé selon une revendication précédente quelconque dans lequel le combiné mobile reçoit le sous-ensemble de mots du serveur distant et exécute la procédure de reconnaissance vocale sur la base du sous-ensemble reçu de mots, réduisant de ce fait l'utilisation du réseau de communication et du serveur distant.

6. Système de reconnaissance vocale comprenant :

un combiné mobile équipé pour communiquer avec un serveur distant sur un réseau sans fil ;
un mécanisme d'acquisition de données (101) pour déterminer des données géographiques pour le combiné mobile suite à une entrée de demande ; et
un mécanisme de sélection (102) pour utiliser les données géographiques déterminées afin de sélectionner automatiquement un sous-ensemble de mots (106) pour une utilisation avec une procédure de reconnaissance vocale ;

dans lequel le sous-ensemble de mots est sélectionné par le serveur distant à partir d'une base de données d'éléments (500) à laquelle ledit serveur distant a accès, ladite base de données d'éléments comprenant des mots qui décrivent des éléments pour chacun d'une pluralité d'emplacements définis géographiquement, réduisant de ce fait le nombre de mots utilisés dans la procédure de reconnaissance vocale.

7. Système de reconnaissance vocale selon la revendication 6 dans lequel les données géographiques sont déduites d'un signal reçu du combiné mobile, ledit signal étant lié à une position actuelle dudit combiné mobile ou à une entrée manuelle entrée dans ledit combiné mobile.

8. Système de reconnaissance vocale selon la revendication 6 dans lequel les données géographiques déterminées comprennent au moins une parmi des données de position du système de positionnement mondial (GPS), des données d'identité de cellule (identifiant de cellule), des données d'identification d'appelant (identifiant d'appelant), des données de nom de lieu ou des données de code postal.

9. Système de reconnaissance vocale selon l'une quelconque des revendications 6 à 8 dans lequel les éléments comprennent au moins un parmi des noms de rues, des entreprises, des commerçants, des centres d'intérêt, des installations de transport, des maisons individuelles, des activités ou des points de repère.

10. Système de reconnaissance vocale pour associer une voix reçue à des mots récupérés d'une base de données d'éléments, le système comprenant :

des moyens pour déterminer une zone géographique d'intérêt suite à une demande dans laquelle une voix doit être reçue, lesdits moyens comprenant en outre :

(i) un mécanisme de détermination d'emplacement de portable associé à un combiné mobile pour générer un signal d'indication révélateur d'un emplacement géographique actuel, et
(ii) un serveur, en communication avec le mécanisme de détermination d'emplacement de portable, programmé pour déterminer une zone géographique d'intérêt à partir du signal d'indication ;

des moyens pour sélectionner un sous-ensemble de mots parmi la base de données d'éléments sur la base de la zone géographique d'intérêt ; lesdits moyens comprenant le serveur programmé pour extraire de la base de données d'éléments seulement des mots qui sont associés à la zone géographique d'intérêt ; et
des moyens pour associer la voix reçue à des mots sélectionnés parmi le sous-ensemble de mots, lesdits moyens étant mis en oeuvre sur le combiné mobile.

11. Système de reconnaissance vocale selon la revendication 10 comprenant en outre :

des moyens pour accepter un signal d'au moins une interface utilisateur graphique pour sélectionner une partie de la zone géographique d'intérêt déterminée afin de spécifier de ce fait une zone géographique d'intérêt limitée supplémentaire ; et
des moyens pour réduire en outre le sous-ensemble de mots selon la zone géographique

Amazon/VB Assets
Exhibit 1003
Page 296

d'intérêt limitée supplémentaire afin de générer un sous-ensemble supplémentaire de mots, de telle sorte que les moyens pour associer une voix reçue à des mots sélectionnent seulement des mots parmi le sous-ensemble supplémentaire de mots.

12. Système de reconnaissance vocale selon la revendication 11 dans lequel les moyens pour accepter un signal comprennent :

un affichage électronique pour afficher une carte de la zone géographique d'intérêt déterminée ; et
un mécanisme de traitement pour combiner le signal accepté avec la carte de la zone géographique d'intérêt déterminée de façon à provoquer un affichage de la zone géographique d'intérêt limitée supplémentaire sur l'affichage électronique.

13. Combiné mobile (250) comprenant des moyens de reconnaissance vocale pour une association d'une voix reçue à des mots reçus d'une base de données distante, le combiné mobile comprenant :

des moyens pour recevoir une demande ;
des moyens pour fournir un signal représentant une zone géographique d'intérêt suite à ladite demande ;
des moyens pour transmettre ledit signal à un serveur distant et recevoir du serveur distant un sous-ensemble de mots provenant de la base de données distante sur la base de la zone géographique d'intérêt ; et
des moyens pour associer la voix reçue à des mots sélectionnés parmi le sous-ensemble de mots.

14. Combiné mobile selon la revendication 13 dans lequel les moyens pour fournir un signal comprennent des moyens pour déterminer une zone géographique d'intérêt dans laquelle une voix doit être reçue.

15. Combiné mobile selon la revendication 13 ou 14 comprenant en outre:

des moyens (201) pour accepter au moins une entrée utilisée pour déterminer une zone géographique affinée d'intérêt dans la zone géographique d'intérêt ;
des moyens (205) pour réduire le sous-ensemble de mots selon la zone géographique affinée d'intérêt;

et dans lequel :

les moyens pour associer la voix reçue sont agencés pour une association d'une voix seulement avec le sous-ensemble réduit de mots.

Amazon/VB Assets
Exhibit 1003
Page 297

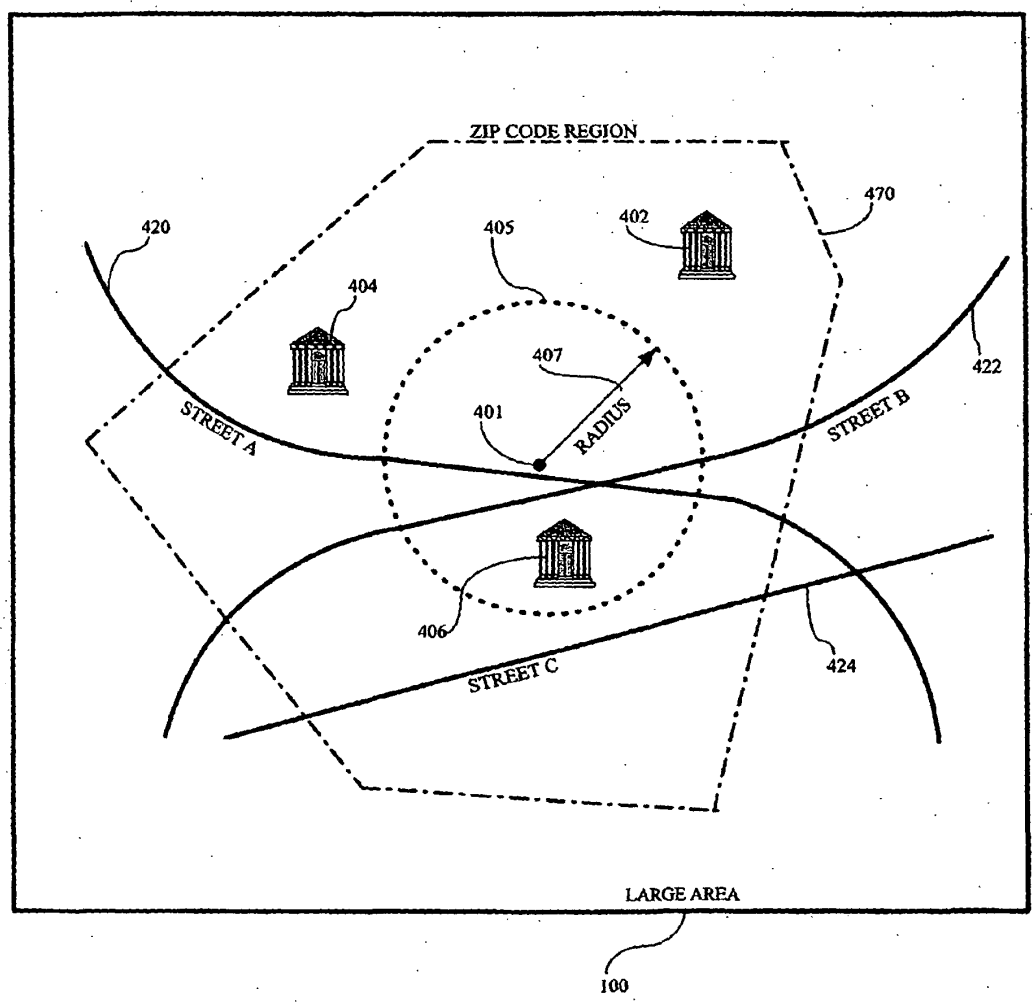

ZIP CODE REGION
470
420
405
402
404
407
422
401
RADIUS
STREET A
STREET B
406
424
STREET C
LARGE AREA
100

FIG. 1

500

ELEMENT DATABASE

ELEMENT NAME FIELD 501	PARAMETER FIELD FOR THE REFINED GEOGRAPHIC AREA OF INTEREST 503	PARAMETER FIELD FOR THE COARSE GEOGRAPHIC AREA OF INTEREST 505
FIRST BUILDING	{[603001, 2428833]}	10017
SECOND BUILDING	{[602298, 2428830]}	10017
THIRD BUILDING	{[603010, 2428836]}	10017
STREET A	{[603012, 2428836],[603013, 24428836]}	10017, 10036
STREET B	{[603014, 2428836],[603015, 24428836]}	10017, 10055
STREET C	{[679010, 2428836],[680000, 24428836]}	10017, 12533
...	...	...
DEVON FARMS ROAD	{[673010, 2428836],[675000, 24428836]}	12582

FIG. 2

600

WORD SUBSET LIST

ELEMENT NAME FIELD 601
FIRST BUILDING
STREET A
STREET B

FIG. 3

Amazon/VB Assets
Exhibit 1003
Page 299

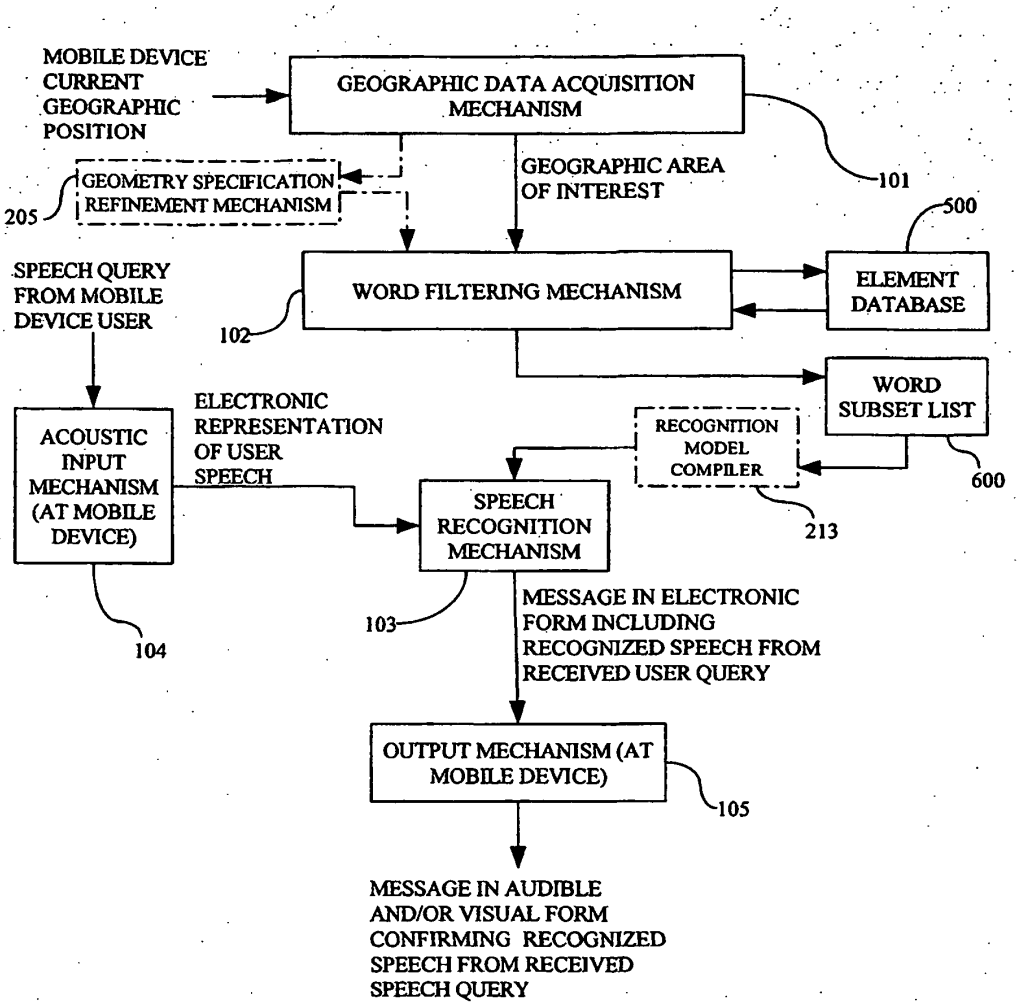


FIG. 4

Amazon/VB Assets
Exhibit 1003
Page 300

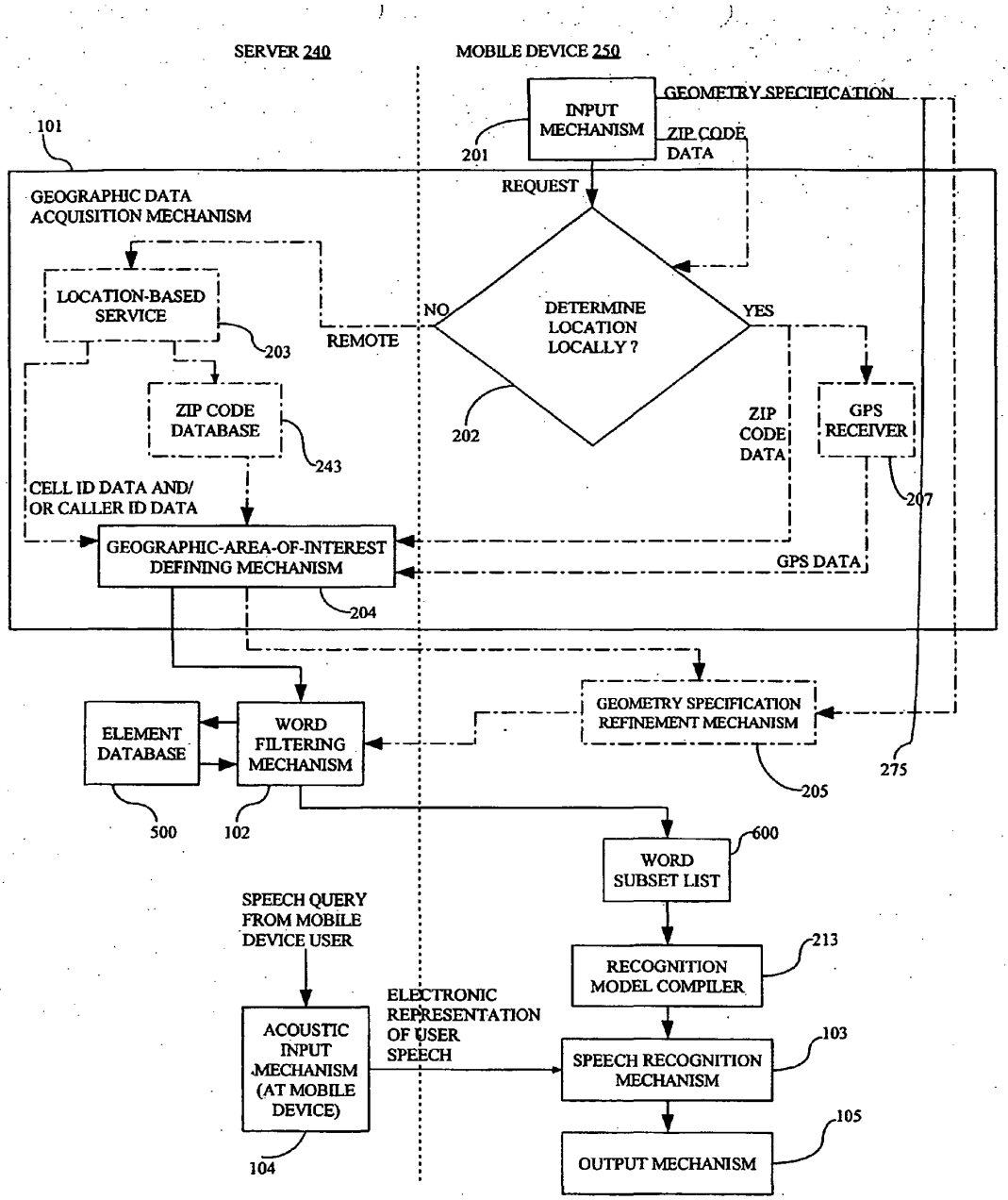


FIG. 5

Amazon/VB Assets
Exhibit 1003
Page 301

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 2003236099 A **[0007] [0009] [0010]**
- EP 1233407 A **[0011]**
- US 6230132 B **[0012]**
- EP 0661688 A **[0013]**

Amazon/VB Assets
Exhibit 1003
Page 302

Electronic Acknowledgement Receipt

EFS ID:	14833684
Application Number:	12943699
International Application Number:	
Confirmation Number:	4185
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Customer Number:	909
Filer:	Ngai Bing Zhang/Miriam Nelsen Awan
Filer Authorized By:	Ngai Bing Zhang
Attorney Docket Number:	090611-0384422
Receipt Date:	30-JAN-2013
Filing Date:	10-NOV-2010
Time Stamp:	16:25:36
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		090611-0384422_Response.pdf	577797 4f9e5dcbebdd97b07fe69c0fe3f7c124bc570362	yes	14

	Multipart Description/PDF files in .zip description				
	Document Description		Start	End	
	Amendment/Req. Reconsideration-After Non-Final Reject		1	1	
	Claims		2	8	
	Applicant Arguments/Remarks Made in an Amendment		9	14	
Warnings:					
Information:					
2	Extension of Time	090611-0384422_Extension.pdf	495240 7ae042d01267c039fb3fde986ae9181dce29d495	no	1
Warnings:					
Information:					
3	Transmittal Letter	0384422_SIDS.pdf	3998724 d8e170bb990b2606fc9fbf0535c62ea5d8bdce45	no	7
Warnings:					
Information:					
4	Information Disclosure Statement (IDS) Form (SB08)	0384422_IDSPTO1449.pdf	1301425 96985df0451c3a73292b88972c2aaa09105b4483	no	11
Warnings:					
Information:					
This is not an USPTO supplied IDS fillable form					
5	Foreign Reference	EP1320043A2.pdf	1416212 2ec7d64c9ab7063100d70c545035fd72a13ded77	no	28
Warnings:					
Information:					
6	Foreign Reference	EP1646037.pdf	262902 56537431a1754beec4b827cf69e772a2da14c1a8	no	20
Warnings:					
Information:					
7	Non Patent Literature	Statement_in_Accordance_with_EPO_Notice.pdf	86084 a41f2c134966e666f56e1c9c6e421bc9a3281ef0	no	1
Warnings:					
Information:					
		Total Files Size (in bytes):	8138384		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

IN THE UNITED STATES PATENT & TRADEMARK OFFICE

IN RE PATENT APPLICATION OF:	Mike KENNEWICK *et al.*
SERIAL NO.:	12/943,699
FILING DATE:	November 10, 2010
ATTORNEY DOCKET NO.:	090611-0384422
CONFIRMATION NO.:	4185
EXAMINER:	Tim B. Hale
ART UNIT:	3685
FOR:	System and Method for Providing A Natural Language Content Dedication Service

RESPONSE TO NON-FINAL OFFICE ACTION mailed 01 OCTOBER 2012

MAIL STOP AMENDMENT
Commissioner for Patents
P.O. Box 1450
Alexandria, VA. 22313-1450

Dear Sir:

In response to the Non-Final Office Action mailed **October 1, 2012**, please amend the above- identified application as follows:

Amendments to the Claims begin on page **2** of this paper.

Remarks begin on page **9** of this paper.

It is not believed that extensions of time or fees for net addition of claims are required beyond those that may otherwise be provided for in documents accompanying this paper. However, if additional extensions of time are necessary to prevent abandonment of this application, then such extensions of time are hereby petitioned for under 37 C.F.R. § 1.136(a), and any fees required therefore (including fees for net addition of claims) are hereby authorized to be charged to Deposit Account No. 033975 (**Ref. No. 090611-0384422**).

This listing of claims will replace all prior versions and listings of claims in the Application.

LISTING OF CLAIMS:

1. (**Currently Amended**) A system for providing a natural language content dedication service, comprising:

~~an electronic device configured to~~ one or more processors; and

one or more non-transitory computer readable mediums storing executable instructions that when executed by the one or more processors cause the one or more processors to:

receive a first utterance that includes a natural language utterance; ~~a multi-modal interaction that includes a request to dedicate content, wherein the multi-modal interaction includes a natural language utterance relating to the request to dedicate the content; and~~

~~capture an utterance dedicating the content for a recipient; and~~

~~a content dedication system in communication with the electronic device through a message interface, wherein the content dedication system is configured to:~~

determine, based on processing of the first utterance by a speech recognition engine, one or more words or phrases of the first utterance;

identify, based on a natural language processing of the one or more words or phrases of the first utterance, content to dedicate to a recipient;

initiate, based on the identified content, a dedication to the recipient;

receive a second utterance to be associated with the dedication; and

~~identify the content to dedicate for the recipient based on an interpretation of the natural language utterance included in the multi-modal interaction;~~

~~process a transaction for the identified content to dedicate for the recipient; and~~

send ~~a content dedication message to the recipient, wherein the content dedication message includes~~ information to enable ~~for~~ the recipient to access the content and the second utterance. ~~, and wherein the content dedication message further includes~~

~~information for the recipient to access the utterance dedicating the content for the recipient.~~

2-4. (**Canceled**)

5. (**Currently Amended**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to ~~wherein the tags dedicating the content for the recipient include one or more of an image or a picture to insert within album art metadata for the content~~ provide an image as one or more tags within metadata of the content.

6. (**Currently Amended**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to ~~the content dedication system is further configured to~~ annotate the content with the second utterance. ~~utterance dedicating the content for the recipient.~~

7. (**Currently Amended**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to ~~the content dedication system is further configured to~~ purchase the ~~identified~~ content from a content provider ~~to process the transaction for the identified content~~ to enable the recipient to access the content.

8. (**Currently Amended**) The system of claim 7, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to: ~~the content dedication system is further configured to:~~

identify, for the dedication, a purchase option ~~for the request to dedicate the content, wherein the purchase option~~ that includes one or more of a buy-to-own purchase option, a pay-to-pay purchase option, or a paid subscription purchase option; and

bill a user associated with the first utterance ~~of the electronic device for the natural~~

~~content dedication service~~ based on ~~identified~~ the purchase option.

9. (**Currently Amended**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to: ~~the content dedication message sent to the recipient includes an electronic link for the recipient to access the content and the utterance dedicating the content for the recipient.~~

send, to the recipient, an electronic link to access the content as the information to enable the recipient to access the content; and

send, to the recipient, the second utterance in an encoded audio format as the information to enable the recipient to access the second utterance.

10. (**Original**) The system of claim 1, wherein the content includes one or more of a song or a video.

11. (**Currently Amended**) A method for providing a natural language content dedication service, comprising:

receiving, by one or more processors, a first utterance that includes a natural language utterance; ~~at an electronic device, a multi-modal interaction that includes a request to dedicate content, wherein the multi-modal interaction includes a natural language utterance relating to the request to dedicate the content;~~

~~capturing an utterance dedicating the content for a recipient;~~

determining, by the one or more processors, one or more words or phrases of the first utterance based on processing of the first utterance by a speech recognition engine;

identifying, by the one or more processors, content to dedicate to a recipient based on a natural language processing of the one or more words or phrases of the first utterance;

initiating, by the one or more processors, a dedication to the recipient based on the identified content;

receiving, by the one or more processors, a second utterance to be associated with the dedication; and

~~identifying the content to dedicate for the recipient based on an interpretation of the natural language utterance included in the multi-modal interaction;~~

~~processing at a content dedication system, a transaction for the identified content to dedicate for the recipient; and~~

sending, by the one or more processor, ~~a content dedication message from the content dedication system to the recipient, wherein the content dedication message includes~~ information to enable ~~for~~ the recipient to access the content and the second utterance. ~~, and wherein the content dedication message further includes information for the recipient to access the utterance dedicating the content for the recipient.~~

12-14. (**Canceled**)

15. (**Currently Amended**) The method of claim 11, further comprising providing an image as one or more tags within metadata of the content. ~~wherein the tags dedicating the~~

~~content for the recipient include one or more of an image or a picture to insert within art metadata for the content.~~

16. (**Currently Amended**) The method of claim 11, further comprising annotating the content with the second utterance. ~~utterance dedicating the content for the recipient.~~

17. (**Currently Amended**) The method of claim 11, further comprising purchasing~~, by the content dedication system,~~ the ~~identified~~ content from a content provider ~~to process the transaction for the identified content~~ to enable the recipient to access the content.

18. (**Currently Amended**) The method of claim 17, further comprising:

identifying, ~~at the content dedication system,~~ for the dedication, a purchase option ~~for the request to dedicate the content, wherein the purchase option~~ that includes one or more of a buy-to-own purchase option, a pay-to-pay purchase option, or a paid subscription purchase option; and

billing~~, by the content dedication system,~~ a user associated with the first utterance ~~of the electronic device for the natural content dedication service~~ based on ~~identified~~ the purchase option.

19. (**Currently Amended**) The method of claim 11, further comprising: ~~wherein the content dedication message sent to the recipient includes an electronic link for the recipient to access the content and the utterance dedicating the content for the recipient.~~

sending, to the recipient, an electronic link to access the content as the information to enable the recipient to access the content; and

sending, to the recipient, the second utterance in an encoded audio format as the information to enable the recipient to access the second utterance.

20. (**Original**) The method of claim 11, wherein the content includes one or more of a song or a video.

21. (**New**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to:

determine that a user is interacting with a media item; and

identify, based on the natural language processing and the interaction of the user with the media item, the media item as the content to dedicate to the recipient.

22. (**New**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to:

determine a media item that is a user is playing;

determine that the user provided the first utterance during the playing of the media item by the user; and

identify, based on the natural language processing and the determination that the user provided the first utterance during the playing of the media item, the media item as the content to dedicate to the recipient.

23. (**New**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to:

determine, based on processing of the second utterance by the speech recognition engine, one or more words or phrases of the second utterance; and

provide the one or more words or phrases of the second utterance as textual annotations within metadata of the content.

24. (**New**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to provide an electronic link associated with the content as one or more tags within metadata of the content.

25. (**New**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to provide a voice tag as

one or more tags within metadata of the content.

26. (**New**) The system of claim 1, wherein the executable instructions, when executed by the one or more processors, cause the one or more processors to provide the second utterance within the content.

REMARKS

In response to the Non-Final Office Action mailed **October 1, 2012** (hereinafter "Office Action"), claims 1, 5-9, 11, and 15-19 have been amended, and claims 2-4 and 12-14 have been canceled without prejudice or disclaimer. Claims 21-26 have been newly added. Therefore, claims 1, 5, 6-11, and 15-26 are pending. Support for the instant amendments is provided throughout the as-filed Specification. Thus, no new matter has been added. In view of the foregoing amendments and following comments, allowance of all the claims pending in the application is respectfully requested.

INFORMATION DISCLOSURE STATEMENT

Applicants are submitting herewith a Supplemental Information Disclosure Statement and respectfully request that the Examiner consider the cited references and provide a signed copy of the Form PTO-1449 for this submission with the next Office Action.

EXAMINER INTERVIEW

Applicants thank the Examiner for the courtesies extended to Applicants' representative (Ngai Zhang) during the telephonic interviews conducted on January 15, 2013 and on January 23, 2013. During the interviews, the subject matter of the application was explained and outstanding issues were discussed. Without acquiescing in the propriety of the Examiner's preliminary comments regarding alleged functional-language issues or the propriety of the alleged 35 U.S.C. §§ 101 and 112 rejections, it is noted that the Examiner indicated that amending the claims as presented herein would overcome the alleged issues in the preliminary comments and the alleged 35 U.S.C. §§ 101 and 112 rejections.

REJECTION UNDER 35 U.S.C. § 102

Claims 1-10 stand rejected under 35 U.S.C. § 102(b) as allegedly being anticipated by U.S. Patent Application Publication No. 2004/0193420 to Kennewick *et al.* ("Kennewick") [Office Action, pg. 7]. In particular, with respect to the alleged rejection, the Office Action alleges that a substantial portion of the recitations of independent claim 1 and all of the recitations of dependent claims 2-10 are directed to functional language. Based on this contention, the Office Action argues that it is not necessary that Kennewick teaches such claim recitations to reject the claims under 35 U.S.C. § 102(b). [Office Action, pg. 8]. Such a contention is misguided. There is no prohibition on functional claiming.

Nonetheless, *solely* in an effort to expedite prosecution, and in *no way* acquiescing in the propriety of the alleged rejection, claims 1, 5-9, 11, and 15-19 have been amended as suggested by the Examiner during the interviews. As indicated by the Examiner during the interviews, the amendment overcomes the alleged functional-language issues. In addition, because the Office Action fails to address each and every feature in independent claim 1, and the claims depending therefrom, in the alleged rejection, the alleged rejection is improper and should be withdrawn.

Moreover, claims 2-4 have been canceled without prejudice or disclaimer. As such, the alleged rejection with respect to claims 2-4 is now moot and should be withdrawn.

REJECTIONS UNDER 35 U.S.C. § 103

A. Claims 1, 2, 7, 9, 10, 11, 12, 17, 19, and 20

Claims 1, 2, 7, 9, 10, 11, 12, 17, 19, and 20 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over Kennewick in view of U.S. Patent Application Publication No. 2002/0032752 to Gold *et al*. ("Gold") [Office Action, pg. 8, ¶11].

Applicants disagree with the propriety of the alleged rejection. However, *solely* in an effort to expedite prosecution, and in *no way* acquiescing in the propriety of the alleged rejection, independent claim 1 has been amended to clarify aspects of the invention. In particular, claim 1 recites, *inter alia*, the features of:

> determine, based on a processing of the first utterance by a speech recognition engine, one or more words or phrases of the first utterance [that includes a natural language utterance];
>
> identify, based on a natural language processing of the one or more words or phrases of the first utterance, content to dedicate to a recipient;
>
> initiate, based on the identified content, a dedication to the recipient . . . . [Emphasis added].

Independent claim 11 includes similar recitations.

Even assuming arguendo that it were deemed legally proper to combine Kennewick and Gold in the manner alleged by the Examiner (which Applicants do not concede), the alleged combination still fails to teach or suggest *at least* the foregoing claim recitation.

In the Office Action, the Examiner relies on Kennewick to teach the foregoing claim recitation. [Office Action, pg. 9]. Applicants note, however, that while Kennewick discloses natural language processing, Kennewick does not describe *at least* all of the features of "identify, based on a natural language processing of the one or more words or phrases of the first utterance, content to dedicate to a recipient; [and] initiate, based on the identified content, a dedication to the recipient." Gold, which provides for a website that enables users to select a song (e.g., by clicking an electronic link) and a recipient to whom the song is to be sent (e.g., by typing in the recipient's email address), also does not provide for such features.

Accordingly, for at least the foregoing reasons, Kennewick and Gold, either alone or in combination, fail to teach or suggest each and every feature recited in independent claims 1 and 11. The alleged rejection under 35 U.S.C. § 103 is therefore improper and should be withdrawn. Dependent claims 7, 9, 10, 17, 19, and 20 are also allowable because they each ultimately depend from an allowable claim, as well as for the further features they recite.

Moreover, claims 2 and 12 have been canceled without prejudice or disclaimer. As such, the alleged rejection with respect to claims 2 and 12 is now moot and should be withdrawn.

B. Claims 3, 4, 6, 13, 14 and 16

Claims 3, 4, 6, 13, 14 and 16 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick and Gold, further in view of U.S. Patent Application Publication No. 2008/0140385 to Mahajan *et al*. ("Mahajan") [Office Action, pg. 12, ¶12].

Applicants disagree with the propriety of the alleged rejection. Nonetheless, based on the foregoing reasons in Section A, dependent claims 6 and 16 are allowable because they reach ultimately depend from an allowable claim. Moreover, dependent claims 6 and 16 are allowable for the further features they recite. For example, dependent claim 6 recites the feature of:

> annotate the content [that is identified as content to be dedicated to the recipient based on a processing of the first utterance] with the second utterance [that is to be associated with the content]. [Emphasis added]

Dependent claim 16 includes similar recitations.

Even assuming arguendo that it were deemed legally proper to combine Kennewick, Gold, and Mahajan in the manner alleged by the Examiner (which Applicants do not concede), the alleged combination still fails to teach or suggest *at least* the foregoing claim recitation.

In the Office Action, the Examiner acknowledges that neither Kennewick nor Gold, either alone or in combination, describe such annotations of content (that is identified as content to be dedicated to a recipient based on a processing of a first utterance) with a second utterance that is to be associated with the content. Instead, the Examiner relies on Mahajan to provide for such features. [Office Action, pg. 12]. However, Mahajan (at best) describes recognizing textual words of A/V content that is currently playing and displaying the recognized textual words as the A/V content is played. Thus, Mahajan also does not provide for the recited features of claims 6 and 16. Accordingly, the alleged rejection under 35 U.S.C. § 103 is therefore improper and should be withdrawn.

Moreover, claims 3, 4, 13, and 14 have been canceled without prejudice or disclaimer. As such, the alleged rejection with respect to claims 3, 4, 13, and 14 is now moot and should be withdrawn.

C. Claims 5, 8, 15, and 18

Claims 5 and 15 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick and Gold, further in view of U.S. Patent Application Publication No. 2010/0036967 to Caine *et al*. ("Caine") [Office Action, pg. 13, ¶13]. Claims 8 and 18 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the alleged combination of Kennewick and Gold, further in view of U.S. Patent No. 5,892,900 to Golden *et al*. ("Golden") [Office Action, pg. 14, ¶14].

Applicants disagree with the propriety of the alleged rejections. Nonetheless, based on the foregoing reasons in Section A, dependent claims 5, 8, 15, and 18 are allowable because they reach ultimately depend from an allowable claim, as well as for the further features they recite. Accordingly, the alleged rejections under 35 U.S.C. § 103 are therefore improper and should be withdrawn.

NEW CLAIMS

New claims 21-26 are allowable *at least* by virtue of their dependence on independent claim 1, as well as for the further features they recite.

CONCLUSION

Having addressed each of the foregoing rejections, it is respectfully submitted that a full and complete response has been made to the outstanding Office Action and, as such, the application is in condition for allowance. Notice to that effect is respectfully requested.

If the Examiner believes, for any reason, that personal communication will expedite prosecution of this application, the Examiner is invited to telephone the undersigned at the number provided.

Respectfully submitted,

Date: **January 30, 2013**

PILLSBURY WINTHROP SHAW PITTMAN LLP

By: /Ngai Zhang/
Ngai B. Zhang
Registration No. 65,473

Customer No. 00909

P.O. Box 10500
McLean, Virginia 22102
Direct Dial: 703.770.7532
Main: 703.770.7900
Fax: 703.770.7901

PTO/SB/22 (10-12)
Approved for use through 1/31/2013. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARMENT OF COMMERCE
Under the paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless if displays a valid OMB control number.

PETITION FOR EXTENSION OF TIME UNDER 37 CFR 1.136(a)

Docket Number (Optional): 090611-0384422

Application Number	Filed
12/943,699	November 10, 2010

For: SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE

Art Unit	Examiner
3685	Tim B. Hale

This is a request under the provisions of 37 CFR 1.136(a) to extend the period for filing a reply in the above identified application.

The requested extension and fee are as follows (check time period desired and enter the appropriate fee below):

		Fee	Small Entity Fee	
☒	One month (37 CFR 1.17(a)(1))	$150	$75	$ 75.00
☐	Two months (37 CFR 1.17(a)(2))	$570	$285	$
☐	Three months (37 CFR 1.17(a)(3))	$1,290	$645	$
☐	Four months (37 CFR 1.17(a)(4))	$2,010	$1,005	$
☐	Five months (37 CFR 1.17(a)(5))	$2,730	$1,365	$

☒ Applicant claims small entity status. See 37 CFR 1.27.

☐ A check in the amount of the fee is enclosed.

☐ Payment by credit card. Form PTO-2038 is attached.

☐ The Director has already been authorized to charge fees in this application to a Deposit Account.

☒ The Director is hereby authorized to charge any fees which may be required, or credit any overpayment, to Deposit Account Number 033975.

☒ Payment made via EFS-Web.

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

I am the

☐ applicant/inventor.

☐ assignee of record of the entire interest. See 37 CFR 3.71. 37 CFR 3.73(b) statement is enclosed (Form PTO/SB/96).

☒ attorney or agent of record. Registration Number 65,473.

☐ attorney or agent acting under 37 CFR 1.34. Registration number _____.

/Ngai Zhang/	January 30, 2013
Signature	Date
Ngai Zhang	(703) 770-7532
Typed or printed name	Telephone Number

NOTE: This form must be signed in accordance with 37 CFR 1.33. See 37 CFR 1.4 for signature requirements and certifications. Submit multiple forms if more than one signature is required, see below*.

☒ *Total of 1 forms are submitted.

This collection of information is required by 37 CFR 1.136(a). The information is required to obtain or retain a benefit by the public, which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 6 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Mail Stop PCT, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

American LegalNet, Inc.
www.FormsWorkFlow.com

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANTS :	Mike KENNEWICK *et al.*	CONFIRMATION NO.:	4185
SERIAL NUMBER :	12/943,699	EXAMINER:	Tim B. Hale
FILING DATE :	November 10, 2010	ART UNIT:	3685
FOR :	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE		

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT

Commissioner for Patents
P. O. Box 1450
Alexandria, VA 22313-1450

Sir:

Pursuant to the duty of disclosure under 37 C.F.R. §§ 1.56, 1.97 and 1.98, Applicants hereby make of record in the above-identified application the documents listed on the attached Form PTO-1449.

Applicants have listed a publication date on the attached PTO-1449 based on information presently available to the undersigned. However, the listed publication date should not be construed as an admission that the information was actually published on the date indicated.

Applicants reserve the right to establish the patentability of the claimed invention over any of the information provided herewith, and/or to prove that this information may not be prior art, and/or to prove that this information may not be enabling for the teachings purportedly offered.

This Supplemental Information Disclosure Statement is being filed more than three months after the filing date of a national application other than a continued prosecution application under 37 C.F.R. § 1.53(d), or more than three months after the date of entry of the national stage as set forth in 37 C.F.R. § 1.491 in an international application, and after the mailing date of the first Office Action on the merits, but before the mailing date of any of a Final Action under 37 C.F.R. § 1.113, a Notice of Allowance under 37 C.F.R. § 1.311, or an action that otherwise closes prosecution in the application. 37 C.F.R. § 1.97(c).

Please charge the $180.00 fee [37 C.F.R. § 1.17(p)] to Deposit Account No. 03-3975. In addition, please charge any fee deficiencies, or credit any overpayment, to Deposit Account No. 03-3975 as needed to ensure consideration of the Supplemental Information Disclosure Statement.

Document 1Z was cited in a search report by a foreign patent office in a counterpart foreign application. Submission of an English language version of the search report that indicates the degree of relevance found by the foreign patent office is provided.

It is respectfully requested that the Examiner initial and return a copy of the enclosed PTO-1449.

Respectfully submitted,

PILLSBURY WINTHROP SHAW PITTMAN LLP

By /Ngai Zhang/
Ngai B. Zhang
Registration No.: 65,473

Customer No.
00909

P. O. Box 10500
McLean, Virginia 22102
Telephone: (703) 770-7900
Telefax: (703) 770-7901

Dated: January 30, 2013

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	1A	5,027,406	06/25/1991	Roberts *et al.*	381	43	
	1B	5,208,748	05/04/1993	Flores *et al.*	364	419	
	1C	5,357,596	10/18/1994	Takebayashi *et al.*	395	2.84	
	1D	5,577,165	11/19/1996	Takebayashi et al.	395	2.84	
	1E	5,608,635	03/04/1997	Tamai	364	449.3	
	1F	5,740,256	04/14/1998	Castello Da Costa *et al.*	361	94.7	
	1G	5,748,974	05/05/1998	Johnson	395	759	
	1H	5,839,107	11/17/1998	Gupta *et al.*	704	270	
	1I	5,848,396	12/08/1998	Gerace	705	10	
	1J	5,855,000	12/29/1998	Waibel *et al.*	704	235	
	1K	5,897,613	04/27/1999	Chan	704	210	
	1L	5,960,397	09/28/1999	Rahim	704	244	
	1M	5,960,399	09/28/1999	Barclay *et al.*	704	270	
	1N	5,960,447	09/28/1999	Holt *et al.*	707	500	
	1O	5,995,928	11/30/1999	Nguyen *et al.*	704	251	
	1P	6,018,708	01/25/2000	Dahan *et al.*	704	244	
	1Q	6,035,267	03/07/2000	Watanabe *et al.*	704	1	
	1R	6,049,607	04/11/2000	Marash *et al.*	379	410	
	1S	6,067,513	05/23/2000	Ishimitsu	704	233	
	1T	6,076,059	06/13/2000	Glickman *et al.*	704	260	
	1U	6,085,186	07/04/2000	Christianson *et al.*	707	3	
	1V	6,108,631	08/22/2000	Ruhl	704	270	
	1W	6,167,377	12/26/2000	Gillick *et al.*	704	240	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO
	1X	1 320 043 A2	06/18/2003	Europe				
	1Y	1 646 037	04/12/2006	Europe				

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

	1Z	"Statement in Accordance with the Notice from the European Patent Office dated 1 October 2007 Concerning Business Methods (OJ EPO 11/2007, 592-593), XP002456252

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	2A	6,173,266	01/09/2001	Marx *et al.*	704	270	
	2B	6,188,982	02/13/2001	Chiang	704	256	
	2C	6,199,043	03/06/2001	Happ	704	272	
	2D	6,208,964	03/27/2001	Sabourin	704	244	
	2E	6,233,561	05/15/2001	Junqua *et al.*	704	277	
	2F	6,236,968	05/22/2001	Kanevsky *et al.*	704	275	
	2G	6,246,990	06/12/2001	Happ	704	275	
	2H	6,266,636	07/24/2001	Kosaka *et al.*	704	244	
	2I	6,269,336	07/31/2001	Ladd *et al.*	704	270	
	2J	6,275,231	08/14/2001	Obradovich	345	349	
	2K	6,278,968	08/21/2001	Franz *et al.*	704	3	
	2L	6,301,560	10/09/2001	Masters	704	251	
	2M	6,362,748	03/26/2002	Huang	340	901	
	2N	6,366,882	04/02/2002	Bijl *et al.*	704	235	
	2O	6,377,913	04/23/2002	Coffman *et al.*	704	8	
	2P	6,393,403	05/21/2002	Majaniemi	704	275	
	2Q	6,397,181	05/28/2002	Li *et al.*	704	256	
	2R	6,405,170	06/11/2002	Phillips *et al.*	704	270	
	2S	6,429,813	08/06/2002	Feigen	342	357.13	
	2T	6,434,529	08/13/2002	Walker *et al.*	704	275	
	2U	6,456,974	09/24/2002	Baker *et al.*	704	270.1	
	2V	6,510,417	01/21/2003	Woods *et al.*	704	275	
	2W	6,522,746	02/18/2003	Marchok *et al.*	379	406.03	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	3A	6,549,629	04/15/2003	Finn *et al.*	381	92	
	3B	6,560,590	05/06/2003	Shwe *et al.*	706	55	
	3C	6,571,279	05/27/2003	Herz *et al.*	709	217	
	3D	6,578,022	06/10/2003	Foulger *et al.*	706	45	
	3E	6,584,439	06/24/2003	Geilhufe *et al.*	704	270	
	3F	6,587,858	07/01/2003	Strazza	707	102	
	3G	6,591,239	07/08/2003	McCall *et al.*	704	275	
	3H	6,594,367	07/15/2003	Marash *et al.*	381	92	
	3I	6,601,026	07/29/2003	Appelt *et al.*	704	9	
	3J	6,606,598	08/12/2003	Holthouse *et al.*	704	275	
	3K	6,622,119	09/16/2003	Ramaswamy *et al.*	704	9	
	3L	6,631,351	10/07/2003	Ramachandran *et al.*	704	270	
	3M	6,633,846	10/14/2003	Bennett *et al.*	704	257	
	3N	6,681,206	01/20/2004	Gorin *et al.*	704	243	
	3O	6,701,294	03/02/2004	Ball *et al.*	704	257	
	3P	6,704,396	03/09/2004	Parolkar *et al.*	379	88.17	
	3Q	6,704,708	03/09/2004	Pickering	704	235	
	3R	6,745,161	06/01/2004	Arnold *et al.*	704	7	
	3S	6,751,612	06/15/2004	Schuetze *et al.*	707	4	
	3T	6,754,485	06/22/2004	Obradovich *et al.*	455	414.1	
	3U	6,757,544	06/29/2004	Rangarajan *et al.*	455	456.1	
	3V	6,832,230	12/14/2004	Zilliacus *et al.*	707	203	
	3W	6,850,603	02/01/2005	Eberle *et al.*	379	88.16	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	4A	6,868,380	03/15/2005	Kroeker	704	240	
	4B	6,877,001	04/05/2005	Wolf *et al.*	707	3	
	4C	6,910,003	06/21/2005	Arnold *et al.*	704	4	
	4D	6,912,498	06/28/2005	Stevens *et al.*	704	235	
	4E	6,934,756	08/23/2005	Maes	709	227	
	4F	6,954,755	10/11/2005	Reisman	707	10	
	4G	6,959,276	10/25/2005	Droppo *et al.*	704	226	
	4H	6,968,311	11/22/2005	Knockeart *et al.*	704	270	
	4I	6,975,993	12/13/2005	Keiller	704	275	
	4J	6,983,055	01/03/2006	Luo	381	313	
	4K	7,003,463	02/21/2006	Maes *et al.*	704	270.1	
	4L	7,024,364	04/04/2006	Guerra *et al.*	704	270	
	4M	7,027,586	04/11/2006	Bushey *et al.*	379	265.09	
	4N	7,035,415	04/25/2006	Belt *et al.*	381	92	
	4O	7,043,425	05/09/2006	Pao	704	211	
	4P	7,054,817	05/30/2006	Shao	704	270	
	4Q	7,058,890	06/06/2006	George *et al.*	715	728	
	4R	7,069,220	06/27/2006	Coffman *et al.*	704	275	
	4S	7,072,834	07/04/2006	Zhou	704	244	
	4T	7,107,210	09/12/2006	Deng *et al.*	704	226	
	4U	7,107,218	09/12/2006	Preston	704	270	
	4V	7,110,951	09/19/2006	Lemelson *et al.*	704	270	
	4W	7,130,390	10/31/2006	Abburi	379	88.17	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	5A	7,143,037	11/28/2006	Chestnut	704	251	
	5B	7,143,039	11/28/2006	Stifelman *et al.*	704	270	
	5C	7,165,028	01/16/2007	Gong	704	233	
	5D	7,184,957	02/27/2007	Brookes *et al.*	704	246	
	5E	7,190,770	03/13/2007	Ando *et al.*	379	88.01	
	5F	7,197,069	03/27/2007	Agazzi *et al.*	375	233	
	5G	7,206,418	04/17/2007	Yang *et al.*	381	92	
	5H	7,207,011	04/17/2007	Mulvey *et al.*	715	812	
	5I	7,236,923	06/26/2007	Gupta	704	9	
	5J	7,272,212	09/18/2007	Eberle *et al.*	379	88.17	
	5K	7,277,854	10/02/2007	Bennett *et al.*	704	257	
	5L	7,283,829	10/16/2007	Christenson *et al.*	455	461	
	5M	7,283,951	10/16/2007	Marchisio *et al.*	704	9	
	5N	7,299,186	11/20/2007	Kuzunuki *et al.*	704	270.1	
	5O	7,305,381	12/04/2007	Poppink *et al.*	1	1	
	5P	7,321,850	01/22/2008	Wakita	704	10	
	5Q	7,328,155	02/05/2008	Endo *et al.*	704	251	
	5R	7,337,116	02/26/2008	Charlesworth *et al.*	704	254	
	5S	7,366,285	04/29/2008	Parolkar *et al.*	379	88.17	
	5T	7,366,669	04/29/2008	Nishitani *et al.*	704	256	
	5U	7,406,421	07/29/2008	Odinak *et al.*	704	275	
	5V	7,415,414	08/19/2008	Azara *et al.*	704	270	
	5W	7,447,635	11/04/2008	Konopka *et al.*	704	275	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	6A	7,454,608	11/18/2008	Gopalakrishnan *et al.*	713	100	
	6B	7,472,020	12/30/2008	Brulle-Drews	701	211	
	6C	7,472,060	12/30/2008	Gorin *et al.*	704	240	
	6D	7,478,036	01/13/2009	Shen *et al.*	704	9	
	6E	7,493,259	02/17/2009	Jones *et al.*	704	257	
	6F	7,516,076	04/07/2009	Walker *et al.*	704	275	
	6G	7,529,675	05/05/2009	Maes	704	270.1	
	6H	7,536,297	05/19/2009	Byrd *et al.*	704	10	
	6I	7,536,374	05/19/2009	Au	706	55	
	6J	7,542,894	06/02/2009	Murata	704	9	
	6K	7,546,382	06/09/2009	Healey *et al.*	709	246	
	6L	7,558,730	07/07/2009	Davis *et al.*	704	235	
	6M	7,574,362	08/11/2009	Walker *et al.*	704	275	
	6N	7,577,244	08/18/2009	Taschereau	379	218.01	
	6O	7,606,708	10/20/2009	Hwang	704	257	
	6P	7,640,006	12/29/2009	Portman *et al.*	455	412.1	
	6Q	7,676,365	03/09/2010	Hwang *et al.*	704	240	
	6R	7,676,369	03/09/2010	Fujimoto *et al.*	704	270	
	6S	7,684,977	03/23/2010	Morikawa	704	211	
	6T	7,729,916	06/01/2010	Coffman *et al.*	704	270	
	6U	7,729,918	06/01/2010	Walker *et al.*	704	275	
	6V	7,729,920	06/01/2010	Chaar *et al.*	704	275	
	6W	7,748,021	06/29/2010	Obradovich	725	105	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	7A	7,788,084	08/31/2010	Brun *et al.*	704	7	
	7B	7,801,731	09/21/2010	Odinak *et al.*	704	275	
	7C	7,831,426	11/09/2010	Bennett	704	252	
	7D	7,831,433	11/09/2010	Belvin *et al.*	704	275	
	7E	7,873,519	01/18/2011	Bennett	704	257	
	7F	7,873,523	01/18/2011	Potter *et al.*	704	275	
	7G	7,881,936	02/01/2011	Longe *et al.*	704	257	
	7H	7,894,849	02/22/2011	Kass *et al.*	455	550.1	
	7I	7,902,969	03/08/2011	Obradovich	340	439	
	7J	7,917,367	03/29/2011	Di Cristo *et al.*	704	270.1	
	7K	7,920,682	04/05/2011	Byrne *et al.*	379	88.18	
	7L	7,949,529	05/24/2011	Weider *et al.*	704	270	
	7M	7,949,537	05/24/2011	Walker *et al.*	704	275	
	7N	7,953,732	05/31/2011	Frank *et al.*	707	724	
	7O	7,974,875	07/05/2011	Quilici *et al.*	705	14.4	
	7P	7,983,917	07/19/2011	Kennewick *et al.*	704	257	
	7Q	7,984,287	07/19/2011	Gopalakrishnan *et al.*	713	100	
	7R	8,015,006	09/06/2011	Kennewick *et al.*	704	236	
	7S	8,060,367	11/15/2011	Keaveney	704	247	
	7T	8,069,046	11/29/2011	Kennewick *et al.*	704	257	
	7U	8,073,681	12/06/2011	Baldwin *et al.*	704	9	
	7V	8,077,975	12/13/2011	Ma *et al.*	382	187	
	7W	8,082,153	12/20/2011	Coffman *et al.*	704	270	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	8A	8,086,463	12/27/2011	Ativanichayaphong *et al.*	704	275	
	8B	8,112,275	02/07/2012	Kennewick *et al.*	704	240	
	8C	8,140,327	03/20/2012	Kennewick *et al.*	704	226	
	8D	8,140,335	03/20/2012	Kennewick *et al.*	704	257	
	8E	8,145,489	03/27/2012	Freeman *et al.*	704	257	
	8F	8,150,694	04/03/2012	Kennewick *et al.*	704	257	
	8G	8,155,962	04/10/2012	Kennewick *et al.*	704	257	
	8H	8,170,867	05/01/2012	Germain	704	9	
	8I	8,195,468	06/05/2012	Weider *et al.*	704	275	
	8J	8,219,599	07/10/2012	Tunstall-Pedoe	707	955	
	8K	8,224,652	07/17/2012	Wang *et al.*	704	275	
	8L	8,326,627	12/04/2012	Kennewick *et al.*	704	257	
	8M	8,326,634	12/04/2012	Di Cristo *et al.*	704	270.1	
	8N	8,326,637	12/04/2012	Baldwin *et al.*	704	275	
	8O	8,332,224	12/11/2012	Di Cristo *et al.*	704	257	
	8P	2001/0049601	12/06/2001	Kroeker *et al.*	704	254	
	8Q	2001/0054087	12/20/2001	Flom *et al.*	709	218	
	8R	2002/0015500	02/07/2002	Belt *et al.*	381	66	
	8S	2002/0022927	02/21/2002	Lemelson *et al.*	701	301	
	8T	2002/0035501	03/21/2002	Handel *et al.*	705	10	
	8U	2002/0069071	06/06/2002	Knockeart et al.	704	275	
	8V	2002/0087326	07/04/2002	Lee *et al.*	704	270.1	
	8W	2002/0133402	09/19/2002	Faber *et al.*	705	14	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	9A	2002/0161646	10/31/2002	Gailey *et al.*	705	14	
	9B	2002/0173961	11/21/2002	Guerra	704	258	
	9C	2002/0184373	12/05/2002	Maes	709	228	
	9D	2003/0016835	01/23/2003	Elko *et al.*	381	92	
	9E	2003/0065427	04/03/2003	Funk *et al.*	701	1	
	9F	2003/0110037	06/12/2003	Walker *et al.*	704	257	
	9G	2003/0115062	06/19/2003	Walker *et al.*	704	258	
	9H	2003/0120493	06/26/2003	Gupta	704	270.1	
	9I	2003/0144846	07/31/2003	Denenberg *et al.*	704	277	
	9J	2003/0158731	08/21/2003	Falcon *et al.*	704	231	
	9K	2003/0161448	08/28/2003	Parolkar *et al.*	379	88.17	
	9L	2003/0204492	10/30/2003	Wolf *et al.*	707	3	
	9M	2003/0206640	11/06/2003	Malvar *et al.*	381	93	
	9N	2003/0212550	11/13/2003	Ubale	704	215	
	9O	2003/0212558	11/13/2003	Matula	704	260	
	9P	2003/0212562	11/13/2003	Patel *et al.*	704	275	
	9Q	2003/0225825	12/04/2003	Healey *et al.*	709	203	
	9R	2004/0098245	05/20/2004	Walker *et al.*	704	1	
	9S	2004/0117804	06/17/2004	Scahill *et al.*	719	320	
	9T	2004/0140989	07/22/2004	Papageorge	345	700	
	9U	2004/0167771	08/26/2004	Duan *et al.*	704	10	
	9V	2004/0205671	10/14/2004	Sukehiro *et al.*	715	532	
	9W	2005/0021470	01/27/2005	Martin *et al.*	705	51	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	10A	2005/0125232	06/09/2005	Gadd	704	270.1	
	10B	2005/0137877	06/23/2005	Oesterling *et al.*	704	275	
	10C	2005/0143994	06/30/2005	Mori *et al.*	704	235	
	10D	2005/0216254	09/29/2005	Gupta *et al.*	704	9	
	10E	2005/0234727	10/20/2005	Chiu	704	270.1	
	10F	2005/0283752	12/22/2005	Fruchter *et al.*	717	100	
	10G	2006/0041431	02/23/2006	Maes	704	270.1	
	10H	2006/0047509	03/02/2006	Ding *et al.*	704	231	
	10I	2006/0217133	09/28/2006	Christenson *et al.*	455	461	
	10J	2006/0285662	12/21/2006	Yin *et al.*	379	88.16	
	10K	2007/0033020	02/08/2007	(Kelleher) Francois *et al.*	704	226	
	10L	2007/0043868	02/22/2007	Kumar *et al.*	709	226	
	10M	2007/0073544	03/29/2007	Millett *et al.*	704	277	
	10N	2007/0078708	04/05/2007	Yu *et al.*	705	14	
	10O	2007/0078709	04/05/2007	Rajaram	705	14	
	10P	2007/0118357	05/24/2007	Kasravi *et al.*	704	10	
	10Q	2007/0135101	06/14/2007	Ramati *et al.*	455	414.1	
	10R	2007/0162296	07/12/2007	Altberg *et al.*	705	1	
	10S	2007/0198267	08/23/2007	Jones *et al.*	704	257	
	10T	2008/0034032	02/07/2008	Healey *et al.*	709	203	
	10U	2008/0065386	03/13/2008	Cross *et al.*	704	270	
	10V	2008/0109285	05/08/2008	Reuther *et al.*	705	7	
	10W	2008/0147410	06/19/2008	Odinak	704	270.1	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	3685

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	11A	2009/0052635	02/26/2009	Jones *et al.*	379	88.02	
	11B	2009/0067599	03/12/2009	Agarwal *et al.*	379	114.21	
	11C	2009/0106029	04/23/2009	DeLine *et al.*	704	275	
	11D	2009/0271194	10/29/2009	Davis *et al.*	704	235	
	11E	2009/0273563	11/05/2009	Pryor	345	157	
	11F	2009/0313026	12/17/2009	Coffman *et al.*	704	275	
	11G	2010/0029261	02/04/2010	Mikkelsen *et al.*	455	419	
	11H	2010/0185512	07/22/2010	Borger *et al.*	705	14.49	
	11I	2010/0312566	12/09/2010	Odinak *et al.*	704	275	
	11J	2011/0112827	05/12/2011	Kennewick *et al.*	704	9	
	11K	2011/0131036	06/02/2011	Di Cristo *et al.*	704	9	
	11L	2011/0131045	06/02/2011	Cristo *et al.*	704	249	
	11M	2011/0231182	09/22/2011	Weider *et al.*	704	9	
	11N	2011/0231188	09/22/2011	Kennewick *et al.*	704	236	
	11O	2012/0022857	01/26/2012	Baldwin *et al.*	704	9	
	11P	2012/0101809	04/26/2012	Kennewick *et al.*	704	9	
	11Q	2012/0101810	04/26/2012	Kennewick *et al.*	704	9	
	11R	2012/0109753	05/03/2012	Kennewick *et al.*	705	14.58	
	11S	2012/0150636	06/14/2012	Freeman *et al.*	705	14.49	
	11T	2012/0278073	11/01/2012	Weider et al.	704	235	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 320 043 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
18.06.2003 Bulletin 2003/25

(51) Int Cl.[7]: **G06F 17/30**

(21) Application number: **02258126.8**

(22) Date of filing: **26.11.2002**

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR IE IT LI LU MC NL PT SE SK TR
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **13.12.2001 GB 0129788**

(71) Applicant: **Hewlett-Packard Company**
Palo Alto, CA 94304 (US)

(72) Inventor: **Belrose, Guillaume**
Bristol BS2 8HT (GB)

(74) Representative: **Squibbs, Robert Francis**
Hewlett-Packard Limited,
IP Section,
Building 3
Filton Road
Stoke Gifford, Bristol BS 34 8QZ (GB)

(54) Multi-modal picture

(57) A system for presenting a multi-modal picture includes picture presentation equipment (37) for displaying an image (38) of the picture and for enabling a user to interact with the picture by selecting specific picture features and asking a specific query relating to the feature. A voice browser system (3) controlled according to dialog scripts (46) associated with the picture, determines an appropriate response having regard to the spoken user query and the picture feature being pointed to. Each picture can have multiple narrators associated with it and the can choose which narrator is currently active. Picture authoring apparatus is also provided.

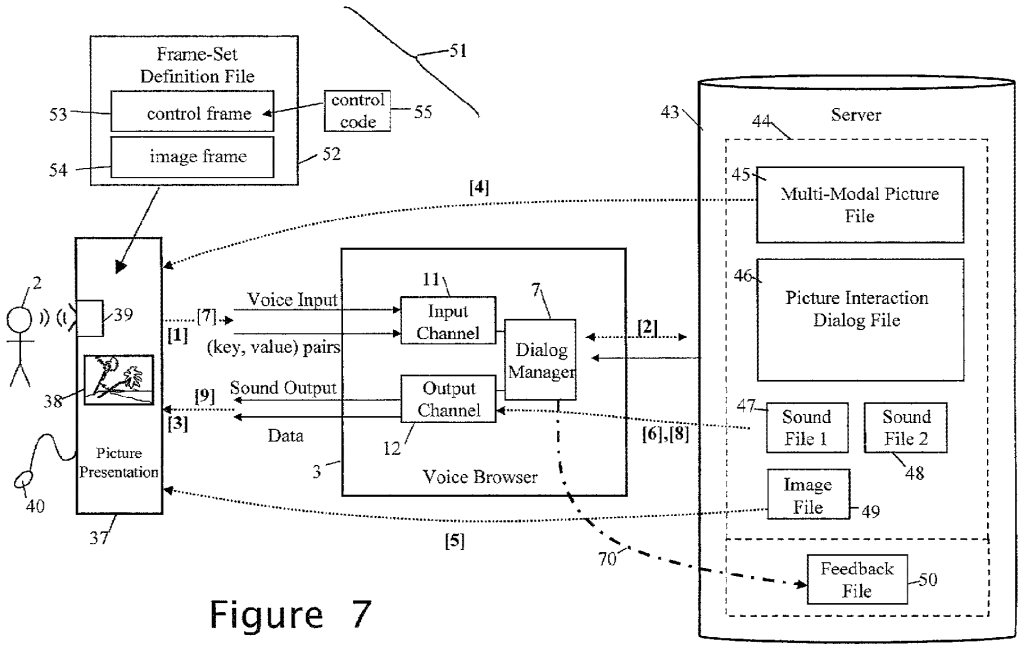


Figure 7

EP 1 320 043 A2

Amazon/VB Assets
Exhibit 1003
Page 335

Description

Field of the Invention

[0001] The present invention relates to multi-modal pictures with which a use can interact by spoken dialog exchanges.

Background of the Invention

[0002] In recent years there has been an explosion in the number of services available over the World Wide Web on the public internet (generally referred to as the "web"), the web being composed of a myriad of pages linked together by hyperlinks and delivered by servers on request using the HTTP protocol. Each page comprises content marked up with tags to enable the receiving application (typically a GUI browser) to render the page content in the manner intended by the page author; the markup language used for standard web pages is HTML (HyperText Markup Language).

[0003] However, today far more people have access to a telephone than have access to a computer with an Internet connection. Sales of cellphones are outstripping PC sales so that many people have already or soon will have a phone within reach where ever they go. As a result, there is increasing interest in being able to access web-based services from phones. 'Voice Browsers' offer the promise of allowing everyone to access web-based services from any phone, making it practical to access the Web any time and any where, whether at home, on the move, or at work.

[0004] Voice browsers allow people to access the Web using speech synthesis, pre-recorded audio, and speech recognition. Figure 1 of the accompanying drawings illustrates the general role played by a voice browser. As can be seen, a voice browser is interposed between a user 2 and a voice page server 4. This server 4 holds voice service pages (text pages) that are marked-up with tags of a voice-related markup language (or languages). When a page is requested by the user 2, it is interpreted at a top level (dialog level) by a dialog manager 7 of the voice browser 3 and output intended for the user is passed in text form to a Text-To-Speech (TTS) converter 6 which provides appropriate voice output to the user. User voice input is converted to text by speech recognition module 5 of the voice browser 3 and the dialog manager 7 determines what action is to be taken according to the received input and the directions in the original page. The voice input / output interface can be supplemented by keypads and small displays.

[0005] In general terms, therefore, a voice browser can be considered as a largely software device which interprets a voice markup language and generate a dialog with voice output, and possibly other output modalities, and / or voice input, and possibly other modalities (this definition derives from a working draft, dated September 2000, of the Voice browser Working Group of the World Wide Web Consortium).

[0006] Voice browsers may also be used together with graphical displays, keyboards, and pointing devices (e.g. a mouse) in order to produce a rich "multimodal voice browser". Voice interfaces and the keyboard, pointing device and display maybe used as alternate interfaces to the same service or could be seen as being used together to give a rich interface using all these modes combined.

[0007] Some examples of devices that allow multimodal interactions could be multimedia PC, or a communication appliance incorporating a display, keyboard, microphone and speaker/headset, an in car Voice Browser might have display and speech interfaces that could work together, or a Kiosk.

[0008] Some services may use all the modes together to provide an enhanced user experience, for example, a user could touch a street map displayed on a touch sensitive display and say "Tell me how I get here?". Some services might offer alternate interfaces allowing the user flexibility when doing different activities. For example while driving speech could be used to access services, but a passenger might used the keyboard.

[0009] Figure 2 of the accompanying drawings shows in greater detail the components of an example voice browser for handling voice pages 15 marked up with tags related to four different voice markup languages, namely:

- tags of a dialog markup language that serves to specify voice dialog behaviour;
- tags of a multimodal markup language that extends the dialog markup language to support other input modes (keyboard, mouse, etc.) and output modes (large and small screens);
- tags of a speech grammar markup language that serve to specify the grammar of user input; and
- tags of a speech synthesis markup language that serve to specify voice characteristics, types of sentences, word emphasis, etc.

[0010] When a page 15 is loaded into the voice browser, dialog manager 7 determines from the dialog tags and multimodal tags what actions are to be taken (the dialog manager being programmed to understand both the dialog and multimodal languages 19). These actions may include auxiliary functions 18 (available at any time during page processing) accessible through APIs and including such things as database lookups, user identity and validation, telephone call control etc. When speech output to the user is called for, the semantics of the output is passed, with any

Amazon/VB Assets
Exhibit 1003
Page 336

associated speech synthesis tags, to output channel 12 where a language generator 23 produces the final text to be rendered into speech by text-to-speech converter 6 and output to speaker 17. In the simplest case, the text to be rendered into speech is fully specified in the voice page 15 and the language generator 23 is not required for generating the final output text; however, in more complex cases, only semantic elements are passed, embedded in tags of a natural language semantics markup language (not depicted in Figure 2) that is understood by the language generator. The TTS converter 6 takes account of the speech synthesis tags when effecting text to speech conversion for which purpose it is cognisant of the speech synthesis markup language 25.

[0011] User voice input is received by microphone 16 and supplied to an input channel of the voice browser. Speech recogniser 5 generates text which is fed to a language understanding module 21 to produce semantics of the input for passing to the dialog manager 7. The speech recogniser 5 and language understanding module 21 work according to specific lexicon and grammar markup language 22 and, of course, take account of any grammar tags related to the current input that appear in page 15. The semantic output to the dialog manager 7 may simply be a permitted input word or may be more complex and include embedded tags of a natural language semantics markup language. The dialog manager 7 determines what action to take next (including, for example, fetching another page) based on the received user input and the dialog tags in the current page 15.

[0012] Any multimodal tags in the voice page 15 are used to control and interpret multimodal input/output. Such input/output is enabled by an appropriate recogniser 27 in the input channel 11 and an appropriate output constructor 28 in the output channel 12.

[0013] Whatever its precise form, the voice browser can be located at any point between the user and the voice page server. Figures 3 to 5 illustrate three possibilities in the case where the voice browser functionality is kept all together; many other possibilities exist when the functional components of the voice browser are separated and located in different logical/physical locations.

[0014] In Figure 3, the voice browser 3 is depicted as incorporated into an end-user system 8 (such as a PC or mobile entity) associated with user 2. In this case, the voice page server 4 is connected to the voice browser 3 by any suitable data-capable bearer service extending across one or more networks 9 that serve to provide connectivity between server 4 and end-user system 8. The data-capable bearer service is only required to carry text-based pages and therefore does not require a high bandwidth.

[0015] Figure 4 shows the voice browser 3 as co-located with the voice page server 4. In this case, voice input/output is passed across a voice network 9 between the end-user system 8 and the voice browser 3 at the voice page server site. The fact that the voice service is embodied as voice pages interpreted by a voice browser is not apparent to the user or network and the service could be implemented in other ways without the user or network being aware.

[0016] In Figure 5, the voice browser 3 is located in the network infrastructure between the end-user system 8 and the voice page server 4, voice input and output passing between the end-user system and voice browser over one network leg, and voice-page text data passing between the voice page server 4 and voice browser 3 over another network leg. This arrangement has certain advantages; in particular, by locating expensive resources (speech recognition, TTS converter) in the network, they can be used for many different users with user profiles being used to customise the voice-browser service provided to each user.

[0017] It is known to enhance pictures by providing associated speech annotations. It is an object of the present invention to provide further enhancements to such pictures.

Summary of the Invention

[0018] According to one aspect of the present invention, there is provided a system for presenting information concerning a picture to a user, the system comprising:

- a data store for holding responses, specific to said picture, in respect of specific user queries concerning particular picture features;
- a manually-operable feature-selection arrangement for enabling a user to select a feature in a displayed image of the picture, and for providing an output indication regarding what said particular feature, if any, the user has thereby selected;
- a voice dialog input-output subsystem including a speech recogniser for interpreting queries from a user;
- a control arrangement responsive to a user selecting a said particular feature and asking a specific query regarding that feature, to output the corresponding stored response.

[0019] Advantageously, each response is associated with a specified narrator and the user can opt to receive only the response of a selected narrator for either a current query or for all queries.

[0020] Preferably, provision is made for providing feedback to an author of the picture or its associated responses in the event that a user makes a query about a picture feature that is not a said particular picture feature or, although

Amazon/VB Assets
Exhibit 1003
Page 337

a query about such a feature, is not a said specific query for which a response is available.

[0021] According to another aspect of the present invention, there is provided a multi-modal picture specified by data held on at least one data carrier, this data comprising:

- picture image data for displaying a picture image;
- response data indicative of voice responses intended to be given to specific user queries concerning particular picture features of the picture;
- first control data for enabling a determination to be made as to which said particular feature in the picture image, if any, a user is selecting when using a selection arrangement to indicate a feature in the displayed image; and
- second control data for determining, on the basis of a spoken user query and on which said particular picture feature is selected by the user using the selection arrangement, which said response is to be used to reply to the user query.

[0022] According to a further aspect of the present invention, there is provided a multi-modal picture comprising a hard-copy picture image, and data held on at least one data carrier, this data comprising:

- response data indicative of voice responses intended to be given to specific user queries concerning particular picture features;
- first control data for enabling a determination to be made as to which said particular feature in the picture image, if any, a user is selecting when using a selection arrangement to indicate a feature of the image; and
- second control data for determining, on the basis of a spoken user query and on which said particular picture feature is selected by the user using the selection arrangement, which said response is to be used to reply to the user query.

[0023] The present invention also encompasses the overall process of generating a multi-modal picture and using it to convey information about particular picture features to a user. Thus, according to a still further aspect of the present invention, there is provided a method of conveying information about particular features in a picture, the method comprising the steps of:

(a) creating the following specifically-associated data:

- picture image data for displaying a picture image;
- response data indicative of voice responses intended to be given to specific user queries concerning particular picture features of the picture;
- first control data for enabling a determination to be made as to which said particular feature in the picture image, if any, a user is selecting when using a selection arrangement to indicate a feature in the displayed image; and
- second control data for determining, on the basis of a spoken user query and on which said particular picture feature is selected by the user, which said response is to be used to reply to the user query;

(b) using the image data to display an image of the picture;
(c) having a user use a manually-operated selection arrangement to select a feature in the displayed image and using the first control data to determine which said particular feature in the picture image, if any, the user is selecting;
(d) receiving and interpreting a spoken query from the user to determine if a said specific query is being asked; and
(e) using the second control data to determine, on the basis of the said particular feature determined as being selected in step (c) and the said specific query determined as being asked in step (d), which said response is to be used to reply and thereupon using the response data to output the corresponding voice response.

[0024] Similarly, according to a yet further aspect of the present invention, there is provided a method of conveying information about particular features in a hard-copy picture, the method comprising the steps of:

(a) creating the following specifically-associated data:

- response data indicative of voice responses intended to be given to specific user queries concerning particular picture features in said picture;
- first control data for enabling a determination to be made as to which said particular feature in the picture, if any, a user is selecting when using a selection arrangement to indicate a feature of the picture; and
- second control data for determining, on the basis of a spoken user query and on which said particular picture

Amazon/VB Assets
Exhibit 1003
Page 338

feature is selected by the user, which said response is to be used to reply to the user query;

(b) making the picture and data available to a user;
(c) having the user use a manually-operated selection arrangement to select a feature in the picture and using the first control data to determine which said particular feature in the picture, if any, the user is selecting;
(d) receiving and interpreting a spoken query from the user to determine if a said specific query is being asked; and
(e) using the second control data to determine, on the basis of the said particular feature determined as being selected in step (c) and the said specific query determined as being asked in step (d), which said response is to be used to reply and thereupon using the response data to output the corresponding voice response.

[0025] According to another aspect of the present invention, there is provided apparatus for authoring a multi-modal picture, comprising:

- a first tool for defining image hotspots associated with particular picture-image features;
- a second tool with speech recognition capability, for recording user responses input by voice, to user-specified queries each associated with a particular said picture-image feature; and
- means for automatically generating control data for determining, on the basis of a spoken user query and on which said particular picture feature is selected by a user, which said response is to be used to reply to the user query.

[0026] With respect to the above-mentioned feature regarding the provision of feedback on user queries for which responses are not available, a further independent aspect of the present invention provides a method of collecting user-interest information regarding a picture, the method comprising the steps of:

(a) presenting a picture image to a user, particular picture features in the image each having associated information that is presented to the user upon the user requesting such information;
(b) recognising a user information request, in respect of a picture feature, made by the user at least selecting that picture feature using a feature-selection arrangement; and
(c) in response to recognition of a user information request in step (b) concerning a picture feature that is other than a said particular picture feature, outputting an identifier of the concerned picture feature independently of the presented image.

[0027] This aspect of the invention may be implemented in the form of a system for collecting user-interest information regarding a picture, the system comprising:

- a data store for holding information concerning particular picture features;
- a feature-selection arrangement for enabling a user to select a feature in a displayed view of the picture, and for providing an output indication regarding what said particular feature, if any, the user has thereby selected;
- an information-output subsystem for outputting information about picture features to the user;
- a control arrangement for recognising a user information request, in respect of a picture feature, made by the user at least selecting that picture feature using the feature-selection arrangement, the control arrangement including:
    - first output means arranged to respond to a user making an information request concerning a said particular feature, by outputting the corresponding stored information via the information-output subsystem; and
    - second output means arranged to respond to a user making an information request concerning a picture feature that is other than a said particular picture feature, by outputting an identifier of the concerned picture feature for informing an author of the picture or its associated information regarding the unsatisfied information request.

Brief Description of the Drawings

[0028] A method and apparatus embodying the invention will now be described, by way of nonlimiting example, with reference to the accompanying diagrammatic drawings, in which:

. **Figure 1** is a diagram illustrating the role of a voice browser;
. **Figure 2** is a diagram showing the functional elements of a voice browser and their relationship to different types of voice markup tags;
. **Figure 3** is a diagram showing a voice service implemented with voice browser functionality located in an end-user system;

Amazon/VB Assets
Exhibit 1003
Page 339

. **Figure 4** is a diagram showing a voice service implemented with voice browser functionality co-located with a voice page server;
. **Figure 5** is a diagram showing a voice service implemented with voice browser functionality located in a network between the end-user system and voice page server; and
. **Figure 6** shows an example picture image of a multi-modal picture embodying the invention;
. **Figure 7** is a diagram of a system for presenting a multi-modal picture to a user;
. **Figure 8** is a diagram showing constituent dialog blocks of a picture interaction dialog file of the Figure 6 multi-modal picture;
. **Figure 9** is a diagram of apparatus for authoring a multi-modal picture;
. **Figure 10** shows, in respect of the Figure 6 image, a user seeking information from two picture features for which there is no associated information; and
. **Figure 11** shows the Figure 6 picture image enhanced with upper and lower information bars.

Best Mode of Carrying Out the Invention

[0029] In the following description, voice dialog interaction with a user is described based on a voice page server serving a dialog page with embedded voice markup tags to a multi-modal voice browser. Unless otherwise indicated, the foregoing description of voice browsers, and their possible locations and access methods is to be taken as applying also to the described embodiments ofthe invention. Furthermore, although voice-browser based forms of voice dialog services are preferred, the present invention in its widest conception, is not limited to these forms of voice dialog service system and other suitable systems will be apparent to persons skilled in the art.

Multi-Modal Picture

[0030] Figure 6 depicts a multi-modal picture comprising a displayed picture image 30 (here shown as being of a holiday island taken, for example, whilst the author of the picture was on holiday) with which a recipient (referred to below as the "user") can interact using multiple modalities and, in particular, by spoken dialogues and the use of a pointing arrangement, such as a cursor controlled by a mouse, stylus or keyboard keys, a touch-screen detection arrangement, etc. Thus, the user can query the picture using speech input, the query being either a general query about the picture or a specific query about a particular feature (item or area) of the picture indicated by use of the pointing arrangement (for example, the user uses a mouse to move a cursor over a picture feature and then clicks a mouse button). Of course, only certain picture features will have been set up to be queried and in the Figure 6 example, there are three such features, namely a coconut tree 31, a first person 32, and a second person 33; the dashed boxes around these features represent the "hotspot" image areas set up to encompass the features, the dashed lines generally not being visible (though displaying a hotspot boundary can be used as one way of indicating the hotspot to the user - a more typical way would be, in the case of the pointing arrangement being a mouse-controlled cursor, to change the cursor image as it moved into and out of a hotspot).

[0031] Processing functionality associated with the multi-modal picture is arranged to recognize one or more one or more general queries and, for each picture feature set up for selection, one or more specific queries. For each query set to be recognized, there is a corresponding response which is output to the user when the query is recognized, this output generally being in spoken form, generated using either speech synthesis or pre-recorded audio content. Thus, for example:

General Queries - available general queries can include picture location, the date of the picture was taken, and a description of the general subject of the picture. For instance:

User: "Describe the picture"
System: "This is a picture of the XYZ beach"
User: "What is the date?"
System: "The picture was taken November last year"
User: "What is the location?"
System: "XYZ beach is in Martinique, an island that is part of the French West Indies."

(In this and other dialog examples given in the present specification, the "System" is the functionality used to present the multi-modal picture to the user and to provide for multi-modal interaction).

Specific Queries - typically, the same specific queries are available for all the selectable picture features; example specific queries are: "What it is?", "Any story?", etc. For instance:

Amazon/VB Assets
Exhibit 1003
Page 340

User: "What is it"? "
System: "This is my cousin John".
User: "Any story?"
System: "John goes fishing quite often in the Caribbean Sea. One day, he just escaped a white shark attack."

[0032] The responses from the system can, of course, be more than just a simple single turn response and may involve a multi-turn structured dialog. Provision may also be made for different levels of content, for instance an initial level of public information and a lower level of private information only accessible under password control (or by some other security check procedure). Alternatively, there may be both a brief and a much fuller response .

[0033] A further possibility is to have responses from more than one narrator. Thus, a user can first obtain a response from one person associated with the picture and then a response from a different person also associated with the picture. For example, the user receives a picture from his/her parents. In the picture, the user's mother appears to be talking to someone, but the user's father looks bored. The user interaction might therefore proceed as follows:

User: "Mother, who is she?" [User uses pointing arrangement to indicate a person in the picture].
System: [In pre-recorded voice of user's mother] "This is my colleague from work. We had such a good time."
User: "Father, can you please describe the picture?"
System: [In pre-recorded voice of user's father] "We went to this rather boring evening with your mother's colleagues. They ended up talking about clothes all the time."

[0034] From the foregoing general description of a multi-modal picture, it will be appreciated that the term "multi-modal picture" covers more than just the picture image and includes also the associated dialog and the behaviour of the picture in response to user input.

Example Implementation

[0035] Figure 7 depicts an example implementation of a system for presenting a multi-modal picture.

[0036] The multi-modal picture is specified by a set 44 of related files held on a server 43, these files specifying both the image associated with the multi-modal picture and also the picture interaction dialogue associated with the picture. At least the dialogue and image files of set 44 are specifically associated with each other- that is, the association of these files is prespecified (in advance of user interaction) with the contents of the dialogue file being specifically tailored to the picture represented by the image file of the same set 44.

[0037] User 2 has local picture-presentation equipment 37; in the present case, this equipment is standard processing platform, such as a PC or portable computer, arranged to run a graphical web browser such as Microsoft's Internet Explorer. Other implementations of the picture presentation equipment 37 can alternatively be provided such as a mobile phone with a graphical display and suitable programs.

[0038] The picture-presentation equipment 37 displays the picture image of the multi-modal picture on display 38 and also provides for the input and output audio data via audio functionality 39, this audio input and output being transferred to and from a voice browser 3 (here shown as located separately from the picture picture-presentation equipment - for example, a network-based voice browser - but alternatively, integrated with the picture-presentation equipment 37). The picture-presentation equipment also has an associated pointing arrangement 40 depicted in Figure 7 as a mouse for controlling a cursor movable over the displayed picture image; other forms of pointing arrangements for selecting features of the picture image are, of course, also possible such as the use of touch pads, tracker balls or joysticks for moving an image cursor; touch-sensitive displays and other arrangements (such as a matrix of infra-red beams immediately overlying the display) for enabling a user to use a finger or stylus to point directly at a feature of the displayed image; etc..

[0039] The voice browser 3 comprises input channel 11, dialog manager 7 and output channel 12. The voice browser provides interactive dialog with the user, via the picture-presentation equipment 37, in respect of the currently-presented multi-modal picture, this dialog being in accordance with picture interaction dialog data retrieved from server 43.

[0040] The Figure 7 voice browser 3, as well as providing for voice input and output exchanges with the picture-presentation equipment 37, is also arranged to receive input data from the equipment in the form of key,value pairs indicating, for example, the selection of a particular picture feature by the user using the pointing arrangement 40 of the picture-presentation equipment. This data input is used by the dialog manager 7 in determining the course of the dialog with the user. The voice browser 3 can also provide data back to the picture-presentation equipment 37.

[0041] The picture-presentation equipment 37, voice browser 3 and server 3 inter-communicate via any suitable communications infrastructure or direct links.

[0042] Considering in more detail the files involved in the presentation of a multi-modal picture, these comprise, in addition to the set 44 of files that specify a multi-modal picture, a set of generic files 51. The multi-modal picture files

Amazon/VB Assets
Exhibit 1003
Page 341

44 comprise:

- the picture file 45, this file including a source reference for the picture image file 49 to be displayed and map data defining the image hotspots;
- a picture interaction dialog file 46 containing dialog scripts;
- one or more sound files 47, 48 (such as " .wav" files) containing audio data;
- the image file 49 (such as a ".jpg" or ".gif' file) containing the image data to be displayed.

[0043] The generic files 51 can be stored locally in the picture-presentation equipment 37 or retrieved from a remote location such as the server 43. The generic files comprise:

- a frame-set definition file 52 defining two frames 53, 54 into which page files can be independently loaded; one frame 53 is used to hold a file 55 containing control code (the contents of this frame not being visible), and the other frame 54 being used to hold the picture file 45 for the multi-modal picture to be presented.
- the control code file 55 to be loaded into frame 53, the control code being in the form of a number of scripts the main purpose of which is to provide key,value pairs to the voice browser according to events detected by the browser software run by the picture-presentation equipment 37 - in particular, clicking on an image hotspot as defined in file 45 is arranged to trigger a corresponding script in the control code file 55 whereby to cause a corresponding key,value pair to be passed to the voice browser 3 to inform it that a particular picture feature (corresponding to the activated hotspot) has been selected by the user.

[0044] To use the generic files to present a particular multi-modal picture, it is necessary to provide a reference to the multi-modal picture. This reference can, for example, be manually input by the user into an initial form displayed in the image frame 54 and then used by a script to request the loading of a file into frame 54. Alternatively, the multi-modal picture reference can be included as data in a query string attached to the URL of the frame-set definition file 52 (this URL and query string being, for example, provided to the user by the author of the multi-modal picture); in this case, in response to a request for the frame-set definition file 52, server-side code could, for example, extract the data from the query string and place it in the file source reference in the definition line for frame 54 in the frame-set definition file before that file is returned to the user.

[0045] In the present example, the multi-modal picture reference used to initiate presentation of the multi-modal picture is a reference to the picture interaction dialog file 46 to be loaded into the voice browser, rather than a reference to the picture file 45 that is to be loaded into the image frame 54. Thus, the multi-modal picture reference, however obtained, is passed in a key,value pair to the voice browser 3; voice browser 3 thereupon retrieves the picture interaction dialog file 46 to the dialog manager 7 of the voice browser. The file 46 includes a reference to the picture image file 45 to be loaded into the frame 54 and this reference is returned to the picture-presentation equipment 37 where it is used to retrieve the picture file 45. Where the picture-presentation equipment 37 comprises a standard web browser, one way of achieving the above is to have the dialog file reference sent (either as a source reference in the frame-set definition file or by a script in, for example, file 55) to the voice browser in a request for a file to load into frame 54, the voice browser sending back the picture file reference as a redirection.

[0046] It will be appreciated that the multi-modal picture reference passed into the generic files could have been that of the picture file 45, the latter then being retrieved into frame 54 and including an "onLoad" event script for passing to the voice browser a reference to the interaction dialog file.

[0047] The various steps involved in presenting a multi-modal picture according to the Figure 7 implementation are summarized below with reference to the reference numerals in square brackets in Figure 7:

[1] the generic files 51 have been loaded into the picture-presentation equipment 37 and a multi-modal picture reference has been provided; as a result, a key,value pair including a reference to the picture interaction dialog file 46 is passed to the voice browser 3;

[2] voice browser retrieves the dialog file 46 from the server 43 and dialog manger 7 uses the file to control further interaction with the user;

[3] the first action taken by the dialog manager 7 under the control of dialog file 46 is to pass to the equipment the reference for the picture file 45 along with a voice greeting to the user;

[4] the picture file 45 (including a reference to the image file 49) is retrieved from server 43 and loaded into the image frame 54;

[5] the image file 49 is retrieved from the server 43 and displayed on display 38;

[6] meanwhile, the dialog manager 7 causes a sound (in sound file 47) to be played to the user to indicate that the picture is ready to receive user input (this sound can simply be an appropriate background sound such as, for the Figure 6 picture image, the sound of the sea);

Amazon/VB Assets
Exhibit 1003
Page 342

[7] the user queries the picture by voice input (and possibly also by pointing to a particular area of the picture, this being indicated by a corresponding key,value pair sent to the voice browser along with the user voice input);
[8] dialog manager 7 acknowledges the receipt of the user query by causing an acknowledgement sound (in sound file 48) to be played back to the user;
[9] the dialog manager 7, having determined the appropriate response to the user query, outputs this response.

Steps [8] and [9] are repeated as many times as required by the user. In due course the user asks to exit and the dialog is terminated by the dialog manager.

[0048] Figure 8 illustrates the contents of the picture interaction dialog file 46. This file contains a number of dialog blocks 60 to 73 that contain dialog elements and/or control structures relating to dialog progression. Thus, dialog block 60 provides the initial greeting and causes the picture file reference to be passed to the equipment 37 (in step [3] above). Block 61 defines the query grammar and represents a waiting state for the dialog pending the receipt of a query from the user.

[0049] Block 62 carries out an analysis of a recognized query to determine whether it is an exit request (if so, an exit dialog block 63 is entered), a generic request, or a specific request; generic and specific requests are further analyzed to determine the nature of the query (that is, what "action" - type of information - is being requested). For a general query, the available actions are, in the present example, "date", "description", and "location; for a specific query, the action types are, in the present example, "what" and "story". Depending on the outcome of the action analysis, the dialog manager proceeds to one of blocks 64 - 66 (for a general query) or one of blocks 67 and 68 (for a specific query). The analysis carried out by dialog block 62 is on the basis of voice input only.

[0050] If the query was a general one, then block 64 is used to answer a date query, block 65 is used to respond to a description query, and block 66 is used to respond to a location query.

[0051] If the query was a specific "what" query, block 67 determines the identity of the picture feature (object) involved using the key,value pair provided to the voice browser; depending on the object identity, the dialog manager proceeds either to a "what" dialog block 70 for a coconut tree or to a "what" dialog block 71 for a French tourist. Similarly if the query was a specific "story" query, block 68 determines the identity of the picture feature (object) involved using the key,value pair provided to the voice browser; depending on the object identity, the dialog manager proceeds either to a "story" dialog block 72 for a coconut tree or to a "story" dialog block 73 for a French tourist.

[0052] After a response is provided by any one of the dialog blocks 64-66 or 70-73, the dialog manager returns to dialog block 61.

[0053] An example dialog interaction based on the Figure 6 picture and the Figure 8 dialog structure is given below; in this example, there is one narrator called Roger.

Example Interaction:

[0054]

System:	"This picture was sent to you by Guillaume. You can talk to Roger to ask for information about it."
User:	"Roger, Describe the picture"
System:	"This is a picture of the XYZ beach."
User:	"What is the location?"
System:	"This picture was taken in Martinique, an island located in the French West Indies."
User:	"What is it" [and clicks on feature 31]
System:	"This is a coconut tree. This kind of trees is very common in the Caribbean."
User:	"Any story?"
System:	"When I was little, I used to get coconuts from my garden. There is a sweet liquid inside that is very tasty. We use coconut to make ice creams and all sorts of nice desserts."
User:	"What is it?" [and clicks on 32]
System:	"I guess this is a French tourist"
User:	"Any story?"
System:	"Tourism is one of the main source of revenue in Martinique. A lot of people come from France when it is winter over there."
User:	"Exit"
System:	"Thanks for using this digital picture frame."

[0055] The Appendix to the present description includes a detailed script example of the Figure 8 dialog interaction as well as the HTML source of a picture file 45.

Amazon/VB Assets
Exhibit 1003
Page 343

Multi-modal picture authoring tool.

[0056] Figure 9 illustrates apparatus for enabling an author 80 to author a multi-modal picture. The apparatus comprises a computing platform 81 providing a graphical user interface and running a tool, such as Microsoft's FrontPage product, for authoring mark-up language pages and for creating image hotspot maps. The apparatus further comprises a speech interface system 82 (here shown as a voice browser though other forms of speech interface system can be used). The speech interface system 82 permits the author 80 to interact with the apparatus by voice and is set up to recognize command words such as "Record".

[0057] Different people, known as narrators, can author different aspects of the same picture. The apparatus keeps a record of narrators known to it.

[0058] The apparatus is arranged to interact with one or more narrators to build up, in memory, the set of files 44 that specify a multi-modal picture, it being assumed that the picture image file 47 around which the multi-modal picture is to be built has already been loaded into memory (for example, from a digital camera) and is displayed via the graphical user interface (GUI) of platform 81. The process of building the required files is controlled by a top-level authoring program 90 that has three main steps 91-93 as will be more fully explained below.

Identifying the narrator - (step 91)

[0059] The first step of the authoring program is to identify the current narrator. The narrator speaks his/her name into the speech interface system 82; if the name is known to the apparatus, the system replies with a recognition greeting. However, if the narrator's name is not already known to the apparatus, the system asks the narrator to create a new profile (basically, input his/her name and any other required information), using appropriate data collection screens displayed on the graphical user interface of the computing platform 81.

[0060] Example: the apparatus knows the following names: "Lawrence" and "Marianne".

Apparatus: "What is your name?"
Narrator: "Steve".
Apparatus: "Sorry, I do not know this name, please write it down."
Narrator: [inputs "Steve" via a data collection screen of the GUI].

[0061] The authoring program uses the narrator's name to customize a greeting dialog block of a template picture interaction dialog file 46.

Adding general information about the picture - (step 92)

[0062] After identification, the narrator can input general information concerning the picture image such as the date, the location or the description, via a spoken dialogue. In the following example, the command words recognized by the speech interface system 82 are shown in bold whilst the nature of information being recorded (corresponding to the query "action" type of Figure 8) is indicated by underlining. The words indicating the nature of the information are either pre-designated to the system (effectively limiting the classification of information to be input) or else the system can be arranged to analyze narrator "Record" commands to determine the nature of the information to be recorded.

Example:

[0063]

Narrator: **"Record** description"
Apparatus: [Plays a beep].
Narrator: "This is a picture of me and John fishing in the Caribbean sea". [The apparatus records this input, either directly as sound data or as a text data after the input has been subject to speech recognition by the system 82]
Narrator: **"Write** date".
Apparatus: [displays date capture screen on GUI].
Narrator: [inputs date information via GUI].
Narrator: **"Record** story"
Apparatus: [Plays a beep.]
Narrator: "This day, John was attacked by a white shark." [The apparatus records this input]

Amazon/VB Assets
Exhibit 1003
Page 344

[0064] The authoring program uses the input from the narrator to create corresponding dialog blocks, similar to those described above with reference to Figure 8, in dialog file 46.

Adding specific information - (step 93)

[0065] The narrator can also input information concerning a specific feature of the picture image. To do this, the narrator indicates the picture feature of interest by using the GUI to draw a "hotspot" boundary around the feature. The apparatus responds by asking the narrator to input a label for the feature via an entry screen displayed on the GUI.. The authoring program uses the input from the narrator to create the multi-media picture file 45 holding an image hotspot map with appropriate links to the control code scripts.

[0066] The narrator can then enter further information using the speech interface system or the GUI. The narrator can record or write multiple descriptions or stories for a single area of the picture, for example, to give different level of details.

Narrator: **"Record** description".
Apparatus: [Plays beep.]
Narrator: "It is a whale" [The apparatus records this input]
Narrator: **"Record** story"
Apparatus: [Plays beep.]
Narrator: "We saw this whale on the way to Dominica." [The apparatus records this input]
Narrator: **"Record next"** (indicates that further details for the story are to be recorded).
Narrator: "We were crossing the Guadeloupe channel when we saw it." [The apparatus records this input]

[0067] Again, the authoring program uses the input from the narrator to create corresponding dialog blocks; thus, for the above example, where a "whale" hotspot has been designated by the user, the authoring program generates a set of dialogs blocks: 'whaleDescription', 'whaleStoryl', 'whaleStory2', etc.

[0068] After the first narrator has finished inputting information, other narrators can enter information in the same manner.

User Feedback

[0069] It will be appreciated that the authoring of a multi-modal picture, and in particular the adding of the dialog data, can be quite involved. As a result, it is quite likely that the author will not always include information that a recipient may be interested in.

[0070] It is therefore useful to be able to monitor a user's interaction with a multi modal picture to see if the user tries to access missing information. For example, a user receives the Figure 6 multi modal picture from a friend. Two objects 35,36 (a building and an island) in the picture image intrigues the user (see Figure 9) who therefore clicks on each picture feature concerned and asks for more information ("What is it?"). Unfortunately, there are no hotspots associated with either picture feature and therefore in each case the voice browser comes back with the response "Sorry, there is no information about this item."

[0071] However, the coordinates corresponding to each picture feature the user clicked are known to the browser used to display the picture image and the control script (for example, in file 55) can be used to pass the coordinates as key,value pairs to the voice browser 3. At the same time, a user verbal query is also passed to the voice browser. The voice browser first determines whether the query is a general one and if it is, the voice browser ignores the received coordinate values; however, if the voice browser determines that the query is a specific one, then it determines from the key,value pairs received that the user has indicated a picture feature for which there is no corresponding information available. In this case, the voice browser logs the received coordinate values and the associated "action" type (in the example given above, "what" or "story") in a feedback file 50 that forms part of the set 44 of related files associated with the multimodal picture (see arrow 70 in Figure 7). For example, upon the user clicking on the island feature 36 in Figure 9 and asking "what is this?", the data (action='what'; coordx =400;coordy=300) is logged to file 50. Such logging functionality is, for example, provided by a further dialog block of Figure 8. The logged coordinates provide, together with an indication of the picture concerned, a picture-feature identifier that identifies the picture feature about which information has been requested by the user.

[0072] Alternatively or additionally to logging the "desired-information" feedback data in file 50, the author of the multi-modal picture can be sent a message (for example, an e-mail message) explaining the query from the user such as "John wants a description of this object.". This message includes a picture feature identifier that identifies the picture feature concerned. The picture feature identifier can take the form of explicit data items indicative of the picture concerned and the coordinates of the feature in the picture or may more directly indicate the feature by including either

Amazon/VB Assets
Exhibit 1003
Page 345

image data representing the relevant portion of the picture image or image data showing the picture image with a marking indicating the position in the image of the feature concerned (both such forms of picture-feature indication can also be included the feedback file 50 additionally or alternatively to the feature coordinates). The picture-feature indication need not be sent to the file 50 (or included in a message to the author) at the time the system detects that the user has asked for information about a non-hotspot picture feature; instead, the indication and related query can be temporarily stored and output, together with other accumulated similar indications and queries, either at a specified time or event (such as termination of viewing of the picture by the user) or upon request from the picture author.

[0073] The author can then decide whether or not to add in further information by adding an additional hotspot and additional dialogs (for example, using the above-described authoring apparatus).

[0074] Of course, the same general feedback process can be used where although a selected picture feature is associated with an existing hotspot, there is no existing query "action" corresponding to a user's query. Furthermore, a similar feedback process can be used where user queries are input by means other than speech (such as, for example, via a keyboard or by a hand-writing or stroke recognition system) and, indeed, where there are no explicit user queries such as when the selecting of picture feature is taken as a request for information about that feature.

Variants

[0075] Many variants are, of course, possible to the arrangements described above. For example, Figure 11 illustrates a variant form of the multi-modal picture in which the picture image 30 is accompanied by upper and lower information bars 100 and 101 respectively. The upper information bar 100 indicates the narrators associated with the picture whilst the lower information bar 101 indicates what types of general and specific queries are available for use. These information bars assists the user in appreciating what queries can be put and to whom. To further assist the user, speaking (or clicking on) a narrator's name is preferably arranged to indicate what hotspots are associated with that narrator and where these hotspots are located in the image - thus, in Figure 11, narrator "Vivian" has been selected and hotspot 32 is indicated on the picture image by a dashed hotspot boundary line. Query types used by "Vivian" can also be indicated by highlighting these types (in the Figure 11 example, "Vivian" has used general query type 'description' and specific query types 'what' and 'story'). Preferably, once a narrator has been selected, that narrator remains selected until a different (or all) narrators is subsequently selected, whereby with a narrator selected, only the responses of that narrator will be used in responding to the users' queries. However, it is also possible to arrange for selection of a narrator to be effective only for a single query. In this latter case, a convenient way of providing for user selection of a narrator to answer a query about a particular feature is for a list of narrators of responses about that feature to be displayed whenever the user points to that feature.

[0076] Instead of identifying narrators by name as shown in Figure 11, other forms of identifier can be used such as an image of each narrator.

[0077] As regards the generic control code, this can be provided in the form of a Java applet or any other suitable form and is not limited to the use of client-side scripts of the form described above. Furthermore, rather than the frame-set definition file 52 and control code file 55 being generic in form, they can be made specific to each multi-modal picture.

[0078] As already indicated, picture invocation can be initiated in a variety of ways. As a further example, picture invocation can be arranged to be effected by the creator sending the user a reference to the picture interaction dialog file, the user sending this reference to the voice browser to enable the latter to retrieve the dialog file; the dialog file is arranged to cause the return to the user of the frame-set definition file (or a reference to it) with the latter already including the correct source reference to the picture file as well as to the control code file.

[0079] Persons skilled in the art will appreciate that there are very many ways of implementing multi-modal pictures and the supporting functionality ranging from the multiple file approach described above to having just a single data file containing all the necessary data and arranged to be executed by specifically adapted software. It will also be appreciated that whilst in the described embodiments the course of the system-user interaction is controlled by the control logic embedded in the picture interaction dialog file and interpreted by the dialog manager 7 of the voice browser, it is possible to provide this functionality separate from the voice browser and the response scripts (thus, in general terms, the dialog block 62, 67 and 68 would be used by a separate control arrangement for determining, on the basis of the user voice and pointing-arrangement inputs, which of multiple stored responses is the appropriate one to play back to the user). Furthermore, it is not necessary to explicitly identify a selected feature from the coordinates output by the pointing arrangement as a separate step to choosing an appropriate response to a particular user query; thus, a picture feature could simply be specified in the control logic of the dialog interaction file by its coordinate values (of range of values) whereby this control logic tests coordinate values output by the pointing arrangement against coordinate values of particular features in the course of selecting a response to a user query.

[0080] With respect to the selection of a picture feature of interest to a user, in the described embodiments this has been carried out by means of a user-operated pointing arrangement whereby data input about the picture feature of interest is generated through a manual operation. In addition to the described arrangements for manually effecting

Amazon/VB Assets
Exhibit 1003
Page 346

feature selection, it is possible to use an arrangement in which the coordinates of the feature of interest are manually input using a keypad (a coordinate system being displayed as part of the image presented to the user, or being presented around the boundary of the display area). Another possible feature-selection arrangement is one based on specifying, via keypad input, a particular area of the display where either the display is divided into labelled areas or there is a direct mapping between keypad keys and display areas. It is also possible to label each feature of interest in the image with a reference and have the user effect feature selection by keypad input of the appropriate reference.

[0081] Further as regards the determination of which picture feature is being selected by a user, it should be noted that determining the picture feature of interest from the image coordinates or image area identified by whatever selection arrangement is being used, can be done in ways other than that described above in which image coordinates generated by the feature-selection arrangement are mapped to picture features using predetermined mapping data as described above. For example, the image can have data encoded into it that labels particular picture features, the pointing arrangement being arranged to read this label data when over the corresponding picture feature. Technology for embedding auxiliary data in picture image data in a way that does not degrade the image for the average human viewer is already known in the art.

[0082] The picture image can be a hard-copy image carrying markings (such as infra-red ink markings) intended to be read by a sensor of a suitable pointing arrangement whereby to determine what particular picture feature is being pointed to; the markings can be a pattern of markings enabling the pointing arrangement to determine the position of its sensor on the image (in which case, an image map can be used to translate the coordinates into picture features) or the markings can be feature labels appropriately located on the image. Even without special markings added to the image, the image can still be a hard-copy image provided the image is located in a reference position relative to which the pointing arrangement can determine the position of a pointing element (so that an image map can be used to translate pointing-element position into picture features). Other manually-operated selection arrangements, such as those based on explicit coordinate input via a keypad, can also be used.

[0083] Where image map data is used to translate image coordinates to picture features, the image map data can be held separately from the picture file and only accessed when needed; this facilitates updating of the image map data. A reference (e.g. URL) to the image map data can be included in the picture file or, where the image is a hard-copy image, in markings carried by the image.

[0084] Preferably, the described embodiments are applied to pictures of scenes and places in the real world such as a tourist might take with a camera. However, it is also possible to apply the embodiments to topographic pictures that are primarily intended to convey map-type information.

Amazon/VB Assets
Exhibit 1003
Page 347

APPENDIX

This Appendix forms an integral part of the description.

Example of Multi-Modal Picture file.

The following HTML file was generated using Microsoft FrontPage and is the Multi-Modal Picture file 45 for the picture depicted in Figure 6; the HTML code pulls in image file "beachPicture.jpg" and defines three "hotspots" which when clicked result in a call to a function (method) of the control code file 55 with the passing of a parameter value indicating which hotspot has been clicked on.

File Name: MMPicture1.htm

```
<html>
<head>
<meta http-equiv="Content-Type" content="text/html; charset=windows-1252">
<meta name="GENERATOR" content="Microsoft FrontPage 4.0">
<meta name="ProgId" content="FrontPage.Editor.Document">
<title>Frame Page</title>
</head>
<body>
<p><map name="FPMap0">
<area href="http://localhost/mmpic1/dpfBlank.htm"
onclick="parent.parent.scriptFrame.pushParams('object=coconutTree')"
target="invisible" coords="128, 84, 253, 236" shape="rect">
<area href="http://localhost/mmpic1/dpfBlank.htm"
onclick="parent.parent.scriptFrame.pushParams('object=frenchTourist')"
target="invisible" shape="rect" coords="204, 266, 219, 282">
<area href="http://localhost/mmpic1/dpfBlank.htm"
onclick="parent.parent.scriptFrame.pushParams('object=frenchTourist')"
target="invisible" shape="rect" coords="309, 257, 324, 269"></map>
<img border="0" src="beachPicture.jpg" align="center" usemap="#FPMap0"
width="569" height="375"></p>
</body>
</html>
```


Amazon/VB Assets
Exhibit 1003
Page 348

Example of Picture Interaction Dialogue file

[0085] The following is an example implementation of the Figure 8 picture interaction dialog file 46. In this example, there is one narrator called Roger. The user can say things like:"Roger, what is the date?".

```
<talkml>
    <!--variables declaration-->
    <create name="object"          expr="'undefined'"/>
    <create name="action"          expr="'what'"/>
    <create name="type"        expr="'generic'"/>
    <!-- Specify here the url of the display webserver-->
    <display-server url="localhost"/>

    <dialog name="welcome">
        <say> Welcome to the digital picture frame. </say>
        <!-- Display the Multi Modal picture of beach. -->
        <display-show url="MMPicture1.htm"/>
        <say>
            This picture was sent to you by Guillaume. You can talk to Roger to
            ask for information about it.
        </say>
        <goto next="getUserRequest"/>
    </dialog>

    <dialog name="getUserRequest">
        <play src="sound1.wav"/>
        <display-listen wait_after="2000"/>
        <speech wait_after="2000">
            sentence =  narrator question | question | exit {action='exit'};
            question =  genericQuestion {type='generic'} | specificQuestion
                        {type='specific'};
```


Amazon/VB Assets
Exhibit 1003
Page 349

```
            narrator = roger {narrator='roger'};
            roger    = "roger";
            genericQuestion = date {action='date'}| location {action='location'}|
                              description {action='description'};
            date     = "what is the date";
            location = "what is the location";
            description = "describe the picture";
            specificQuestion= what {action='what'} | story {action='story'};
            what     = "what is it";
            story    = "any story";
            exit     = "exit";
        </speech>

        <recognised>
            <play src="sound2.wav"/>
            <goto next="filterRequest"/>
        </recognised>

        <unrecognised>
            <play src="sound2.wav"/>
            <say> I am sorry, I have some trouble understanding you.</say>
            <retry/>
        </unrecognised>

</dialog>


<dialog name="filterRequest">
    <if cond="action='exit'">
        <goto next="exit"/>
    </if>
```


Amazon/VB Assets
Exhibit 1003
Page 350

```
        <if cond="type='generic'">
                <if cond="action='description'"><goto next="description"/></if>
                <if cond="action='date'"><goto next="date"/></if>
                <if cond="action='location'"><goto next="location"/></if>
        </if>

        <if cond="type='specific'">
                <if cond="action='what'">    <goto next="what"/> </if>
                <if cond="action='story'">    <goto next="story"/></if>
        </if>

</dialog>


<dialog name="description">
        <say>
                This is a picture of the beach.
        </say>
        <goto next="getUserRequest"/>
</dialog>


<dialog name="date">
        <say>
                Guillaume took this picture during his last trip home to
                Martinique last November
        </say>
        <goto next="getUserRequest"/>
</dialog>


<dialog name="location">
```


Amazon/VB Assets
Exhibit 1003
Page 351

```
        <say>
            This picture was taken in Martinique, an island located in the
            French West Indies.
        </say>
        <goto next="getUserRequest"/>
    </dialog>

    <dialog name="what">
        <if cond="object='coconutTree'"><goto next="coconutTreeWhat"/></if>
        <if cond="object='frenchTourist'"><goto next="frenchTouristWhat"/></if>
        <say>
            Apparently, you did not click on any parts of the picture. Try
            again.
        </say>
        <goto next="getUserRequest"/>
    </dialog>

    <dialog name="story">
        <if cond="object='coconutTree'"><goto next="coconutTreeStory"/></if>
        <if cond="object='frenchTourist'"><goto next="frenchTouristStory"/></if>
        <say>
            Apparently, you did not click on any parts of the picture. Try
            again.
        </say>
        <goto next="getUserRequest"/>
    </dialog>

    <dialog name="frenchTouristWhat">
        <say>
            I guess this is a French tourist.
        </say>
```


Amazon/VB Assets
Exhibit 1003
Page 352

```
        <goto next="getUserRequest"/>
</dialog>

<dialog name="coconutTreeWhat">
        <say>
                This is a coconut tree. This kind of trees is very common in the
                Caribbean.
        </say>
        <goto next="getUserRequest"/>
</dialog>

<dialog name="frenchTouristStory">
        <say>
                Tourism is one of the main source of revenue in Martinique. A lot
                of people come from France when it is winter over there.
        </say>
        <goto next="getUserRequest"/>
</dialog>

<dialog name="coconutTreeStory">
        <say>
                When I was little,I used to get coconuts from my garden. There is
                a sweet liquid inside that is very tasty. We use coconut to make
                ice creams and all sorts of nice desserts.
        </say>
        <goto next="getUserRequest"/>
</dialog>
```


Amazon/VB Assets
Exhibit 1003
Page 353

```
<dialog name="exit">
        <say> Thanks for using this digital picture frame. </say>
</dialog>

</talkml>
```

Claims

1. A system for presenting information concerning a picture to a user, the system comprising:

   - a data store for holding responses, specific to said picture, in respect of specific user queries concerning particular picture features;
   - a manually-operable feature-selection arrangement for enabling a user to select a feature in a displayed image of the picture, and for providing an output indication regarding what said particular feature, if any, the user has thereby selected;
   - a voice dialog input-output subsystem including a speech recogniser for interpreting queries from a user;
   - a control arrangement responsive to a user selecting a said particular feature and asking a specific query regarding that feature, to output the corresponding stored response.

2. A system according to claim 1, wherein image-map data is associated with the picture image for mapping image coordinates to said particular features, the selection arrangement being arranged to use the image-map data to determine what picture feature is selected by the user.

3. A system according to claim 1, wherein said image includes label data positioned in the region of a said particular picture feature to indicate the identity of that feature, the selection arrangement being arranged to read the label data to determine what picture feature is selected by the user.

4. A system according to any one of claims 1 to 3, further comprising a display subsystem for displaying said image provided to it in the form of digital image data.

5. A system according to any one of claims 1 to 3, wherein the picture image is a hard-copy image.

6. A system according to any one of the preceding claims, wherein the data store is arranged also to hold responses concerning general queries that are not associated with any particular picture feature, the control arrangement being arranged to respond to the user voice input of a general query by returning the appropriate response.

7. A system according to any one of the preceding claims, wherein the voice interface subsystem and control arrangement are arranged to cooperate in recognising multiple different queries in respect of a particular picture feature selected by the user.

8. A system according to claim 1, wherein each of at least some of the responses is associated with a specified narrator, the system being arranged to permit a user to receive only the response of a user-selected narrator in respect of at least one query.

9. A system according to claim 8, wherein user selection of a narrator is arranged to be effected by voice input, the control arrangement being arranged to respond to the selection of a particular specified narrator by user voice input, by using only the said responses associated with that narrator in providing a response to a user query.

10. A system according to claim 9, further comprising means for displaying, along with said picture image, identifiers of narrators associated with the picture.

11. A system according to claim 9, further comprising a display subsystem for displaying said image provided to it in

Amazon/VB Assets
Exhibit 1003
Page 354

the form of digital image data and identifiers of narrators associated with the picture; the display subsystem being arranged to respond to selection of a particular specified narrator by user voice input by indicating on the displayed image the said particular features for which responses are available concerning that narrator.

12. A system according to claim 8, further comprising a display subsystem for displaying said image provided to it in the form of digital image data, and identifiers of narrators associated with the picture; user selection of a narrator being arranged to be effected by the user using said selection arrangement to select a displayed narrator identifier, and the control arrangement being arranged to respond to the selection of a particular narrator by using only the said responses associated with that narrator in providing a response to a user query.

13. A system according to claim 12, wherein the display subsystem is arranged to respond to selection of a particular specified narrator to indicate on the displayed image the said particular features for which responses are available concerning the currently-selected narrator.

14. A system according to any one of the preceding claims, wherein said selection arrangement is a pointing arrangement usable by the user to point to a feature of interest in a displayed view of the picture.

15. A system according to any one of the preceding claims, further comprising a feedback arrangement arranged to respond to a user query concerning a picture feature that is other than a said particular picture feature, by outputting at least an identifier of the concerned picture feature for informing an author of the picture or its associated responses regarding the unsatisfied query.

16. A system according to any one of the preceding claims, further comprising a feedback arrangement arranged to respond to a user query that although concerning a said particular picture feature is not a said specific query, by outputting an indication of the query and an identifier of the concerned picture feature for informing an author of the picture or its associated responses regarding the unsatisfied query.

17. A system according to claim 15, wherein the feedback arrangement is further arranged to output an indication of the user query together with said picture feature identifier.

18. A system according to any one of claims 15 to 17, wherein the picture feature identifier arranged to be output by the feedback arrangement comprises an indication of the picture concerned and image coordinates of the picture feature.

19. A system according to any one of claims 15 to 17, further comprising means for converting digital picture data into said displayed image of the picture, the picture feature identifier arranged to be output by the feedback arrangement comprising digital data representing a corresponding portion of the image.

20. A system according to any one of claims 15 to 17, further comprising means for converting digital picture data into said displayed image of the picture, the picture feature identifier arranged to be output by the feedback arrangement comprising digital data representing the image with a marking indicating the position of the feature in the image.

21. A system according to any one of claims 15 to 20, wherein the feedback arrangement is operative to output said picture feature identifier to a log file.

22. A system according to any one of claims 15 to 20, wherein the feedback arrangement is operative to output said picture feature identifier in the form of an electronic message addressed to a person associated with the generation of the picture or its associated responses.

23. A multi-modal picture specified by data held on at least one data carrier, this data comprising:

- picture image data for displaying a picture image;
- response data indicative of voice responses intended to be given to specific user queries concerning particular picture features of the picture;
- first control data for enabling a determination to be made as to which said particular feature in the picture image, if any, a user is selecting when using a selection arrangement to indicate a feature in the displayed image; and
- second control data for determining, on the basis of a spoken user query and on which said particular picture

Amazon/VB Assets
Exhibit 1003
Page 355

feature is selected by the user using the selection arrangement, which said response is to be used to reply to the user query.

24. A multi-modal picture according to claim 23, wherein the first control data comprises image-map data mapping image coordinates to said particular features.

25. A multi-modal picture according to claim 23, wherein the first control data comprises label data arranged to be positioned in the displayed image in the region of each said particular picture feature to indicate the identity of that feature.

26. A multi-modal picture according to claim 23, wherein at least some of the said responses are associated with a narrator identified in the response data, the second control data enabling the determination of which said response is to be used to reply to the user query to be restricted to those responses associated with a said narrator that has been selected by the user.

27. A multi-modal picture according to any one of claims 23 to 26, wherein the picture is of a non-topographic real-world scene.

28. A multi-modal picture comprising a hard-copy picture image, and data held on at least one data carrier, this data comprising:

- response data indicative of voice responses intended to be given to specific user queries concerning particular picture features;
- first control data for enabling a determination to be made as to which said particular feature in the picture image, if any, a user is selecting when using a selection arrangement to indicate a feature of the image; and
- second control data for determining, on the basis of a spoken user query and on which said particular picture feature is selected by the user using the selection arrangement, which said response is to be used to reply to the user query.

29. A multi-modal picture according to claim 28, wherein the first control data comprises image-map data mapping image coordinates to said particular features.

30. A multi-modal picture according to claim 28, wherein the first control data comprises label data arranged to be positioned in or on the image in the region of each said particular picture feature to indicate the identity of that feature.

31. A multi-modal picture according to claim 28, wherein at least some of the said responses are associated with a narrator identified in the response data, the second control data enabling the determination of which said response is to be used to reply to the user query to be restricted to those responses associated with a said narrator that has been selected by the user.

32. A multi-modal picture according to any one of claims 28 to 31, wherein the picture is of a non-topographic real-world scene.

33. Apparatus for authoring a multi-modal picture, comprising:

- a first tool for defining image hotspots associated with particular picture-image features;
- a second tool with speech recognition capability, for recording user responses input by voice, to user-specified queries each associated with a particular said picture-image feature; and
- means for automatically generating control data for determining, on the basis of a spoken user query and on which said particular picture feature is selected by a user, which said response is to be used to reply to the user query.

Amazon/VB Assets
Exhibit 1003
Page 356

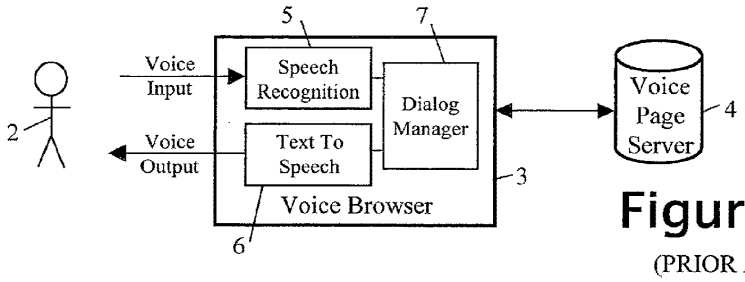


Figure 1
(PRIOR ART)

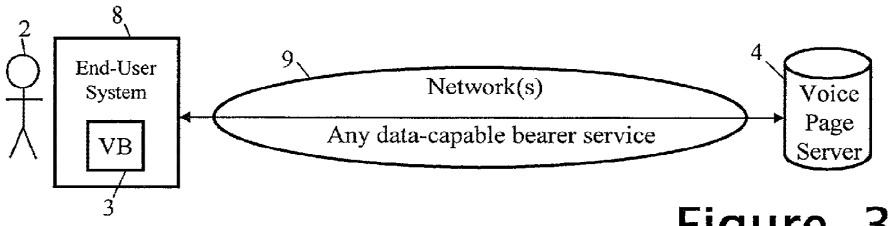


Figure 3
(PRIOR ART)

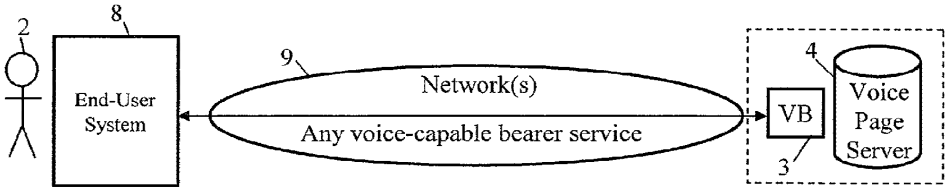


Figure 4
(PRIOR ART)

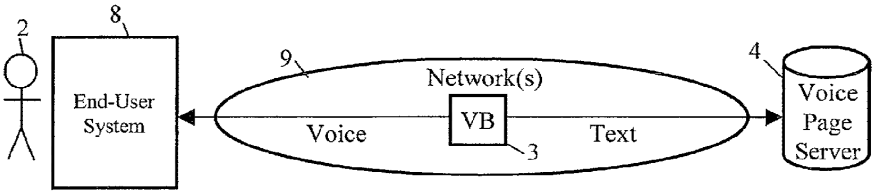


Figure 5
(PRIOR ART)

Amazon/VB Assets
Exhibit 1003
Page 357

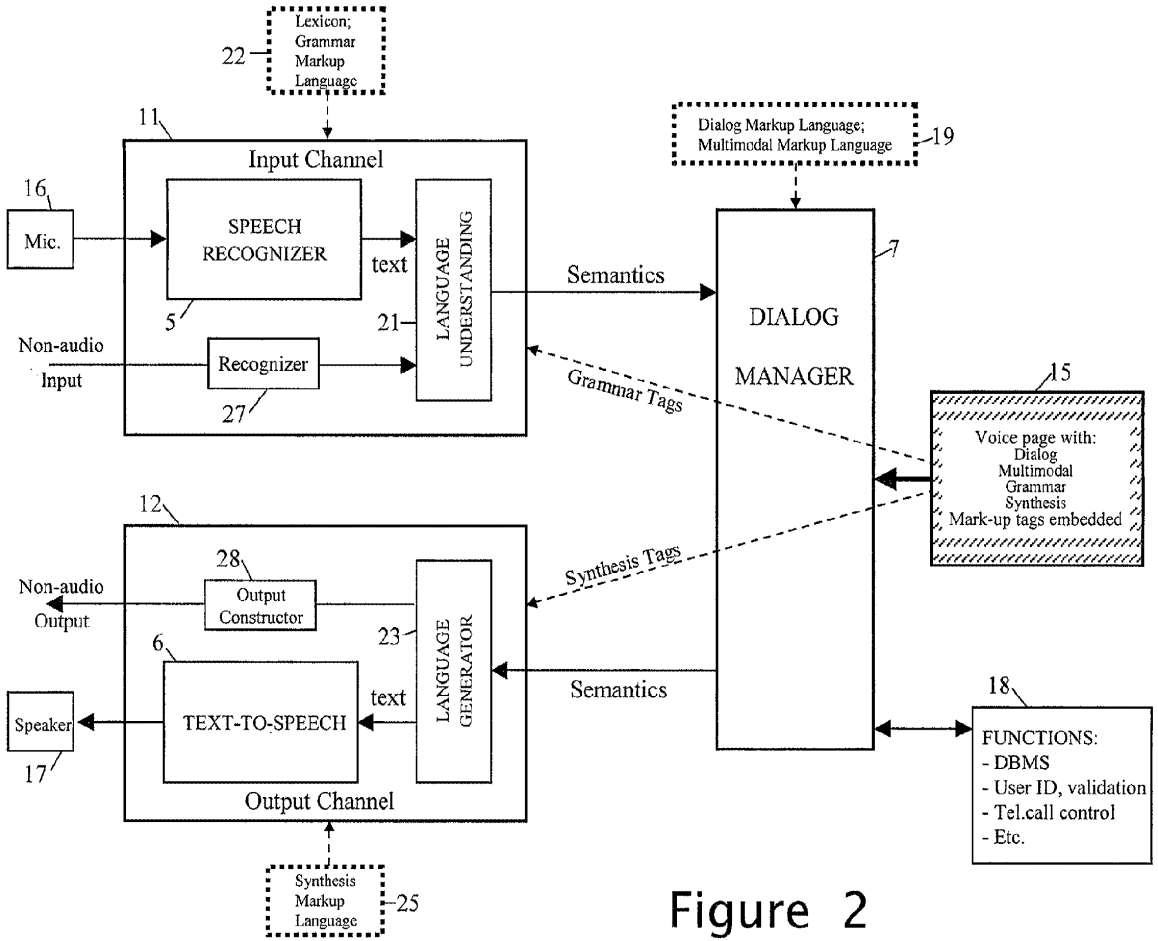


Figure 2
(PRIOR ART)

Amazon/VB Assets
Exhibit 1003
Page 358

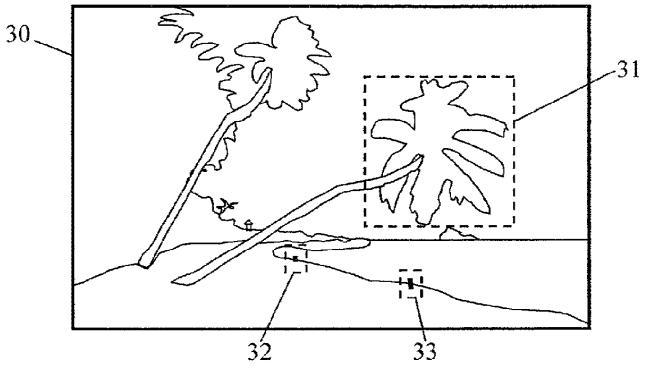


Figure 6

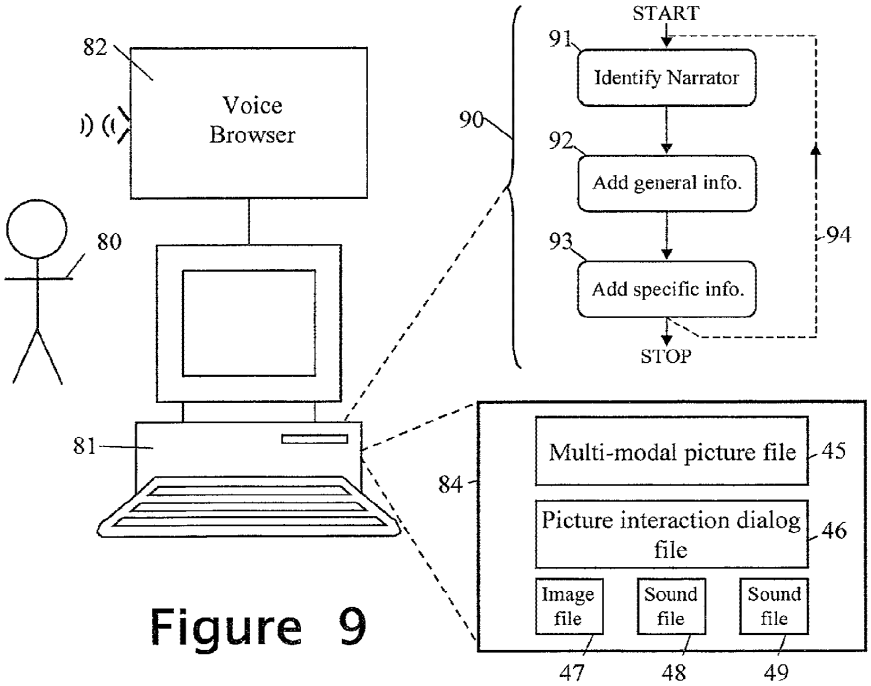


Figure 9

Amazon/VB Assets
Exhibit 1003
Page 359

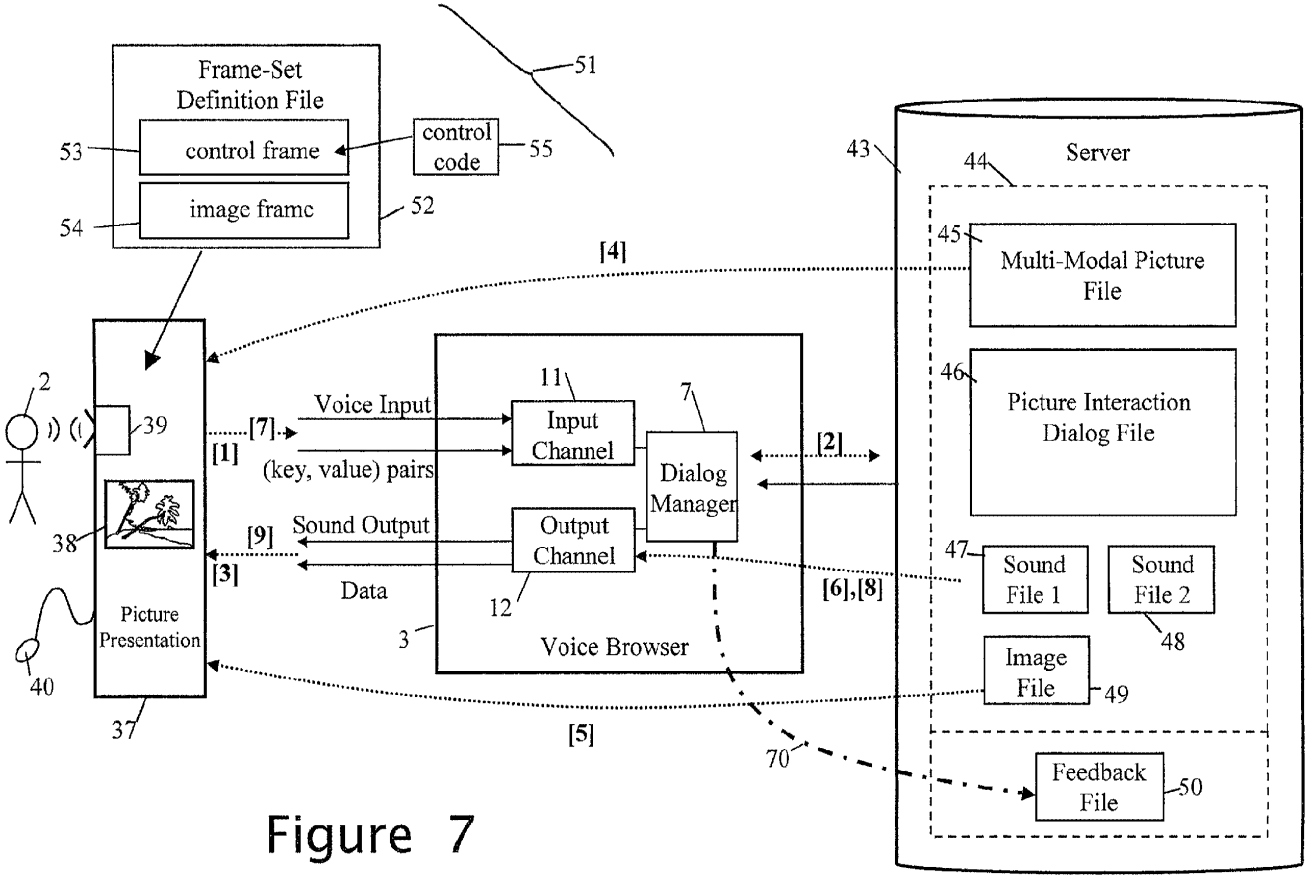
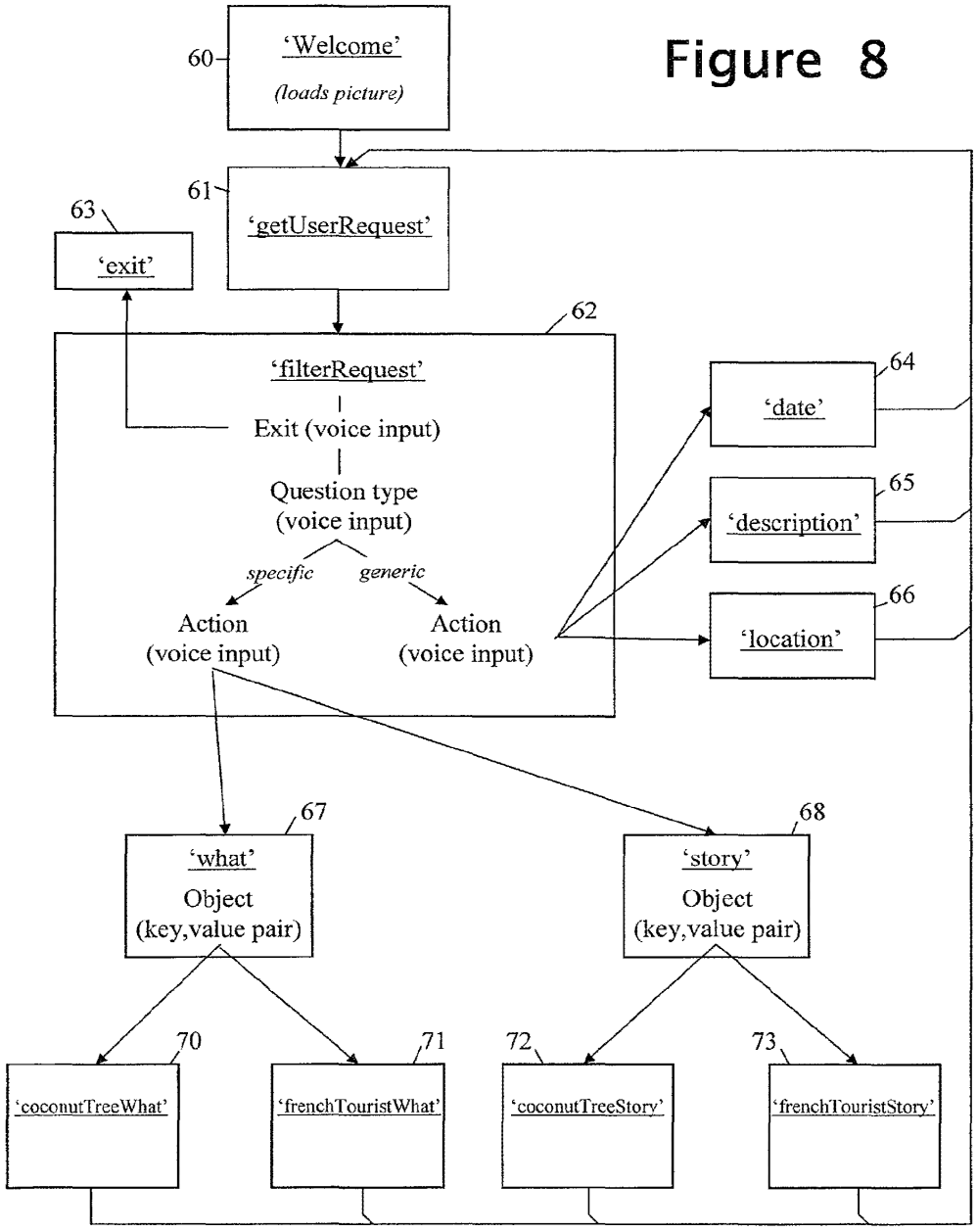


Figure 7

EP 1 320 043 A2

Amazon/VB Assets
Exhibit 1003
Page 360

Figure 8



Amazon/VB Assets
Exhibit 1003
Page 361

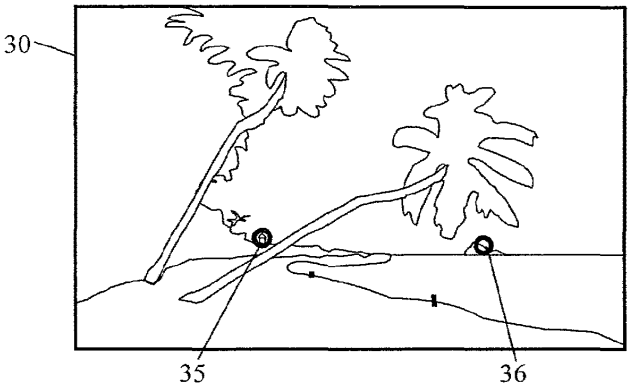


Figure 10

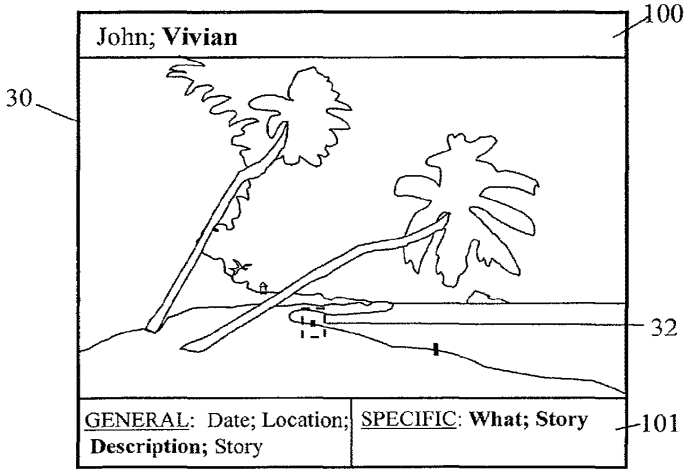


Figure 11

Amazon/VB Assets
Exhibit 1003
Page 362

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 646 037 B1

(12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention of the grant of the patent:
26.03.2008 Bulletin 2008/13

(51) Int Cl.:
G10L 15/26 *(2006.01)*

(21) Application number: **05291989.1**

(22) Date of filing: **23.09.2005**

(54) Method and apparatus for enhancing speech recognition accuracy by using geographic data to filter a set of words

Verfahren und Vorrichtung für erhöhte Spracherkennungspräzision unter Verwendung geografischer Daten zur Filterung einer Wortgruppe

Procédé et appareil pour améliorer la précision de la reconnaissance vocale en utilisant des données géographiques pour filtrer un lot de mots

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU LV MC NL PL PT RO SE SI SK TR

(30) Priority: **29.09.2004 US 953671**

(43) Date of publication of application:
12.04.2006 Bulletin 2006/15

(73) Proprietor: **FRANCE TELECOM**
75015 Paris (FR)

(72) Inventors:
- **Waters, Keith**
  West Newton
  MA 02465 (US)
- **Simon, Benoît Benjamin**
  Boston
  MA 02114 (US)

(56) References cited:
EP-A- 0 661 688
EP-A- 1 233 407
US-B1- 6 230 132

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Printed by Jouve, 75001 PARIS (FR)

EP 1 646 037 B1

Description

FIELD OF THE INVENTION

[0001] The invention relates generally to speech recognition systems and methods and, more particularly, to systems and methods for enhancing speech recognition accuracy by using geographic data to reduce the size of a set of words that must be checked to effect speech recognition.

BACKGROUND OF THE INVENTION

[0002] While visiting an unfamiliar city, a mobile handset user may wish to locate a business, landmark, street, or other place of interest. It would be quite advantageous if the user could place a voice call to a remote server at a designated telephone number, speaking his or her request for information into the handset. The handset would execute a speech recognition procedure, transforming the spoken request into a server query, and receiving the requested information from the remote server in the form of an electronic signal for producing audible speech. For example, a user in Manhattan may wish to locate the closest Neiman Marcus store. She dials a designated telephone number and speaks a request such as "where is the closest Neiman Marcus?" The handset uses speech recognition to convert the request into an electronic query for receipt by the remote server. Based upon the current location of the handset, the remote server accesses a database to retrieve the address and phone number of the nearest Neiman Marcus. The address and phone number are sent from the server to the user's handset in the form of an electronic signal, whereupon the handset converts the electronic signal into audible speech.

[0003] Unfortunately, there are significant hurdles to overcome if one wishes to use existing speech recognition algorithms to carry out the foregoing procedures. More specifically, it is a challenge to develop speech recognition algorithms, which provide a desired level of accuracy in certain operational environments, such as that of mobile telephonic devices. In general, the processing power and memory capabilities of mobile telephonic devices are limited by size and cost constraints, thereby placing limits on the extent to which the complexity of existing algorithms can be increased. Moreover, the existence of ambient noise severely degrades the performance of many speech recognition algorithms. In view of these shortcomings, prior art techniques have been developed to improve the performance of speech recognition algorithms in applications where processing power may be limited. These prior art techniques are directed toward enhancing a recognition model by utilizing improved acoustic models. The following definitions are presented herein for purposes of understanding and clarity:

[0004] *Lexical Model:* A lexical model breaks down individual words of a syntactic model (defined hereinafter) into a sequence of phonemes (also defined hereinafter).

[0005] *Acoustic Model:* An acoustic model breaks down the sequence of phonemes determined by the lexical model into a Markov model of cepstrum coefficients, also termed vector sequences. Cepstrum coefficients are generated by taking the Fourier Transform of the power spectrum of a speech signal.

[0006] *Recognition Model:* A recognition model is a set of rules and operations that are used by a recognition engine to recognize individual words. The recognition model includes a fully deployed speech recognition system which encompasses the acoustic model, the lexical model and the syntactic model as follows: recognition model = (acoustic model) x (lexical model) x (syntactic model).

[0007] U.S. Patent Application Publication No. US2003236099 discloses a technique for improving the accuracy of a speech recognition algorithm by selecting an appropriate acoustic model in accordance with the location of a mobile telephonic device ("mobile handset", or just "handset" for the sake of convenience). As indicated above, an acoustic model includes a set of cepstrum coefficient statistics for each phoneme in a language. Several different acoustic models can be precompiled, each representing a unique acoustical environment, such that a speech recognition system can select an acoustic model appropriate for a given set of environmental conditions. For example, each acoustic model is prepared by individually recording several hundred speakers in a controlled, model-specific laboratory environment that has known acoustical properties illustratively specified in terms of noise, echo, reverberation, and frequency response. Each model-specific laboratory environment is intended to represent a real-world environment, such as a warehouse, an office, a shopping mall, an airport, or the interior of an automobile. Hence, acoustic models incorporate properties of specific environments as well as characteristics of the human voice. The laboratory environment may include artificially generated sounds to simulate, for example, traffic, conversations, computer keyboard clicks, and forklift trucks.

[0008] Pursuant to this prior art technique, position information is retrieved from a location-based service (LBS), as defined hereinafter, based upon the current location of the mobile handset. The retrieved position information is then employed to select the acoustical properties of the environment surrounding the handset. An acoustic model that approximates the acoustical properties of the type of environment for the current location of the handset is selected from a set of previously prepared acoustic models. The selected acoustic model is then utilized to improve the accuracy of the speech recognition algorithm.

[0009] The technique disclosed in U.S. Patent Application Publication No. US2003236099 presents many practical shortcomings. Existing location-based services, using the Global Positioning System (GPS), will not

Amazon/VB Assets
Exhibit 1003
Page 364

determine a user's position to a requisite accuracy of, say, one to two meters, nor will such services determine a user's altitude with sufficient accuracy for many speech recognition applications. The impact of insufficient position accuracy is that, if a user is situated near the entrance of a building, it is not possible to accurately determine whether the user is, in fact, inside the building. The impact of insufficient altitude accuracy is that it is not possible to determine with certainty whether a user is within the building or standing on the roof. An acoustic model that is appropriate for an interior building lobby is unlikely to be totally appropriate for use in an outdoor setting. At the very least, selection of an inappropriate acoustic model based upon an inaccurate determination of a user's actual position will degrade the accuracy of many speech recognition algorithms.

[0010] Another shortcoming of the technique disclosed in U.S. Patent Application Publication No. US2003236099 is the need for creating an acoustic model. First of all, it is not feasible to create a unique acoustic model for every type of location. Secondly, even if an acoustic model is developed for a certain type of location, an actual location may not possess the acoustical characteristics embodied in the model. For example, due to variations in construction materials, furnishings, and room configurations, an acoustic model that attempts to represent a typical conference room, for example, may not adequately represent a specific conference room in a specific office building. Accordingly, what is needed is an improved technique for improving the accuracy of speech recognition algorithms.

[0011] EP1233407 recites a navigation system with speech recognition capacities for a moving vehicle. Whenever the distance between the vehicle and known active locations becomes greater then a given threshold, a pre-existing word list, corresponding to the active location, is acquired by the navigation system for subsequent use with the speech recognition engine.

[0012] US6230132 recites another navigation system with speech recognition functions. A lexicon is generated at periodic intervals with a limited number of words so as to facilitate the speech recognition functions. Furthermore, when a user needs to use the navigation system, he/she has to make a vocal request with the name of a location so that a lexicon can be generated.

[0013] EP0661688 recites a single device navigation apparatus, on board a car, with speed recognition functions. Pre-existing libraries are used and updated at regular intervals to facilitate the speech recognition engine.

SUMMARY OF THE INVENTION

[0014] The present invention is defined by the appended independent claims.

[0015] Other aspects and features of the present invention will become apparent from the following detailed description considered in conjunction with the accompanying drawings. It is to be understood, however, that the drawings are designed solely for purposes of illustration and not as a definition of the limits of the invention, for which reference should be made to the appended claims. It should be further understood that, unless otherwise indicated, the drawings are merely intended to conceptually illustrate the structures and procedures described herein.

BRIEF DESCRIPTION OF THE DRAWINGS

[0016] In the drawings:

[0017] FIG. 1 is a diagrammatic representation of a geographic area of interest.

[0018] FIG. 2 is a data structure diagram setting forth a portion of an illustrative element database.

[0019] FIG. 3 is a data structure diagram setting forth an illustrative subset of words selected from the element database of FIG. 2.

[0020] FIG. 4 is a generalized information flow diagram showing a set of fundamental operations for the present invention.

[0021] FIG. 5 is a detailed block diagram setting forth information flow for a speech recognition system constructed in accordance with a preferred embodiment of the present invention.

DETAILED DESCRIPTION OF THE DRAWINGS

[0022] The following terms used herein are defined as follows:

[0023] *Geographic Area of Interest:* Illustratively defined in terms of a zip code, group of zip codes, cell identification (Cell-ID) information, caller ID (Call-ID) information, the boundaries of a state, county, city, town, place, or other governmental or municipal subdivision, standard USGS (United States Geological Survey) grids, locations within a fixed range of latitude and longitude, or locations within a specified geometric boundary. The geographic area of interest may, but need not, be set automatically based upon the present location of a handset. The geographic area of interest may, but need not, be designated by a user based upon manual, verbal, or user-specified input. The geographic area of interest can be a coarse geographic area of interest or a refined geographic area of interest, with the latter being a portion of the former.

[0024] *E911:* As mandated by the FCC (Federal Communications Commission), a standard for a system by which wireless carriers are able to pinpoint the location of a mobile device in an emergency.

[0025] Element: Name of at least one of a street, a business, a merchant, an individual's name, a place, a point of interest, a transportation facility, an individual household, an activity, and a landmark.

[0026] *GPS*: Global Positioning System -- Enables geographic position determination at a land-based portable device equipped to receive signals from three or more orbiting GPS satellites.

[0027] *LBS:* Location-Based Service -- Any service

Amazon/VB Assets
Exhibit 1003
Page 365

that obtains information concerning the current geographical location of a mobile device user. Illustrative examples of location-based services include maps displayed in an automobile that track a motorists' position using GPS, or automatic localization services provided by dialing 911 on an E911-compliant mobile telephone network.

[0028] *Mobile Device:* Any portable device enabled for wireless communication, such as a wireless telephone, a personal digital assistant, or a laptop computer equipped with a wireless modem.

[0029] *Phoneme:* A phoneme represents a unique human uttered sound. Individual words, as pronounced by a human speaker, consist of a sequence of phonemes.

[0030] *Syntactic Model:* A set of rules and operations that specify allowed word sequences (phrases) of one or more words for subsequent use in a speech recognition procedure. Allowed word sequences are set forth in a list of static allowed phrases. In addition, the phrases can be optionally weighted by occurrence probabilities. Otherwise, word sequences can be in the form of stochastic or probabilistic lists, such as a Markov model of word sequences. Syntactic models are usually designed to be specific to a particular speech engine.

[0031] *Word:* A unit of language that native speakers can identify, and which communicates a meaning. Words are the building blocks from which phrases and sentences are constructed. Illustratively, for elements, "The Empire State Building" is the name of an element that contains four words. A street with the name "125th Street" includes the six words "one hundred and twenty fifth street". The street name "K Street" has the words "K" (which is a word for purposes of this invention) and "street".

[0032] The invention presumes that a very large element database is available for some sizeable, designated geographic area, such as a city, state, county, or country. One example of a known element database is TIGER (Topologically Integrated Geographic Encoding and Referencing system), available from the United States Census Bureau. TIGER compounds localized street names, rivers and railroads into a series of large regional databases.

[0033] The element database contains a set of words for elements at locations that are within the designated geographic area. In accordance with an aspect of the invention, enhanced speech recognition accuracy is provided by using geographic data to automatically select a subset of the words in the element database for use with a speech recognition procedure. The subset of words is selected by selecting only those words in the element database that are situated within a geographic area of interest. The geographic area of interest, either coarse or refined (depending on the circumstances, as explained below) is illustratively defined in terms of a zip code, group of zip codes, cell identification (Cell-ID) information, caller ID (Call-ID) information, the boundaries of a state, county, city, town, place, or other governmental or municipal subdivision, standard USGS (United States Geological Survey) grids, locations within a fixed range of latitude and longitude, or locations within a specified geometric boundary.

[0034] By selecting a subset of words, the total number of words used in the speech recognition procedure is substantially reduced. For example, a large element database containing street and business names used throughout all five boroughs of New York City can be reduced to a substantially shorter list of street and business names covering only the Grand Central Station area by specifying zip code 10017. Such a reduction in the total number of potential street names to be recognized, significantly improves speech recognition accuracy.

[0035] Pursuant to a further aspect of the invention, position information is obtained from an LBS based upon the current location of a mobile handset, illustratively using any of GPS coordinates, zip codes, and cellular base station triangulation. The obtained position information can be utilized to automatically define a coarse geographic area of interest. Let us say, as an example, that zip codes are used for this purpose. Thus, once the current location of the handset is determined, the zip code which includes that location automatically becomes the coarse geographic area of interest. Typically, the system can determine which localization method to use based on a knowledge of call configuration. For example, when a call is initiated, the Cell-ID can be used to determine the approximate location of the device. Alternatively, the application itself can choose which methods to employ, or the application can prompt the user to select which method to use. An illustrative prompt might include "*Do you want to use Cell-ID, or GPS to determine you location?*"

[0036] In accordance with a further aspect of the invention, an interface enables a user to refine the coarse geographic area of interest. For example, the user can select a specified geometrical shape within the coarse geographic area centered about the current location of the handset. This can be accomplished using key-based entries. Also, a graphical user interface can be used to draw a geometrical shape within the coarse geographic area. Additionally or alternatively, a Cell-ID can be specified within the geographic area zip code, thus defining a refined geographic area of interest for which the subset of words is to be selected. Combining automatic definition of a coarse geographic area of interest with additional, user-selectable geometric refinement is advantageous because it further reduces the total number of words that must be used in the speech recognition procedure. For example, a user-selectable rectangular region of 300 x 300 feet may contain only a few dozen elements that need to be recognized.

[0037] FIG. 1 is a diagrammatic representation of a geographic region of interest contained within a large area 100, such as a city. The current geographic location of a mobile device is determined to be point 401. In this illustrative example, the coarse geographic area of inter-

Amazon/VB Assets
Exhibit 1003
Page 366

est is defined by the zip code. Thus, the coarse geographic area of interest is defined by Zip code region 470 within which point 401 is located. A mobile device user has selected a specific radius 407 around point 401 for defining the refined geographic area of interest. Elements contained (in whole or in part) within this coarse geographic area include Street A 420, Street B 422, and a first building 406. Elements contained within this coarse geographic area of interest but which are not within this refined geographic area include Street C 424, second building 404, and third building 402.

[0038] An element database (FIG. 2, 500) includes a list of element names for at least one of streets, businesses, merchants, places, points of interest, transportation facilities, individual households, activities, or landmarks in large area 100 (FIG. 1). This database can be part of the LBS database. Each element name is associated with one or more corresponding parameters for identifying or associating the element with a refined geographic area of interest, and a parameter for identifying or associating the element with a coarse geographic area of interest. A set of phonemes corresponding to the element name can be computed in real time by recognition model compiler 213, using the outputted element name (501) in conjunction with a phonetizer. A phonetizer is a mechanism that provides phonemic transcriptions for each of a plurality of element names. Phonemic transcriptions comprise a phoneme, or a sequences of two or more phonemes, which indicate the manner in which an element name may be spoken. Illustratively, the element name "Madison" has a corresponding phonemic transcription "mäd è sun".

[0039] Element names are stored in an element name field 501. Each name in element name field 501 is associated with a corresponding parameter field (x,y) for the refined geographic area of interest 503, and a corresponding parameter field for the coarse geographic area of interest 505. These element names may or may not include elements that are within whatever refined geographic area of interest is selected by the user during operation of the invention (in this example, defined by radius 407 around point 401 of FIG. 1).

[0040] Refined geographic parameters are stored in the parameter field for the refined geographic area of interest 503 (FIG. 2). A refined geographic parameter sets forth a characteristic of an element in terms of how the refined geographic area of interest is defined. For example, if the refined geographic area is defined by the user as a shape (e.g., a circle) around the current location of the handset, the refined geographic parameters comprise (x,y) coordinates. In the illustrative example of FIG. 2, the refined geographic parameters are (x,y) coordinates determined with respect to an arbitrarily assigned reference point. When a first set and a second set of (x, y) coordinates are present in the parameter field for the refined geographic area of interest 503, this represents a line segment between the first set of (x,y) coordinates and the second set of (x,y) coordinates. Such a segment is used, for example, to define a street that extends from the first set of (x,y) coordinates to the second set of (x, y) coordinates.

[0041] Coarse geographic parameters are stored in the parameter field for the coarse geographic area of interest 505. A coarse geographic parameter sets forth a characteristic of an element in terms of how the coarse geographic area of interest is defined. For example, if the coarse geographic area of interest is defined in terms of one or more zip codes, the coarse geographic parameters in the parameter field for the coarse geographic area of interest 505 comprise one or more zip codes. For example, if Street A 420 (FIG. 1) crosses into two different zip code regions, then the parameter field for the coarse geographic area of interest 505 includes two zip codes, such as 10017 and 10036. However, other types of coarse geographic parameters could be employed to relate the element to a zip code system, such as standard USGS (United States Geological System) grids, regions within a fixed range of latitude and longitude, state boundaries, county boundaries, city boundaries, and/or boundaries of urbanized areas. For example, a lookup table could be used to determine the foregoing relationship.

[0042] Returning now to FIG. 1, and in accordance with one feature of the invention, only words corresponding to elements within the geographic area of interest are extracted from element database 500 (FIG. 2) to form word subset list 600 (FIG. 3). If the user defines the refined geographic area of interest as shown in FIG. 1, the limited set of words does not include Street C 424, second building 404, and third building 402. Use of this limited set of words enhances the accuracy of a subsequently performed speech recognition procedure because words pertaining to each and every element of large area 100 of FIG. 1 need not be considered.

[0043] FIG. 3 is a data structure diagram setting forth a portion of an illustrative word subset list 600. The names of one or more elements within the geographic area of interest are stored in an element name field 601. Optionally, word subset list 600 may include additional information, such as the probability that a particular element will be incorporated into a user's query based upon the distance of that element from the current position of the mobile device, so as to provide an enhanced level of performance.

[0044] FIG. 4 is a generalized information flow diagram describing a set of fundamental operations for the present invention. Geographic data acquisition mechanism 101 (the details of which are shown in FIG. 5) is equipped to collect geographic data corresponding to a mobile device's current geographic position and, based thereon, to determine and output a geographic area of interest. Geographic data includes, but is not limited to, any of: global positioning system (GPS) coordinates in latitude and longitude, and also altitude if needed, [x,y, z] coordinates, city name, place name, zip code area, geographic area, type of area (urban, suburban or coun-

Amazon/VB Assets
Exhibit 1003
Page 367

try), one or more cell identification (Cell-ID) parameters from a cellular network, or a caller identification (Call-ID) parameter from a terrestrial telephone network.

[0045] Word filtering mechanism 102 is equipped to receive a geographic area of interest from geographic data acquisition mechanism 101 and, based thereupon, select a subset of words from element database 500 (FIGs. 2 and 4). If a coarse geographic area of interest is being used, then the subset of words is selected by comparing coarse geographic parameters in the parameter field for the coarse geographic area of interest 505 corresponding to the elements in element name field 501 with coarse geographic parameters used to characterize the geographic area of interest. More specifically, if the coarse geographic area of interest is zip code 10017, then all the element names in element name field 501 will be selected which are associated with zip code 10017 in the parameter field for the coarse geographic area of interest 505. A similar operation is performed if a refined geographic area of interest is used. Of course, the parameter field for the refined geographic area of interest 503 will be used instead of the parameter field for the coarse geographic area of interest 505. The refined geographic parameters are combined with the coarse geographic parameters to enhance the word filtering mechanism 102 by using a two-step filtering process. A first step uses the parameter field for the coarse geographic area of interest 505, and a second step uses the parameter field for the refined geographic area of interest 503.

[0046] Word filtering mechanism 102 places the selected subset of words into a word subset list 600 (FIGs. 3 and 4). Speech recognition mechanism 103 (FIG. 4) is equipped to receive word subset list 600 (FIGS. 3 and 4) from word filtering mechanism 102 (FIG. 4) and, based thereupon, perform a speech recognition procedure. Illustratively, speech recognition mechanism 103 may be equipped to implement a speech recognition procedure by utilizing a (syntactic model) x (lexical model) x (acoustic model) paradigm in conjunction with Markov models. Alternatively or additionally, speech recognition mechanism 103 may utilize a neural network. More generally, the invention is applicable to any speech recognition mechanism that accepts and processes words or groups of words referred to as "grammar". Accordingly, word filtering mechanism 102 includes a grammar formatting mechanism that receives the word subset list 600 and formats the list into a grammar format usable by recognition model compiler 213.

[0047] Recognition model compiler 213 may or may not be required, depending upon the specific design characteristics of speech recognition mechanism 103. Recognition model compiler 213 parses (or analyzes) all of the entries in a formatted grammar, one after the other, and then, in one or more successive stages or "passes", constructs an output code readable by speech recognition mechanism 103. Since the speech recognition procedure is based upon a geographically refined subset of words, the overall accuracy of this procedure is enhanced relative to prior art designs. Speech recognition mechanism 103 can be implemented by a server or by a mobile handset.

[0048] A speech query from a mobile device user is received by an acoustic input mechanism 104 at the mobile device. Acoustic input mechanism 104 may be implemented using a microphone or other transducer equipped to transform acoustical vibrations into electronic signals. Acoustic input mechanism 104 feeds an electronic representation of user speech to speech recognition mechanism 103. Speech recognition mechanism 103 attempts to recognize one or more words in the electronic representation of user speech that are included in word subset list 600 retrieved from element database 500. If one or more words are recognized, speech recognition mechanism 103 sends an electronic message to an output mechanism 105 at the mobile device. The electronic message includes any recognized words in the speech query received from the mobile device user that were also included in word subset list 600. Output mechanism 105 converts the electronic message into audible and/or visual form for perception by the mobile device user. Illustratively, output mechanism 105 may be implemented using a loudspeaker coupled to audio amplifier circuitry and voice synthesizer circuitry for producing an audio message including one or more recognized words. Alternatively or additionally, output mechanism 105 may be implemented using a visual display device for displaying alphanumeric characters including one or more recognized words.

[0049] FIG. 5 is a detailed block diagram setting forth information flow for a speech recognition system constructed in accordance with an embodiment of the present invention. FIG. 5 shows details of geographic data acquisition mechanism 101. The depicted operations are performed either by a remote server 240 or by mobile device 250. Information flow commences upon receipt of a "start" signal at input mechanism 201 of mobile device 250. In practice, this "start" signal can be a request issued from a multimodal web page, an HTML page, or just a phone call. Alternatively, a user may place a call on a mobile handset to a designated telephone number to access a zip code form displayed on the mobile handset. The user fills in the zip code form using DTMF keypresses on the handset and validates the entered zip code by pressing DTMF digit "1" as a "start" signal. Basically, the "start" signal triggers a request for information.

[0050] Next, at mechanism 202, a decision is made as to whether the geographic area of interest will be determined locally (i.e., at mobile device 250), or at a remote location (i.e., at remote server 240). Depending upon the specifics of a given system application, this decision can be made by issuing a query to mobile device 250 in order to ascertain the configuration and capabilities of mobile device 250.

[0051] In cases where the geographic area of interest is determined locally, any one of two approaches may be employed. Pursuant to a first approach, mobile device

Amazon/VB Assets
Exhibit 1003
Page 368

250 may include an optional GPS receiver 207, whereupon acquired GPS data are forwarded to a geographic-area-of-interest defining mechanism 204 (server side, see FIG. 5). Mechanism 204 uses the GPS data corresponding to the current location of mobile device 250 to obtain coarse geographic parameters for characterizing the coarse geographic area of interest. According to a second approach, zip code data, for example, can be entered by the user at input mechanism 201 and forwarded to geographic-area-of-interest defining mechanism 204. Illustratively, zip code data are entered into input mechanism 201 by sequentially pressing keys on a DTMF (dual-tone, multifrequency) keypad. Mechanism 204 uses the entered zip code data as a coarse geographic parameter for characterizing the geographic area of interest.

[0052] In cases where the geographic area of interest is determined remotely, mechanism 202 forwards the request received from input mechanism 201 to a location-based service 203 at server 240. Alternatively or additionally, GPS data from GPS receiver 207 are conveyed to location-based service 203, if needed. In response to the received request, location-based service 203 retrieves, for example, zip code information or cell identification (Cell ID) data for mobile device 250. Cell-IDs data can be retrieved in a mobile communications network using the Automatic Number Identification (ANI) retrieved from the mobile device 250. ANI is a service that provides the receiver of a telephone call with the number of the calling phone. The method of providing this information is determined by the service provider. Cell ID data specifies the geographic location of the wireless base station currently in communication with mobile device 250 by identifying at least one of the telephone number and the Internet Protocol (IP) address assigned to mobile device 250. For example, mobile device 250 may send to location-based service 203 Cell ID data in the form of an automated number identification (ANI) parameter. Optionally or alternatively, location-based service 203 determines the geographic location of mobile device 250 by performing a triangulation procedure among a plurality of wireless base stations. The Cell ID data, Caller ID data, and/or triangulation data retrieved by location-based service 203 (ultimately through the telecommunication provider) are forwarded to geographic-area-of-interest defining mechanism 204.

[0053] Optionally, in cases where the geographic area of interest is determined remotely, LBS service 203 utilizes at least one of Caller ID data and Cell ID data to access a zip code database 243. Based upon the Caller ID data or the Cell ID data, a zip code is retrieved from zip code database 243 and then sent to geographic-area-of-interest defining mechanism 204. Although the illustrative embodiment of FIG. 5 shows zip code database 243 and location-based service 203 as separate elements, those skilled in the relevant art will easily recognize that zip code database 243 could, but need not, be incorporated into location-based service 203. Likewise, although element database 500 and location-based service 203 are shown as separate elements, those skilled in the relevant art will easily recognize that element database 500 could, but need not, be incorporated into location-based service 203.

[0054] Geographic-area-of-interest defining mechanism 204 processes incoming geographic data, illustratively comprising at least one of incoming GPS data, zip code data, Caller ID data, or Cell ID data, to generate one or more geographic parameters defining the geographic area of interest for mobile device 250. For example, as discussed above, if the incoming data are based upon zip code data, then the output of mechanism 204 is a zip code. This parameter, e.g., a zip code, is used to specify a geographic area of interest. The defined geographic area of interest may combine a coarse geographic area of interest with a refined geographic area of interest, as is illustrated in the following example. One or more zipcode(s) may be deduced from GPS coordinates, thus defining a coarse geographic area of interest. The GPS coordinates themselves, associated with an area of uncertainty, define a refined geographic area of interest. Such a defined geographic area of interest may be used to implement the two-step filtering process described above.

[0055] Optionally, as indicated by the dotted lines of FIG. 5, a user at mobile device 250 may wish to further refine the geographic area of interest defined by geographic-area-of-interest defining mechanism 204. If so, the geographic area of interest defined by geographic-area-of-interest defining mechanism 204 is referred to as a coarse geographic area of interest, and the area of interest as refined by the user is referred to as a refined geographic area of interest. The refined geographic area of interest comprises a subset of the coarse geographic area of interest. Input mechanism 201 accepts user input setting forth a geometric specification of a refined geographic area of interest. User input is illustratively received over connection line 275 in the form of keypad entries, graphical objects, sketches, textual descriptions, or vocal input. Examples of geometric specifications include squares, rectangles, octagons, triangles, circles, and freehand sketches. Geometry specification refinement mechanism 205 accepts user input from input mechanism 201, and refines the set of geographic parameters generated by geographic-area-of-interest defining mechanism 204.

[0056] The one or more geographic parameters generated by geographic-area-of-interest defining mechanism 204, as optionally refined by geometry specification refinement mechanism 205, are received by word filtering mechanism 102. Word filtering mechanism 102 is enabled to carry out its functions when two conditions are met. Firstly, mobile device 250 should be executing a speech-enabled application where the user has issued a request for information. This requirement is met upon generation of a request from input mechanism 201. Secondly, the geographic area of interest should be defined.

Amazon/VB Assets
Exhibit 1003
Page 369

This requirement is met once geographic-area-of-interest defining mechanism 204 generates the above-described geographic parameters.

[0057] When enabled, word filtering mechanism 102 uses the one or more geographic parameters received from mechanism 204 or 205 to determine whether or not each of a plurality of elements in element database 500 is situated within the geographic area of interest. The element names stored in element name field 501 (FIG. 2) may include elements that are within the geographic area of interest, as well as elements that do not lie within the geographic area of interest. Accordingly, word filtering mechanism 102 (FIG. 5) compares the one or more received parameters defining the geographic area of interest against one or more parameters stored in at least one of parameter field for the refined geographic area of interest 503 or parameter field for the coarse geographic area of interest 505 (FIG. 2), to determine whether or not each of a plurality of elements in element database 500 (FIGs. 2 and 5) is situated within the geographic area of interest. For example, Street A is characterized in the parameter field for the refined geographic area of interest 503 by a first and a second set of (x,y) coordinates. Accordingly, Street A is conceptualized as including a straight line segment connecting the first and second sets of (x,y) coordinates. If any portion of this straight line segment intersects the geographic area of interest, then Street A is included in word subset list 600 (FIGs. 3 and 5).

[0058] Word filtering mechanism 102 appends to word subset list 600 only those element names within the geographic area of interest. Optionally, if the user has made any refinements to the coarse geographic area of interest, only elements that are within the refined geographic area of interest are appended to word subset list 600.

[0059] The word subset list 600 of FIGs. 3 and 5 represents an illustrative output format for word filtering mechanism 102. Alternatively or additionally, the output of word filtering mechanism 102 may be conceptualized as a computer-readable stream of words representing one or more elements in a geographic area of interest. This stream, illustratively in text format, is compatible with at least one of a speech recognition engine or a voice browser embedding speech recognition mechanism 103 (FIG. 5). Optionally, word filtering mechanism 102 may reformat at least one of word subset list 600 or the computer-readable stream of words into a format such as an XML document, to provide a subset of words compatible with speech recognition mechanism 103.

[0060] Speech recognition mechanism 103 may be implemented at server 240 or at mobile device 250. In cases where speech recognition is performed at mobile device 250, the output from word filtering mechanism 102 (illustratively in the form of a stream of words corresponding to grammar subset list 600, FIG. 3) is sent from server 240 to mobile device 250. Word subset list 600 is then compiled by the recognition model compiler 213 at mobile device 250 to provide a speech recognition model for use by a speech recognition mechanism 103 responsive to the acoustic input mechanism 104. Speech recognition mechanism 103 then forwards recognized speech to the output mechanism 105 for generating an electronic output signal representing the recognized speech. Such an arrangement is advantageous because it reduces the burden on wireless communication resources. Transmissions between the server and the mobile device are reduced because only word subset list 600 needs to be communicated to the mobile device. The speech recognition steps can be performed at the mobile device with reference to word subset list 600, thereby enhancing the speed and efficiency of the speech recognition process.

[0061] If speech recognition is performed at server 240, then recognition model compiler 213 and speech recognition mechanism 103 are implemented at server 240 instead of at mobile device 250, wherein the output of speech recognition mechanism 103 is sent from server 240 to output mechanism 105 at mobile device 250.

[0062] The operational principles of the configurations shown in FIGs. 4 and 5 are best illustrated with reference to several illustrative scenarios. These scenarios clarify the location resolution process that will govern the filtering process performed by word filtering mechanism 102 in accordance with location data to be resolved. The scenarios consider use of any of several different types of inputs by which a location context will be determined.

[0063] Scenario One: Geographic area of interest to be determined by entered zip code and/or city. Zip code and/or city are entered into mobile device 250 by the user. Possible clients used by mobile device 250 to accept the entered zip code and/or city include a voice browser and/or a multimodal client. In the case of a phone call, the client is not the phone itself, but the voice browser (from the server point of view). Location data are provided in the form of a city and/or a zip code. The geographic area of interest corresponding to this location data is determined in the form of a city or zip code. In this case, location resolution is performed locally. The zip code and/or city are entered into mobile device 250 using DTMF keypresses and/or voice. An HTTP request is issued in response to the entered information after the information is received at input mechanism 201. In this case, word subset list 600 is generated by selecting a set of words and/or grammar pertaining to the entered city and/or zip code. The following steps are performed:

[0064] 1. User connects to geographic data acquisition mechanism 101 using mobile device 250. If mobile device 250 represents a multimodal client, an HTTP request is made by the client to a document server of word filtering mechanism 102 (FIGs. 4 and 5). If mobile device 250 represents a wireless telephone, an http request is made by the voice browser ("speaking" with the telephone) to the document server of word filtering mechanism 102.

[0065] 2. A first page is loaded into mobile device 250 or loaded onto the voice browser. The page is prompting the user to enter a zip code. The zip code by either be entered by DTMF (touch tones) or by voice. The grammar used to recognize the zip code is static.

Amazon/VB Assets
Exhibit 1003
Page 370

[0066] 3. The user enters the zip code into input mechanism 201. A second HTTP request is made by the client, with the zip code value as a parameter. As location-based service 203 is not employed in this scenario to resolve the location, location resolution is "local". The geographic area of interest is evaluated at geographic area of interest defining mechanism 204, and no further refinement of the geographical area of interest takes place.

[0067] 4. A second page is loaded into mobile device 250 (the "reduced" page). The document server of word filtering mechanism 102 computes reduced grammars associated with the page in the form of word subset list 600.

[0068] 5. The user can vocally interact with the page over mobile device 250, as the page is optimized for its location

[0069] Note: The geometry specification refinement mechanism 205 is not used in this scenario. If this mechanism were to be used, a "remote" location resolution must be performed so as to enable sending a map (for example) back to the multimodal client of mobile device 250.

[0070] Scenario Two: Geographic area of interest to be determined automatically using Caller ID or Cell ID.

[0071] Possible clients used by mobile device 250 for Caller Id or Cell ID include a voice browser and/or a multimodal client. Location data are provided in the form of a caller ID or cell ID parameter. The geographic area of interest is determined as a zip code or as an area ID that identifies a geographic area. Location resolution is performed remotely, not locally. Input mechanism 201 is not employed for purposes of location resolution. Mobile device 250 is programmed to issue an HTTP request while a first page is retrieved from a document server of word filtering mechanism 102. In this scenario, words and/or grammar from element database 500 are filtered to prepare word subset list 600 based upon ANI or Cell-ID information. Input mechanism 201 is completely transparent, in the sense that the request that trims the word filtering process appears at the first stage of the interaction, and there is no need of a data collecting stage, as in the previous scenario. The HTTP request can be issued by a voice browser (that provides ANI info) or a multimodal client at mobile device 250.

[0072] The following operational sequence is performed:

[0073] 1. User connects to geographic data acquisition mechanism 101 using mobile device 250. If the user is using a multimodal client at mobile device 250, an HTTP request is made by the client to the document server of word filtering mechanism 102. If mobile device 250 is a wireless telephone, an http request is made by the voice browser ("speaking" with the telephone) to the document server of word filtering mechanism 102. The ANI and/or the Cell ID are passed to the document server as part of the HTTP request. Location-based service 203 determines the geographic area of interest, in this case following the remote path indicated in FIG. 5 which commences at mechanism 202. Location-based service also determines geographic position of the user, an area of uncertainty with respect to the user's position, and one or more associated zip codes.

[0074] 2. A page is downloaded to mobile device 250 or loaded onto the voice browser. The geographic context is used by word filtering mechanism 102 to compute a word subset list 600 associated with the page in accordance with the previously determined geographic area of interest.

[0075] 3. The user can vocally interact with the page using mobile device 250. The page is optimized for the user's geographic area of interest.

[0076] Scenario Three: Geographic area of interest to be determined automatically using GPS data. Possible clients used by mobile device 250 for GPS include a multimodal client. Location data are provided in the form of GPS coordinates. The geographic area of interest is determined as a zip code, as (x,y) or (x,y,z) coordinates, optionally incorporating an area of uncertainty. Location resolution is performed locally, not remotely. Input mechanism 201 is not employed to determine geographic area of interest. An HTTP request is issued by mobile device 250 during the first page fetch. In this case, word filtering mechanism 102 is preparing a word subset list 600 by filtering element database 500 using GPS data. Input mechanism 201 is transparent, in the sense that the HTTP request that results in word filtering appears at the first stage of the interaction. Unlike the first scenario, there is no need for a data collecting stage. The following steps are performed:

[0077] 1. User connects to geographic data acquisition mechanism 101 using mobile device 250. If mobile device 250 utilizes a multimodal client, an HTTP request is made by the client to a document server of word filtering mechanism 102. The GPS information is sent as part of the http request. Location-based service 203 determines the geographic area of interest by adding coarse localization to the GPS information, and the coarse localization is sent to geographic area of interest defining mechanism 204.

[0078] 2. A page is downloaded to mobile device 250: The geographic area of interest in the form of the coarse and/or refined localization received at geographic area of interest defining mechanism 204 is used by word filtering mechanism 102 to compute a word subset list 600 and to place the word subset list into a page that is sent to mobile device 250, using the two-step filtering process described above.

[0079] 3. The user can vocally interact with the page, optimized for the geographic area of interest. Note:

[0080] 1. In cases where geometry specification refinement mechanism 205 is employed, an intermediate page is downloaded to mobile device 250, asking to the user to precise its area of interest (before step 2):

[0081] 1.a - An intermediate (i.e., "localization") page is downloaded to mobile device 250.

[0082] 1.b - The user inputs an area of interest into

Amazon/VB Assets
Exhibit 1003
Page 371

mobile device 250.

[0083] 1.c - A refined geographic area of interest is sent by mobile device 250 and incorporated into the geographic area of interest defined by geographic area of interest defining mechanism 204.

[0084] In contrast to prior art techniques, the present invention does not employ acoustic models to improve speech recognition accuracy. Instead, the invention reduces the total number of words that need to be considered by a speech recognition algorithm. The number of words is reduced by excluding words that do not relate to a geographic area of interest. In this manner, the reduced processing burden makes it possible to handle speech recognition at the handset. Moreover, the accuracy of the speech recognition algorithm is improved.

[0085] To summarize, the advantages of the invention include more accurate speech recognition due to utilization of a database that has been reduced in size (i.e., word subset list 600). The burden on communication resources is decreased by sending the reduced database to the handset, whereupon the handset performs speech recognition using only the reduced size database instead of a much larger database. The decreased burden on communication resources results in faster, more efficient speech recognition.

[0086] Thus, while there have been shown and described novel features of the invention as applied to embodiments thereof, it will be understood that various omissions and substitutions and changes in the form and details of the devices illustrated, and in their operation, may be made by those skilled in the art without departing from the scope of the invention, which is solely defined by the appended claims. For example, although an embodiment of the invention was described in terms of a user refining an automatically determined geographic area of interest, it is also possible for a user to augment an automatically determined coarse geographic area of interest.

[0087] In view of the foregoing, it is expressly intended that all combinations of those elements and/or method steps which perform substantially the same function in substantially the same way to achieve the same results may be incorporated in any other disclosed or described or suggested form or embodiment as a general matter of design choice in as much as they fall within the scope of the claims appended hereto.

Claims

1. A method for use with a mobile handset equipped to communicate with a remote server over a wireless communications network, the remote server equipped to access an element database (500) including words that describe elements at each of a plurality of geographically defined locations, the method comprising the steps of:

   determining (202) geographic data for the mobile handset following a request; and
   using (102) the determined geographic data to automatically select by the remote server a subset (600) of words for use with a speech recognition procedure;

   wherein the subset of words is selected from the element database, thereby reducing the number of words used in the speech recognition procedure.

2. The method of claim 1 wherein the determined geographic data includes at least one of global positioning system (GPS) position data, cell identity (Cell-ID) data, caller identification (Caller ID) data, place name data, or zip code data.

3. The method of claim 1 wherein the geographic data are derived from a signal received from the mobile handset, said signal being related to a current position of said mobile handset or to a manual input entered into said mobile handset.

4. The method of any preceding claim wherein the elements include at least one of street names, businesses, merchants, points of interest, transportation facilities, individual households, activities, or landmarks.

5. The method of any preceding claim wherein the mobile handset receives the subset of words from the remote server and executes the speech recognition procedure based upon the received subset of words, thereby reducing use of the communications network and the remote server.

6. A speech recognition system comprising:

   a mobile handset equipped to communicate with a remote server over a wireless network,
   a data acquisition mechanism (101) for determining geographic data for the mobile handset following a request input; and
   a selection mechanism (102) for using the determined geographic data to automatically select a subset of words (106) for use with a speech recognition procedure;

   wherein the subset of words is selected by the remote server from an element database (500) said remote server has access to, said element database including words that describe elements at each of a plurality of geographically defined locations, thereby reducing the number of words used in the speech recognition procedure.

7. The speech recognition system of claim 6 wherein the geographic data are derived from a signal re-

Amazon/VB Assets
Exhibit 1003
Page 372

ceived from the mobile handset, said signal being related to a current position of said mobile handset or to a manual input entered into said mobile handset.

8. The speech recognition system of claim 6 wherein the determined geographic data includes at least one of global positioning system (GPS) position data, cell identity (Cell-ID) data, caller identification (Caller ID) data, place name data, or zip code data.

9. The speech recognition system of any of claims 6 to 8 wherein the elements include at least one of street names, businesses, merchants, points of interest, transportation facilities, individual households, activities, or landmarks.

10. A speech recognition system for associating received speech with words retrieved from an element database, the system comprising:

means for determining a geographic area of interest following a request wherein speech is to be received, said means further comprising:

(i) a portable location determining mechanism associated with a mobile handset for generating an indication signal indicative of current geographic location, and
(ii) a server, in communication with the portable location determining mechanism, programmed to determine a geographic area of interest from the indication signal;

means for selecting a subset of words from the element database based upon the geographic area of interest; said means comprising the server programmed to extract from the element database only words that are associated with the geographic area of interest; and,
means for associating the received speech with words selected from the subset of words, said means being implemented on the mobile handset.

11. The speech recognition system of claim 10 further comprising:

means for accepting a signal from at least one graphical user interface for selecting a portion of the determined geographic area of interest to thereby specify a further limited geographic area of interest; and
means for further reducing the subset of words according to the further limited geographic area of interest to generate a further subset of words, such that the means for associating received speech with words only selects words from the further subset of words.

12. The speech recognition system of claim 11 wherein the means for accepting a signal includes:

an electronic display for displaying a map of the determined geographic area of interest; and
a processing mechanism for combining the accepted signal with the map of the determined geographic area of interest so as to cause a display of the further limited geographic area of interest on the electronic display.

13. A mobile handset (250) comprising speech recognition means for association of received speech with words received from a remote database, the mobile handset comprising:

means for receiving a request;
means for providing a signal representing a geographic area of interest following said request;
means for transmitting said signal to a remote server and for receiving from the remote server a subset of words from the remote database based upon the geographic area of interest; and
means for associating the received speech with words selected from the subset of words.

14. The mobile handset of claim 13 wherein the means for providing a signal comprises means for determining a geographic area of interest wherein speech is to be received.

15. The mobile handset of claim 13 or 14 further comprising:

means (201) for accepting at least one input used to determine a refined geographic area of interest in the geographic area of interest;
means (205) for reducing the subset of words according to the refined geographic area of interest;

and wherein:

the means for associating the received speech is arranged for association of speech only with the reduced subset of words.

Patentansprüche

1. Verfahren zur Verwendung mit einem mobilen Handapparat, der dafür ausgestattet ist, über ein drahtloses Kommunikationsnetz mit einem abgesetzten Server zu kommunizieren, wobei der abgesetzte Server dafür ausgestattet ist, auf eine Elementdatenbank (500) zuzugreifen, die Wörter enthält, die

Amazon/VB Assets
Exhibit 1003
Page 373

Elemente an jedem von mehreren geographisch definierten Orten beschreiben, wobei das Verfahren die folgenden Schritte umfaßt:

Bestimmen (202) geographischer Daten für den mobilen Handapparat nach einer Anforderung und Verwenden (102) der bestimmten geographischen Daten zum automatischen Auswählen einer Teilmenge (600) von Wörtern zur Verwendung mit einer Spracherkennungsprozedur durch den abgesetzten Server;

wobei die Teilmenge von Wörtern aus der Elementdatenbank ausgewählt wird, wodurch die Anzahl der in der Spracherkennungsprozedur verwendeten Wörter reduziert wird.

2. Verfahren nach Anspruch 1, wobei die bestimmten geographischen Daten Positionsdaten des globalen Navigationssystems (GPS) und/oder Daten der Zellenidentität (Cell-ID) und/oder Daten der Anruferidentifikation (Caller ID) und/oder Städtenamendaten und/oder Postleitzahldaten umfassen.

3. Verfahren nach Anspruch 1, wobei die geographischen Daten aus einem aus dem mobilen Handapparat empfangenen Signal abgeleitet werden, wobei das Signal eine aktuelle Position des mobilen Handapparats oder eine in den mobilen Handapparat eingegebene manuelle Eingabe betrifft.

4. Verfahren nach einem der vorhergehenden Ansprüche, wobei die Elemente Straßennamen und/oder Unternehmen und/oder Händler und/oder interessierende Punkte und/oder Transporteinrichtungen und/oder individuelle Haushalte und/oder Aktivitäten und/oder Sehenswürdigkeiten umfassen.

5. Verfahren nach einem der vorhergehenden Ansprüche, wobei der mobile Handapparat die Teilmenge von Wörtern von dem abgesetzten Server empfängt und auf der Basis der empfangenen Teilmenge von Wörtern die Spracherkennungsprozedur ausführt und **dadurch** die Benutzung des Kommunikationsnetzes und des abgesetzten Servers reduziert.

6. Spracherkennungssystem, umfassend:

einen mobilen Handapparat, der dafür ausgestattet ist, über ein drahtloses Netz mit einem abgesetzten Server zu kommunizieren,
einen Datenbeschaffungsmechanismus (101) zum Bestimmen von geographischen Daten für den mobilen Handapparat nach einer Anforderungseingabe und
einen Auswahlmechanismus (102) zum Verwenden der bestimmten geographischen Daten zum automatischen Auswählen einer Teilmenge von Wörtern (106) zur Verwendung mit einer Spracherkennungsprozedur;

wobei die Teilmenge von Wörtern durch den abgesetzten Server aus einer Elementdatenbank (500) ausgewählt wird, zu der der abgesetzte Server Zugang hat, wobei die Elementdatenbank Wörter enthält, die Elemente an jedem von mehreren geographisch definierten Orten beschreiben, wodurch die Anzahl der bei der Spracherkennungsprozedur verwendeten Wörter reduziert wird.

7. Spracherkennungssystem nach Anspruch 6, wobei die geographischen Daten aus einem aus dem mobilen Handapparat empfangenen Signal abgeleitet werden, wobei das Signal eine aktuelle Position des mobilen Handapparats oder eine in den mobilen Handapparat eingegebene manuelle Eingabe betrifft.

8. Spracherkennungssystem nach Anspruch 6, wobei die bestimmten geographischen Daten Positionsdaten des globalen Navigationssystems (GPS) und/oder Daten der Zellenidentität (Cell-ID) und/oder Daten der Anruferidentifikation (Caller ID) und/oder Städtenamendaten und/oder Postleitzahldaten umfassen.

9. Spracherkennungssystem nach einem der Ansprüche 6 bis 8, wobei die Elemente Straßennamen und/oder Unternehmen und/oder Händler und/oder interessierende Punkte und/oder Transporteinrichtungen und/oder individuelle Haushalte und/oder Aktivitäten und/oder Sehenswürdigkeiten umfassen.

10. Spracherkennungssystem zum Assoziieren von empfangener Sprache mit aus einer Elementdatenbank abgerufenen Wörtern, wobei das System folgendes umfaßt:

Mittel zum Bestimmen eines interessierenden geographischen Gebiets nach einer Aufforderung, wobei Sprache zu empfangen ist, wobei die Mittel ferner folgendes umfassen:

i) einen mit einem mobilen Handapparat assoziierten tragbaren Ortsbestimmungsmechanismus zum Erzeugen eines Anzeigesignals, das einen aktuellen geographischen Ort anzeigt und
ii) einen mit dem tragbaren Ortsbestimmungsmechanismus kommunizierenden Server, der dafür programmiert ist, aus dem Anzeigesignal ein interessierendes geographisches Gebiet zu bestimmen;

Mittel zum Auswählen einer Teilmenge von Wörtern aus der Elementdatenbank auf der Ba-

Amazon/VB Assets
Exhibit 1003
Page 374

sis des interessierenden geographischen Gebiets; wobei die Mittel den Server umfassen, der dafür programmiert ist, aus der Elementdatenbank nur Wörter zu extrahieren, die mit dem interessierenden geographischen Gebiet assoziiert sind und
Mittel zum Assoziieren der empfangenen Sprache mit aus der Teilmenge von Wörtern ausgewählten Wörtern, wobei die Mittel auf dem mobilen Handapparat implementiert sind.

11. Spracherkennungssystem nach Anspruch 10, ferner umfassend:

Mittel zum Annehmen eines Signals von mindestens einer graphischen Benutzeroberfläche zum Auswählen eines Teils des bestimmten interessierenden geographischen Gebiets, um **dadurch** ein weiter eingeschränktes interessierendes geographisches Gebiet zu spezifizieren und
Mittel zum weiteren Reduzieren der Teilmenge von Wörtern gemäß dem weiter eingeschränkten interessierenden geographischen Gebiet, um eine weitere Teilmenge von Wörtern zu erzeugen, dergestalt, daß die Mittel zum Assoziieren empfangener Sprache mit Wörtern nur Wörter aus der weiteren Teilmenge von Wörtern auswählen.

12. Spracherkennungssystem nach Anspruch 11, wobei die Mittel zum Annehmen eines Signals folgendes umfassen:

ein elektronisches Display zum Anzeigen einer Karte des bestimmten interessierenden geographischen Gebiets und
einen Verarbeitungsmechanismus zum Kombinieren des angenommenen Signals mit der Karte des bestimmten interessierenden geographischen Gebiets, um so eine Anzeige des weiter eingeschränkten interessierenden geographischen Gebiets auf dem elektronischen Display zu bewirken.

13. Mobiler Handapparat (150) mit Spracherkennungsmitteln zum Assoziieren empfangener Sprache mit aus einer abgesetzten Datenbank empfangenen Wörtern, wobei der mobile Handapparat folgendes umfaßt:

Mittel zum Empfangen einer Anforderung;
Mittel zum Bereitstellen eines Signals, das ein interessierendes geographisches Gebiet repräsentiert, nach der Anforderung;
Mittel zum Senden des Signals zu einem abgesetzten Server und zum Empfangen einer Teilmenge von Wörtern aus der abgesetzten Datenbank von dem abgesetzten Server auf der Basis des interessierenden geographischen Gebiets und
Mittel zum Assoziieren der empfangenen Sprache mit aus der Teilmenge von Wörtern ausgewählten Wörtern.

14. Mobiler Handapparat nach Anspruch 13, wobei die Mittel zum Bereitstellen eines Signals Mittel zum Bestimmen eines interessierenden geographischen Gebiets umfaßt, worin Sprache empfangen werden soll.

15. Mobiler Handapparat nach Anspruch 13 oder 14, ferner umfassen:

Mittel (210) zum Annehmen mindestens einer Eingabe, mit der ein verfeinertes interessierendes geographisches Gebiet in dem interessierenden geographischen Gebiet bestimmt wird;
Mittel (205) zum Reduzieren der Teilmenge von Wörtern gemäß dem verfeinerten interessierenden geographischen Gebiet und wobei

die Mittel zum Assoziieren der empfangenen Sprache dafür ausgelegt sind, Sprache nur mit der reduzierten Teilmenge von Wörtern zu assoziieren.

Revendications

1. Procédé pour une utilisation avec un combiné mobile équipé pour communiquer avec un serveur distant sur un réseau de communication sans fil, le serveur distant étant équipé pour accéder à une base de données d'éléments (500) comprenant des mots qui décrivent des éléments pour chacun d'une pluralité d'emplacements définis géographiquement, le procédé comprenant les étapes de :

détermination (202) de données géographiques pour le combiné mobile suite à une demande ; et utilisation (102) des données géographiques déterminées pour sélectionner automatiquement par le serveur distant un sous-ensemble (600) de mots pour une utilisation avec une procédure de reconnaissance vocale ;

dans lequel le sous-ensemble de mots est sélectionné parmi la base de données d'éléments, réduisant de ce fait le nombre de mots utilisés dans la procédure de reconnaissance vocale.

2. Procédé selon la revendication 1 dans lequel les données géographiques déterminées comprennent au moins une parmi des données de position du système de positionnement mondial (GPS), des données d'identité de cellule (identifiant de cellule), des

Amazon/VB Assets
Exhibit 1003
Page 375

données d'identification d'appelant (identifiant d'appelant), des données de nom de lieu ou des données de code postal.

3. Procédé selon la revendication 1 dans lequel les données géographiques sont déduites d'un signal reçu du combiné mobile, ledit signal étant lié à une position actuelle dudit combiné mobile ou à une entrée manuelle entrée dans ledit combiné mobile.

4. Procédé selon une revendication précédente quelconque dans lequel les éléments comprennent au moins un parmi des noms de rues, des entreprises, des commerçants, des centres d'intérêt, des installations de transport, des maisons individuelles, des activités ou des points de repère.

5. Procédé selon une revendication précédente quelconque dans lequel le combiné mobile reçoit le sous-ensemble de mots du serveur distant et exécute la procédure de reconnaissance vocale sur la base du sous-ensemble reçu de mots, réduisant de ce fait l'utilisation du réseau de communication et du serveur distant.

6. Système de reconnaissance vocale comprenant :

un combiné mobile équipé pour communiquer avec un serveur distant sur un réseau sans fil ;
un mécanisme d'acquisition de données (101) pour déterminer des données géographiques pour le combiné mobile suite à une entrée de demande ; et
un mécanisme de sélection (102) pour utiliser les données géographiques déterminées afin de sélectionner automatiquement un sous-ensemble de mots (106) pour une utilisation avec une procédure de reconnaissance vocale ;

dans lequel le sous-ensemble de mots est sélectionné par le serveur distant à partir d'une base de données d'éléments (500) à laquelle ledit serveur distant a accès, ladite base de données d'éléments comprenant des mots qui décrivent des éléments pour chacun d'une pluralité d'emplacements définis géographiquement, réduisant de ce fait le nombre de mots utilisés dans la procédure de reconnaissance vocale.

7. Système de reconnaissance vocale selon la revendication 6 dans lequel les données géographiques sont déduites d'un signal reçu du combiné mobile, ledit signal étant lié à une position actuelle dudit combiné mobile ou à une entrée manuelle entrée dans ledit combiné mobile.

8. Système de reconnaissance vocale selon la revendication 6 dans lequel les données géographiques déterminées comprennent au moins une parmi des données de position du système de positionnement mondial (GPS), des données d'identité de cellule (identifiant de cellule), des données d'identification d'appelant (identifiant d'appelant), des données de nom de lieu ou des données de code postal.

9. Système de reconnaissance vocale selon l'une quelconque des revendications 6 à 8 dans lequel les éléments comprennent au moins un parmi des noms de rues, des entreprises, des commerçants, des centres d'intérêt, des installations de transport, des maisons individuelles, des activités ou des points de repère.

10. Système de reconnaissance vocale pour associer une voix reçue à des mots récupérés d'une base de données d'éléments, le système comprenant :

des moyens pour déterminer une zone géographique d'intérêt suite à une demande dans laquelle une voix doit être reçue, lesdits moyens comprenant en outre :

(i) un mécanisme de détermination d'emplacement de portable associé à un combiné mobile pour générer un signal d'indication révélateur d'un emplacement géographique actuel, et
(ii) un serveur, en communication avec le mécanisme de détermination d'emplacement de portable, programmé pour déterminer une zone géographique d'intérêt à partir du signal d'indication ;

des moyens pour sélectionner un sous-ensemble de mots parmi la base de données d'éléments sur la base de la zone géographique d'intérêt ; lesdits moyens comprenant le serveur programmé pour extraire de la base de données d'éléments seulement des mots qui sont associés à la zone géographique d'intérêt ; et
des moyens pour associer la voix reçue à des mots sélectionnés parmi le sous-ensemble de mots, lesdits moyens étant mis en oeuvre sur le combiné mobile.

11. Système de reconnaissance vocale selon la revendication 10 comprenant en outre :

des moyens pour accepter un signal d'au moins une interface utilisateur graphique pour sélectionner une partie de la zone géographique d'intérêt déterminée afin de spécifier de ce fait une zone géographique d'intérêt limitée supplémentaire ; et
des moyens pour réduire en outre le sous-ensemble de mots selon la zone géographique

Amazon/VB Assets
Exhibit 1003
Page 376

d'intérêt limitée supplémentaire afin de générer un sous-ensemble supplémentaire de mots, de telle sorte que les moyens pour associer une voix reçue à des mots sélectionnent seulement des mots parmi le sous-ensemble supplémentaire de mots.

12. Système de reconnaissance vocale selon la revendication 11 dans lequel les moyens pour accepter un signal comprennent :

un affichage électronique pour afficher une carte de la zone géographique d'intérêt déterminée ; et
un mécanisme de traitement pour combiner le signal accepté avec la carte de la zone géographique d'intérêt déterminée de façon à provoquer un affichage de la zone géographique d'intérêt limitée supplémentaire sur l'affichage électronique.

13. Combiné mobile (250) comprenant des moyens de reconnaissance vocale pour une association d'une voix reçue à des mots reçus d'une base de données distante, le combiné mobile comprenant :

des moyens pour recevoir une demande ;
des moyens pour fournir un signal représentant une zone géographique d'intérêt suite à ladite demande ;
des moyens pour transmettre ledit signal à un serveur distant et recevoir du serveur distant un sous-ensemble de mots provenant de la base de données distante sur la base de la zone géographique d'intérêt ; et
des moyens pour associer la voix reçue à des mots sélectionnés parmi le sous-ensemble de mots.

14. Combiné mobile selon la revendication 13 dans lequel les moyens pour fournir un signal comprennent des moyens pour déterminer une zone géographique d'intérêt dans laquelle une voix doit être reçue.

15. Combiné mobile selon la revendication 13 ou 14 comprenant en outre:

des moyens (201) pour accepter au moins une entrée utilisée pour déterminer une zone géographique affinée d'intérêt dans la zone géographique d'intérêt ;
des moyens (205) pour réduire le sous-ensemble de mots selon la zone géographique affinée d'intérêt;

et dans lequel :

les moyens pour associer la voix reçue sont agencés pour une association d'une voix seulement avec le sous-ensemble réduit de mots.

Amazon/VB Assets
Exhibit 1003
Page 377

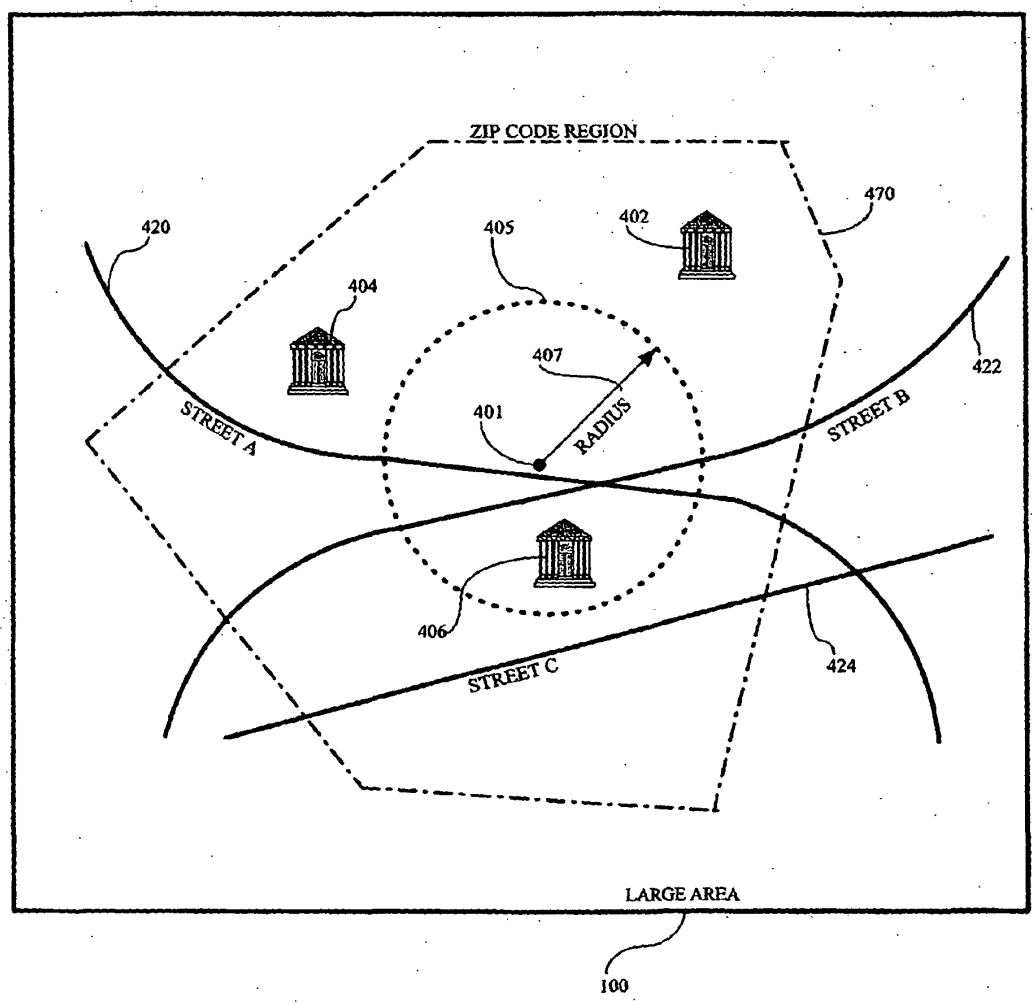

ZIP CODE REGION
470
420
405
402
404
407
422
401
RADIUS
STREET A
STREET B
406
424
STREET C
LARGE AREA
100

FIG. 1

500

ELEMENT DATABASE

ELEMENT NAME FIELD 501	PARAMETER FIELD FOR THE REFINED GEOGRAPHIC AREA OF INTEREST 503	PARAMETER FIELD FOR THE COARSE GEOGRAPHIC AREA OF INTEREST 505
FIRST BUILDING	{[603001, 2428833]}	10017
SECOND BUILDING	{[602298, 2428830]}	10017
THIRD BUILDING	{[603010, 2428836]}	10017
STREET A	{[603012, 2428836],[603013, 24428836]}	10017, 10036
STREET B	{[603014, 2428836],[603015, 24428836]}	10017, 10055
STREET C	{[679010, 2428836],[680000, 24428836]}	10017, 12533
...	...	...
DEVON FARMS ROAD	{[673010, 2428836],[675000, 24428836]}	12582

FIG. 2

600

WORD SUBSET LIST

ELEMENT NAME FIELD 601
FIRST BUILDING
STREET A
STREET B

FIG. 3

Amazon/VB Assets
Exhibit 1003
Page 379

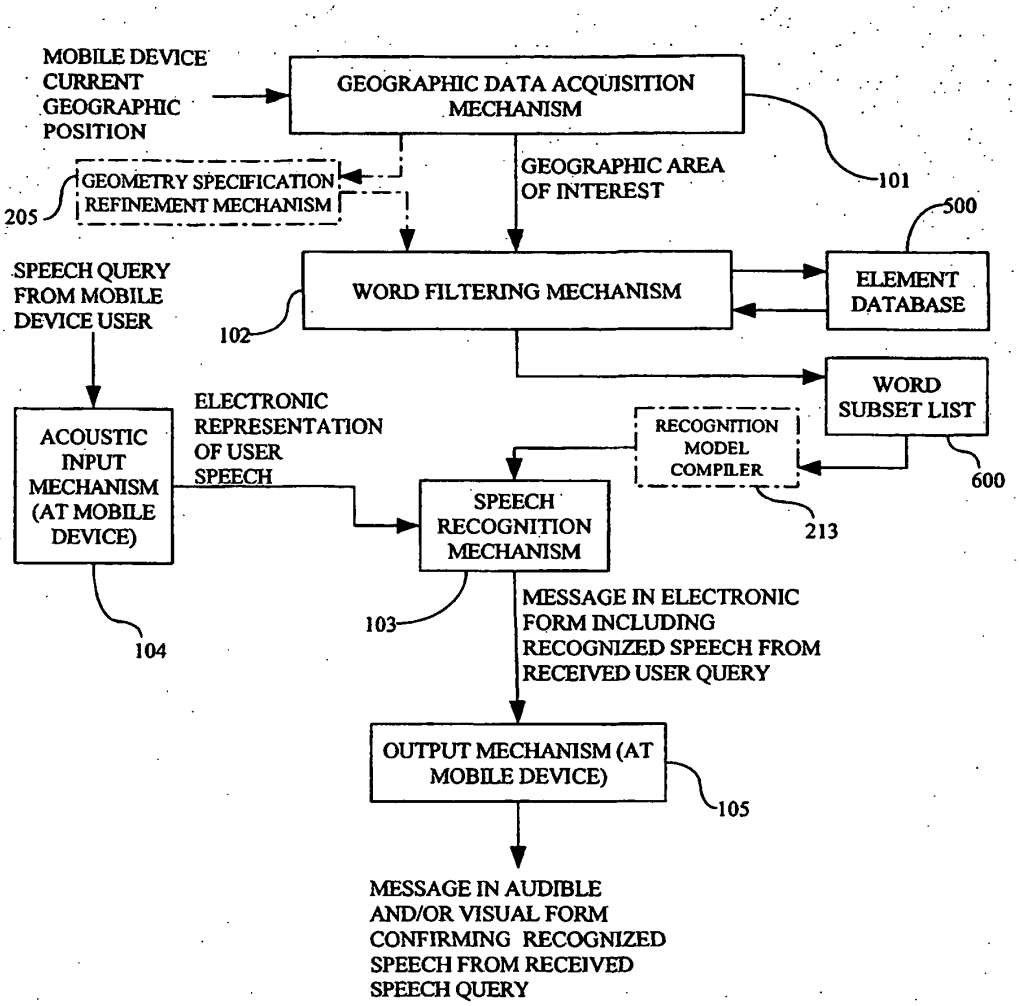


FIG. 4

Amazon/VB Assets
Exhibit 1003
Page 380

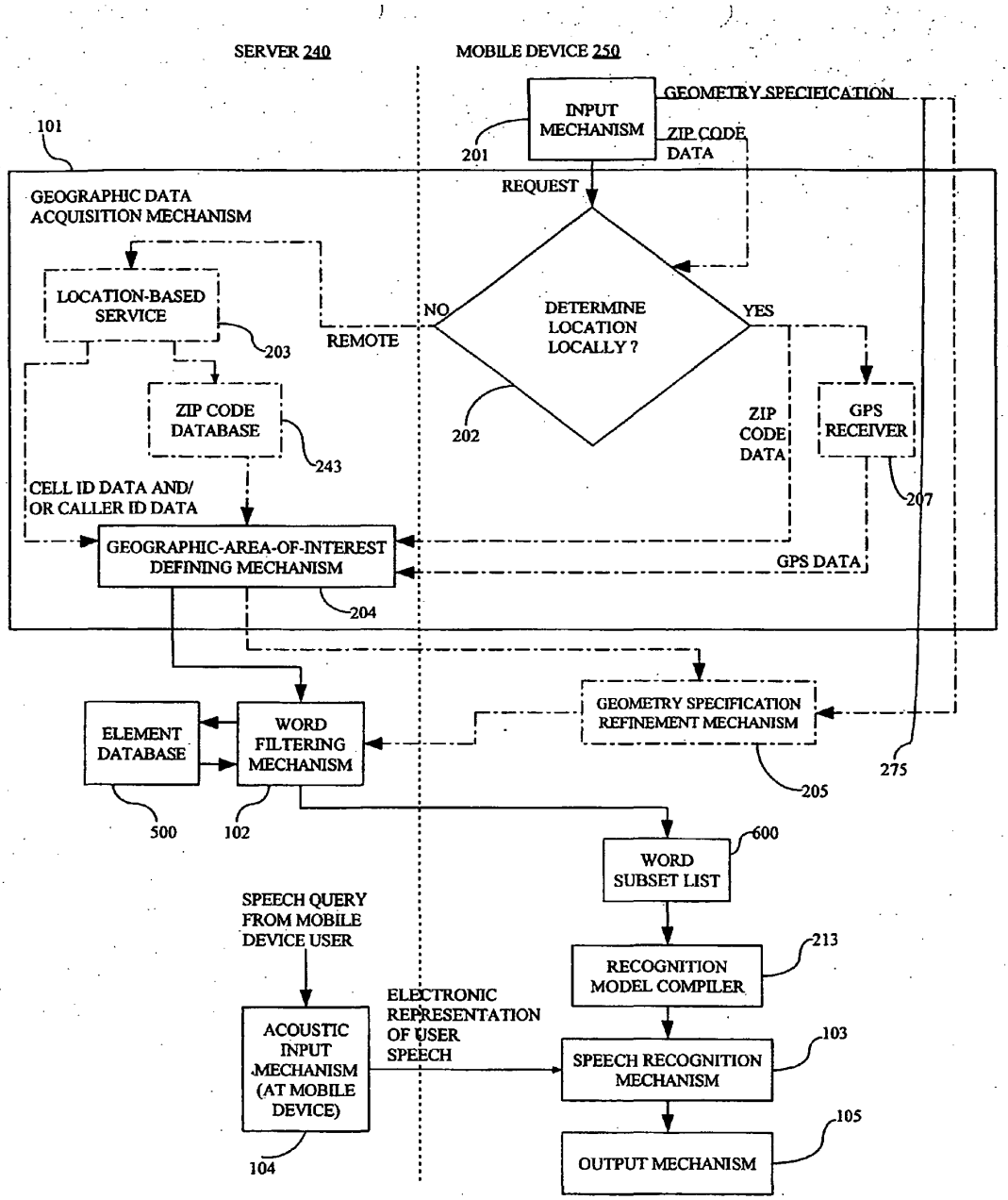


FIG. 5

Amazon/VB Assets
Exhibit 1003
Page 381

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 2003236099 A **[0007] [0009] [0010]**
- EP 1233407 A **[0011]**
- US 6230132 B **[0012]**
- EP 0661688 A **[0013]**

Amazon/VB Assets
Exhibit 1003
Page 382

Electronic Patent Application Fee Transmittal

Application Number:	12943699
Filing Date:	10-Nov-2010
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Filer:	Ngai Bing Zhang/Miriam Nelsen Awan
Attorney Docket Number:	090611-0384422

Filed as Small Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				
Extension - 1 month with $0 paid	2251	1	75	75

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
Submission- Information Disclosure Stmt	1806	1	180	180
		Total in USD ($)		**255**

Electronic Acknowledgement Receipt	
EFS ID:	14831291
Application Number:	12943699
International Application Number:	
Confirmation Number:	4185
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Customer Number:	909
Filer:	Ngai Bing Zhang/Miriam Nelsen Awan
Filer Authorized By:	Ngai Bing Zhang
Attorney Docket Number:	090611-0384422
Receipt Date:	30-JAN-2013
Filing Date:	10-NOV-2010
Time Stamp:	14:46:32
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		090611-0384422_Response.pdf	577797 4f9e5dcbebdd97b07fe69c0fe3f7c124bc570362	yes	14

	Multipart Description/PDF files in .zip description				
	Document Description		Start	End	
	Amendment/Req. Reconsideration-After Non-Final Reject		1	1	
	Claims		2	8	
	Applicant Arguments/Remarks Made in an Amendment		9	14	
Warnings:					
Information:					
2	Extension of Time	090611-0384422_Extension.pdf	495240 7ae042d01267c039fb3fde986ae9181dce29d495	no	1
Warnings:					
Information:					
3	Transmittal Letter	0384422_SIDS.pdf	3998724 d8e170bb990b2606fc9fbf0535c62ea5d8bdce45	no	7
Warnings:					
Information:					
4	Information Disclosure Statement (IDS) Form (SB08)	0384422_IDSPTO1449.pdf	1301425 96985df0451c3a73292b88972c2aaa09105b4483	no	11
Warnings:					
Information:					
This is not an USPTO supplied IDS fillable form					
5	Foreign Reference	EP1320043A2.pdf	1416212 2ec7d64c9ab7063100d70c545035fd72a13ded77	no	28
Warnings:					
Information:					
6	Foreign Reference	EP1646037.pdf	262902 56537431a1754beec4b827cf69e772a2da14c1a8	no	20
Warnings:					
Information:					
7	Non Patent Literature	Statement_in_Accordance_with_EPO_Notice.pdf	86084 a41f2c134966e666f56e1c9c6e421bc9a3281ef0	no	1
Warnings:					
Information:					
8	Fee Worksheet (SB06)	fee-info.pdf	32590 78a6683ff79183dc3a0f0f93a06e36ef3d28905d	no	2

Warnings:	
Information:	
Total Files Size (in bytes):	8170974

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
12/943,699	11/10/2010	Mike Kennewick	090611-0384422	4185

909 7590 10/01/2012

Pillsbury Winthrop Shaw Pittman, LLP (NV)
PO Box 10500
McLean, VA 22102

EXAMINER	
HALE, TIM B	
ART UNIT	PAPER NUMBER
3685	

NOTIFICATION DATE	DELIVERY MODE
10/01/2012	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

docket_ip@pillsburylaw.com
margaret.drosos@pillsburylaw.com

Office Action Summary	Application No.	Applicant(s)	
	12/943,699	KENNEWICK ET AL.	
	Examiner	Art Unit	
	TIM HALE	3685	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

4)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

5)☒ Claim(s) *1-20* is/are pending in the application.

5a) Of the above claim(s) _____ is/are withdrawn from consideration.

6)☐ Claim(s) _____ is/are allowed.

7)☒ Claim(s) *1-20* is/are rejected.

8)☐ Claim(s) _____ is/are objected to.

9)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

10)☐ The specification is objected to by the Examiner.

11)☒ The drawing(s) filed on *10 November 2010* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

12)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

13)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date *12-06-2010 and 11/10/2010*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

Status of Claims

1. Claims 1-20 have been presented.

2. Claims 1-20 are rejected.

Examiner's Comments

3. Per claims 1, 2, 5, 9, 11, 12, 15 and 20 the type of data being collected by the system is non-functional as it does not alter how the machine functions (E.g., the content dedication message(1, 9, 11, 19), messages (2, 12), album art metadata (5, 15), and content (10, 20) is data stored in memory.) It has been held that nonfunctional descriptive material cannot render nonobvious an invention that would have otherwise been obvious (In re Gulack, 217 USPQ 401 (Fed. Cir. 1983), In re Ngai, 70 USPQ2d (Fed. Cir. 2004), In re Lowry, 32 USPQ2d 1031 (Fed. Cir. 1994); MPEP 2106.01)

Per claims 1,2, 3, 4, 6, 7, 8, the "configured to" language is functional and does not have patentable weight (See MPEP 2114). While features of an apparatus may be recited either structurally or functionally, claims directed to an apparatus must be

distinguished from the prior art in terms of structure rather than function alone (MPEP 2214; In re Swineheart, 169 USPQ 226; In re Schreiber, 44 USPQ2d 1429 (Fed. Cir. 1997))

Functional recitation(s) using the word "to(1, 11)", "for(8, 15, 17, 18)", or other functional language have been considered but are given little patentable weight because they fail to add any structural limitations and are thereby regarded as intended use language. A recitation of the intended use of the claimed product must result in a structural difference between the claimed product and the prior art in order to patentably distinguish the claimed product from the prior art. If the prior art structure is capable of performing the intended use, then it reads on the claimed limitation. In re Casey, 370 F.2d 576, 152 USPQ 235 (CCPA 1967) ("The manner or method in which such machine is to be utilized is not germane to the issue of patentability of the machine itself."); *In re Otto,* 136 USPQ 458,459 (CCPA 1963). See also MPEP §§ 2114 and 2115

Claim Rejections - 35 USC § 101

4. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

5. Claims 1-20 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter.

Claim 1 is directed to software per se as it recites a "an electronic device." The electronic device can be interpreted to be software because one of ordinary skill would consider these components are broad enough such that their functionality can be implemented in software alone. Software is defined by the Microsoft Press Computer Dictionary Third Edition (1997) as "computer programs; instructions that make computers work." p. 441. According to MPEP 2106 II IV, however, there are four categories of invention: process, machine, article of manufacture or composition of matter. Therefore, as "software" is neither a category of invention nor a subset of one of the categories it does not represent patent eligible subject matter (In re Nuijten, 500 F.3d 1346 (Fed. Cir. 2007)).

Claims 2-10 depend from claim 1 and add no new material that would render the claim statutory. Hence, claims 2-10 are similarly rejected.

Based on Supreme Court precedent (See also Diamond v. Diehr, 450 U.S. 175, 184 (1981); Parker v. Flook, 437 U.S. 584, 588 n.9 (1978); Gottschalk v. Benson, 409 U.S. 63, 70 (1972); Cochrane v. Deener, 94 U.S. 780, 787-88 (1876)) and recent Federal Circuit decisions, a §101 process must (1) be tied to a particular machine or apparatus (machine implemented) or (2) transform underlying subject matter (such as an article or materials)

to a different state or thing. In addition, the tie to a particular apparatus, for example, cannot be mere extra-solution activity. See Bilski v. Kappos, 95 USPQ2d 1001 (US 2010).

An example of a method claim that would not qualify as a statutory process would be a claim that recited purely mental steps.

To meet prong (1), the method step should positively recite the particular apparatus (the thing or product) to which it is tied. This may be accomplished by having the claim positively recite the machine that accomplishes the method steps. Alternatively or to meet prong (2), the method step should positively recite identifying the material that is being changed to a different state or positively recite the subject matter that is being transformed.

Based upon consideration of all of the relevant factors with respect to the claim as a whole, claim 11 is held to claim an abstract idea, and is therefore rejected as ineligible subject matter under 35 U.S.C. 101. The rationale for this finding is explained below:

The claim includes an insufficient recitation of a machine or transformation because the involvement of machine, or transformation, with the steps merely involves insignificant extra-solution activity or is merely nominally, insignificantly, or tangentially related to the performance of the steps (e.g., identifying the content to dedicate for the

recipient based on an interpretation of the natural language utterance included in the multi-modal interaction). Claim 11 is therefore held to claim an abstract idea, and is/are therefore rejected as ineligible subject matter under 35 U.S.C. 101.

Dependent claim(s) 12-20 when analyzed as a whole are held to be ineligible subject matter and are rejected under 35 U.S.C. § 101 because the additional recited limitation(s) fail(s) to establish that the claim is not directed to an abstract idea.

Claim Rejections - 35 USC § 112

6. The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

7. Claims 3 and 13 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Claims 3 and 13 contain the language "one or more tags." However, the tags appear to be already in existence. Hence, one of ordinary skill in the art could interpret the claim as the system creates the tags, the tags are already on the system, or the system obtains the tags from a third party.

8. Claims 5, 8, 14 and 18 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Claim 5 and 14 recite the limitation "album art metadata." There is insufficient antecedent basis for this limitation in the claim.

Claim 8 and 18 recite the limitation "the natural content." There is insufficient antecedent basis for this limitation in the claim.

Claim Rejections - 35 USC § 102

9. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

10. Claims 1-10 are rejected under 35 U.S.C. 102(b) as being anticipated by US Patent Publication 20040193420 to Kennewick (hereafter "Kennewick).

Per claim 1, Kennewick teaches:

"A system for providing a natural language content dedication service, comprising:

an electronic device" Figs. 1-3, 5-6; [0013][0014][0016]]

and "a content dedication system in communication with the electronic device "[0013][0016][0017][0022] Figs. 5-6.

The remainder of claim 1 is directed to functional language (e.g, configured to:). Similarly, claims 2-10 are also directed to functional language. Therefore, as Kennewick teaches the claimed system for providing a natural language content dedication service and content dedication system in communication with the electronic device, it is sufficient in terms of art.

Claim Rejections - 35 USC § 103

11. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1, 2, 7, 9, 10 , 11, 12, 17, 19 and 20 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent Publication 20040193420 to Kennewick (hereafter "Kennewick") in view of US Patent Publication 20020032752 to Gold (hereafter "Gold").

Per claim 1, Kennewick teaches

"A system for providing a natural language content dedication service, comprising:

an electronic device configured to:" [0013][0014][0016][0095][0096][0097][0111]Figs. 1-3, 5-6

"receive a multi-modal interaction that includes a request to dedicate content, wherein the multi-modal interaction includes a natural language utterance relating to the request to dedicate the content;[0028][0034][0038[0101]]" and

and

"a content dedication system in communication with the electronic device through a message interface,]wherein the content dedication system is configured to:" [0013][0016][0017][0022][0096][0097][0118]

"identify the content to dedicate for the recipient based on an interpretation of the natural language utterance included in the multi-modal interaction;"[0018][0028][0123]

"process a transaction for the identified content to dedicate for the recipient;"[0018][0028][0123]

Kennewick does not teach, but Gold teaches:

"capture an utterance dedicating the content for a recipient"[0038][0039];Abstract

"send a content dedication message to the recipient, wherein the content dedication message includes information for the recipient to access the content and wherein the content dedication message further includes information for the recipient to access the utterance dedicating the content for the recipient." [0042][0043]

Per claims 2 and 12, Gold teaches: "receive[ing [at the content dedication system] one or more messages from the electronic device through the messaging interface, wherein one or more messages contain information relating to the multi-modal interaction that includes the request to dedicate the content, and wherein the one or more messages further contain information corresponding to the utterance dedicating the content for the recipient." [0038]-[0039]; Abstract

Per claims 7 and 17, Gold teaches:" purchase[ing], [by the content dedication system] the identified content from a content provider to process the transaction for the identified content." [0009]

Per claims 9 and 19, Gold teaches, "wherein the content dedication message sent to the recipient includes an electronic link for the recipient to access the content and the utterance dedicating the content for the recipient." [0042]; Abstract; Claim 8

Per claims 10 and 20, Gold teaches, "wherein the content includes one or more of a song or a video." [0007]

Per claim 11, Kennewick teaches: "A method for providing a natural language content dedication service, comprising:

receiving, at an electronic device, a multi-modal interaction that includes a request to dedicate content, wherein the multi-modal interaction includes a natural language utterance relating to the request to dedicate the content";

[0013][0014][0016][0095][0096][0097][0111]Figs. 1-3, 5-6

"identifying the content to dedicate for the recipient based on an interpretation of the natural language utterance included in the multi-modal interaction;"[0018][0028][0123]

"processing, at a content dedication system, a transaction for the identified content to dedicate for the recipient;" and [0018][0028][0123]

Gold teaches:

"capturing an utterance dedicating the content for a recipient;" "[0038][0039]; Abstract

"sending a content dedication message from the content dedication system to the recipient, wherein the content dedication message includes information for the recipient to access the content, and wherein the content dedication message further includes

information for the recipient to access the utterance dedicating the content for the recipient." [0042][0043]

It would have been obvious to combine Kennewick and Gold to produce a system for distributing recorded messages associated with content.

12. Claims 3, 4, 6, 13, 14 and 16 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent Publication 20040193420 to Kennewick (hereafter "Kennewick") in view of US Patent Publication 20020032752 to Gold (hereafter "Gold") and US Patent Publication 20080140385 to Mahajan (hereafter "Mahajan").

Per claims 3 and 13, neither Kennewick nor Gold teach, but Mahajan teaches :

"receive [receiving at the electronic device] information corresponding to one or more tags dedicating the content for the recipient." [0027]

Per claims 6 and 16, Mahajan teaches: "annotate[ing] the content with the utterance dedicating the content for the recipient." [0040]

It would have been obvious to combine Kennewick, Gold and Mahajan to produce a system for distributing A/V content based on speech recognition.

Per claims 4 and 14, Neither Kennewick nor Mahajan teaches, but Gold teaches: "insert[ing] the utterance dedicating the content for the recipient within the content;"[0041] and

Neither Kennewick nor Gold teaches, but Mahajan teaches: "insert[ing] the tags dedicating the content for the recipient within metadata for the content." [0033][0038]

It would have been obvious to combine Kennewick, Gold and Mahajan to produce a system for modifying A/V content based on user inputs.

13. Claims 5 and 15 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent Publication 20040193420 to Kennewick (hereafter "Kennewick") in view of US Patent Publication 20020032752 to Gold (hereafter "Gold") and US Patent Publication 20100036967 to Caine (hereafter "Caine").

Per claims 5 and 15, neither Kennewick nor Gold teach, but Caine teaches: "wherein the tags dedicating the content for the recipient include one or more of an image or a picture to insert within album art metadata for the content." [0079][0080]

It would have been obvious to combine Kennewick, Gold and Caine to produce a system for modifying visual content associated with digital content.

14. Claims 8 and 18 are rejected under 35 U.S.C. 103(a) as being unpatentable over US Patent Publication 20040193420 to Kennewick (hereafter "Kennewick") in view of US Patent Publication 20020032752 to Gold (hereafter "Gold") and US Patent 5892900 to Ginter (hereafter "Ginter").

Per claims 8 and 18, Kennewick teaches:

"identify[ing] a purchase option for the request to dedicate the content, wherein the

purchase option includes one or more of a buy-to-own purchase option, a pay-to-pay purchase option, or a paid subscription purchase option;" [0019]

Neither Kennewick nor Gold teach, but Ginter teaches:

"bill[ing] a user of the electronic device for the natural content dedication service based on identified purchase option" col. 33 ln 1-35 col. 47 ln 20-55

It would have been obvious to combine Kennewick, Gold and Ginter to produce a system for distributing digital content

Conclusion

15. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure:

- US Pat App Pub 20020029261 to Shibata

Any inquiry concerning this communication or earlier communications from the examiner should be directed to TIM HALE whose telephone number is (571)270-3794. The examiner can normally be reached on Monday - Friday 9am - 5pm EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Calvin Hewitt II can be reached on (571)272-6709. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO

Customer Service Representative or access to the automated information system, call

800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/CALVIN L HEWITT II/
Supervisory Patent Examiner, Art Unit 3685

Notice of References Cited	Application/Control No. 12/943,699	Applicant(s)/Patent Under Reexamination KENNEWICK ET AL.	
	Examiner TIM HALE	Art Unit 3685	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
*	A	US-5,892,900 A	04-1999	Ginter et al.	726/26
*	B	US-2002/0032752 A1	03-2002	Gold et al.	709/218
*	C	US-2002/0029261 A1	03-2002	Shibata, Junji	709/219
*	D	US-2004/0158555 A1	08-2004	Seedman et al.	707/003
*	E	US-2004/0193420 A1	09-2004	Kennewick et al.	704/257
*	F	US-7,082,469 B2	07-2006	Gold et al.	709/231
*	G	US-2007/0146833 A1	06-2007	Satomi et al.	358/537
*	H	US-2008/0140385 A1	06-2008	Mahajan et al.	704/9
*	I	US-2009/0216540 A1	08-2009	TESSEL et al.	704/275
*	J	US-7,640,272 B2	12-2009	Mahajan et al.	1/1
*	K	US-7,693,720 B2	04-2010	Kennewick et al.	704/275
*	L	US-8,005,683 B2	08-2011	Tessel et al.	704/275
*	M	US-2010/0036967	02-2010	Caine et al.	709/236

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	
	V	
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	12943699	KENNEWICK ET AL.
	Examiner	Art Unit
	TIM HALE	3685

SEARCHED			
Class	Subclass	Date	Examiner

SEARCH NOTES		
Search Notes	Date	Examiner
East text search	9/12/2012	TBH

INTERFERENCE SEARCH			
Class	Subclass	Date	Examiner

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	2447

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	A	2010/0204986	08/12/2010	Kennewick *et al.*	704	226	
	B	2010/0204994	08/12/2010	Kennewick *et al.*	704	257	
	C	2010/0217604	08/26/2010	Baldwin *et al.*	704	275	
	D	2010/0286985	11/11/2010	Kennewick *et al.*	704	257	
	E	2010/0299142	11/25/2010	Freeman *et al.*	704	9	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO
	F	WO 2007/027546	03/08/2007	International				
	G	WO 2007/027989	03/08/2007	International				
	H	WO 2008/098039	08/14/2008	International				
	I	WO 2008/118195	10/02/2008	International				
	J	WO 2009/075912	06/18/2009	International				
	K	WO 2009/145796	12/03/2009	International				
	L	WO 2010/096752	08/26/2010	International				

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

09/18/2012

EXAMINER	/T.H./ /Tim Hale/	DATE CONSIDERED	09/18/2012

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

402527207v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

EAST Search History

EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	2	"20040193420"	US-PGPUB; USPAT	AND	OFF	2012/09/19 16:44
L2	2	"20020032752"	US-PGPUB; USPAT	AND	OFF	2012/09/19 16:45
L3	2	"20080140385"	US-PGPUB; USPAT	AND	OFF	2012/09/19 16:45
L4	1	"5892900".pn.	US-PGPUB; USPAT	AND	OFF	2012/09/19 16:46
L5	1	"20020029261"	US-PGPUB; USPAT	AND	OFF	2012/09/19 16:46
S1	0	("(outputdevice).ti.").PN.	US-PGPUB	OR	OFF	2012/07/03 14:03
S2	1910	(output device).ti.	US-PGPUB	AND	OFF	2012/07/03 14:03
S3	1753	(touch panel).ti.	US-PGPUB	AND	OFF	2012/07/03 14:09
S4	111	second same verification same after first verification same digest	US-PGPUB	AND	OFF	2012/07/03 14:15
S5	2	"20040193420"	US-PGPUB; USPAT	AND	OFF	2012/09/18 11:38
S6	2	"20090216540"	US-PGPUB; USPAT	AND	OFF	2012/09/18 11:43
S7	2	"20080140385"	US-PGPUB; USPAT	AND	OFF	2012/09/18 11:46
S8	10761	voice replay	US-PGPUB; USPAT	AND	OFF	2012/09/18 15:27
S9	207	(voice NEAR replay)	US-PGPUB; USPAT	AND	OFF	2012/09/18 15:27
S10	177	@ay<"2009" (voice NEAR replay)	US-PGPUB; USPAT	AND	OFF	2012/09/18 15:27
S11	40	@ay<"2009" (voice NEAR replay) and query	US-PGPUB; USPAT	AND	OFF	2012/09/18 15:27

S12	32	@ay<"2009" (voice NEAR replay) and query and (link or html)	US-PGPUB; USPAT	AND	OFF	2012/09/18 15:28
S13	0	@ay<"2009" ((voice NEAR replay) with (link or html)) and query	US-PGPUB; USPAT	AND	OFF	2012/09/18 15:28
S14	3	@ay<"2009" ((voice NEAR replay) with (link or html))	US-PGPUB; USPAT	AND	OFF	2012/09/18 15:28
S15	3	("20020029261" OR "20040193420" OR "20080140385") and html	US-PGPUB; USPAT	AND	OFF	2012/09/18 16:31
S16	5	("20020029261" OR "20040193420" OR "20080140385") and link	US-PGPUB; USPAT	AND	OFF	2012/09/18 16:33
S17	0	("20020029261" OR "20040193420" OR "20080140385") and (bill or charge)	US-PGPUB; USPAT	AND	OFF	2012/09/18 17:01
S18	0	@ay<"2009" and (insert WITH metadata WITH art WITH album)	US-PGPUB; USPAT	AND	OFF	2012/09/18 17:06
S19	0	@ay<"2009" and (insert WITH metadata WITH picture WITH album)	US-PGPUB; USPAT	AND	OFF	2012/09/18 17:07
S20	0	@ay<"2009" and (insert WITH metadata WITH image WITH album)	US-PGPUB; USPAT	AND	OFF	2012/09/18 17:07
S21	38	@ay<"2009" and (insert WITH image WITH album)	US-PGPUB; USPAT	AND	OFF	2012/09/18 17:08
S22	1512	"5892900"	US-PGPUB; USPAT	AND	OFF	2012/09/18 17:14
S23	1	"5892900".pn.	US-PGPUB; USPAT	AND	OFF	2012/09/18 17:14
S24	7239768	@ay<"2009"	US-PGPUB; USPAT	AND	OFF	2012/09/19 13:50
S25	46	@ay<"2009" AND (voice dedication replay)	US-PGPUB; USPAT	AND	OFF	2012/09/19 13:51
S26	2	@ay<"2009" AND (voice SAME dedication SAME replay)	US-PGPUB; USPAT	AND	OFF	2012/09/19 13:51
S27	1	"20040158555"	US-PGPUB; USPAT	AND	OFF	2012/09/19 14:38

9/19/2012 4:47:14 PM
C:\Users\thale\Documents\EAST\Workspaces\13153706.wsp

Amazon/VB Assets
Exhibit 1003
Page 409

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	*Unassigned*
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	Herewith (November 10, 2010)	*Unassigned*

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	1A	4,430,669	02/07/1984	Cheung	358	122	
	1B	5,155,743	10/13/1992	Jacobs	375	28	
	1C	5,274,560	12/28/1993	LaRue	364	444	
	1D	5,377,350	12/27/1994	Skinner	395	600	
	1E	5,386,556	01/31/1995	Hedin et al.	395	600	
	1F	5,424,947	06/13/1995	Nagao et al.	364	419.08	
	1G	5,471,318	11/28/1995	Ahuja et al.	358	400	
	1H	5,475,733	12/12/1995	Eisdorfer et al.	379	52	
	1I	5,488,652	01/30/1996	Bielby *et al.*	379	88	
	1J	5,499,289	03/12/1996	Bruno et al.	379	220	
	1K	5,500,920	03/19/1996	Kupiec	395	2.79	
	1L	5,517,560	05/14/1996	Greenspan	379	114	
	1M	5,533,108	07/02/1996	Harris et al.	379	201	
	1N	5,537,436	07/16/1996	Bottoms et al.	375	222	
	1O	5,539,744	07/23/1996	Chu et al.	370	60	
	1P	5,557,667	09/17/1996	Bruno et al.	379	201	
	1Q	5,563,937	10/08/1996	Bruno et al.	379	201	
	1R	5,590,039	12/31/1996	Ikeda et al.	395	759	
	1S	5,617,407	04/01/1997	Bareis	369	275.3	
	1T	5,633,922	05/27/1997	August et al.	379	220	
	1U	5,675,629	10/07/1997	Raffel et al.	379	58	
	1V	5,696,965	12/09/1997	Dedrick	395	610	
	1W	5,708,422	01/13/1998	Blonder et al.	340	825.34	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO
	1X	WO 99/46763	09/16/1999	International				
	1Y	WO 00/21232	04/13/2000	International				
	1Z	WO 00/46792	08/10/2000	International				
	1AA	WO 01/78065	10/18/2001	International				
	1BB	WO 2004/072954	08/26/2001	International				

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

	1CC	REUTERS, "IBM to Enable Honda Drivers to Talk to Cars", Charles Schwab & Co., Inc., July 28, 2002, 1 page.
	1DD	LIN, Bor-shen, *et al.*, "A Distributed Architecture for Cooperative Spoken Dialogue Agents with Coherent Dialogue State and History", ASRU'99, 1999, 4 pages.

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

402465373v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	*Unassigned*
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	Herewith (November 10, 2010)	*Unassigned*

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	2A	5,721,938	02/24/1998	Stuckey	395	754	
	2B	5,722,084	02/24/1998	Chakrin et al.	455	551	
	2C	5,742,763	04/21/1998	Jones	395	200.3	
	2D	5,748,841	05/05/1998	Morin *et al.*	395	2.66	
	2E	5,752,052	05/12/1998	Richardson et al.	395	759	
	2F	5,754,784	05/19/1998	Garland et al.	395	200.49	
	2G	5,761,631	06/02/1998	Nasukawa	704	9	
	2H	5,774,859	06/30/1998	Houser *et al.*	704	275	
	2I	5,794,050	08/11/1998	Dahlgren et al.	395	708	
	2J	5,794,196	08/11/1998	Yegnanarayanan *et al.*	704	255	
	2K	5,797,112	08/18/1998	Komatsu *et al.*	701	201	
	2L	5,799,276	08/25/1998	Komissarchik *et al.*	704	251	
	2M	5,802,510	09/01/1998	Jones	707	2	
	2N	5,832,221	11/03/1998	Jones	375	200.36	
	2O	5,867,817	02/02/1999	Catallo *et al.*	704	255	
	2P	5,878,385	03/02/1999	Bralich *et al.*	704	9	
	2Q	5,878,386	03/02/1999	Coughlin	704	10	
	2R	5,892,813	04/06/1999	Morin *et al.*	379	88.01	
	2S	5,895,464	04/20/1999	Bhandari *et al.*	707	3	
	2T	5,895,466	04/20/1999	Goldberg et al.	707	5	
	2U	5,902,347	05/11/1999	Backman *et al.*	701	200	
	2V	5,911,120	06/08/1999	Jarett et al.	455	417	
	2W	5,918,222	06/29/1999	Fukui *et al.*	707	1	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO
	2X	WO 2007/019318	02/15/2007	International				
	2Y	WO 2007/021587	02/22/2007	International				

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

	2Z	KUHN, Thomas, *et al.*, "Hybrid In-Car Speech Recognition for Mobile Multimedia Applications", Vehicular Technology Conference, IEEE, July 1999, pages 2009-2013.
	2AA	BELVIN, Robert, *et al.*, "Development of the HRL Route Navigation Dialogue System", Proceedings of the First International Conference on Human Language Technology Research, San Diego, 2001, pages 1-5.

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

402465373v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	*Unassigned*
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	Herewith (November 10, 2010)	*Unassigned*

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	3A	5,926,784	07/20/1999	Richardson et al.	704	9	
	3B	5,933,822	08/03/1999	Braden-Harder et al.	707	5	
	3C	5,953,393	09/14/1999	Culbreth et al.	379	88.25	
	3D	5,963,894	10/05/1999	Richardson et al.	704	9	
	3E	5,963,940	10/05/1999	Liddy et al.	707	5	
	3F	5,987,404	11/16/1999	Della Pietra et al.	704	9	
	3G	5,991,721	11/23/1999	Asano et al.	704	257	
	3H	5,995,119	11/30/1999	Cosatto et al.	345	473	
	3I	6,009,382	12/28/1999	Martino et al.	704	1	
	3J	6,014,559	01/11/2000	Amin	455	413	
	3K	6,021,384	02/01/2000	Gorin et al.	704	1	
	3L	6,044,347	03/28/2000	Abella et al.	704	272	
	3M	6,049,602	04/11/2000	Foladare et al.	379	265	
	3N	6,058,187	05/02/2000	Chen	380	21	
	3O	6,078,886	06/20/2000	Dragosh et al.	704	270	
	3P	6,081,774	06/27/2000	De Hita et al.	704	9	
	3Q	6,101,241	08/08/2000	Boyce et al.	379	88.01	
	3R	6,119,087	09/12/2000	Kuhn et al.	704	270	
	3S	6,134,235	10/17/2000	Goldman et al.	370	352	
	3T	6,144,667	11/07/2000	Doshi et al.	370	401	
	3U	6,144,938	11/07/2000	Surace *et al.*	704	257	
	3V	6,154,526	11/28/2000	Dahlke *et al.*	379	88.03	
	3W	6,160,883	12/12/2000	Jackson et al.	379	230	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

	3X	LIND, R., *et al.*, "The Network Vehicle—A Glimpse into the Future of Mobile Multi-Media", IEEE Aerosp. Electron. Systems Magazine, Volume 14, No. 9, September 1999, pages 27-32.
	3Y	ZHAO, Yilin, "Telematics: Safe and Fun Driving", IEEE Intelligent Systems, Volume 17, Issue 1, 2002, pages 10-14.
	3Z	CHAI *et al.*, "MIND: A Semantics-Based Multimodal Interpretation Framework for Conversational System", Proceedings of the International CLASS Workshop on Natural, Intelligent and Effective Interaction in Multimodal Dialogue Systems, June 2002, pages 37-46.

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

402465373v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	*Unassigned*
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	Herewith (November 10, 2010)	*Unassigned*

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	4A	6,173,279	01/09/2001	Levin et al.	707	5	
	4B	6,175,858	01/16/2001	Bulfer et al.	709	206	
	4C	6,185,535	02/06/2001	Hedin *et al.*	704	270	
	4D	6,192,110	02/20/2001	Abella et al.	379	88.01	
	4E	6,192,338	02/20/2001	Haszto et al.	704	257	
	4F	6,195,634	02/27/2001	Dudemaine *et al.*	704	231	
	4G	6,195,651	02/27/2001	Handel *et al.*	707	2	
	4H	6,208,972	03/27/2001	Grant *et al.*	704	275	
	4I	6,219,346	04/17/2001	Maxemchuk	370	338	
	4J	6,219,643	04/17/2001	Cohen et al.	704	257	
	4K	6,226,612	05/01/2001	Srenger *et al.*	704	256	
	4L	6,233,556	05/15/2001	Teunen et al.	704	250	
	4M	6,233,559	05/15/2001	Balakrishnan	704	275	
	4N	6,246,981	06/12/2001	Papineni et al.	704	235	
	4O	6,272,455	08/07/2001	Hoshen et al.	704	1	
	4P	6,288,319	09/11/2001	Catona	84	609	
	4Q	6,292,767	09/18/2001	Jackson et al.	704	1	
	4R	6,308,151	10/23/2001	Smith	704	235	
	4S	6,314,402	11/06/2001	Monaco et al.	704	275	
	4T	6,366,886	04/02/2002	Dragosh et al.	704	270.1	
	4U	6,374,214	04/16/2002	Friedland *et al.*	704	235	
	4V	6,381,535	04/30/2002	Durocher *et al.*	701	202	
	4W	6,385,596	05/07/2002	Wiser *et al.*	705	51	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

	4X	CHEYER *et al.*, "Multimodal Maps: An Agent-Based Approach", International Conference on Cooperative Multimodal Communication (CMC/95), May 24-26, 1995, pages 111-121.
	4Y	ELIO *et al.*, "On Abstract Task Models and Conversation Policies" in Workshop on Specifying and Implementing Conversation Policies, Autonomous Agents '99, Seattle, 1999, 10 pages.
	4Z	TURUNEN, "Adaptive Interaction Methods in Speech User Interfaces", Conference on Human Factors in Computing Systems, Seattle, Washington, 2001, pages 91-92.

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

402465373v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	*Unassigned*
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	Herewith (November 10, 2010)	*Unassigned*

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	5A	6,385,646	05/07/2002	Brown et al.	709	217	
	5B	6,393,428	05/21/2002	Miller et al.	707	102	
	5C	6,404,878	06/11/2002	Jackson et al.	379	221.01	
	5D	6,408,272	06/18/2002	White *et al.*	704	270.1	
	5E	6,411,810	06/25/2002	Maxemchuk	455	453	
	5F	6,415,257	07/02/2002	Junqua *et al.*	704	275	
	5G	6,418,210	07/09/2002	Sayko	379	142.15	
	5H	6,420,975	07/16/2002	DeLine *et al.*	340	815.4	
	5I	6,430,285	08/06/2002	Bauer et al.	379	265.01	
	5J	6,430,531	08/06/2002	Polish	704	257	
	5K	6,434,523	08/13/2002	Monaco	704	257	
	5L	6,434,524	08/13/2002	Weber	704	257	
	5M	6,442,522	08/27/2002	Carberry et al.	704	257	
	5N	6,446,114	09/03/2002	Bulfer et al.	709	206	
	5O	6,453,153	09/17/2002	Bowker et al.	455	67.4	
	5P	6,453,292	09/17/2002	Ramaswamy et al.	704	235	
	5Q	6,456,711	09/24/2002	Cheung et al.	379	265.09	
	5R	6,466,654	10/15/2002	Cooper *et al.*	379	88.01	
	5S	6,466,899	10/15/2002	Yano *et al.*	704	1	
	5T	6,470,315	10/22/2002	Netsch *et al.*	704	256	
	5U	6,498,797	12/24/2002	Anerousis et al.	370	522	
	5V	6,499,013	12/24/2002	Weber	704	257	
	5W	6,501,833	12/31/2002	Phillips *et al.*	379	88.07	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

	5X	MAO, Mark Z., "Automatic Training Set Segmentation for Multi-Pass Speech Recognition", Department of Electrical Engineering, Stanford University, CA, copyright 2005, IEEE, pages I-685 to I-688.
	5Y	VANHOUCKE, Vincent, "Confidence Scoring and Rejection Using Multi-Pass Speech Recognition", Nuance Communications, Menlo Park, CA, 2005, 4 pages.

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

402465373v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	*Unassigned*
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	Herewith (November 10, 2010)	*Unassigned*

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	6A	6,501,834	12/31/2002	Milewski et al.	379	93.24	
	6B	6,513,006	01/28/2003	Howard *et al.*	704	257	
	6C	6,523,061	02/18/2003	Halverson et al.	709	202	
	6D	6,532,444	03/11/2003	Weber	704	257	
	6E	6,539,348	03/25/2003	Bond et al.	704	9	
	6F	6,553,372	04/22/2003	Brassell et al.	707	5	
	6G	6,556,970	04/29/2003	Sasaki *et al.*	704	257	
	6H	6,556,973	04/29/2003	Lewin	704	277	
	6I	6,560,576	05/06/2003	Cohen et al.	704	270	
	6J	6,567,778	05/20/2003	Chao Chang et al.	704	257	
	6K	6,567,797	05/20/2003	Schuetze *et al.*	707	2	
	6L	6,570,555	05/27/2003	Prevost *et al.*	345	156	
	6M	6,570,964	05/27/2003	Murveit et al.	379	67.1	
	6N	6,574,597	06/03/2003	Mohri *et al.*	704	251	
	6O	6,574,624	06/03/2003	Johnson et al.	707	5	
	6P	6,581,103	06/17/2003	Dengler	709	231	
	6Q	6,594,257	07/15/2003	Doshi et al.	370	352	
	6R	6,598,018	07/22/2003	Junqua	704	251	
	6S	6,604,075	08/05/2003	Brown *et al.*	704	270.1	
	6T	6,604,077	08/05/2003	Dragosh et al.	704	270.1	
	6U	6,611,692	08/26/2003	Raffel et al.	455	552	
	6V	6,614,773	09/02/2003	Maxemchuk	370	337	
	6W	6,615,172	09/02/2003	Bennett *et al.*	704	257	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

	6X	WENG, Fuliang, *et al.*, "Efficient Lattice Representation and Generation", Speech Technology and Research Laboratory, SRI International, Menlo Park, CA, 1998, 4 pages.
	6Y	EL MELIANI *et al.*, "A Syllabic-Filler-Based Continuous Speech Recognizer for Unlimited Vocabulary", Canadian Conference on Electrical and Computer Engineering, Volume 2, September 5-8, 1995, pages 1007-1010.

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

402465373v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	*Unassigned*
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	Herewith (November 10, 2010)	*Unassigned*

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	7A	6,629,066	09/30/2003	Jackson et al.	704	9	
	7B	6,631,346	10/07/2003	Karaorman *et al.*	704	9	
	7C	6,643,620	11/04/2003	Contolini *et al.*	704	270	
	7D	6,650,747	11/18/2003	Bala et al.	379	265.06	
	7E	6,658,388	12/02/2003	Kleindienst *et al.*	704	275	
	7F	6,678,680	01/13/2004	Woo	707	6	
	7G	6,691,151	02/10/2004	Cheyer et al.	709	202	
	7H	6,708,150	03/16/2004	Hirayama *et al.*	704	243	
	7I	6,721,001	04/13/2004	Berstis	348	231.3	
	7J	6,721,706	04/13/2004	Strubbe *et al.*	704	275	
	7K	6,735,592	05/11/2004	Neumann *et al.*	707	101	
	7L	6,741,931	05/25/2004	Kohut *et al.*	701	209	
	7M	6,742,021	05/25/2004	Halverson *et al.*	709	218	
	7N	6,751,591	06/15/2004	Gorin *et al.*	704	257	
	7O	6,757,718	06/29/2004	Halverson *et al.*	709	218	
	7P	6,795,808	09/21/2004	Strubbe *et al.*	704	275	
	7Q	6,801,604	10/05/2004	Maes *et al.*	379	88.17	
	7R	6,801,893	10/05/2004	Backfried *et al.*	704	257	
	7S	6,829,603	12/07/2004	Chai *et al.*	707	5	
	7T	6,833,848	12/21/2004	Wolff *et al.*	345	719	
	7U	6,856,990	02/15/2005	Barile *et al.*	707	10	
	7V	6,865,481	03/08/2005	Kawazoe *et al.*	701	211	
	7W	6,877,134	04/05/2005	Fuller *et al.*	715	500.1	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

	7X	ARRINGTON, Michael, "Google Redefines GPS Navigation Landscape: Google Maps Navigation for Android 2.0", TechCrunch, printed from the Internet <http://www.techcrunch.com/2009/10/28/google-redefines-car-gps-navigation-google-maps-navigation-android/>, October 28, 2009, 4 pages.
	7Y	BAZZI, Issam *et al.*, "Heterogeneous Lexical Units for Automatic Speech Recognition: Preliminary Investigations", Processing of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Volume 3, June 5-9, 2000, XP010507574, pages 1257-1260.

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

402465373v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	*Unassigned*
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	Herewith (November 10, 2010)	*Unassigned*

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	8A	6,901,366	05/31/2005	Kuhn *et al.*	704	275	
	8B	6,937,977	08/30/2005	Gerson	704	201	
	8C	6,944,594	09/13/2005	Busayapongchai *et al.*	704	275	
	8D	6,950,821	09/27/2005	Faybishenko *et al.*	707	10	
	8E	6,973,387	12/06/2005	Masclet *et al.*	701	211	
	8F	6,980,092	12/27/2005	Turnbull *et al.*	340	425.5	
	8G	6,990,513	01/24/2006	Belfiore *et al.*	709	203	
	8H	6,996,531	02/07/2006	Korall *et al.*	704	270	
	8I	7,016,849	03/21/2006	Arnold *et al.*	704	275	
	8J	7,020,609	03/28/2006	Thrift *et al.*	704	270.1	
	8K	7,027,975	04/11/2006	Pazandak *et al.*	704	9	
	8L	7,062,488	06/13/2006	Reisman	707	8	
	8M	7,082,469	07/25/2006	Gold *et al.*	709	231	
	8N	7,092,928	08/15/2006	Elad *et al.*	706	60	
	8O	7,127,400	10/24/2006	Koch	704	270.1	
	8P	7,136,875	11/14/2006	Anderson *et al.*	707	104.1	
	8Q	7,137,126	11/14/2006	Coffman *et al.*	719	328	
	8R	7,146,319	12/05/2006	Hunt	704	254	
	8S	7,203,644	04/10/2007	Anderson *et al.*	704	246	
	8T	7,228,276	06/05/2007	Omote *et al.*	704	243	
	8U	7,231,343	06/12/2007	Treadgold *et al.*	704	9	
	8V	7,289,606	10/30/2007	Sibal *et al.*	379	52	
	8W	7,301,093	11/27/2007	Sater *et al.*	84	615	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

	8X	O'SHAUGHNESSY, Douglas, "Interacting with Computers by Voice: Automatic Speech Recognition and Synthesis", Proceedings of the IEEE, Volume 91, No. 9, September 1, 2003, XP011100665, pages 1272-1305.

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

402465373v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	*Unassigned*
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	Herewith (November 10, 2010)	*Unassigned*

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	9A	7,340,040	03/04/2008	Saylor *et al.*	379	67.1	
	9B	7,376,645	05/20/2008	Bernard	707	3	
	9C	7,386,443	06/10/2008	Parthasarathy *et al.*	704	201	
	9D	7,398,209	07/08/2008	Kennewick *et al.*	704	255	
	9E	7,424,431	09/09/2008	Greene *et al.*	704	270	
	9F	7,461,059	12/02/2008	Richardson *et al.*	707	5	
	9G	7,487,088	02/03/2009	Gorin *et al.*	704	240	
	9H	7,493,559	02/17/2009	Wolff *et al.*	715	727	
	9I	7,502,738	03/10/2009	Kennewick *et al.*	704	257	
	9J	7,620,549	11/17/2009	Di Cristo *et al.*	704	257	
	9K	7,634,409	12/15/2009	Kennewick *et al.*	704	257	
	9L	7,640,160	12/29/2009	Di Cristo *et al.*	704	257	
	9M	7,693,720	04/06/2010	Kennewick *et al.*	704	275	
	9N	7,809,570	10/05/2010	Kennewick *et al.*	704	257	
	9O	7,818,176	10/19/2010	Freeman *et al.*			
	9P	2001/0041980	11/15/2001	Howard *et al.*	704	270	
	9Q	2002/0049805	04/25/2002	Yamada *et al.*	709	202	
	9R	2002/0065568	05/30/2002	Silfvast *et al.*	700	94	
	9S	2002/0069059	06/06/2002	Smith	704	257	
	9T	2002/0082911	06/27/2002	Dunn *et al.*	705	14	
	9U	2002/0087525	07/04/2002	Abbott *et al.*	707	3	
	9V	2002/0120609	08/29/2002	Lang *et al.*	707	1	
	9W	2002/0124050	09/05/2002	Middeljans	709	203	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

402465373v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	*Unassigned*
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	Herewith (November 10, 2010)	*Unassigned*

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	10A	2002/0138248	09/26/2002	Corston-Oliver *et al.*	704	1	
	10B	2002/0143535	10/03/2002	Kist *et al.*	704	251	
	10C	2002/0188602	12/12/2002	Stubler *et al.*	707	3	
	10D	2002/0198714	12/26/2002	Zhou	704	252	
	10E	2003/0014261	01/16/2003	Kageyama	704	275	
	10F	2003/0046346	03/06/2003	Mumick *et al.*	709	205	
	10G	2003/0064709	04/03/2003	Gailey *et al.*	455	412	
	10H	2003/0088421	05/08/2003	Maes *et al.*	704	270.1	
	10I	2003/0097249	05/22/2003	Walker *et al.*	704	1	
	10J	2003/0112267	06/19/2003	Belrose	345	728	
	10K	2003/0135488	07/17/2003	Amir *et al.*	707	3	
	10L	2003/0182132	09/25/2003	Niemoeller	704	275	
	10M	2003/0236664	12/25/2003	Sharma	704	251	
	10N	2004/0006475	01/08/2004	Ehlen *et al.*	704	270.1	
	10O	2004/0025115	02/05/2004	Sienel *et al.*	715	513	
	10P	2004/0044516	03/04/2004	Kennewick *et al.*	704	5	
	10Q	2004/0166832	08/26/2004	Portman *et al.*	455	412.1	
	10R	2004/0193408	09/30/2004	Hunt	704	209	
	10S	2004/0193420	09/30/2004	Kennewick *et al.*	704	257	
	10T	2004/0199375	10/07/2004	Ehsani *et al.*	704	4	
	10U	2004/0243417	12/02/2004	Pitts, III *et al.*	704	276	
	10V	2005/0015256	01/20/2005	Kargman	704	272	
	10W	2005/0021334	01/27/2005	Iwahashi	704	240	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

402465373v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	*Unassigned*
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	Herewith (November 10, 2010)	*Unassigned*

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	11A	2005/0021826	01/27/2005	Kumar	709	232	
	11B	2005/0033574	02/10/2005	Kim *et al.*	704	251	
	11C	2005/0043940	02/24/2005	Elder	704	9	
	11D	2005/0114116	05/26/2005	Fiedler	704	201	
	11E	2005/0137850	06/23/2005	Odell	704	4	
	11F	2005/0246174	11/03/2005	DeGolia	704	270	
	11G	2006/0206310	09/14/2006	Ravikumar *et al.*	704	9	
	11H	2007/0033005	02/08/2007	Cristo *et al.*	704	9	
	11I	2007/0038436	02/15/2007	Cristo *et al.*	704	9	
	11J	2007/0043574	02/22/2007	Coffman *et al.*	704	275	
	11K	2007/0050191	03/01/2007	Weider *et al.*	704	275	
	11L	2007/0055525	03/08/2007	Kennewick *et al.*	704	257	
	11M	2007/0179778	08/02/2007	Gong *et al.*	704	9	
	11N	2007/0186165	08/09/2007	Maislos *et al.*	715	728	
	11O	2007/0214182	09/13/2007	Rosenberg	707	104.1	
	11P	2007/0250901	10/25/2007	McIntire *et al.*	725	146	
	11Q	2007/0265850	11/15/2007	Kennewick *et al.*	704	257	
	11R	2007/0299824	12/27/2007	Pan *et al.*	707	3	
	11S	2008/0091406	04/17/2008	Baldwin *et al.*	704	4	
	11T	2008/0103761	05/01/2008	Printz *et al.*	704	9	
	11U	2008/0115163	05/15/2008	Gilboa *et al.*	725	34	
	11V	2008/0133215	06/05/2008	Sarukkai	704	2	
	11W	2008/0177530	07/24/2008	Cross *et al.*	704	4	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

402465373v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	*Unassigned*
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	Herewith (November 10, 2010)	*Unassigned*

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	12A	2008/0189110	08/07/2008	Freeman *et al.*	704	251	
	12B	2008/0235023	09/25/2008	Kennewick *et al.*	704	257	
	12C	2008/0235027	09/25/2008	Cross	704	270.1	
	12D	2008/0319751	12/25/2008	Kennewick *et al.*	704	257	
	12E	2009/0117885	05/07/2009	Roth	455	414.3	
	12F	2009/0144271	06/04/2009	Richardson *et al.*	707	5	
	12G	2009/0150156	06/11/2009	Kennewick *et al.*	704	257	
	12H	2009/0171664	07/02/2009	Kennewick *et al.*	704	257	
	12I	2009/0299745	12/03/2009	Kennewick *et al.*	704	257	
	12J	2010/0023320	01/28/2010	Di Cristo *et al.*	704	9	
	12K	2010/0049501	02/25/2010	Kennewick *et al.*	704	9	
	12L	2010/0049514	02/25/2010	Kennewick *et al.*	704	233	
	12M	2010/0057443	03/04/2010	Di Cristo *et al.*	704	9	
	12N	2010/0063880	03/11/2010	Atsmon *et al.*	705	14.53	
	12O	2010/0145700	06/10/2010	Kennewick *et al.*	704	257	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED
/Tim Hale/	/T.H./

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

402465373v1

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /T.H./

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	12943699	KENNEWICK ET AL.
	Examiner	**Art Unit**
	TIM HALE	3685

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	09/18/2012								
	1	✓								
	2	✓								
	3	✓								
	4	✓								
	5	✓								
	6	✓								
	7	✓								
	8	✓								
	9	✓								
	10	✓								
	11	✓								
	12	✓								
	13	✓								
	14	✓								
	15	✓								
	16	✓								
	17	✓								
	18	✓								
	19	✓								
	20	✓								

Amazon/VB Assets
Exhibit 1003
Page 422

UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

PILLSBURY WINTHROP SHAW PITTMAN, LLP (NV)
PO BOX 10500
MCLEAN VA 22102

MAILED
JUL 02 2012
OFFICE OF PETITIONS

In re Application of
Kennewick et al.
Application No. 12/943,699
Filed: November 10, 2010
Attorney Docket No. 090611-0384422

DECISION ON PETITION
TO WITHDRAW FROM RECORD

This is a decision on the Request to Withdraw as attorney or agent of record under 37 C.F.R. § 1.36(b), filed June 6, 2012.

The request is **APPROVED**.

A grantable request to withdraw as attorney/agent of record must be signed by every attorney/agent seeking to withdraw or contain a clear indication that one attorney is signing on behalf of another/others. The Office will require the practitioner(s) to certify that he, she or they have: (1) given reasonable notice to the client, prior to the expiration of the reply period, which the practitioner(s) intends to withdraw from employment; (2) delivered to the client or a duly authorized representative of the client all papers and property (including funds) to which the client is entitled; and (3) notified the client of any replies that may be due and the time frame within which the client must respond, pursuant to 37 CFR 10.40 (c).

The request was signed by Timothy C. Rooney on behalf of himself. Accordingly, Timothy C. Rooney associated with the above-identified application has been withdrawn as attorney of record in the above-identified application.

The correspondence address of record remains unchanged.

Additionally, the address given on the petition differs from the address of record. A courtesy copy of this decision is being mailed to the address given on the petition; however, the Office will mail all future correspondence solely to the address of record.

Currently, there is no outstanding Office that requires a reply.

Telephone inquiries concerning this decision should be directed to the undersigned at (571) 272-7751. All other inquiries concerning either the examination or status of the application should be directed to the Technology Center.

/Joan Olszewski/
Joan Olszewski
Petitions Examiner
Office of Petitions

cc: Timothy C. Rooney
600 Dulany St.
Mail Stop OED
Alexandria, Virginia 22314

OPAP JUN 06 2012 PATENT & TRADEMARK OFFICE

Doc Code: PET.POA.WDRW

Document Description: Petition to withdraw attorney or agent (SB83)

PTO/SB/83 (11-08)
Approved for use through 11/30/2011. OMB 0651-0035
U.S. Patent and Trademark Office, U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

REQUEST FOR WITHDRAWAL AS ATTORNEY OR AGENT AND CHANGE OF CORRESPONDENCE ADDRESS

Application Number	12/943,699
Filing Date	November 10, 2010
First Named Inventor	Mike Kennewick
Art Unit	3625
Examiner Name	Adam Levine
Attorney Docket Number	090611-0384422

To: Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Please withdraw me as attorney or agent for the above identified patent application, and

- ☐ all the practitioners of record;
- ☐ the practitioners (with registration numbers) of record listed on the attached paper(s); or
- ☐ the practitioners of record associated with Customer Number: _____

NOTE: The immediately preceding box should only be marked when the practitioners were appointed using the listed Customer Number.

The reason(s) for this request are those described in 37 CFR :

☐ 10.40(b)(1)	☐ 10.40(b)(2)	☐ 10.40(b)(3)	☐ 10.40(b)(4)
☐ 10.40(c)(1)(i)	☐ 10.40(c)(1)(ii)	☐ 10.40(c)(1)(iii)	☐ 10.40(c)(1)(iv)
☐ 10.40(c)(1)(v)	☐ 10.40(c)(1)(vi)	☐ 10.40(c)(2)	☐ 10.40(c)(3)
☐ 10.40(c)(4)	☐ 10.40(c)(5)	☑ 10.40(c)(6) Please explain below:	

Departed handling firm for employment at USPTO.

Certifications

Check each box below that is factually correct. WARNING: If a box is left unchecked, the request will likely not be approved.

1. ☑ I/We have given reasonable notice to the client, prior to the expiration of the response period, that the practitioner(s) intend to withdraw from employment.

2. ☑ I/We have delivered to the client or a duly authorized representative of the client all papers and property (including funds) to which the client is entitled.

3. ☑ I/We have notified the client of any responses that may be due and the time frame within which the client must respond.

Please provide an explanation, if necessary:

This collection of information is required by 37 CFR 1.36. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

PTO/SB/83 (11-08)
Approved for use through 11/30/2011. OMB 0651-0035
U.S. Patent and Trademark Office, U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

REQUEST FOR WITHDRAWAL AS ATTORNEY OR AGENT AND CHANGE OF CORRESPONDENCE ADDRESS

Complete the following section only when the correspondence address will change. *Changes of address will only be accepted to an inventor or an assignee that has properly made itself of record pursuant to 37 CFR 3.71.*

Change the correspondence address and direct all future correspondence to:

A. ☐ The address of the inventor or assignee associated with Customer Number: _____

OR

B. ☐ Inventor or Assignee name	

Address			
City	State	Zip	Country
Telephone		Email	

I am authorized to sign on behalf of myself and all withdrawing practitioners.

Signature			
Name	Timothy C. Rooney	Registration No.	55,878
Address	600 Dulany St., Mail Stop OED		
City Alexandria	State Virginia	Zip 22314	Country United States
Date	June 6, 2012	Telephone No.	571.270.7965

NOTE: ***Withdrawal is effective when approved rather than when received.***

This collection of information is required by 37 CFR 1.36. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY.DOCKET.NO	TOT CLAIMS	IND CLAIMS
12/943,699	11/10/2010	3625	527	090611-0384422	20	2

CONFIRMATION NO. 4185

CORRECTED FILING RECEIPT

909
PILLSBURY WINTHROP SHAW PITTMAN, LLP
P.O. BOX 10500
MCLEAN, VA 22102

OC000000048223795

Date Mailed: 06/16/2011

Receipt is acknowledged of this non-provisional patent application. The application will be taken up for examination in due course. Applicant will be notified as to the results of the examination. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please submit a written request for a Filing Receipt Correction. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections**

Applicant(s)

Mike Kennewick, Bellevue, WA;
Lynn Elise Armstrong, Woodinville, WA;

Assignment For Published Patent Application

VoiceBox Technologies, Inc., Bellevue, WA

Power of Attorney:
James Gatto--32694
Bradford Blaise--47429
Timothy Rooney--55878
Syed Ali--58780

Domestic Priority data as claimed by applicant

This appln claims benefit of 61/259,820 11/10/2009

Foreign Applications (You may be eligible to benefit from the **Patent Prosecution Highway** program at the USPTO. Please see http://www.uspto.gov for more information.)

If Required, Foreign Filing License Granted: 11/22/2010

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 12/943,699**

Projected Publication Date: Not Applicable

Non-Publication Request: No

Early Publication Request: No
**** SMALL ENTITY ****

Title

SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE

Preliminary Class

705

PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at http://www.uspto.gov/web/offices/pac/doc/general/index.html.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, http://www.stopfakes.gov. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4158).

LICENSE FOR FOREIGN FILING UNDER

Title 35, United States Code, Section 184

Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where

the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign AssetsControl, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANTS :	Mike KENNEWICK *et al.*	CONFIRMATION NO. :	4185
SERIAL NUMBER :	12/943,699	EXAMINER :	Adam L. Levine
FILING DATE :	November 10, 2010	ART UNIT :	3625
FOR :	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE		

REQUEST FOR CORRECTED FILING RECEIPT

Office of Initial Patent Examination's
Filing Receipt Corrections
Commissioner for Patents
P. O. Box 1450
Alexandria, VA 22313-1450

Sir:

Attached is a copy of the Updated Filing Receipt received from the U.S. Patent and Trademark Office in the above-referenced application for which issuance of a corrected filing receipt is respectfully requested. There is an error with respect to the following data which is inaccurate:

Error in	**Correct Data**
Second Applicant's Name	Lynn Elise Armstrong

The correct middle name of the second inventor is "Elise", as can be seen from the Application Data Sheet filed on November 10, 2010, and the Combined Declaration and Power

of Attorney for Patent Application filed January 26, 2011. Accordingly, a Corrected Filing Receipt is respectfully requested. A marked-up copy of the Updated Filing Receipt is attached.

Although Applicants believe no fee is due in connection with this filing, the Commissioner is authorized to charge any additional fees that may be due to Deposit Account No. 03-3975, Ref. No. 090611-0384422.

Respectfully submitted,

PILLSBURY WINTHROP SHAW PITTMAN LLP

By _____
Syed Jafar Ali
Registration No.: 58,780

Customer No.
00909

P. O. Box 10500
McLean, Virginia 22102
Telephone: (703) 770-7900
Telefax: (703) 770-7901

Dated: June 14, 2011

Amazon/VB Assets
Exhibit 1003
Page 431

COPY

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY.DOCKET.NO	TOT CLAIMS	IND CLAIMS
12/943,699	11/10/2010	2447	527	090611-0384422	20	2

CONFIRMATION NO. 4185

909
PILLSBURY WINTHROP SHAW PITTMAN, LLP
P.O. BOX 10500
MCLEAN, VA 22102

UPDATED FILING RECEIPT

OC00000045799799

Date Mailed: 02/04/2011

Receipt is acknowledged of this non-provisional patent application. The application will be taken up for examination in due course. Applicant will be notified as to the results of the examination. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please submit a written request for a Filing Receipt Correction. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections**

Applicant(s)

Mike Kennewick, Bellevue, WA;
~~Lynn Elis Armstrong~~, Woodinville, WA; Lynn Elise Armstrong

Assignment For Published Patent Application

VoiceBox Technologies, Inc., Bellevue, WA

Power of Attorney:
James Gatto--32694
Bradford Blaise--47429
Timothy Rooney--55878
Syed Ali--58780

Domestic Priority data as claimed by applicant

This appln claims benefit of 61/259,820 11/10/2009

Foreign Applications (You may be eligible to benefit from the **Patent Prosecution Highway** program at the USPTO. Please see http://www.uspto.gov for more information.)

If Required, Foreign Filing License Granted: 11/22/2010

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 12/943,699**

Projected Publication Date: 05/12/2011

Non-Publication Request: No

Early Publication Request: No
**** SMALL ENTITY ****

Amazon/VB Assets
Exhibit 1003
Page 432

Title

SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE

Preliminary Class

709

PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at http://www.uspto.gov/web/offices/pac/doc/general/index.html.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, http://www.stopfakes.gov. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4158).

LICENSE FOR FOREIGN FILING UNDER

Title 35, United States Code, Section 184

Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where

Amazon/VB Assets
Exhibit 1003
Page 433

the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign AssetsControl, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

Electronic Acknowledgement Receipt	
EFS ID:	10296335
Application Number:	12943699
International Application Number:	
Confirmation Number:	4185
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Customer Number:	00909
Filer:	Syed Jafar Ali/Susan Trader
Filer Authorized By:	Syed Jafar Ali
Attorney Docket Number:	090611-0384422
Receipt Date:	14-JUN-2011
Filing Date:	10-NOV-2010
Time Stamp:	11:14:25
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Request for Corrected Filing Receipt	0384422_Request.pdf	201361 c25f6def20e387fb502b692d295591482ab65a3f	no	5
Warnings:					
Information:					

Total Files Size (in bytes):	201361

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
12/943,699	11/10/2010	Mike Kennewick	090611-0384422

CONFIRMATION NO. 4185

909
PILLSBURY WINTHROP SHAW PITTMAN, LLP
P.O. BOX 10500
MCLEAN, VA 22102

PUBLICATION NOTICE

OC000000047631223

Title:SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE

Publication No.US-2011-0112921-A1
Publication Date:05/12/2011

NOTICE OF PUBLICATION OF APPLICATION

The above-identified application will be electronically published as a patent application publication pursuant to 37 CFR 1.211, et seq. The patent application publication number and publication date are set forth above.

The publication may be accessed through the USPTO's publically available Searchable Databases via the Internet at www.uspto.gov. The direct link to access the publication is currently http://www.uspto.gov/patft/.

The publication process established by the Office does not provide for mailing a copy of the publication to applicant. A copy of the publication may be obtained from the Office upon payment of the appropriate fee set forth in 37 CFR 1.19(a)(1). Orders for copies of patent application publications are handled by the USPTO's Office of Public Records. The Office of Public Records can be reached by telephone at (703) 308-9726 or (800) 972-6382, by facsimile at (703) 305-8759, by mail addressed to the United States Patent and Trademark Office, Office of Public Records, Alexandria, VA 22313-1450 or via the Internet.

In addition, information on the status of the application, including the mailing date of Office actions and the dates of receipt of correspondence filed in the Office, may also be accessed via the Internet through the Patent Electronic Business Center at www.uspto.gov using the public side of the Patent Application Information and Retrieval (PAIR) system. The direct link to access this status information is currently http://pair.uspto.gov/. Prior to publication, such status information is confidential and may only be obtained by applicant using the private side of PAIR.

Further assistance in electronically accessing the publication, or about PAIR, is available by calling the Patent Electronic Business Center at 1-866-217-9197.

Office of Data Managment, Application Assistance Unit (571) 272-4000, or (571) 272-4200, or 1-888-786-0101

Amazon/VB Assets
Exhibit 1003
Page 437

PATENT APPLICATION FEE DETERMINATION RECORD

Substitute for Form PTO-875

Application or Docket Number: 12/943,699

APPLICATION AS FILED - PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	SMALL ENTITY RATE($)	SMALL ENTITY FEE($)	OR	OTHER THAN SMALL ENTITY RATE($)	OTHER THAN SMALL ENTITY FEE($)
BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A	82		N/A	
SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A	270		N/A	
EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A	110		N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	20 minus 20 =	*	x 26 =	0.00	OR		
INDEPENDENT CLAIMS (37 CFR 1.16(h))	2 minus 3 =	*	x 110 =	0.00			
APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $270 ($135 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).			0.00			
MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))				0.00			
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL	462		TOTAL	

APPLICATION AS AMENDED - PART II

AMENDMENT A

	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	SMALL ENTITY RATE($)	SMALL ENTITY ADDITIONAL FEE($)	OR	OTHER THAN SMALL ENTITY RATE($)	OTHER THAN SMALL ENTITY ADDITIONAL FEE($)
Total (37 CFR 1.16(i))	*	Minus	**	=	x =		OR	x =	
Independent (37 CFR 1.16(h))	*	Minus	***	=	x =		OR	x =	
Application Size Fee (37 CFR 1.16(s))									
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
					TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

AMENDMENT B

	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE($)	ADDITIONAL FEE($)		RATE($)	ADDITIONAL FEE($)
Total (37 CFR 1.16(i))	*	Minus	**	=	x =		OR	x =	
Independent (37 CFR 1.16(h))	*	Minus	***	=	x =		OR	x =	
Application Size Fee (37 CFR 1.16(s))									
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
					TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".
The "Highest Number Previously Paid For" (Total or Independent) is the highest found in the appropriate box in column 1.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY.DOCKET.NO	TOT CLAIMS	IND CLAIMS
12/943,699	11/10/2010	2447	527	090611-0384422	20	2

CONFIRMATION NO. 4185

909
PILLSBURY WINTHROP SHAW PITTMAN, LLP
P.O. BOX 10500
MCLEAN, VA 22102

UPDATED FILING RECEIPT

OC000000045799799

Date Mailed: 02/04/2011

Receipt is acknowledged of this non-provisional patent application. The application will be taken up for examination in due course. Applicant will be notified as to the results of the examination. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please submit a written request for a Filing Receipt Correction. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections**

Applicant(s)

Mike Kennewick, Bellevue, WA;
Lynn Elis Armstrong, Woodinville, WA;

Assignment For Published Patent Application

VoiceBox Technologies, Inc., Bellevue, WA

Power of Attorney:
James Gatto--32694
Bradford Blaise--47429
Timothy Rooney--55878
Syed Ali--58780

Domestic Priority data as claimed by applicant

This appln claims benefit of 61/259,820 11/10/2009

Foreign Applications (You may be eligible to benefit from the **Patent Prosecution Highway** program at the USPTO. Please see http://www.uspto.gov for more information.)

If Required, Foreign Filing License Granted: 11/22/2010

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 12/943,699**

Projected Publication Date: 05/12/2011

Non-Publication Request: No

Early Publication Request: No
**** SMALL ENTITY ****

Title

SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE

Preliminary Class

709

PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at http://www.uspto.gov/web/offices/pac/doc/general/index.html.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, http://www.stopfakes.gov. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4158).

LICENSE FOR FOREIGN FILING UNDER

Title 35, United States Code, Section 184

Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where

the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign AssetsControl, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

Attorney Docket No. 090611-0384422

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of: Mike Kennewick, et al.

Application No.:	12/943,699	Confirmation No.:	4185
Filed:	November 10, 2010	Group No.:	2447
		Examiner:	Unassigned

For: SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE

RESPONSE TO NOTICE TO FILE MISSING PARTS

Mail Stop Missing Parts
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

In response to the Notice to File Missing Parts of Application mailed November 29, 2010, a copy of which is attached, Applicants herewith submit the following:

☒ **Executed Combined Declaration and Power of Attorney of Inventors**

Please charge any fees associated with the submission of this paper to Deposit Account Number 033975. The Commissioner for Patents is also authorized to credit any over payments to the above-referenced Deposit Account.

Date: January 26, 2011

Respectfully submitted,

By _____

Syed Jafar Ali
Registration No. 58,780
Telephone: (703) 770-7540
PILLSBURY WINTHROP SHAW PITTMAN LLP
P.O. Box 10500
McLean, VA 22102
Customer No.: 00909

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
12/943,699	11/10/2010	Mike Kennewick	090611-0384422

CONFIRMATION NO. 4185

909
PILLSBURY WINTHROP SHAW PITTMAN, LLP
P.O. BOX 10500
MCLEAN, VA 22102

FORMALITIES LETTER

OC000000044624812

Date Mailed: 11/29/2010

NOTICE TO FILE MISSING PARTS OF NONPROVISIONAL APPLICATION

FILED UNDER 37 CFR 1.53(b)

Filing Date Granted

Items Required To Avoid Abandonment:

An application number and filing date have been accorded to this application. The item(s) indicated below, however, are missing. Applicant is given **TWO MONTHS** from the date of this Notice within which to file all required items below to avoid abandonment. Extensions of time may be obtained by filing a petition accompanied by the extension fee under the provisions of 37 CFR 1.136(a).

- The oath or declaration is missing.
  A properly signed oath or declaration in compliance with 37 CFR 1.63, identifying the application by the above Application Number and Filing Date, is required.
  Note: If a petition under 37 CFR 1.47 is being filed, an oath or declaration in compliance with 37 CFR 1.63 signed by all available joint inventors, or if no inventor is available by a party with sufficient proprietary interest, is required.

The applicant needs to satisfy supplemental fees problems indicated below.

The required item(s) identified below must be timely submitted to avoid abandonment:

- A surcharge (for late submission of filing fee, search fee, examination fee or oath or declaration) as set forth in 37 CFR 1.16(f) of **$65** for a small entity in compliance with 37 CFR 1.27, must be submitted.

SUMMARY OF FEES DUE:

Total fee(s) required within **TWO MONTHS** from the date of this Notice is **$65** for a small entity
- **$65** Surcharge.

Amazon/VB Assets
Exhibit 1003
Page 443

Replies should be mailed to:

Mail Stop Missing Parts
Commissioner for Patents
P.O. Box 1450
Alexandria VA 22313-1450

Registered users of EFS-Web may alternatively submit their reply to this notice via EFS-Web.
https://sportal.uspto.gov/authenticate/AuthenticateUserLocalEPF.html

For more information about EFS-Web please call the USPTO Electronic Business Center at **1-866-217-9197** or visit our website at http://www.uspto.gov/ebc.

If you are not using EFS-Web to submit your reply, you must include a copy of this notice.

/atesfaye/

Office of Data Management, Application Assistance Unit (571) 272-4000, or (571) 272-4200, or 1-888-786-0101

Amazon/VB Assets
Exhibit 1003
Page 444

COMBINED DECLARATION AND POWER OF ATTORNEY FOR PATENT APPLICATION

As the below named inventors, we hereby declare that:

Our residences, post office addresses, and citizenship are as stated next to our names;

We believe that we are the original, first and joint inventors of the subject matter which is claimed and for which a patent is sought on the invention entitled:

"SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE"

the specification of which was filed on November 10, 2010, as U.S. Patent Application Serial No. 12/943,699.

We have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above;

We acknowledge the duty to disclose to the U.S. Patent and Trademark Office all information which is known to us to be material to the patentability of said invention in accordance with 37 C.F.R. § 1.56;

We do not know and do not believe said invention was ever known or used in the United States of America before our invention thereof, or patented or described in any printed publication in any country before our invention thereof or more than one year prior to said application; that said invention was not in public use or on sale in the United States of America more than one year prior to said application; that said invention has not been patented or made the subject of an inventor's certificate issued before the date of said application in any country foreign to the United States of America on any application filed by us or our legal representatives or assigns more than six months prior to said application;

We hereby claim the benefit under 35 U.S.C. § 119(e) of any United States provisional application(s) listed below:

61/259,820	November 10, 2009
(Application Number)	(Filing Date)

401509814v1

We hereby appoint the following attorneys and/or agents to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith:

James G. Gatto	Registration No. 32,694
Bradford C. Blaise	Registration No. 47,429
Timothy C. Rooney	Registration No. 55,878
Syed Jafar Ali	Registration No. 58,780

Address all telephone calls to James G. Gatto at telephone number 703.770.7754. Address all correspondence to Customer Number:

00909

We hereby declare that all statements made herein of our own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of first inventor: **Mike Kennewick**

Inventor's signature: _____ Date: _____

Residence: 13407 37th Place, NE
Bellevue, Washington 98005

Citizen of: United States

Post Office Address: Same as residence

Full name of second inventor: **Lynn Elise Armstrong**

Inventor's signature: _____ Date: 1-25-2018

Residence: 19807 183rd Place, NE
Woodinville, Washington 98077

Citizen of: United States

Post Office Address: Same as residence

Electronic Patent Application Fee Transmittal

Application Number:	12943699
Filing Date:	10-Nov-2010
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Filer:	Syed Jafar Ali/Miriam E. Nelsen
Attorney Docket Number:	090611-0384422

Filed as Small Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Late filing fee for oath or declaration	2051	1	65	65
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
			Total in USD ($)	**65**

Electronic Acknowledgement Receipt	
EFS ID:	9309838
Application Number:	12943699
International Application Number:	
Confirmation Number:	4185
Title of Invention:	SYSTEM AND METHOD FOR PROVIDING A NATURAL LANGUAGE CONTENT DEDICATION SERVICE
First Named Inventor/Applicant Name:	Mike Kennewick
Customer Number:	00909
Filer:	Syed Jafar Ali/Miriam E. Nelsen
Filer Authorized By:	Syed Jafar Ali
Attorney Docket Number:	090611-0384422
Receipt Date:	26-JAN-2011
Filing Date:	10-NOV-2010
Time Stamp:	13:33:58
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	$65
RAM confirmation Number	678
Deposit Account	033975
Authorized User	
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: Charge any Additional Fees required under 37 C.F.R. Section 1.21 (Miscellaneous fees and charges)	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		090611-0384422_ResponseNTFMPs.pdf	163432 6a7f3a31837a4a16f5ea493b884bdac4795f05f0	yes	5

Multipart Description/PDF files in .zip description		
Document Description	**Start**	**End**
Applicant Response to Pre-Exam Formalities Notice	1	3
Oath or Declaration filed	4	5

Warnings:

Information:

2	Fee Worksheet (PTO-875)	fee-info.pdf	30410 6da4508d2f6922b51822f535c08a4b3cbadf1ed3	no	2

Warnings:

Information:

Total Files Size (in bytes):	193842

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

INFORMATION DISCLOSURE CITATION IN AN APPLICATION	Docket Number	Application Number
	090611-0384422	12/943,699
	Applicant(s)	
	Mike KENNEWICK *et al.*	
	Filing Date	Group Art Unit
Form PTO-1449	November 10, 2010	2447

U.S. PATENT DOCUMENTS

*EXAMINER INITIAL	REF	DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE
	A	2010/0204986	08/12/2010	Kennewick *et al.*	704	226	
	B	2010/0204994	08/12/2010	Kennewick *et al.*	704	257	
	C	2010/0217604	08/26/2010	Baldwin *et al.*	704	275	
	D	2010/0286985	11/11/2010	Kennewick *et al.*	704	257	
	E	2010/0299142	11/25/2010	Freeman *et al.*	704	9	

FOREIGN PATENT DOCUMENTS

	REF	DOCUMENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translations YES	Translations NO
	F	WO 2007/027546	03/08/2007	International				
	G	WO 2007/027989	03/08/2007	International				
	H	WO 2008/098039	08/14/2008	International				
	I	WO 2008/118195	10/02/2008	International				
	J	WO 2009/075912	06/18/2009	International				
	K	WO 2009/145796	12/03/2009	International				
	L	WO 2010/096752	08/26/2010	International				

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

EXAMINER	DATE CONSIDERED

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP Section 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

402527207v1

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 March 2007 (08.03.2007)

PCT

(10) International Publication Number
WO 2007/027546 A2

(51) International Patent Classification:
G10L 21/00 (2006.01)

(21) International Application Number:
PCT/US2006/033378

(22) International Filing Date: 29 August 2006 (29.08.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/212,693 29 August 2005 (29.08.2005) US

(71) Applicant *(for all designated States except US)*: **VOICEBOX TECHNOLOGIES, INC.** [US/US]; 10505 NE 38th Place, Building 9, Kirkland, WA 98033 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **WEIDER, Chris** [US/US]; 2818 Grand Ave. #B-508, Everett, WA 98201 (US). **KENNEWICK, Rich** [US/US]; 14222 227th Ave., NE, Woodinville, WA 98077 (US). **KENNEWICK, Mike** [US/US]; 13407 37th Place, NE, Bellevue, WA 98005 (US). **DI CRISTO, Phillippe** [FR/US]; 177 107th Ave. #1007, NE, Bellevue, WA 98004 (US). **KENNEWICK, Robert, A.** [US/US]; 2204 S. Eddy St., Seattle, WA 98108 (US).

(74) Agents: **GATTO, James, G.** et al.; Pillsbury Winthrop Shaw Pittman LLP, P.O. Box 10500, McLean, VA 22102 (US).

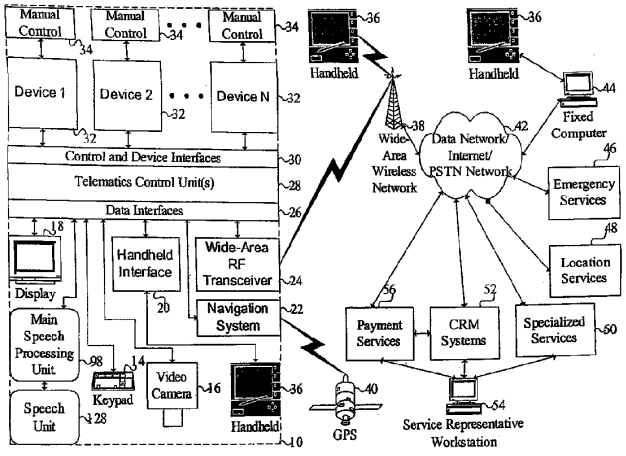
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

[Continued on next page]

(54) Title: MOBILE SYSTEMS AND METHODS OF SUPPORTING NATURAL LANGUAGE HUMAN-MACHINE INTERACTIONS



(57) Abstract: A mobile system is provided that includes speech-based and non-speech-based interfaces for telematics applications. The mobile system identifies and uses context, prior information, domain knowledge, and user specific profile data to achieve a natural environment for users that submit requests and/or commands in multiple domains. The invention creates, stores and uses extensive personal profile information for each user, thereby improving the reliability of determining the context and presenting the expected results for a particular question or command. The invention may organize domain specific behavior and information into agents, that are distributable or updateable over a wide area network.

WO 2007/027546 A2

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

MOBILE SYSTEMS AND METHODS OF SUPPORTING NATURAL LANGUAGE HUMAN-MACHINE INTERACTIONS

FIELD OF THE INVENTION

[001] The invention provides mobile devices having a speech interface and/or a combination of a speech interface and a non-speech interface to enable natural language human-machine interactions. More specifically, the invention enables mobile users to submit natural language speech and/or non-speech questions or commands in a wide range of domains. The mobile device is configured to present responses in a natural manner.

BACKGROUND OF THE INVENTION

[002] Telematic systems are systems that bring human-computer interfaces to mobile environments. Conventional computer interfaces use some combination of keyboards, keypads, point and click techniques and touch screen displays. These conventional interface techniques are generally not suitable for a mobile environments, due at least in part to the speed of interaction and the inherent danger and distraction. Therefore, speech interfaces are being adopted in many telematic applications.

[003] However, creating a natural language speech interface that is suitable for use in the mobile environment has proved difficult. A general-purpose telematics system should accommodate commands and requests from a wide range of domains and from many users with diverse preferences and needs. Further, multiple mobile users may want to use such systems, often simultaneously. Finally, most mobile environments are relatively noisy, making speech recognition inherently difficult.

[004] Retrieval of both local and network hosted online information and processing of commands in a natural manner remains a difficult problem in any environment, especially a mobile environment. Cognitive research on human interaction shows that verbal communication, such as a person asking a question or giving a command, typically relies heavily on context and domain knowledge of the target person. By contrast, machine-based requests (a request may be a question, a command, and/or other types of communications) may be highly structured and may not be inherently natural to the human user. Thus, verbal communications and machine processing of requests that are extracted from the verbal communications may be fundamentally

incompatible. Yet the ability to allow a person to make natural language speech-based requests remains a desirable goal.

[005] Research has been performed on multiple fields of natural language processing and speech recognition. Speech recognition has steadily improved in accuracy and today is successfully used in a wide range of applications. Natural language processing has previously been applied to the parsing of speech queries. Yet, a limited number of systems have been developed that provide a complete environment for users to make natural language speech requests and/or commands and receive natural sounding responses in a mobile environment. There remain a number of significant barriers to creation of a complete natural language verbal and/or textual-based query and response environment.

[006] The fact that most natural language requests and commands are incomplete in their definition is a significant barrier to natural language query-response interaction. Further, some questions can only be interpreted in the context of previous questions, knowledge of the domain, or the user's history of interests and preferences. Thus, some natural language questions and commands may not be easily transformed to machine processable form. Compounding this problem, many natural language questions may be ambiguous or subjective. In these cases, the formation of a machine processable query and returning of a natural language response is difficult at best.

[007] Even once a question is asked, parsed and interpreted, machine processable requests and commands must be formulated. Depending on the nature of the question, there may not be a simple set of requests that return an adequate response. Several requests may need to be initiated and even these requests may need to be chained or concatenated to achieve a complete result. Further, no single available source may include the entire set of results required. Thus, multiple requests, perhaps with several parts, may need to be made to multiple data sources, which can be located both locally or remotely. Not all of these sources and requests may return useful results or any results at all.

[008] In a mobile environment, the use of wireless communications may further reduce the chances that requests will be complete or that successful results will be returned. Useful results that are returned are often embedded in other information and may need to be extracted therefrom. For example, a few key words or numbers often need to be "scraped" from a larger amount of other information in a text string, table,

Amazon/VB Assets
Exhibit 1003
Page 455

list, page, or other information. At the same time, other extraneous information such as graphics or pictures may need to be removed to process the response in speech. In any case, the multiple results should be evaluated and combined to form the best possible answer, even in the case where some requests do not return useful results or fail entirely. In cases where the question is ambiguous or the result inherently subjective, determining the best result to present is a complex process. Finally, to maintain a natural interaction, responses should be returned rapidly to the user. Managing and evaluating complex and uncertain requests, while maintaining real-time performance, is a significant challenge.

[009] These and other drawbacks exist in existing systems.

SUMMARY OF THE INVENTION

[010] The invention overcomes these and other drawbacks of prior telematic systems.

[011] According to one aspect of the invention, speech-based and non-speech-based systems are provided that act on commands and retrieve information. The invention uses context, prior information, domain knowledge, and user specific profile data to achieve a natural environment for users that submit requests and/or commands in multiple domains. At each step in the process, accommodation may be made for full or partial failure and graceful recovery. The robustness to partial failure may be achieved through the use of probabilistic and fuzzy reasoning at several stages of the process. This robustness to partial failure promotes the feeling of a natural response to questions and commands.

[012] According to another aspect of the invention, the interactive natural language system (herein "the system") may be incorporated into mobile devices or may be connected to the mobile device via a wired or wireless connection. The mobile device may interface with computers or other electronic control systems through wired or wireless links. The mobile device may also operate independently of a mobile structure and may be used to remotely control devices through a wireless local area connection, a wide area wireless connection or through other communication links.

[013] According to one aspect of the invention, software may be installed onto the mobile device that includes an input module that captures the user input; a parser that parses the input, a text to speech engine module for converting text to speech; a network interface for enabling the device to interface with one or more networks; a

Amazon/VB Assets
Exhibit 1003
Page 456

non-speech interface module, an event manager for managing events and/or other modules. In some embodiments, the event manager may be in communication with a context description grammar, a user profile module that enables user profiles to be created, modified and/or accessed, a personality module that enables various personalities to be created and/or used, an agent module, an update manager and one or more databases. It will be understood that this software may be distributed in any way between a mobile device, a computer attached to a mobile structure, a desktop computer or a server without altering the function, features, scope, or intent of the invention.

[014] According to one aspect of the invention, the system may include a speech unit interface device that receives spoken natural language requests, commands and/or other utterances from users, and a computer device or system that receives input from the speech unit and processes the input and responds to the user with a natural language speech response.

[015] According to another aspect of the invention, the system may be interfaced by wired or wireless connections to one or more other systems. The other systems may themselves be distributed between electronic controls or computers that are attached to a mobile structure or are located external to the mobile structure. The other systems may include, electronic control systems, entertainment devices, navigation equipment, measurement equipment or sensors, or other systems. External systems may also be provided with features that include payment systems, emergency assistance networks, remote ordering systems, automated or attended customer service functions, or other features.

[016] According to another aspect of the invention, the system may be deployed in a network of devices that share a common base of agents, data, information, user profiles, histories or other components. Each user may interact with, and receive, the same services and applications at any location equipped with the mobile device on the network. For example, multiple mobile devices may be placed at different locations throughout a home, place of business, vehicle or other location. In such a case, the system can use the location of the particular device addressed by the user as part of the context for the questions asked.

[017] According to one embodiment of the invention, processing may be performed at the mobile devices. The commands may be processed on-board to enable the

Amazon/VB Assets
Exhibit 1003
Page 457

mobile devices to control themselves and/or to control other mobile devices, fixed computers, mobile telephones, and other devices. Additionally, mobile devices may track context.

[018] According to one embodiment of the invention, infrastructure may be provided to maintain context information during multimodal interactions, such as speech and/or non-speech interactions. According to one exemplary embodiment of the invention, context information may be maintained in a multimodal environment by providing communication channels between mobile device, or multimodal devices, and the system. The communication channels allow the system to receive multimodal input such as text-based commands and questions and/or voice-based commands and questions. According to another embodiment of the invention, the multimodal input may include a string of text, such as keywords, that are received as commands or questions. According to yet another embodiment of the invention, the system may synchronize the context between the mobile devices and the speech-based units. In order to send a response to the corresponding mobile device, the system may track the source and send the response to the corresponding speech interface or the non-speech interface.

[019] According to an alternative embodiment of the invention, context information may be maintained using a context manager that may be centrally positioned to receive input from multiple mobile devices and to provide output to multiple mobile devices. According to one embodiment, the mobile devices that communicate with the context manager may register through a registration module and may subscribe to one or more events. According to another embodiment of the invention, the context manager may receive input in Context XML form, for example. The other registered mobile devices may be informed of context changes through a context tracking module to enable synchronizing of context across the registered modules. According to one embodiment of the invention, registered modules may be added or removed from the system. The registered modules may include dynamic link libraries (DLLs), or other information sources, that are specific to multimodal devices.

[020] According to yet another alternative embodiment of the invention, context information may be determined from a command or request that is presented in a textual format and/or a command or request that is presented as an utterance and processed using a multi-pass automatic speech recognition module that transcribes the

Amazon/VB Assets
Exhibit 1003
Page 458

utterance to a text message. The command or request may be compared against a context description grammar to identify a match. Any active grammars in the context description grammar may be scored against the command or request and a best match may be sent to a response generator module. Agents may be associated with corresponding response generator modules and may retrieve the requested information for generation of a response. The agents may update a context stack to enable follow-up requests.

[021] According to another embodiment of the invention, mobile devices may be configured to allow verbal annotations of objects stored thereon. Mobile devices may transcribe the verbal annotation to text and store the textual annotation with the object. Alternatively, mobile devices may be configured to enable users to manually input textual descriptions that are stored along with the objects. According to one embodiment of the invention, the textual annotations and/or textual descriptions may be classified and searched. In an alternative embodiment, mobile devices may classify and search the verbal annotations rather than textual annotations. However, classifying and searching verbal annotations may be considerably more difficult than classifying and searching textual annotations and/or textual descriptions.

[022] According to one embodiment of the invention, the textual annotations and textual descriptions may be communicated using a short message service on the mobile telephone or other device. Short message service is a text message service that enables sending and receiving of short textual messages. The textual messages may be stored at data centers for forwarding to intended recipients. Other configurations may be used.

[023] According to another aspect of the invention, domain specific behavior and information may be organized into data managers. Data managers are autonomous executables that receive, process and respond to user questions, queries and commands. The data managers provide complete, convenient and re-distributable packages or modules of functionality that are typically directed to a specific domain of applications. Data managers may be complete packages of executable code, scripts, links to information, and other forms of communication data that provide a specific package of functionality, usually in a specific domain. In other words, data managers may include components for extending the functionality to a new domain. Further, data managers and their associated data may be updated remotely over a

Amazon/VB Assets
Exhibit 1003
Page 459

network as new behavior is added or new information becomes available. Data managers may use system resources and the services of other, typically more specialized, data managers. Data managers may be distributed and redistributed in a number of ways including on removable storage media, transfer over networks or attached to emails and other messages. An update manager may be used to add new data managers to the system or update existing data managers.

[024] In order to enhance the natural query and response environment, the system may format results to increase understandability to users. Formatting and presentation of results may be based on the context of the questions, the contents of the response being presented, the history of the interaction with the user, the user's preferences and interests and the nature of the domain. By contrast, rigid, highly formatted, or structured presentation of results may be deemed unnatural by many users.

[025] According to another embodiment of the invention, the system may simulate some aspects of a human "personality". In some cases, the presentation of the response and the terms that are used to provide the response may be randomized to avoid the appearance of rigidly formatted or mechanical responses. The use of other simulated personality characteristics is also desirable. For example, a response that may be upsetting to the user may be presented in a sympathetic manner. Furthermore, results of requests may be long text strings, lists, tables or other lengthy sets of data. Natural presentation of this type of information presents particular challenges because simply reading the long response is generally not preferred. Instead, the system may parse important sections from the response and may initially provide only reports. Determining what parts of a long response are presented may be based on the context of the questions, the contents of the response being presented, the history of the interaction with the user, the user's preferences and interests and the nature of the domain. At the same time, the system may give the user interactive control over what information to present and how much information to present, to stop the response all together, or to take other actions.

[026] The invention may be applied as a user interface to telematics systems in a wide variety of environments. These environments may include, but are not limited to, the following: 1) personal automobiles, rented automobiles, or fleet automobiles; 2) motorcycles, scooters, and other two wheeled or open-air vehicles; 3)commercial long-haul and short haul trucks; 4) delivery service vehicles; 5) fleet service vehicles;

Amazon/VB Assets
Exhibit 1003
Page 460

6) industrial vehicles; 7) agricultural and construction machinery; 8) water-borne vehicles; 9) aircraft; and, 10) specialized military, law enforcement and emergency vehicles.

[027] According to another aspect of the invention, the system may process and respond to questions, requests and/or commands. Keywords or context may be used to determine whether the received utterance and/or textual message includes a request or command. For example, utterances may include aspects of questions, requests and/or commands. For example, a user may utter "tune in my favorite radio station." A request is processed to determine the name, the channel, and time for the users favorite radio station. If the programming on that station is of a type the user generally does not listen to, the system can suggest using an alterative option, such as listening to a CD more likely to please the user. A command must be executed to set a tune the radio station.

[028] The invention can be used for generalized local or network information query, retrieval and presentation in a mobile environment. For each user utterance, including a question or query or set of questions or queries, the system may perform multiple steps possibly including: 1) capturing the user's question or query through speech recognition operating in a variety of real-world environments; 2) parsing and interpreting the question or query; 3) determining the domain of expertise required and the context to invoke the proper resources, including agents; 4) formulating one or more queries to one or more local and/or network data sources or sending appropriate commands to local or remote devices or the system itself; 5) performing required formatting, variable substitutions and transformations to modify the queries to a form most likely to yield desired results from the available sources; 6) executing the multiple queries or commands in an asynchronous manner and dealing gracefully with failures; 7) extracting or scraping the desired information from the one or more results, which may be returned in any one of a number of different formats; 8) evaluating and interpreting the results, including processing of errors, gathered and combine them into a single best result judged to be "best" even if the results are ambiguous, incomplete, or conflicting; 9) performing required formatting, variable substitutions and transformations to modify the results to a form most easily understood by the user; 10) presenting the compound result, through a text to speech engine or multimodal interface, to the user in a useful and/or expected manner; 11)

Amazon/VB Assets
Exhibit 1003
Page 461

optionally, providing a response to users indicating the success or failure of the command, and possibly including state information; or other steps.

[029] The above steps may be performed with knowledge of the domain of expertise, the context for the question or command, domain specific information, the history of the user's interactions, user preferences, available information sources or commands, and responses obtained from the sources.

[030] Probabilistic or fuzzy set decision and matching methods may be applied to deal with inconsistent, ambiguous, conflicting and incomplete information or responses. In addition, asynchronous queries may be used to provide rapid and graceful failure of requests or commands that allow the system to robustly return results quickly and in a manner that seems natural to the user.

[031] Many everyday questions are inherently subjective and result in answers that are a matter of opinion or consensus, as much as fact. Such questions are often ad hoc in their nature, as well. The system, according to another aspect of the invention, may use adaptive, probabilistic, and fuzzy set decision and matching methods to identify the subjective nature of the question and to evaluate a range of possible answers, wherein one or more answers may be selected that most accurately represent the type of result desired by the user.

[032] The context and expected results from a particular question may be highly dependent on the individual asking the question. Therefore, the system may create, store and use personal profile information for each user. Information in the profile may be added and updated automatically as the user uses the system or may be manually added or updated by the user or by others. Domain specific agents may collect, store and use specific profile information, as may be required for optimal operations. Users may create commands for regularly used reports, automatically generated alerts, and other requests and for the formatting and presentation of results. The system may use profile data in interpreting questions, formulating requests, interpreting request results and presenting answers to the user. Examples of information in a user profile includes, history of questions asked, session histories, formatting and presentation preferences, special word spelling, terms of interest, special data sources of interest, age, sex, education, location or address, place of business, type of business, investments, hobbies, sports interests, news interests and other profile data.

Amazon/VB Assets
Exhibit 1003
Page 462

[033] According to one aspect of the invention, the system may attempt to provide rapid responses in order to provide a natural question and response environment. The rapid responses may be provided without obtaining additional information. The system may determine agent composition, context and/or domain for a user's question or command, for example, by using a real-time scoring system or other technique. Based on this determination, the system may trigger one or more agents to respond to the user's question or command. The agents may make one or more requests and rapidly return a formatted response. Thus, users may receive direct responses to a set of questions, each with a different response or context. In some cases, the available information, including the request results, may not adequately answer the questions presented. In such situations, the user may be asked one or more follow-up questions to resolve the ambiguity. Additional requests may then be made before an adequate response is provided. In these cases, the system may use context information, user profile information and/or domain specific information to minimize the interaction with the user required to deliver a response.

[034] If the confidence level of the domain or context score is not high enough to ensure a reliable response, the system may request that the user verify that the question or command is correctly understood. In general, the question may be phrased to indicate the context of the question including all criteria or parameters. If the user confirms that the question is correct, the system may proceed to produce a response. Otherwise, either the user can rephrase the original question, perhaps adding additional information to remove ambiguity, or the system may ask one or more questions to attempt to resolve the ambiguity or other actions may taken.

[035] According to one aspect of the invention, the system may accept any natural language question or command and, as a result, may be subject to ambiguous requests. To assist users in formulating concise questions and commands, the system may support a voice query language. The language may help users clearly specify the keywords or contexts of the question or command along with the parameters or criteria. The system may provide built in training capabilities to help the user learn the best methods to formulate their questions and commands.

[036] In order to make the responses to user's questions and commands seem more natural, the system may employ one or more dynamically invokeable personalities and/or emotional models. Personalities and emotional models have specific

Amazon/VB Assets
Exhibit 1003
Page 463

characteristics that simulate the behavioral characteristics of real humans. Examples of these characteristics include, sympathy, irritation, and helpfulness and associated emotions. The personality also randomizes aspects of responses, just as a real human would do. This behavior includes randomization of terms used and the order of presentation of information. Characteristics of the personality and/or emotions are invoked using probabilistic or fuzzy set decision and matching methods and using criteria, including the context for the question, the history of the user's interaction, user preferences, information sources available, responses obtained from the sources.

[037] According to another aspect of the invention, special procedures may be employed to present information in the form of long text strings, tables, lists or other long response sets. Simply presenting a long set of information in an ordered manner may not be considered natural or what most users have in mind. The system may use probabilistic or fuzzy set matching methods to extract relevant information and present these subsets first. Further the system may provide commands allowing users to skip through the list, find keywords or key information in the list or stop processing the list altogether.

[038] According to one embodiment of the invention, the system may support multiple users that access the system at different times. According to another embodiment of the invention, the system may support multiple users that access the system during a same session, in an interleaved or overlapping manner. The system may recognize the multiple users by name, voice, or other characteristic and may invoke a correct profile for each user. If multiple users are addressing the system in overlapping or interleaved sessions, the system may identify the multiple users and may invoke one or more corresponding profiles. For applications that require security safeguards, the multiple users may be verified using voiceprint matching, password or pass-phrase matching, or other security safeguards.

[039] When multiple users are engaged in interleaved sessions, the system may gracefully resolve conflicts using a probabilistic or fuzzy set decision method for each user. This process may simulate the manner in which a human would address multiple questions from various sources. For example, the system may answer short questions first in time while answering longer questions later in time. Alternatively, the system may answer questions in the order that they are received, among other configurations.

Amazon/VB Assets
Exhibit 1003
Page 464

[040] Since the invention may operate in many environments, including mobile environments with background noise, point noise sources and people holding conversations, filtering of speech input may be advantageous. The invention may use, for example, either one-dimensional or two-dimensional array microphones (or other devices) to receive human speech. The array microphones may be fixed or employ dynamic beam forming techniques. The array pattern may be adjusted to maximize gain in the direction of the user and to null point noise sources. Alternatively, microphones may be placed at particular locations within the mobile environment near where occupants are likely to use the system. These microphones can be single microphones, directional microphones or an array of microphones. Speech received at the microphones may then be processed with analog or digital filters to optimize the bandwidth, cancel echoes, notch-out narrow band noise sources, or perform other functions. Following filtering, the system may use variable rate sampling to maximize the fidelity of the encoded speech, while minimizing required bandwidth. This procedure may be particularly useful in cases where the encoded speech is transmitted over a wireless network or link.

[041] The invention can be applied to a wide range of telematics applications. General applications areas can include, but are not limited to remote or local vehicle control, information query, retrieval and presentation from local or network sources, safety applications, and security applications.

[042] The system may provide local or remote control functions for the system, including devices that are located on the mobile structure or off the mobile structure. Users may initiate commands locally or remotely. Typically, remote operation may be conducted through an IP connection, a telephone connection, or other connections. The user may address spoken commands to a mobile device or desktop unit, which may send the commands to controllers on the vehicle over wireless links. Other remote command techniques may be used. The system may process commands in a nearly identical manner to a request. One difference being that the result of the command is generally an action rather than a response. In many cases, the system may give the user a cue or response to indicate that the command has been successfully executed or has failed. In cases of failure, an interactive session may be started to allow the user to resolve the difficulty or formulate a command more likely to succeed.

Amazon/VB Assets
Exhibit 1003
Page 465

[043] The invention provides users, including mobile structure operators, with the capability to control most any mobile system functions using interactive speech and non-speech commands and/or requests. Generally, controls of a critical nature, or having safety implications, may employ fail-safe checks that, before execution, verify that a command will not create a hazardous condition. Manual overrides may also be provided as an extra precaution. The invention may provide built in help and user guidance for the devices under control. This guidance may include step-by-step training for operators that are learning to use the features of the mobile structure. The system can provide extensive interactive guidance when commands cannot be executed or when commands fail. This advice may include suggestions to reformulate the command to succeed, suggestions to work around a failure, suggestions for alternative commands that may achieve a similar function, or other suggestions. Examples of control functions that can be performed from local or remote locations by the invention include:

1. Control of mobile structure multimedia entertainment electronics, such as radio, CD player, or video player. This control can be based on user specified play lists and may be sensitive to user profile information including, preferences history, or other information. The invention may include the ability to control multiple or individual multimedia entertainment stations.
2. Control of communications devices such as cell phones, voice mail systems, fax systems, text or instant messaging systems, call and message forwarding systems, email systems and other communication devices. This control includes controlling features such as, address books, phone books, call forwarding, conference calling, and voice mail, among others.
3. Local or remote control of mobile structure systems. Most any device on the mobile structure may be controlled including, door locks, window controls, interior temperature controls, shifting of the transmission, turn signals, lights, safety equipment, engine ignition, cruse control, fuel tank switches, seat adjustments, specialized equipment such as winches, lifting systems or loading systems, and other mobile structure systems.
4. Systems that are located external to the mobile structure typically may be controlled through wireless links including, garage door openers, gate

Amazon/VB Assets
Exhibit 1003
Page 466

controllers, mobile structure entry security passes, automated toll collection systems, and mobile structure weighing systems and other external systems.

5. Mobile structure power management and systems control. The invention may provide the mobile structure operator with information on limitations and on tips for better power management or fuel utilization or other systems control.

6. Diagnostic information management. The invention may provide diagnostic information announcements and warnings for the mobile structure operator. These announcements and warnings may be interactive allowing the operator to request additional information or to suggest various courses of action. The invention may mediate a solution to the problem, including providing access to a scheduling service, summoning help or providing instructions for remedial action until a permanent solution can be achieved. The system can ask the operator to authorize the ordering of needed parts and may provide cost estimates. The system may receive data for these announcements and warnings from a wide range of sources including sensors and vehicle control computers. Sensors may include fuel level sensors, coolant temperature sensors, oil temperature sensors, axel temperature sensors, tire pressure sensors, and other sensors.

7. System status inquiry. Mobile structure operators may use the interactive natural language interface of the invention to query and receive reports on the status of systems that are located on the mobile structure including, fuel level, interior temperature, outside temperature, engine or other mobile structure system status. The operator can further query the system to receive more information or determine a course of action if problem is detected.

8. Vehicle service history. The invention can provide the vehicle operator or other personnel with interactive access to the vehicle service history. The invention can provide announcements or warnings as the time for service draws near. The user can interact with the system to schedule the required service, order required parts, receive cost estimates, or update the service history. Users may customize the nature of this interaction to suit their desires or policies.

Amazon/VB Assets
Exhibit 1003
Page 467

9. Diagnostic and service history. The invention may provide diagnostic and service history information to service personnel. This information may include vehicle fault codes and other information for devices that are under control of or measured by the system. Alternatively, the invention may receive information on the state and history of vehicle operation from other control computers. The invention may provide interactive service information and history. The service history may be queried and presented through the speech interface or the non-speech interface. The system may prompt service personnel for more information if the record is deemed incomplete. In other cases, the invention may prompt service personnel for information on their actions if a change is detected in system status, such as replacement of a part.

[044] The invention can provide users or operators of a mobile structure with specialized safety functions through the interactive speech interface and/or non-speech interface. The invention may use a dynamically evocable personality that is capable of creating announcements that are appropriate for the severity of the situation. The announcements and personalities may be under user control and configuration. Some examples of these safety applications can include:

1. The invention can provide automated detection and reporting of accident situations through the wireless communications link. Information on an accident situation can be gathered from airbag control systems or other sensors. Once an accident situation has been detected the invention may use the interactive speech interface and/or non-speech interface to determine the nature of the accident and/or the condition of victims. This information, along with location information and other relevant information, may be reported over the wireless link. Alternatively, the invention may establish voice channel communications or other communications between occupants of the vehicle and emergency personnel.

2. The invention can be used to store and retrieve medical information on vehicle occupants. Following an accident, emergency personnel may query the system for this information. Alternatively, the system may annunciate a warning to emergency personnel if a person has a special medical condition. The system maintains the security of medical information through a number of techniques, including not annunciating medical information unless an accident

Amazon/VB Assets
Exhibit 1003
Page 468

is detected, or not annunciating medical information unless that person or another authorized person gives permission.

3. Occupants of the vehicle can summon help in the event of a crime using the speech interface and/or non-speech interface. Typical crimes may include robberies and hijackings. The invention allows vehicle occupants to set panic or emergency words or phrases that indicate to the system a crime is occurring.

4. The invention can provide the vehicle operator with safety announcements if an unsafe or potentially unsafe situation is detected. The operator can use the interactive speech interface and/or non-speech interface to obtain more information on the situation or dismiss the alert. The operator can annunciate or otherwise provide commands to remedy or mitigate the situation during this dialog. Conditions that can be announced include, following another vehicle too closely, too great a speed for the road or conditions, obstruction on roadway, a fire in some part of the vehicle, high cargo pressure or temperature, leaks, and other information.

5. The interactive speech interface and non-speech interface of the invention can provide the operator with real-time assistance. This assistance may include, aid in parking or backing a vehicle, aid with complex maneuvers, aid with optimal operation of the vehicle, and other operations. The operator can ask the system for advice or assistance with a planned maneuver or operation. Alternatively, the invention can proactively offer assistance if certain situations are detected.

6. The interactive speech interface and non-speech interface may be used to improve vehicle security. Voiceprints or voice authentication can be used to gain access to the vehicle or start the vehicle. Alternatively or in addition, a password or pass-phrase can be used. In another alternative speech security can be used as a supplement to other vehicle security techniques.

7. The invention can provide measurement of operator fatigue and alert the operator or remote personnel if unacceptable levels of fatigue are detected. The interactive speech interface and/or non-speech interface can be used to query the operator to detect fatigue. Alternatively, or in addition, other measurements of operator fatigue can be used. If a fatigue situation is

Amazon/VB Assets
Exhibit 1003
Page 469

detected the invention may initiate a dialog with the operator to determine the extent of the problem and if required, ask the operator to cease operation.

[045] The invention can offer vehicle operators and occupants a variety of services that are useful while in the vehicle and/or while arriving at a destination. Further, users can employ the interactive natural language interface to customize these services to suit each individual. Some examples of services that can be supported by the natural-language interactive speech interface of the invention, include:

1. Providing vehicle operators with interactive directions to a destination or waypoint, wherein the user can specify a desired destination and any preferred waypoints. A destination can be specified in any manner including providing the name of a place, an address, a name of a person, a name of a business, or other type of information. As the trip progresses the system may provide the operator with continued directions and warnings if a mistake has been made. The operator can query the system for additional information or request less information. Generally, the system is interfaced with one or more navigation sensors and local or remote map databases. The invention may provide operators or passengers with alerts of upcoming points or interest, exits or stops, hazards, or other conditions. The users can query the system for more specific information. Alternatively, the invention can provide operators and occupants of the vehicle with an interactive guided tour. The system's information query, retrieval and presentation capability can be employed by users to receive additional information or points or items of interest during the tour and may take into account stored personal profile information for a user.
2. The invention can provide the operator of a vehicle with interactive dynamic routing information. The routing can be updated based on traffic conditions, weather conditions, facilities availability, and information provided by the operator. Generally, the system is interfaced with one or more navigation sensors, local or remote map databases, and sources of traffic, weather, and facilities use data.
3. The direction, routing and communications capabilities of the invention can be combined in an interactive system that helps one or more operators rendezvous at a predetermined destination or any other convenient midpoint. The operators may use the interactive natural language interface to

Amazon/VB Assets
Exhibit 1003
Page 470

communicate with the system to arrange the rendezvous, to receive directions as they travel to the rendezvous point and/or to communicate with the other operators.

4. The navigational capabilities of the invention may be used to place limits on where a vehicle is allowed to go and/or for how long. The system may employ the interactive natural language speech interface and/or non-speech interface to inform the operator when the vehicle is approaching or has exceeded a limit. The operator can query the system to determine the best course of action to return to limits or prevent exceeding them. Alternatively, the system can query the operator to determine why they are exceeding the limits or to mediate a negotiation to extend the limits if this is required by circumstances. This capability is useful in several situations including, keeping a delivery or passenger vehicle on a regular route, setting and enforcing use limits on teenagers, preventing an operator from using the vehicle in an unauthorized manner, or other situations.

5. The interactive natural language interface of the invention can be used to provide Customer Relationship Management (CRM) services to vehicle operators and passengers. The user can interact with the services offered via data networks, video signals, or audio. The interaction can be performed with automated services or a live Customer Service Representative. Interactions with the customer service representatives can be through any combination of possible techniques, such as, live audio, live video, electronic messaging or email, instant messaging, and other techniques. These services can be offered by a number of entities including, vehicle manufacturers, vehicle dealers, vehicle service organizations, automobile or travel clubs, wireless carriers, travel service organizations, or other organizations. The services offered can be personalized to the occupants of the vehicle using a variety of information including, user profile information, history, location, paths traveled, time of day, day of week, or other information. In addition, the system can offer customized services based on information about the vehicle, including paths traveled, distance, service history, type of equipment on vehicle. These services can be accessed while a person is an occupant in a vehicle, while the person is not an occupant but are using a wireless or wired network equipped

Amazon/VB Assets
Exhibit 1003
Page 471

mobile device or while using a wired or wireless network desktop system, among other situations. Examples of these services include:

a. Location based marketing programs wherein occupants of the vehicle receive promotional offers from merchants along a route of travel. Occupants can query the system for offers and promotions for goods, services, or other information, along the travel route. The system may apply other available information to form a response, including, the user profile, history, location, or other information. The system can provide optimized interactive routing assistance to the vehicle operator. Alternatively, the system can provide interactive offers and promotions for goods and services along the route, or in advance of a particular trip. Promotions can be offered for goods and services can include but are not limited to, travel services, groceries, prepared foods, vehicle service, fuel, entertainment, or other goods and services.

b. Remote ordering and payment for goods and services. The system can interactively present the menu or product catalog using the list and table presentation capabilities of the invention. The system may facilitate remote ordering by using location information, customer preferences, customer order history, etc. The system can manage a secure payment wallet for the users. Voiceprints, spoken passwords, and non-speech security methods (i.e. PIN pad, etc.) can be combined to create the appropriate level of security.

c. Travel services for occupants of the vehicle. These services can include, directories of travel and entertainment services, or reservations for entertainment restaurants, hotels and other accommodation. The system may present directories, lists and menus using its interactive list and table presentation capabilities. The travel service capability can be used in conjunction with the remote ordering and payment capabilities and the dynamic interactive routing capability.

d. Answer specialized travel related questions in areas such as vehicle registration, taxes, safety laws, required inspections, weight

Amazon/VB Assets
Exhibit 1003
Page 472

limits, insurance coverage requirements, insurance policy provisions, or other areas.

6. The invention can provide an operator or other occupant of a vehicle with an interactive location sensitive shopping list or a location and time sensitive task reminder list using the natural language interface. Users can create the list while in the vehicle, while on foot using a mobile device, or at a fixed location using a handheld or desktop device. A user may grant permission to other users to add tasks or shopping items to their lists. Once in the vehicle, the system may provide occupants with routing assistance to optimize travel time and may provide reminders of items to be purchased and tasks to be completed as the vehicle comes in close proximity to a particular location, type of merchant or other service provider, or when a set time has been reached, among other criteria.

7. Automatic interactive dispatch and reporting for fleet vehicles having vehicle operators or other vehicle occupants that use the speech interface and/or non-speech interface to interact with these services. These services can include, dynamic optimal routing, inventory of parts and other materials, ordering of required parts and materials, work orders, receipt generation, payments, or other services.

8. Sales force automation, sales reporting, contact database management, calendar management, and call routing. The system may employ its interactive list and table presentation capabilities to supply catalog and pricing information or other information. These services can use local or network data. Additional services can include, memos, reminders, activity lists, or other information.

[046] Vehicle operators and other occupants can use the interactive natural language interactive interface of the invention to perform many types of information query, retrieval and presentation operations. Using the natural language interactive interface, users can modify the parameters of queries or specify the presentation formats for results. Data used to create a response can be from any combination of local and remote data sources. User specific data can be synchronized between systems fixed to one or more vehicles, mobile structures and desktop systems. Some examples of

Amazon/VB Assets
Exhibit 1003
Page 473

information query, retrieval and presentation applications for the invention include, but are not limited to the following:

1. White pages and yellow pages lookups to find, email addresses, telephone numbers, street addresses and other information for businesses and individuals. These services can be used in conjunction with other services, including remote ordering and payment, offers and promotions, mapping, and driving directions;

2. Management and access to personal address book, calendars and reminders for users;

3. Automatic telephone dialing, reading and sending emails, pages, instant messaging by voice, text or video and other communications control functions;

4. Selection, schedules and play list management for television, satellite broadcast, radio or other entertainment schedule. The available information can include reviews and other information on programming. The system may provide device control for users;

5. Weather information for the local area or other locations;

6. Stock and other investment information including, prices, company reports, profiles, company information, business news stories, company reports, analysis, price alerts, news alerts, portfolio reports, portfolio plans, or other information;

7. Local, national and international news information including headlines of interest by subject or location, story summaries, full stories, audio and video retrieval and play for stories;

8. Sports scores, news stories, schedules, alerts, statistics, background and history information, or other information;

9. The ability to subscribe interactively to multimedia information channels, including sports, news, business, different types of music and entertainment, applying user specific preferences for extracting and presenting information;

10. Rights management for information or content that is used or published;

Amazon/VB Assets
Exhibit 1003
Page 474

11. Horoscopes, daily jokes and comics, crossword puzzle retrieval and display and related entertainment or diversions;

12. Interactive educational programs using local and network material, with lesson material level that is set based on user's profile, location of the vehicle, planned route of the vehicle, planned activities during the trip and including, interactive multimedia lessons, religious instruction, calculator, dictionary and spelling, geographic information, instruction for specialized tasks planned during the trip, language training, foreign language translation, presentation of technical manuals, and encyclopedias and other reference material.

[047] It will be appreciated that the foregoing statements of the features of the invention are not intended as exhaustive or limiting, the proper scope thereof being appreciated by reference to this entire disclosure and reasonably apparent variations and extensions thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

[048] The invention will be described by reference to the preferred and alternative embodiments thereof, in conjunction with the drawings in which:

[049] Figure 1 is an overall block diagram of the system according to a first embodiment of the invention;

[050] Figure 2 is an overall block diagram of the system according to a second embodiment of the invention;

[051] Figure 3 is an overall block diagram of a mobile device according to an embodiment of the invention;

[052] Figure 4 is an overall block diagram of a fixed computer according to an embodiment of the invention;

[053] Figure 5 is an overall diagrammatic view of the interactive natural language processing system according to an embodiment of the invention;

[054] Figure 6 is a schematic block diagram showing the agent architecture according to an embodiment of the invention;

[055] Figure 7 illustrates a diagram for centrally maintaining context, according to one embodiment of the invention; and

Amazon/VB Assets
Exhibit 1003
Page 475

[056] Figure 8 illustrates a system diagram for an enhanced conversational voice user interface according to one embodiment of the invention.

DETAILED DESCRIPTION

[057] The following detailed description refers to the accompanying drawings, and describes exemplary embodiments of the invention. Other embodiments are contemplated and modifications may be made to the exemplary embodiments without departing from the spirit, functionality and scope of the invention. Therefore, the following detailed descriptions are not meant to limit the invention.

[058] According to one embodiment of the invention, a telematic natural language speech interface and non-speech interface are provided for use in mobile environments and telematic applications. The system, or portions thereof, may be used in vehicles, while on foot or at a fixed location such as an office or home, or at other locations. An overall block diagram of one embodiment of the invention is illustrated in Figure 1.

[059] A speech unit 128 and/or a keypad 14 may be integrally coupled to a mobile structure 10 or may be part of mobile devices 36, fixed home or office computer systems 44, or other devices. Mobile devices 36 may include mobile telephones, personal digital assistants, digital radios, compact disk players, navigation systems, or other mobile devices. The mobile devices 36 may be configured to integrate with set-top boxes, alarm clocks, radios, or other electronic components. The speech unit 128 and/or keypad 14 may be interfaced to a Telematics Control Unit (TCU) 28 through one or more data interfaces 26. According to some embodiments, the main speech-processing unit 98 may be embedded in one or more TCU 28. In some embodiments, the components of the speech unit 128 may also be distributed between one or more TCUs.

[060] A speech-processing unit may be built into mobile devices 36 and may be coupled with the data interfaces 26 though a wireless, or wired, handheld interface 20. Other user interface peripherals may be connected to the TCU through the data interfaces and may include displays 18, such as touch screen displays for displaying text, graphics and video; keypads 14 for receiving textual data input; video cameras 16 for receiving multimedia communications or conferences; a pointing device or stylus, or other user interface peripherals. Other devices may be connected to the TCU through the data interfaces including wide-area RF transceivers 24, navigation

Amazon/VB Assets
Exhibit 1003
Page 476

system components 22, or other devices. The navigation system may include several components such as, Global Positioning System (GPS) receivers or other radiolocation system receivers, gyroscopes or other inertial measurement equipment, distance measurement sensors, such as odometers, or other components. Radiolocation equipment may receive coded signals from one or more satellites or terrestrial sources 40. The one or more location service servers 48 may assist the navigation system. Other systems that can connect to the TCU through the data interfaces may include automotive control computers, digital control interfaces for devices such as media players or other electronic systems, measurement sensors, or other specialized electronic equipment.

[061] The control and device interfaces 30 may connect the TCU 28 to various devices 32 within the mobile structure 10. The control and device interfaces 30 may be used to execute local or remote commands from users of the natural language interface. In some cases, the control and device interfaces 30 may include specialized hardware for interacting with different types of devices. The hardware interfaces may include analog or digital signal interfaces for device control, along with analog or digital interfaces for measurements that may control devices 32. These interfaces may also include specialized software that encapsulates or abstracts specific behavior of the devices 32. The interface software may include one or more drivers that are specific to the hardware interface and to one or more agents. The domain agents may include the specialized software behavior and data required to control a particular device or class of devices. New or updated behavior may be added to the system by updating data managers that are associated with specific devices or class of devices. The devices 32 may include manual controls or manual overrides 34. For safety reasons, the control and device interfaces 30 may incorporate fail-safe systems that, for example, may verify operating limits before changing settings, ensuring that commands do not conflict with settings from manual controls, and will not, in some combination with other commands or device settings, create an unsafe situation. The software behavior and data that ensures safe operations may be included within the domain agent that is specific to the device or class of devices. Examples of devices and system that can be controlled through the control and device interfaces 30 include, power management systems, measurement sensors, door locks, window controls, interior temperature controls, shifting of the transmission, turn signals,

Amazon/VB Assets
Exhibit 1003
Page 477

lights, safety equipment, engine ignition, cruise control, fuel tank switches, seat adjustments, specialize equipment such as winches, lifting systems or loading system, and other systems.

[062] The wide-area RF transceiver 24 may communicate with one or more wide-area wireless networks 38, which may be connected to data networks 42, including the Internet, the Public Switched Telephone Network (PSTN) 42 or other data networks. The wide-area wireless networks can be of any suitable terrestrial or satellite based type. Mobile devices 36 may communicate with one or more local or wide-area wireless networks. Home or office systems 44, equipped with wired or wireless network interfaces, may communicate through the data networks or PSTN.

[063] According to one embodiment of the invention, data and agents may be stored and synchronized in mobile structures 10, mobile devices 36 and/or fixed systems 44 having one or more main speech-processing units 98. The synchronization between these different systems can occur on the wide area wireless network 38, the data network 42, through the handheld interface 20, or other local data connections. The synchronization may be performed automatically when any two or more of the computers are connected to these networks. Alternatively, the synchronization may be performed on demand under user control. The synchronization process attempts to determine which version of a data element or an agent are the newest or most up-to-date and propagates that element. Thus, synchronization is an incremental change processes. In some cases, a complete replacement of a database, a portion of the database or of one or more agents may be performed rather than performing a series of incremental updates.

[064] The wide-area wireless networks 38, the data networks 42 or PSTN, may connect mobile structures 10, mobile devices 36 and fixed computers 44 to one or more servers that provide one of more services. According to one embodiment of the invention, an interactive natural language user interface may be provided that supports a transfer of data or transmission of speech, text, video and other formats. For data centric applications, a standardized data transfer format may be used including, for example, Hypertext Markup Language over Hypertext Transfer Protocol (HTTP), Extensible Markup Language (XML), and other data formats or schemas, over HTTP or other transfer protocol, Electronic Data Interchange formats over a variety of transport protocols, etc. It will be understood that the exact configuration of the

Amazon/VB Assets
Exhibit 1003
Page 478

servers may be determined on a case-by-case basis with consideration being given to the exact combinations of services being offered, the service providers providing the services, the contractual relationships between the service provider, and other factors. The invention supports any suitable configuration. In each case these servers may themselves be distributed over one or more public or private networks. Some examples of servers that may be used to deliver these services, are given below:

1. One or more payment service providers 56 that supply payment capabilities to users. These payment services may include electronic wallet capabilities for one or more payment accounts and may include payment security information, payment account information, transaction histories, account balance information, and other information. Exemplary payment types include, stored value accounts, promotional accounts, credit accounts, telecommunications billing accounts, debit accounts, or other payment types that are accessed using online or offline methods. Payments can be computed in any manner including payment for a specific good or service, subscription payment, metered payment, or other payments. The payment services can be distributed in a number of ways. Examples of computers and servers used to store and process payment transaction information include, smart cards, main speech processing units 128, mobile devices 36, TCUs 28, fixed personal computers 44, payment gateways, payment servers 56, or other systems.
2. One or more Customer Relationship Management (CRM) systems 52 may supply any number of consumer and business customer services as discussed above. The CRM system can supply automated services or services that are partly or completely manual. For manual services one or more customer service representatives may use one or more service representative workstations 54. The CRM system and the service representative workstation may be connected to one or more data networks 42 or the PSTN. Any other servers may also have connections to one or more service representative workstations, which may be in common or independent of each other.
3. One or more specialized service servers 50 may support specialized consumer and business services.
4. One or more location service servers 48 may supply location information and location based services. The location data is used as a data

Amazon/VB Assets
Exhibit 1003
Page 479

input to the location services that may then be distributed in any suitable manner including, on main speech processing units 128, mobile devices 36, TCUs 28, fixed personal computers 44, other servers (i.e. 46, 50, 52, 56) or the one or more location services servers 48.

5. One or more emergency services servers 46 may supply both public and private emergency services to the users.

[065] Main speech processing unit 98, speech unit 128 and keyboard 14 may be distributed in mobile devices 36 in a number of ways. For example, these units may be attached to the mobile devices 36 as independent components or as a single integrated component. In another embodiment of the invention, some or all of the main speech processing unit 98 and speech unit 128 may be embedded in one or more of the TCUs 28, mobile devices 36, fixed computer systems 44, or other devices.

[066] Figure 2 illustrates an alternative block diagram of a second embodiment of the invention. In this embodiment, the main speech processing unit 98 and speech unit 128 are external to the TCU 28. These components may be housed in one or more packages or may be included in a single integrated package.

[067] In all other respects, the second embodiment resembles the first embodiment illustrated in Figure 1. It will be understood that the exact distribution and packaging of the main speech processing unit 98 and speech unit 128 may be determined by the details of the deployment situation and should not change the functionality, capabilities or spirit of the invention in any way.

[068] Figure 3 illustrates a block diagram of one embodiment of the mobile devices 36 that include main speech processing unit 98 and speech unit 128 embedded in the mobile device 36, mobile telephone or other mobile device. The main speech processing unit may interface to one or more processing units 70 located in the mobile device 36. The processing units 70 may include one or more central processing units, one or more data and address busses, data interfaces, volatile memory, or other components. The processing units 70 may use one of more types of nonvolatile memory 80 for software and data storage. Suitable types of nonvolatile memory 80 may include flash memory and hard disk drives. In some embodiments, the main speech processing unit 98 may be integrated with the one or more processing units 70.

[069] According to one embodiment, users may interact with the mobile device 36 through the speech unit 128, the keypad 74 or keyboard, a display 72 that displays

Amazon/VB Assets
Exhibit 1003
Page 480

text, graphics, video, or other peripheral. In some embodiments the display may be a touch screen type. Alternatively, a pointing device (not shown) may be used, among other devices.

[070] The mobile device 36 may connect to one or more wired or wireless wide-area or local-area networks through one or more interfaces. A wide-area network transceiver 78 may connect to the wide-area wireless network 38 or the data network 42 using a wireless or wired connection, including an IP connection, a dial-up PSTN network connection or other connections. The local-area network transceiver 76 may connect to wired or wireless local area networks. These networks may include the handheld interface 20 or connections to fixed computer systems 44. In both mobile device environments and fixed computer environments, communications between humans and machines may not provide accurate results at least because natural language commands may be incomplete in their definition. The occurrence of inaccurate results may be reduced by leveraging context of previous utterances, knowledge of the domain and/or the user's history of interests and preferences or other factors.

[071] According to one embodiment of the invention, processing may be performed at the mobile devices 36. Alternatively, the processing may be performed on the server side. In yet another embodiment, the processing may occur on both the mobile devices 36 and the server side. The commands may be processed on-board the mobile devices 36 to enable the mobile devices 36 to control themselves and/or to control other mobile devices 36, fixed computers 44, mobile telephones, or other devices. Additionally, mobile devices 36 may track context. According to another embodiment of the invention illustrated in FIG. 7, a context manager 702 may be centrally maintained to allow input to and output from multiple mobile devices 36. Each of the mobile devices may communicate with the context manager 702 to register via registration module 712, wherein registration may indicate events that the mobile devices 36 may subscribe to. The context manager 702 may receive input in Context XML form, for example. The other registered mobile devices 36 may be informed of context changes through context tracking module 714, thereby synchronizing the context across the registered mobile devices 36. According to one embodiment of the invention, registered mobile devices 36 may be added or removed.

Amazon/VB Assets
Exhibit 1003
Page 481

The registered mobile devices 36 may be dynamic link libraries (DLLs), object code, or other data that may be specific to the mobile device 36.

[072] According to yet another alternative embodiment of the invention, context information may be determined from a command or request that is presented as a text message and/or a command or request that is presented as a verbal utterance and processed using a multi-pass automatic speech recognition module that transcribes the verbal utterance to a text message. The multi-pass automatic speech recognition module may use a dictation grammar or a large vocabulary grammar, among other resources, to transcribe the verbal utterance into a text message. On platforms where a dictation grammar is not available, the multi-pass automatic speech recognition may use a virtual dictation grammar that uses decoy words for out-of-vocabulary words. Decoy words may include utility words, nonsense words, isolated syllables, isolated distinct sounds and other decoy words.

[073] According to one embodiment of the invention, the text message may be searched for a particular character, group of characters, word, group of words, and other text combinations. The text combination may be compared against entries in a context description grammar that is associated with each agent 106. If a match is identified between an active grammar in the context description grammar and the command and/or request, then the match may be scored. The agents 106 may be ranked based on the determined score. In generating an aggregate response from the one or more responses received from the agents 106, the ordering of the responses from the individual agents may be determined based on the rank of agents 106. The aggregate response may be generated by a response generator module. Agents 106 may update a context stack, that includes an ordered list of command contexts, to enable follow-up requests.

[074] According to another embodiment of the invention, if a match is not found, or only a partial match is found, between the text message and active grammars, then a knowledge-enhanced speech recognition system may be used to semantically broaden the search. The knowledge-enhanced speech recognition system may be used to determine the intent of the request and/or to correct false recognitions. The knowledge-enhanced speech recognition may access a set of expected contexts that are stored in a context stack to determine a most likely context. The knowledge-enhanced speech recognition may use context specific matchers that are able to

Amazon/VB Assets
Exhibit 1003
Page 482

identify context such as time, location, numbers, dates, categories (e.g., music, movies, television, addresses, etc.) and other context. The matching may be performed by comparing a character, group of characters, a word, group of words, and other text combinations. Alternatively, or in addition to text based matching, the matching may be performed using phonetic matching, among other techniques. The results of any match may be used to generate a command and/or request that is communicated to agents 106 for additional processing. According to one embodiment of the invention, non-speech interface 114 may show system, state and history information in a more concise manner than is possible through the speech interface. Non-speech interface 114 may be accessed to create or extend capabilities of agents 106. These operations may include scripting of agents, adding data to the agent or databases 102 used by the agent, adding links to information sources, among other operations.

[075] According to another embodiment of the invention, mobile devices 36 may be configured to allow speech annotations of objects stored thereon. The objects may include photographs, calendar entries, email messages, instant messages, phonebook entries, voice mail entries, digital movies or other objects. Mobile devices 36 may transcribe the speech annotations to textual annotations and store the textual annotations with the object. Alternatively, mobile devices 36 may be configured to enable users to input non-speech annotations, such as textual descriptions, that are stored along with the objects.

[076] According to one embodiment of the invention, the annotated objects may be stored on a server side, a client side, a combination of server side and client side, or according to other configurations. The invention further contemplates collaboratively exchanging and sharing the annotated objects among distributed workgroups that may include centralized servers having shared workspaces for providing common object storage and retrieval facilities. The shared workspaces may be implemented on the centralized servers and may be accessed from different platforms using the mobile devices 36. The system may include a peer-to-peer system for accessing the annotated objects.

[077] According to one embodiment of the invention, the non-speech annotations may be classified and searched. In an alternative embodiment, the speech annotations may be classified and searched. However, classifying and searching speech

Amazon/VB Assets
Exhibit 1003
Page 483

annotations may be considerably more difficult than classifying and searching non-speech annotations.

[078] According to another embodiment of the invention, the textual annotations and/or textual descriptions that are associated with the objects may be stored as metadata, thereby enabling searching for the objects using the metadata. The metadata may include GPS information, environmental information, geographic information, or other information. For example, proximity to famous landmarks may be determined using GPS information, environmental information, geographic information, or other information, and this information may be integrated into the metadata associated with the objects. According to one embodiment of the invention, GPS coordinates may be stored in the metadata associated with the objects and users may search for selected objects based on the GPS coordinates. A user may provide a speech command such as "show me all the photos of Greece." In this case, the system would limit the type of object to photographs and would determine the GPS coordinates of Greece. The system would then search the metadata for objects that correspond to photographs and that also satisfy GPS coordinates for Greece. According to another embodiment of the invention, including GPS coordinates in the metadata of objects enables post-processing of the objects based on the GPS coordinates. For example, the objects may be subjected to an initial sort based on generalized GPS coordinates that are stored in the metadata and may be subjected to additional sorting based on more particular criteria for the GPS coordinates. Thus, a user may initially search for object metadata that corresponds to a location near a famous landmark and may use image matching to label the objects with the searchable metadata (i.e., textual descriptions). For example, using the GPS coordinates stored in the metadata, users may first determine that photographs were taken at the Jefferson Memorial and may use this information to label the photographs with searchable metadata (i.e., textual descriptions) including "photo of the Jefferson Memorial." One of ordinary skill in the art will readily appreciate that metadata may include various types of information and may be searched using the various types of information.

[079] According to one embodiment of the invention, the textual annotations may be communicated using a short message service on the mobile telephone or other device. Short message service is a text message service that enables sending and receiving of

Amazon/VB Assets
Exhibit 1003
Page 484

short textual messages. The textual messages may be stored at data centers for forwarding to intended recipients. Other configurations may be used.

[080] According to another embodiment of the invention, the mobile devices 36 may support multi-modal communications that enable displaying of non-speech search results on a graphical interface and receipt of speech commands to provide a follow-up search, among other configurations. For example, the user may be presented with textual search results corresponding to a name of a famous person and the user may provide a speech command to find a biography of the famous person. The system may maintain the context of the textual search results to find the biography associated with the famous person. By contrast, known systems may perform a follow-up search on the term "biography" and may present a dictionary definition of the term "biography."

[081] Figure 4 illustrates an embodiment that uses a fixed computer 44. In some embodiments, the main speech processing unit 98 and speech unit 128 may be embedded into the fixed computer 44. The main speech processing unit may interface with one or more processing units 84 that are associated with the fixed computer. The processing units may include one or more central processing units, one or more data and address busses, data interfaces, volatile memory, or other components. The processing unit may use one of more types of nonvolatile memory 94 for software and data storage. Suitable types of nonvolatile memory include, for example, flash memory and hard disk drives. In some embodiments, the main speech processing unit 98 may be integrated with the one or more processing units 84. According to one embodiment of the invention, processing may be performed at the fixed computers 44. The commands may be processed on-board to enable the fixed computers 44 to control themselves and/or to control other fixed computers 44, mobile devices 36, mobile telephones, or other devices. Additionally, fixed computers 44 may track context as described above.

[082] According to another embodiment of the invention, fixed computers 44 may be configured to allow verbal annotations of objects stored thereon. Fixed computers 44 may transcribe the verbal annotation to text and store the textual annotation with the object. Alternatively, fixed computers 44 may be configured to enable users to manually input textual descriptions that are stored along with the objects. According to one embodiment of the invention, the textual annotations and/or textual

Amazon/VB Assets
Exhibit 1003
Page 485

descriptions may be classified and searched. In an alternative embodiment, fixed computers 44 may classify and search the verbal annotations rather than textual annotations. However, classifying and searching verbal annotations may be considerably more difficult than classifying and searching textual annotations and/or textual descriptions.

[083] In another embodiment of the invention, users may interact with fixed computers 44 using speech unit 128, the keyboard 88 or keypad, a display 86 for displaying text, graphics, video, or other peripherals. According to some embodiments of the invention, the display may be a touch screen type. Alternatively, a pointing device (not shown) may be used, along with other devices. Fixed computers 44 may be coupled to one or more wired or wireless wide-area or local-area networks through one or more interfaces. A wide-area transceiver 92 may connect to the wide-area wireless network 38 or the data network 42, using a wireless or wired connection, including an IP network, a dial-up PSTN network connection, or other connections. The local-area network transceiver 90 may connect to wired or wireless local area networks. These networks may include connections to mobile devices 36.

[084] In order for devices to properly respond to requests and/or commands that are submitted in a natural language form, machine processable requests and/or algorithms may be formulated after the natural form questions or commands have been parsed and interpreted. Algorithms describe how the machines should gather data to respond to the questions or commands. Depending on the nature of the requests or commands, there may not be a simple set of requests and/or algorithms that will return an adequate response. Several requests and algorithms may need to be initiated and even these requests and algorithms may need to be chained or concatenated to achieve a complete response. Further, no single available source may contain the entire set of results needed to generate a complete response. Thus, multiple requests and/or algorithms, perhaps with several parts, may be generated to access multiple data sources that are located both locally or remotely. Not all of the data sources, requests and/or algorithms may return useful results or any results at all. Useful results that are returned are often embedded in other information and may need to be extracted from the other information. For example, a few key words or numbers may need to be "scraped" from a larger amount of other information in a text string, table, list, page,

Amazon/VB Assets
Exhibit 1003
Page 486

video stream or other information. At the same time, extraneous information including graphics or pictures may be removed to process the response. In any case, the multiple results must be evaluated and combined to form the best possible response, even in cases where some requests do not return useful results or fail to produce results entirely. In cases where the command is determined to be ambiguous or the result is inherently subjective, determining the results to present in the response is a complex process. Finally, to maintain a natural interaction, responses should be returned to the user rapidly. Managing and evaluating complex and uncertain requests, while maintaining real-time performance, is a significant challenge.

[085] The invention provides a complete speech-based command generation, information query, retrieval, processing and presentation environment or a combination of speech-based and non-speech-based command generation, information query, retrieval, processing and presentation environment for telematic applications. In addition, the invention may be useful for controlling the system itself and/or external devices. This integrated environment makes maximum use of context, prior information and domain and user specific profile data to achieve a natural environment for one or more users submitting requests or commands in multiple domains. Through this integrated approach, a complete speech-based natural language command, algorithm and response environment or a combination of speech-based and non-speech-based command, algorithm and response environment may be created.

[086] The telematic natural language interface may be deployed as part of, or a peripheral to a TCU or other mobile devices 36, as part of a mobile device interfaced to vehicle computers and other mobile system through wired, wireless, optical, or other types of connections or fixed computers interfaced to the vehicle computers or other systems through a combination of wired, wireless, optical and/or other types of connections. Alternatively, the components of the interactive natural language telematic interface may be distributed in any suitable manner between these multiple computing platforms.

[087] Figure 5 illustrates one exemplary schematic diagram view of a system 90 according to an embodiment of the invention. The system 90 may include a main unit 98, speech units 128, and multi-modal devices 155. In an alternative embodiment, the system 98 may include completely separate systems for main unit 98, speech units

Amazon/VB Assets
Exhibit 1003
Page 487

128 and multi-modal devices 155. The event manager 100 may mediate interactions between components of main unit 98. The event manager 100 provides a multi-threaded environment allowing the system 98 to operate on multiple commands or questions from multiple user sessions, without conflict, and in an efficient manner to maintain real-time response capabilities.

[088] FIG. 8 illustrates one exemplary embodiment of an enhanced conversational user interface 800 that receives multi-modal input from users 802 and communicates with a conversational speech analyzer 804. The conversational speech analyzer 804 may be coupled to several components, including a general cognitive model 806, an environmental model 808, a personalized cognitive model 810, and an adaptive misrecognition analysis engine 812 or other components. According to one embodiment of the invention, the conversational speech analyzer 804 may include one or more speech recognition engines that transcribe utterances to textual messages. The transcription may be performed using one or more conversational grammars, a context description grammar 112, or other transcription methods. According to one embodiment of the invention, data that is entered in a text format may be merged with data that is transcribed to a textual format from the utterance. Conversational speech analyzer 804 also may include a semantic knowledge-based model that analyzes the textual message and detects command components.

[089] According to one embodiment of the invention, personalized cognitive model 810 is a model derived from a user's interaction pattern with the system and may be used to predict what actions the user may take next in time, thus assisting with speech recognition and/or question or command recognition. Personalized cognitive model 810 may track actions performed by the user. When the system is attempting to predict user behavior, the personalized cognitive model may be consulted first. The system may have multiple personalized cognitive models, wherein one may correspond to each user.

[090] According to another embodiment of the invention, general cognitive model 806 is a statistical abstract that corresponds to interaction patterns with the system for multiple users. Data stored within general cognitive model 806 may be used to predict a user's next action, thus assisting with speech recognition and/or question or command recognition. The general cognitive model 806 may also track what actions

Amazon/VB Assets
Exhibit 1003
Page 488

a particular user has performed and may be used when the user interacts with the system in a way that is not handled in the personalized cognitive model.

[091] According to one embodiment of the invention, the environmental model 808 may include information associated with the user's environment and surroundings. The information may include the type of environment that a user is in (e.g., quiet or noisy); details of a microphone and/or speaker system; the user's current global position and movement, as may be determined by GPS; current system status, such as what song/movie is playing, is the system in the midst of retrieving something, or other system status; details on all voice-enabled devices in the immediate vicinity, such as a presence of a voice-enabled TV, stereo, and DVD player in the same room; a user's credit card information, such as numbers and current balances, wherein the user may ask a mobile telephone to download and pay for a video and the system may respond that there are insufficient funds; or other information. The information may be accessed to invoke a context, domain knowledge, preferences, and/or other cognitive qualities that enhance the interpretation of questions and/or commands.

[092] The conversational speech analyzer 804 may also access the general cognitive model 806 and/or the personalized cognitive model 810 to further refine a context, domain knowledge, preferences, and/or other cognitive qualities to enhance the interpretation of questions and/or commands. Based on information received from general cognitive model 806, environmental model 808 and/or the personalized cognitive model 810, the system may enhance responses to commands and questions by including a prediction of user behavior.

[093] Adaptive Misrecognition Analysis Engine 812 may analyze and store textual messages, including transcribed utterances, that are identified as being unrecognized or incorrectly recognized by conversational speech analyzer 804. Upon a determination that text is unrecognized, the system may generate an unrecognized event. For example, an unrecognized event may result from not finding a match to text and/or the transcribed utterance.

[094] According to one embodiment of the invention, the system may implement one or more techniques to determine that textual messages are incorrectly recognized. For example, the user may command the system to play a specific song title and the system may misrecognize the requested song title and provide a song having a different title, the system may misrecognize an incorrect song title and

Amazon/VB Assets
Exhibit 1003
Page 489

provide a song having a different than requested title, the system may misrecognize a correct song title and provide a song having an invalid song title, among other variations. When the system misrecognizes a request, the user typically provides immediate feedback such as, overriding the command in a time shorter than the expected execution time of the command, repeating the original request, issuing a stop command, or taking other action, wherein the action may be presented verbally, non-verbally by pushing a selected button on a cell phone or remote control, or both, among other configurations. According to one embodiment of the invention, the system may detect the user action and may prompt the user to re-phrase the request to enable the system to modify words in the query. The user's actions may be analyzed in a statistical model to determine a frequency occurrence of misrecognitions for particular commands, with the results being used to update the corresponding personalized cognitive model 810.

[095] According to another embodiment of the invention, the conversational speech analyzer 804 may access the personalized cognitive model 810 to proactively select a next best (or nth best) match for the received text. A match may be confirmed by user action that include the user not immediately canceling the command or taking other action. The misrecognitions may also be analyzed to potentially determine personalized tuning parameters for the speech recognition components of the system. For example, the system may, over time, tune the speech recognition engine to increase recognition by analyzing how the speech recognition engine misrecognizes the utterance.

[096] The event manager 100 may mediate interactions between other components of the invention. The event manager can provide a multi-threaded environment allowing the system to operate on multiple commands or questions from multiple user sessions without conflict and in an efficient manner, maintaining real-time response capabilities.

[097] Agents 106 may include a collection of grammars, criteria handlers, and algorithms that are accessible to respond to a set of requests and/or commands. Agents 106 may further contain packages of both generic and domain specific behavior for the system 98. Agents 106 may use nonvolatile storage for data, parameters, history information, and locally stored content provided in the system databases 102 or other local sources. One or more user profiles 110 may be provided

Amazon/VB Assets
Exhibit 1003
Page 490

that include user specific data, parameters, and session and history information for determining the behavior of agents 106. One or more personality modules 108 may be provided in a data determining system that include personality characteristics for agents. The update manager 104 manages the automatic and manual loading and updating of agents 106 and their associated data from the Internet 146 or other network through the network interface 116.

[098] According to one embodiment of the invention, the speech-based interface for the system 90 may include one or more speech units 128. Speech units 128 may include one or more microphones, for example array microphone 134, to receive the utterances from the user. The speech received at the microphone 134 may be processed by filter 132 and passed to the speech coder 138 for encoding and compression. In one embodiment, a transceiver module 130 may transmit the coded speech to the main unit 98. Coded speech received from the main unit 98 is detected by the transceiver 130, then decoded and decompressed by the speech coder 138 and annunciated by the speaker 136.

[099] According to one embodiment of the invention, the non-speech-based interface for the system 90 may include one or more multi-modal devices 155 that may include mobile devices, stand alone or networked computers, personal digital assistances (PDAs), portable computer devices, or other multi-modal devices.

[0100] The speech units 128, multi-modal devices 155 and the main unit 98 may communicate over a communication link. The communication link may include a wired or wireless link. According to one embodiment, the communication link may comprise an RF link. The transceiver 130 on the speech unit may communicate coded speech data bi-directionally over the communication link with the transceiver 126 on the main unit 98. According to another embodiment, the RF link may use any standard local area wireless data protocols including the IEEE 802.11, Bluetooth or other standards. Alternatively, an infrared data link conforming to any suitable standard such as IrDA or other infrared standards can be used. In an alternative embodiment, wires may connect the speech unit 128 and the main unit 98, eliminating the need for one speech coder 138. Other wired or wireless analog or digital transmission techniques can be used.

[0101] According to one embodiment of the invention, coded speech received at the transceiver 126 on the main unit 98 is passed to the speech coder 122 for decoding

Amazon/VB Assets
Exhibit 1003
Page 491

and decompression. The decoded speech may be processed by the speech recognition engine 120 using the context description grammar module 112, among other information. Any recognized information may be processed by the parser 118, which transforms information into complete algorithms and questions using data supplied by knowledge agents. Knowledge agents may be selected from the plurality of agents 106 using a grammar stack, wherein the knowledge agents provide information for generating a response to the question or command. The knowledge agents may then process the commands or questions by creating requests that are submitted to local databases 102 or submitted through the network interface 116 to external data sources over the Internet 146 or external other networks. Algorithms typically result in actions taken by the system 90 itself (i.e., pause or stop), or to a remote device or data source (i.e., download data or program, or control a remote device), through the network interface to the Internet or other data interface.

[0102] According to one embodiment of the invention, knowledge agents may return results of questions as responses to users. The responses may be created using the results of information requests, the system personality 108, the user preferences, other data in the user profile 110, and/or other information. Agents 106 may present these results using the speech unit 128. The agents 106 may create a response string, which is sent to the text to speech engine 124. The text to speech engine 124 may generate the utterances, which may be encoded and compressed by the speech coder 122. Once coded, the utterances are transmitted from the main unit 98 by the transceiver 126 to the transceiver 130 on the speech unit 128. The utterance is then decoded and decompressed by the speech coder 138 and output by the speaker 136. Alternatively, agents 106 may present the results using multi-modal devices 155.

[0103] The non-speech interface 114 may be part of, or separate from, the multi-modal devices 155 and may be used as a substitute or complement to the speech interface. For example, non-speech interface 114 may be used to present and interact with non-speech (e.g., graphical or tabular) information in a manner more easily understood by the user. According to one embodiment of the invention, multimodal support may be provided to maintain the context during both voice interaction and interaction through the non-speech interface 114. In one exemplary embodiment, a communication channel may be opened between multimodal devices 155 and the main user interface system 90 to allow multimodal devices 155 to input text

Amazon/VB Assets
Exhibit 1003
Page 492

commands and questions. According to another embodiment of the invention, multimodal devices 155 may send a string of text or keywords for a command or question. The main interface system 90 may synchronize the context between multimodal device 155 and the speech units 128. In order to send a response to the corresponding device, the main user interface system 90 may track where the input came from so that the response may be sent to a TTS or multi-modal device 155.

[0104] According to one embodiment of the invention, non-speech interface 114 may show system, state and history information in a more concise manner than is possible through the speech interface. Non-speech interface 114 may be accessed to create or extend capabilities of agents 106. These operations may include scripting of agents, adding data to the agent or databases 102 used by the agent, adding links to information sources, among other operations.

[0105] According to another embodiment of the invention, system 90 may include different types of agents 106. In some embodiments of the invention, generic and domain specific behavior and information may be organized into domain agents 156. The system agent, on the other hand, may provide default functionality and basic services. The domain specific agents may provide complete, convenient and re-distributable packages or modules for each application area. In other words, the domain agents may include data that is needed to extend or modify the functionality of the system 90 in a current or new domain. Further, domain agents and their associated data can be updated remotely over a network as new behavior is added or new information becomes available. Domain agents may access a plurality of sources that may provide various services. Domain agents may use the services of other, typically more specialized, data managers and system agent. Agents are distributed and redistributed in a number of ways including on removable storage media, transfer over networks or attached to emails and other messages. The invention may provide license management capability allowing the sale of data managers by third parties to one or more users on a one time or subscription basis. In addition, users with particular expertise can create data managers, update existing data managers by adding new behaviors and information and making these data managers available to other users. A block diagram of the agent architecture according to an embodiment of the invention is shown in Figure 6.

Amazon/VB Assets
Exhibit 1003
Page 493

[0106] Agents 106 may receive and return events to the event manager 100. Both system agents 150 and domain agents 156 may receive questions and commands from the parser 118. Based on keywords in the questions and commands and the structures of the questions and commands, the parser may invoke selected agents. Agents use the nonvolatile storage for data, parameters, history information and local content provided in the system databases 102.

[0107] According to one embodiment of the invention, when the system starts-up or boots-up, the agent manager 154 may load and initialize the system agent 150 and the one or more domain agents 156. Agent manager 154 includes knowledge of agents 106 and maps agents 106 to the agent library 158. At shutdown, the agent manager may unload the agents 106. The agent manager 154 also performs license management functions for the domain agents 156 and content in the databases 102.

[0108] The system agent 150 manages the criteria handlers 152 that handle specific parameters or values (criteria) used to determine context for questions and commands. According to one embodiment of the invention, criteria handlers 152 include parsing routines that are specialized to recognize particular parts of speech, such as times, locations, movie titles, and other parts of speech. The criteria handlers 152 may identify matching phrases and extract semantic attributes from the phrases. Both the system agent 150 and the domain agents 156 may use the criteria handlers 152. The various domain agents 156 may use the services of the system agent 150 and of other, typically more specialized, domain agents 156. The system agent 150 and the domain agents 156 may use the services of the agent library 158, which contains utilities for commonly used functions. According to one embodiment of the invention, the agent library may be a dynamic link library that implements one or more agents. The agent library may include utilities for text and string handling, network communications, database lookup and management, fuzzy and probabilistic evaluation, text to speech formats, and other utilities.

[0109] Domain agents 156 may be data-driven, scripted or created with compiled code. A base of generic agent may be used as the starting point for data-driven or scripted agents. Agents created with compiled code are typically built into dynamically linkable or loadable libraries. Developers of agents can add new functionality to the agent library 158 as required. Details of agent distribution and update, and agent creation or modification are discussed in sections below.

Amazon/VB Assets
Exhibit 1003
Page 494

[0110] According to another embodiment of the invention, capabilities are provided to distribute and update system agents 150, domain agents 156, agent library 158 components, databases 102, and context description grammar 112 over wireless or wired networks 136, including IP networks and dial-up networks using the update manager 104. The network interface 116 may provide connections to one or more networks. The update manager 104 may also manage the downloading and installation of core system updates. The agent manager 154 may perform license management functions for the domain agents and the databases. The update manager 104 and agent manager 154 may perform these functions for all agents and database content including, agents and content available to all users or agents and/or content available to selected users. Examples of agent and database components added or updated on a periodic basis include:

- agents for new domains;
- additional domain knowledge for agents;
- new keywords for a domain, which can include names of politicians, athletes, entertainers, names of new movies or songs, etc. who have achieved recent prominence;
- links to a preferred set of information sources for the domains covered including links for, entertainment, news, sports, weather, etc.;
- updates to domain information based on, for example, changes to tax laws, company mergers, changing political boundaries;
- updates to content, including dictionaries, encyclopedias and almanacs; and
- other content and database components.

[0111] When a user requires or selects a new domain agent 156 or database element 102, the update manager 104 may connect to their source on the network 146 though the network interface 116, download and install the agent or data. To save system resources and to comply with any license conditions, the update manger 104 may uninstall agents that are no longer in use. In one embodiment of the invention, the update manager 104 may periodically query one or more sources of the licensed agents and database components to locate and download updates to agent executables, scripts or data as they become available. Alternatively, the agent sources may initiate

Amazon/VB Assets
Exhibit 1003
Page 495

the downloading of agent updates of the registered or licensed agents to the update manager as they become available.

[0112] The agent manager 154 may provide license management clients that are capable of executing most any license terms and conditions. When a particular agent 106 and/or database element 102 is selected based on a submitted command, the agent manager 154 verifies that the use of the agent or data element is within the allowed terms and conditions, and if so, invokes the agent or allows access to the data element. License management schemes that can be implemented through the agent manager 154 include outright purchase, subscription for updates, one time or limited time use. Use of shared agents and data elements (such as those down-loaded from web sites maintained by groups of domain experts) may also be managed by the agent manager 154.

[0113] If questions or commands do not match an agent that is currently loaded on the system, the agent manager 154 may search the network 146 through the network interface 116 to find a source for a suitable agent. This process may be triggered, for example, when a query is made in a domain for which an agent is not available, when a new device is added to a mobile structure, or when the behavior of a mobile device is updated. Once located, the agent can be loaded under the control of the update manager 104, within the terms and conditions of the license agreement, as enforced by the agent manager.

[0114] New commands, keywords, information, or information sources can be added to any domain agent 156 by changing agent data or scripting. These configuration capabilities may allow users and content developers to extend and modify the behavior of existing domain agents 156 or to create new domain agents 156 from a generic agent without the need to create new compiled code. Thus, the modification of the domain agents 156 may range from minor data-driven updates by even the most casual users, such as specifying the spelling of words, to development of complex behavior using the scripting language as would typically be done by a domain expert. The user can create and manage modifications to domain agents 156 through speech interface commands or using non-speech interface 114. User-specific modifications of domain agents 156 are stored in conjunction with the user's profile 110 and accessed by the domain agent 156 at run-time.

Amazon/VB Assets
Exhibit 1003
Page 496

[0115] The data used to configure data driven agents 156 may be structured in a manner to facilitate efficient evaluation and to help developers with organization. This data is used not only by the agent, but also in the speech recognition engine 120, the text to speech engine 124, and the parser 118. Examples of some major categories of data include:

1. Content packages may include questions or commands. Each command or question or group of commands or questions may include contexts used for creation of one or more requests. The domain agent 156 may pass a context description grammar expression to the parser 118 for evaluation of a context or question. An initial or default context is typically supplied for each command or question. The command or question includes a grammar for the management and evaluation of the context stack.
2. Parameters and other operating data on devices that are under control of the natural language speech interface. The domain agent 156 may use this data and parameters to determine how to execute a command, how to formulate the command string for the parser 118, determine if the command is feasible, and determine if the command can be executed within safety and operating limits.
3. Page lists or pointers to other local or network content sources. For each page or content source, there may be a pointer (e.g. URL, URI, or other pointer) to the page or source. Each page may have specific scraping information used to extract the data of interest. The scraping information may include, for example, matching patterns, HTML or other format parsing information.
4. A response list, determining the response of the domain agent 156 to a particular command or question given the context, the user profile and the information retrieved. Responses may include diagnostic error messages or requests for more information if the question or command cannot yet be resolved from the known information. Responses can be based on or be dependent on thresholds or probabilistic or fuzzy weights for the variables.
5. Substitution lists that include variable substitutions and transformations, often applied by the agents 150, 156 in the formatting of requests and results. For example, a stock domain specific agent 156 would

Amazon/VB Assets
Exhibit 1003
Page 497

use a substitution list of company trading symbols, company names and commonly used abbreviations. Substitutions and transformations can be performed on commands and questions to create precise queries, which can be applied against one or more information sources or to results for creating more meaningful output to the user. Substitution lists also include information for optimally dealing with structured information, such as HTML formatted page parsing and evaluation.

6. Personalities used for responses. Personalities can be constructed by combining multiple traits in a weighted manner. Weights can be specified for each agent's domain area to create one or more specific personalities. Examples of personality traits include, sarcasm, humor, irritation, and sympathy, and other traits.

7. Public and user specific parameters for sources, substitutions, transformations, variables or criteria. The public parameter lists are part of the agent package 156. The user specific parameters are included in the user profile 110.

[0116] Commands and questions are interpreted, requests formulated, responses created and results presented can be based on the user's personal or user profile 110 values. Personal profiles may include information specific to the individual, their interests, their special use of terminology, the history of their interactions with the system, domains of interest, or other factors. The personal profile data can be used by the agents 106, the speech recognition engine 120, the text to speech engine 124, and the parser 118. Preferences can include, special (modified) commands, past behavior or history, questions, information sources, formats, reports, alerts or other preferences. User profile data can be manually entered by the user and/or can be learned by the system based on user behavior. User profile values can include: 1) spelling preferences; 2) date of birth for user, family and friends; 3) income level; 4) gender; 5) occupation; 6)location information such as, home address, neighborhood, and business address, paths traveled, locations visited; 7) vehicle type or types; 8) vehicle operator certifications, permits or special certificates; 9) history of commands and queries; 10) telecommunications and other service providers and services; 11) financial and investment information; 12) synonyms (i.e., a nick name for someone,

Amazon/VB Assets
Exhibit 1003
Page 498

different terms for the same item); 13) special spelling; 14) keywords; 15) transformation or substitution variables; 16) domains of interest; or, 17) other values.

[0117] End users may use the data driven agent 156 extension and modification facilities and values stored in user profiles 110 to create special reports, packages of queries, alerts and output formats. A single alert or report can be configured to use multiple data sources and other variables (i.e., time, location, measured value) value to determine when alerts should be sent. For example, an alert can be generated by sampling a stock price every 15 minutes and sending an alert if the price drops below some value. In another example, an alert may be generated when a particular condition or combination of conditions occur for the vehicles. Alerts and reports can be directed to a local or remote output.

[0118] To create a report, the user may first specify a set of commands or questions. Next, the user can create or select a format for the report. Finally the user may name the report. A report can have variable parameters. For example, a user may create a company stock report and execute the report by stating its name and the company name, which gives the user selected information in a specified format for that company. In another example, a user can create a "morning" report, which presents selected multimedia information from different sources (news, sports, traffic, weather) in the order and formats desired. In yet another example, the user can create a report on the status of one or more vehicle systems. Alerts and reports can be created using only voice commands and responses, commands and responses through the graphical user interface 114, or a combination of the two. Reports can be run locally or remotely with respect to the vehicle. To create a report, alert, or other specialized behavior, the user performs a number of steps including: 1) specifying the command to run a report or alert; 2) specifying the question or questions, including keywords, used for a query; 3) setting the criteria for running the report such as on command or when a particular condition is met; 4) defining preferred information sources; 5)defining preferences for order of result evaluation by source, value, and other parameters; 6) specifying the presentation medium for a report or alert, such as an email, the text to speech engine, a message to a pager, or a text and graphics display; and, 7) specifying the preferred format for the report, such as information to be presented, order of information to be presented, preferred abbreviations or other variable substitutions.

Amazon/VB Assets
Exhibit 1003
Page 499

[0119] Filtering and noise elimination may be important in facilitating the various functionalities of system 90 and may improve operation in noisy mobile environments. Recognition and parsing of the user's speech is implemented with good signal to noise ratio at the input to the speech recognition engine 120. To provide acceptable results, a set of acoustic models, an array microphone 134, a filter 132, or other components, may be employed. If a good signal to noise ratio cannot be achieved, a noise identification algorithm may be used and the appropriate acoustic model, for example, one that has been trained in conditions similar to the identified noise, may be selected. According to one embodiment of the invention, the microphone array, filters and speech coder 138 are physically separated from the main unit 98 into a speech unit 128, and connected using a wireless link. Since bandwidth on a wireless connection is at a premium, the speech coder dynamically adapts the digitization rate and compression of the captured speech.

[0120] Some embodiments of the invention may use one or more arrays of microphones 134 to provide better directional signal capture and noise elimination than can be achieved with a single microphone. The microphone array can be one-dimensional (a linear array) or two-dimensional (a circle, square, triangle or other suitable shape). The beam pattern of the array can be fixed or made adaptive though use of analog or digital phase shifting circuitry. The pattern of the active array is steered to point in the direction of the one or more users speaking. At the same time, nulls can be added to the pattern to notch out point or limited area noise sources. The use of the array microphone also helps reduce the cross talk between output from the text to speech engine 124 through the speaker 136 or from another user talking and detection of the user's speech.

[0121] The invention may use an analog or digital filter 132 between the array microphone or conventional microphone 134 and the speech coder 138. The pass band of the filter can be set to optimize the signal to noise ratio at the input to the speech recognition engine 120. In some embodiments, the filter is adaptive, using band shaping combined with notch filtering to reject narrow-band noise. One embodiment employs adaptive echo cancellation in the filter. The echo cancellation helps prevent cross talk between output from the text to speech engine and detection of the user's speech as well as suppression of environmentally caused echoes.

Amazon/VB Assets
Exhibit 1003
Page 500

Algorithms comparing the background noise to the signal received from the users speech may be used to optimize the band-shaping parameters of the adaptive filter.

[0122] The speech received by the array microphone 134 and passed through the filter 132 may be sent to the speech digitizer or coder 138. The speech coder may use adaptive lossy audio compression to optimize bandwidth requirements for the transmission of the coded speech to the speech recognition engine 120 over a wireless link. The lossy coding is optimized to preserve only the components of the speech signal required for optimal recognition. Further, the lossy compression algorithms that may be used are designed to prevent even momentary gaps in the signal stream, which can cause errors in the speech recognition engine. The digitized speech may be buffered in the coder and the coder may adapt the output data rate to optimize the use of the available bandwidth. The use of the adaptive speech coder is particularly advantageous when a band-limited wireless link is used between the coder and the speech recognition engine.

[0123] The microphone can be complemented with an analog or digital (i.e., Voice over IP) speech interface. This interface allows a remote user to connect to the system and interact with it in the same manner possible if they were physically present.

[0124] In an alternative embodiment, the array microphone can be replaced by a set of physically distributed microphones or headsets worn by the users. The distributed microphones can be placed in different parts of a vehicle, different parts of a room or in different rooms of a building. The distributed microphones can create a three-dimensional array to improve signal to noise ratio. The headset can use a wireless or wired connection.

[0125] While the invention is intended to be able to accept most any natural language question or command, ambiguity may be a problem. To assist users formulate concise questions and commands, the system can support a voice query language. The language may be structured to allow a variety of queries and commands with minimal ambiguity. Thus, the voice query language helps users clearly specify the keywords or contexts of the question or command along with the parameters or criteria. The language can provide a grammar to clearly specify the keyword used to determine the context and present a set of one or criteria or parameters. A user asking a question or

Amazon/VB Assets
Exhibit 1003
Page 501

stating a command in the voice query language may nearly always be guaranteed to receive a response.

[0126] The voice query language may be sensitive to the contents of the context stack, wherein a context defines a set of questions that can be activated or deactivated during a conversation. According to one embodiment, each agent may designate one context to be the root context that defines base algorithms that the agent implements. Thus, follow-on questions can be asked using an abbreviated grammar, since key words and criteria can be inherited from the context stack. For example, the user can simply ask about another keyword if the criteria of the question remain constant.

[0127] The system 90, may provide built in training capabilities to help users learn the best methods to formulate their questions and commands. The interactive training allows users to audibly or visibly see the machine interpretation of their queries and provides suggestions on how to better structure a query. Using the interactive training, users can quickly become comfortable with the voice query language and, at the same time, learn how to optimize the amount of information required with each step of a dialog.

[0128] The output of the speech coder 122 may be fed to the speech recognition engine 120. The speech recognition engine 120 recognizes words and phrases, using information in the context description grammar 112 and passes these to the parser 118 for interpretation. The speech recognition engine 120 may determine the user's identity by voice and name for each utterance. Recognized words and phrases may be tagged with this identity in all further processing. Thus, as multiple users engage in overlapping sessions, the tags added by the speech recognition engine 120 to each utterance allows other components of the system 90 to tie that utterance to the correct user and dialog. The user recognition capability may further be used as a security measure for applications, such as auctions or online shopping, where this is required. Voice characteristics of each user may be contained in the user profile 110.

[0129] According to one embodiment of the invention, users may start a dialog with the system 90 when they first address it. This can be done by speaking a generic word ("computer") or addressing a specific name ("Fred"), which may be generally tied to a system personality 108. Once the user starts the dialog, it may be recognized by the speech recognition engine 120, using unique characteristics of the user's speech. At

Amazon/VB Assets
Exhibit 1003
Page 502

the end of a dialog or to interrupt a dialog, the user may utter a dismissal word ("good bye").

[0130] According to another embodiment of the invention, the system 90 may employ a speech recognition engine 124 that gains improved word recognition accuracy using data from context description grammar 112, user profiles 110, and the agents 106, among other components. At the same time, the fuzzy set possibilities or prior probabilities for the words in the context description grammar may be dynamically updated to maximize the probability of correct recognition at each stage of the dialog. The probabilities or possibilities may be dynamically updated based on a number of criteria including the application domain, the questions or commands, contexts, the user profile and preferences, user dialog history, the recognizer dictionary and phrase tables, and word spellings, among other criteria.

[0131] For uncommon words or new vocabulary words, the user may be given the option to spell the words. The spelling can be done by saying the names or the letters or using a phonetic alphabet. The phonetic alphabet can be a default one or one of the user's choosing.

[0132] Alternatively, when a user submits a word that is not recognized at all or is not correctly recognized by the speech recognition engine 120, then the user may be asked to spell the word. The speech recognition engine 120 determines this condition based on confidence level for the scoring process. The word may be looked up in a dictionary and the pronunciation for the word is added to either the dictionary, the agent 106, or the user's profile 110. The word pronunciation can then be associated with the domain, the question, the context and the user. Through this process, the speech recognition engine learns with time and improves in accuracy. To assist users in spelling words, an individualized phonetic alphabet can be used. Each user can modify the standard phonetic alphabets with words, which they can remember more easily.

[0133] Once the words and phrases have been recognized by the speech recognition engine 120, the tokens and user identification may be passed to the parser 118. The parser examines the tokens for the questions or commands, context and criteria. The parser may determine a context for an utterance by applying prior probabilities or fuzzy possibilities to keyword matching, user profile 110, dialog history, and context stack contents. The context of a question or command may determine the domain

Amazon/VB Assets
Exhibit 1003
Page 503

and, thereby, the domain agent 156, if any, to be invoked. For example, a question with the keywords "temperature" implies a context value of weather for the question. Within a different dialog, the keyword "temperature" can imply a context for a measurement. The parser dynamically receives keyword and associated prior probability or fuzzy possibility updates from the system agent 150 or an already active domain agent 156. Based on these probabilities or possibilities the possible contexts are scored and the top one or few are used for further processing.

[0134] The parser 118 uses a scoring system to determine the most likely context or domain for a user's question and/or command. The score is determined from weighing a number of factors including, the user profile 110, the domain agent's data content and previous context. Based on this scoring, the system 90 invokes the correct agent. If the confidence level of the score is not high enough to ensure a reliable response, the system 90 may ask the user to verify whether the question and/or command is correctly understood.

[0135] In general, the question that is asked by the system 90 may be phrased to indicate the context of the question including all criteria or parameters. For example, the question can be in the form of: "Did I understand that you want such-and-such" If the user confirms that the question is correct the system proceeds to produce a response. Otherwise, the user can rephrase the original question, perhaps adding additional information to remove ambiguity, or the system can ask one or more questions to attempt to resolve the ambiguity.

[0136] Once the context for the question or command has been determined, the parser 118 can invoke the correct agent 156, 150. To formulate a question or command in the regular grammar used by agents, the parser will preferably determine required and optional values for the criteria or parameters. These criteria may have been explicitly supplied by the user or may need to be inferred. The parser may make use of the criteria handlers 152 supplied by the system agent. The criteria handlers can provide context sensitive procedures for extracting the criteria or parameters from the user's question or command. Some criteria may be determined by executing algorithms in the agent, while others may be determined by applying probabilistic or fuzzy reasoning to tables of possible values. Prior probabilities or fuzzy possibilities and associated values may be received from a number of sources including, for example, the history of the dialog, the user profile 110, and the agent. Based on user

Amazon/VB Assets
Exhibit 1003
Page 504

responses, the prior probabilities or fuzzy possibilities may be updated as the system learns the desired behavior. For a weather context, examples of criteria include, location, date and time. Other criteria can include command criteria (i.e., yes/no, on/off, pause, stop), and spelling. Special criteria handlers are available from the system agent for processing lists, tables, barge-in commands, long strings of text and system commands.

[0137] The criteria handlers 152 can operate iteratively or recursively on the criteria extracted to eliminate ambiguity. This processing may help reduce the ambiguity in the user's question or command. For example, if the user has a place name (or other proper noun) in their utterance the parser 118 can use services of the domain agent 156 to look up tables in the databases 102 for place names or can attempt to determine which word is the proper noun from the syntax of the utterance. In another example, the user asks, "what about flight one hundred and twenty too?" The parser and domain agent use flight information in the database and network information along with context to determine the most plausible interpretation among; flight 100 and flight 20 also, flight 100 and flight 22, flight 122, etc.

[0138] Once the context and the criteria are determined, the parser 118 may form the question or command in a standard format or hierarchical data structure used for processing by the agents 150, 156. The parser 118 may fill in all required and some optional tokens for the grammar of the context. Often the tokens must be transformed to values and forms acceptable to the agents. The parser obtains the required transformations from the agents, dialog history or user profile 110. Examples of transformations or substitutions performed by the parser on tokens include: 1) substituting a stock symbol for a company name or abbreviation; 2) substituting a numerical value for a word or words; 3) adding a zip code to an address; and, 4) changing a place or other name to a commonly used standard abbreviation.

[0139] The agents 150, 156 may receive a command or question once the parser 118 has placed it in the required standard format. Based on the context, the parser can evoke the correct agent to process the question or command.

[0140] Commands can be directed to the system or to an external entity. System commands are generally directed to the system agent 150. Commands for external entities are generally processed by a domain agent 156, which includes the command context and behavior for the external entity.

Amazon/VB Assets
Exhibit 1003
Page 505

[0141] Specific questions may be generally directed to one of the domain agents 156. The real-time selection of the correct agent allows the invention to dynamically switch contexts. Based on the question, command or context and the parameters or criteria, the domain agent may create one or more queries to one or more local or external information sources. Questions can be objective or subjective in nature. Results for objective questions can often be obtained by structured queries to one or more local or network information sources. Even for objective questions, the system may need to apply probabilistic or fuzzy set analysis to deal with cases of conflicting information or incomplete information. Information to answer subjective questions is generally obtained by one or more ad-hoc queries to local or network data sources, followed by probabilistic or fuzzy set evaluation of the one results to determine a best answer.

[0142] Once the domain agent 156 has formulated the one or more queries, they may be sent to local and/or network information sources. The queries may be performed in an asynchronous manner to account for the fact that sources respond at different speeds or may fail to respond at all. Duplicate queries can be sent to different information sources to ensure that at least one source responds with a useful result in a timely manner. Further, if multiple results are received in a timely manner, they can be scored by the system to determine which data is most reliable or appropriate. Examples of data sources accommodated include, HTTP data sources, sources with meta-data in various formats including XML, measurement data from sensors using various formats, device 32 setting parameters, entertainment audio, video and game files including MP3, databases using query languages and structured responses such as SQL, and other data sources.

[0143] The local information sources can be stored in one or more system databases 102 or can be on any local data storage such as a set of CDs or DVDs in a player or other local data storage. In other cases, local information can be obtained from vehicle system settings or measurement devices. Network information sources can be connected to, the control and device interfaces 30, the data interfaces 26, the Internet 42 or other network and accessed through a series of plug-ins or adaptors, known a pluggable sources, in the network interface 116. The pluggable sources are capable of executing the protocols and interpreting the data formats for the data sources of interest. The pluggable source provides information scrapping forms and procedures

Amazon/VB Assets
Exhibit 1003
Page 506

for each source to the domain agents 156. If a new type of data source is to be used a new plug-in or adaptor can be added to the appropriate interface.

[0144] The domain agent 156 can evaluate the results of the one or more queries as they arrive. The domain agent may score the relevance of the results based on results already received, the context, the criteria, the history of the dialog, the user profile 110 and domain specific information using probabilistic or fuzzy scoring techniques. Part of the dialog history is maintained in a context stack. The weight of each context for the scoring may be based on the relevance of one context to another and the age of the contexts. Other scoring variables can be associated through the context stack. Contexts can also be exclusive, so that previous contexts have no weight in the scoring.

[0145] Based on the on-going scoring processes, the domain agent 156 may determine if a single best answer can be extracted. For most questions, the desired result may include a set of tokens that may be found to formulate an answer. Once a value has been found for each of these tokens, the results are ready for presentation to the user. For example, for a question on weather, the tokens can include the date, day of week, predicted high temperature, predicted low temperature, chance of precipitation, expected cloud cover, expected type of precipitation and other tokens. Results processed in this manner may include error messages. For subjective questions, this determination is made by determining a most likely answer or answers, extracted by matching of the results received. If no satisfactory answer can be inferred from the results of the query, the agent can do one of the following:

1. Ask the user for more information, typically through the speech interface, and based on the results obtained formulate new queries. This approach is applied when an irresolvable ambiguity arises in the formulation of a response.
2. Formulate new queries based on the results received from the first set of queries. This approach is typically applied in cases where the responses received do not include all the required information. Information sources to queries can be inferred from the results already obtained (i.e., links in an HTML document or measurements or settings from other devices 32) or from other sources. Using this approach one or more sets of queries and responses can be chained without the need for action by the user.

Amazon/VB Assets
Exhibit 1003
Page 507

3. Wait for additional queries to return results.

[0146] In any case, the domain agent 156 may continue to make queries and evaluate results until a satisfactory response is constructed. In doing so, the agent can start several overlapping query paths or threads of inquiry, typically mediated by the event manager 100. This technique, combined with the use of asynchronous queries from multiple data sources provides the real-time response performance required for a natural interaction with the user.

[0147] The domain agent 156 may apply conditional scraping operations to each query response as it is received. The conditional scraping actions may depend on the context, the criteria, user profile 110, and domain agent coding and data. For each token to be extracted a scraping criteria 152 can be created using the services of the system agent 150. The scraping criteria may use format specific scraping methods including, tables, lists, text, and other methods. One or more scraping criteria can be applied to a page or results set. Once additional results are received, the domain agent can create new scraping criteria to apply to results already acquired. The conditional scarping process removes extraneous information, such as graphics, which need not be further processed or stored, improving system performance.

[0148] Specific commands are generally directed to one of the domain agents 156. The real-time selection of the correct agent allows the invention to dynamically switch contexts. Command oriented domain agents 156 evaluate the command and the state of vehicle systems, system capabilities, and measurements to determine if the command can be executed at all or if doing so will exceed operating or safety limits. If the command is ambiguous or cannot be executed for some other reason, the system may ask the user for more information or may suggest what the problem is and a likely approach to the solution. The domain agent may format the command for the specific device 32 and control and device interface 30. This formatting may involve variable substitution, inference of missing values and other formatting. Variable substitution and inference depends on the command context, the user profile 110, command history, state of vehicle systems and measured values, and other factors. A complex command can result in more atomic commands being sent to multiple devices, perhaps in a sequence. The sequence and nature of subsequent commands may depend on the previous commands, results of pervious commands, device settings and other measurements. As a command is executed, measurements are made

Amazon/VB Assets
Exhibit 1003
Page 508

and results collected to determine if the execution was correct and the desired state or states were reached.

[0149] Once the domain agent 156 has created a satisfactory response to a question, or to a command, the agent may format that response for presentation. Typically, the domain agent can format the response into the markup format used by the text to speech engine 124. The domain agent may format the result presentation using available format templates and based on the context, the criteria, and the user profile 110. The domain agent may perform variable substitutions and transformations to produce a response best understood and most natural to the user. The domain agent may vary the order of presentation of tokens and the exact terminology used to create a more natural response to the user. The domain agent may also select the presentation personality 108 to be used.

[0150] For both command and query responses, the domain agent 156 may select the presentation template, determine order of presentation for tokens and determine variable substitutions and transformations using probabilistic or fuzzy set decision methods. The template used to form the presentation can be from the domain agent itself or from the user profile 110. The user profile can completely specify the presentation format or can be used to select and then modify an existing presentation format. Selection and formatting of presentation template can also depend on the presentation personality 108. At the same time, the characteristics of the personality used for the response are dynamically determined using probabilities or fuzzy possibilities derived from the context, the criteria, the domain agent itself and the user profile 110.

[0151] The domain agent 156 may apply a number of transformations to the tokens before presentation to the user. These variable substitutions and transformations may be derived from a number of sources including, domain information carried by the agent, the context, the token values, the criteria, the personality 108 to be used, and the user profile 110. Examples of variable substitutions and transformations include: 1) substitution of words for numbers; 2) substitution of names for acronyms or symbols (i.e., trading symbols); 3) use of formatting information derived from the information sources (i.e., HTML tags); 4) nature of the response including, text, long text, list, table; 5) possible missing information or errors; 6) units for measurement

Amazon/VB Assets
Exhibit 1003
Page 509

(i.e., English or metric); and, 7) preferred terminology from the user profile or presentation personality 108.

[0152] The invention may provide special purpose presentation capabilities for long text strings, tables, lists and other large results sets. Domain agents 156 may use special formatting templates for such results. The system agent 150 can provide special criteria handlers 152 for presentation and user commands for large results sets. The presentation templates used by the domain agents for large results sets typically include methods for summarizing the results and then allowing the user to query the result in more detail. For example, initially only short summaries, such as headlines or key numbers, are presented. The user can then query the results set further. The criteria handlers provide users with the capability to browse large results sets. Commands provided by the criteria handlers for large results sets include, stop, pause, skip, rewind, start, and forward.

[0153] Some information, in formats such as video, pictures and graphics, may be best presented in a displayed format. The domain agents 156 apply suitable presentation templates in these cases and present the information through the non-speech interface 114. The system agent 150 provides special criteria handlers 152 for presentation and user commands for display presentation and control.

[0154] Although particular embodiments of the invention have been shown and described, it will be understood that it is not intended to limit the invention to the embodiments that are disclosed and it will be obvious to those skilled in the art that various changes and modifications may be made without departing from the spirit and scope of the invention. Thus, the invention is intended to cover alternatives, modifications, and equivalents, which may be included within the spirit and scope of the invention as defined by the claims.

Amazon/VB Assets
Exhibit 1003
Page 510

We claim:

1. A mobile device for processing speech and non-speech communications, comprising:

an identification module that identifies a user that communicates with the mobile device;

a matching module that matches the identified user to one or more profiles;

a transcription module that transcribes the speech and the non-speech communications to create a speech-based textual message and a non-speech-based textual message, wherein the transcription module accesses the one or more profiles;

a merging module that merges the speech-based textual message and the non-speech-based textual message to generate a query;

a search module that searches the query for text combinations;

a comparison module that compares the text combinations to entries in a context description grammar;

a plurality of domain agents that are associated with the context description grammar;

a scoring module that provides relevance scores based on results from the comparison module;

a domain agent selector that selects domain agents based on results from the scoring module; and

a response generating module that communicates with the selected domain agents to obtain content that is gathered by the selected domain agents and that generates a response from the content, wherein the content is arranged in a selected order based on results from the scoring module.

2. The mobile device according to claim 1, wherein the response generating module generates an aggregate response that includes the content gathered by the selected domain agents.

3. The mobile device according to claim 1, wherein the identification module supports multiple users that access the mobile device during a same session.

Amazon/VB Assets
Exhibit 1003
Page 511

4. The mobile device according to claim 3, wherein the session may include an overlapping session or an interleaved session.

5. The mobile device according to claim 3, wherein the identification module verifies an identify of the multiple users based on voiceprint matching or password matching.

6. The mobile device according to claim 4, wherein, during the interleaved session, the mobile device processes one or more queries in an order of receipt.

7. The mobile device according to claim 1, wherein, during the interleaved session, the mobile device processes one or more queries based on an order determined by a length of the queries.

8. The mobile device according to claim 1, wherein the mobile device includes (i) a personal digital assistant, (ii) a cellular telephone, (iii) a portable computer, (iv) a desktop computer, or any combination of (i) to (iv).

9. The mobile device according to claim 1, wherein the mobile device receives follow-up speech and non-speech communications and wherein the transcription module transcribes the follow-up speech and non-speech communications to create a follow-up speech-based textual message and a follow-up non-speech-based textual message.

10. The mobile device according to claim 9, wherein the merging module merges the follow-up speech-based textual message and the follow-up non-speech-based textual message to generate a follow-up query.

11. The mobile device according to claim 1, further comprising a personality module that facilitates formatting the response.

12. The mobile device according to claim 1, wherein the mobile device is configured to communicate with a context stack that includes one or more contexts that are selected based on the query.

13. The mobile device according to claim 12, wherein the scoring module determines the one or more contexts based on at least applying prior probabilities or fuzzy possibilities to (i) keyword matching, (ii) user profiles, (iii) a dialog history, or any combination of (i) to (iii).

Amazon/VB Assets
Exhibit 1003
Page 512

14. The mobile device according to claim 1, wherein at least one of the domain agents creates and directs a request to a local information source or a network information source.

15. The mobile device according to claim 1, wherein at least one of the domain agents creates and directs a command to a remote or local device.

16. The mobile device according to claim 10, wherein the follow-up query is associated with a same context as the query.

17. The mobile device according to claim 1, wherein the mobile device supports Internet browsing and wherein the generated query is used to conduct the Internet browsing.

18. The mobile device according to claim 1, wherein the mobile device supports multimodal communications.

19. The mobile device, comprising:

an object that is accessible by the mobile device;

an input device that receives speech and non-speech annotations corresponding to the object;

a transcription module that transcribes the speech and the non-speech annotations to create a speech-based textual message and a non-speech-based textual message; and

an association module that associates the object with the speech-based textual message and the non-speech-based textual message.

20. The mobile device according to claim 19, further comprising:

a classification module that classifies the speech-based textual message and the non-speech-based textual message; and

a searching module that searches the speech-based textual message and the non-speech-based textual message based on a search request.

21. The mobile device according to claim 19, wherein the association module associates the object with the speech and non-speech annotations.

22. The mobile device according to claim 21, further comprising:

a classification module that classifies the speech and non-speech annotations; and

Amazon/VB Assets
Exhibit 1003
Page 513

a searching module that searches the speech and non-speech annotations based on a search request.

23. The mobile device according to claim 19, wherein the object includes digital media.

24. The mobile device according to claim 23, wherein the object includes a digital photograph, a calendar entry, an email message, an instant message, a phonebook entry, a voicemail entry or a digital movie.

25. The mobile device according to claim 19, wherein the association module associates the object to the speech and non-speech annotations, the speech-based textual message, and the non-speech-based textual message.

26. The mobile device according to claim 25, wherein the object is retrievable using (i) a voice command, (i) a text command, or a combination of (i) and (ii).

27. The mobile device according to claim 19, wherein the object is stored (i) in the mobile device, (ii) at a server, or a combination of (i) and (ii).

28. The mobile device according to claim 19, wherein the object includes metadata.

29. The mobile device according to claim 28, wherein the metadata includes global positioning system information, environmental information, or geographic information.

30. The mobile device according to claim 28, further comprising a metadata searching module that enables searching of the metadata.

31. A system, comprising:

a plurality of mobile devices that support multimodal communications, wherein two or more mobile devices are associated with a user; and

a context module that is coupled to the plurality of mobile devices, wherein a same context is maintained at the two or more mobile devices that are associated with the user.

32. The system according to claim 31, wherein the plurality of mobile devices comprise:

an identification module that identifies a user that communicates with the mobile device;

Amazon/VB Assets
Exhibit 1003
Page 514

a matching module that matches the identified user to one or more profiles;

a transcription module that transcribes the speech and the non-speech communications to create a speech-based textual message and a non-speech-based textual message, wherein the transcription module accesses the one or more profiles;

a merging module that merges the speech-based textual message and the non-speech-based textual message to generate a query;

a search module that searches the query for text combinations;

a comparison module that compares the text combinations to entries in a context description grammar;

a plurality of domain agents that are associated with the context description grammar;

a scoring module that provides relevance scores based on results from the comparison module;

a domain agent selector that selects domain agents based on results from the scoring module; and

a response generating module that communicates with the selected domain agents to obtain content that is gathered by the selected domain agents and that generates a response from the content, wherein the content is arranged in a selected order based on results from the scoring module.

33. The system according to claim 31, wherein the context module is configured to broadcast the same information to two or more mobile devices and to receive a response from one of the mobile devices.

34. The system according to claim 31, wherein the context module broadcasts a first type of information and receives a second type of information.

35. The system according to claim 34, wherein first type of information and the second type of information include speech information or non-speech information.

36. The system according to claim 35, wherein the first type of information includes non-speech information that is presented on a graphical user interface and the second type of information includes speech information that is applied to search a network.

Amazon/VB Assets
Exhibit 1003
Page 515

37. The system according to claim 36, wherein the first type of information and the second type of information correspond to the same context.

38. A method of processing speech and non-speech communications in a multimodal environment, comprising:

receiving the speech and non-speech communications;

identifying a user that provides the speech and non-speech communications;

matching the identified user to one or more profiles;

transcribing the speech and non-speech communications to create a speech-based textual message and a non-speech-based textual message, wherein the transcribing includes accessing the one or more profiles;

merging the speech-based textual message and the non-speech-based textual message to generate a query;

searching the query for text combinations;

comparing the text combinations to entries in a context description grammar;

accessing a plurality of domain agents that are associated with the context description grammar;

generating a relevance score based on results from comparing the text combinations to entries in the context description grammar;

selecting one or more domain agents based on results from the relevance score;

obtaining content that is gathered by the selected domain agents; and

generating a response from the content, wherein the content is arranged in a selected order based on results from the relevance score.

39. The method according to claim 38, further comprising generating an aggregate response that includes the content that is gathered by the selected domain agents.

40. The method according to claim 38, further comprising:

receiving a follow-up speech and non-speech communications;

transcribing the follow-up speech and non-speech communications to create a follow-up speech-based textual message and a follow-up non-speech-based textual message; and

Amazon/VB Assets
Exhibit 1003
Page 516

merging the follow-up speech-based textual message and the follow-up non-speech-based textual message to generate a follow-up query.

41. The method according to claim 38, further comprising a personality module that communicates the response to a user.

42. The method according to claim 38, wherein the system supports multiple users during a same session.

43. The method according to claim 42, wherein the session may include an overlapping session or an interleaved session.

44. The method according to claim 42, wherein the multiple users are identified based on voiceprint matching or password matching.

45. The method according to claim 43, further comprising processing one or more queries in an order of receipt.

46. The method according to claim 43, further comprising processing one or more queries based on an order determined by a length of the queries.

47. The method according to claim 38, further comprising generating a context stack that includes one or more contexts that are selected based on the query.

48. The method according to claim 47, wherein the one or more contexts are generated based on applying prior probabilities or fuzzy possibilities to (i) keyword matching, (ii) user profiles, (iii) a dialog history, or any combination of (i) to (iii).

49. The method according to claim 38, further comprising conducting Internet browsing with the generated query.

Amazon/VB Assets
Exhibit 1003
Page 517

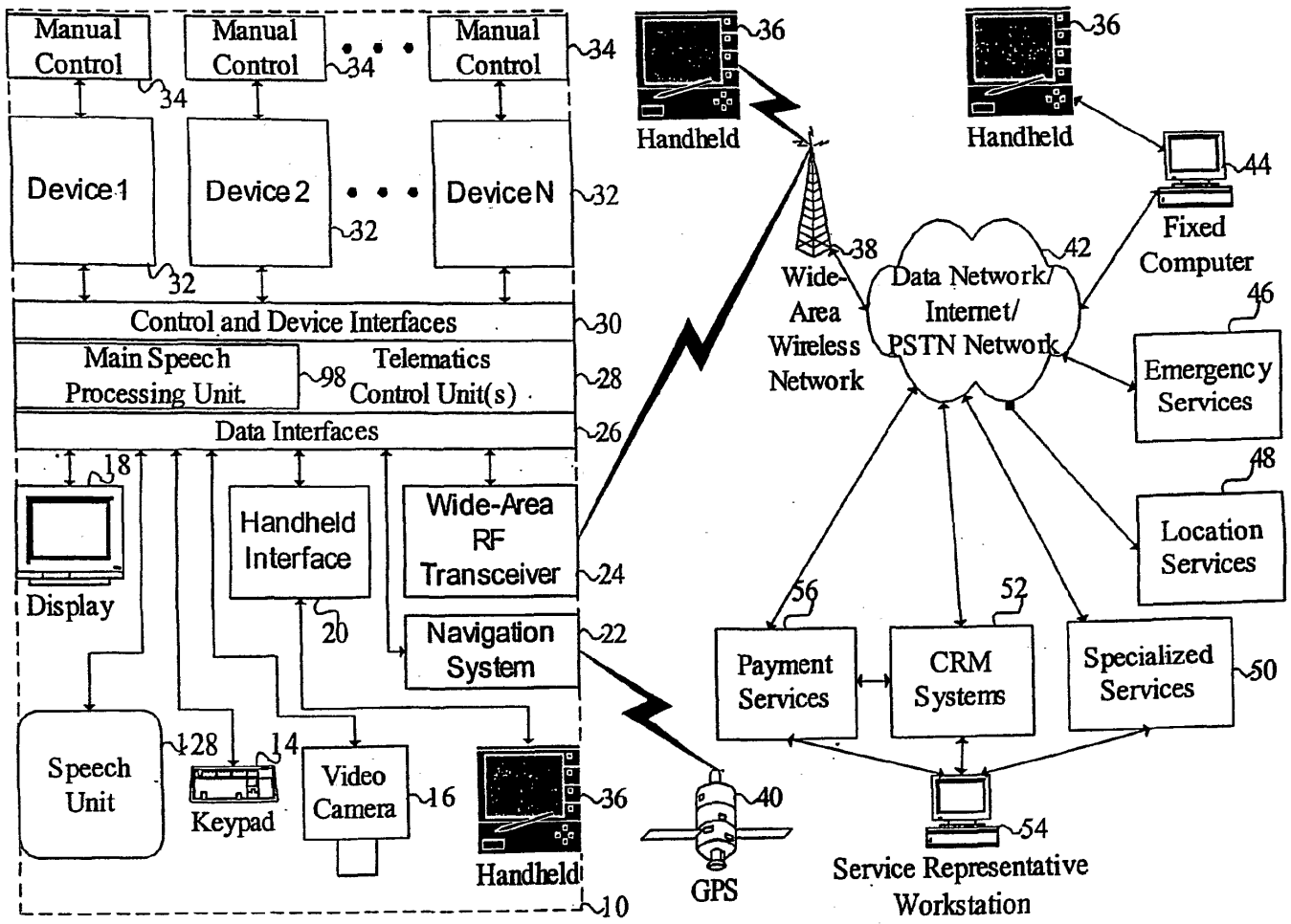

Manual Control
Manual Control
Manual Control
34
Device 1
Device 2
Device N
32
Control and Device Interfaces
30
Main Speech Processing Unit
98
Telematics Control Unit(s)
28
Data Interfaces
26
18
Display
Handheld Interface
20
Wide-Area RF Transceiver
24
Navigation System
22
Speech Unit
128
Keypad
14
Video Camera
16
Handheld
36
10
Handheld
36
Handheld
36
38
Wide-Area Wireless Network
42
Data Network/ Internet/ PSTN Network
44
Fixed Computer
46
Emergency Services
48
Location Services
56
Payment Services
52
CRM Systems
Specialized Services
50
40
GPS
54
Service Representative Workstation

Figure 1

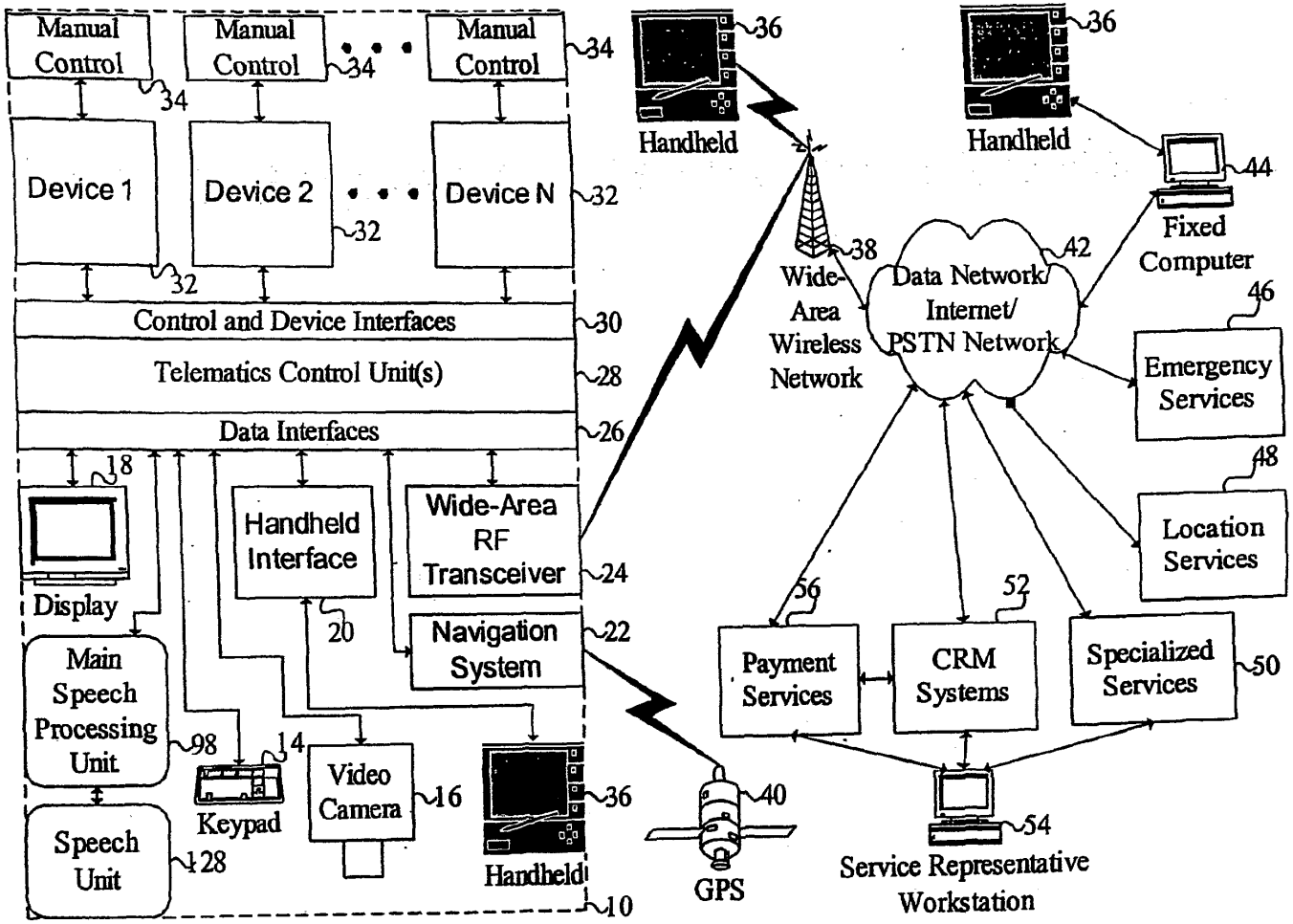


Figure 2

Amazon/VB Assets
Exhibit 1003
Page 519

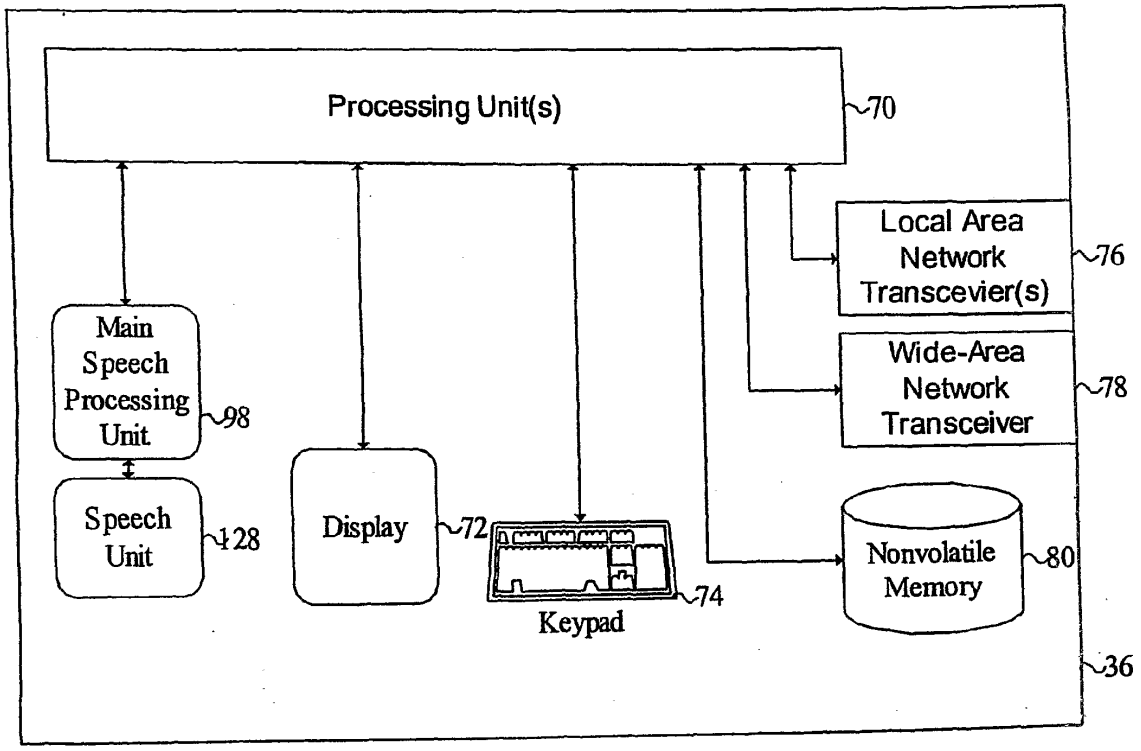

Processing Unit(s)
70
Local Area Network Transcevier(s)
76
Wide-Area Network Transceiver
78
Main Speech Processing Unit
98
Speech Unit
128
Display
72
Keypad
74
Nonvolatile Memory
80
36

Figure 3.

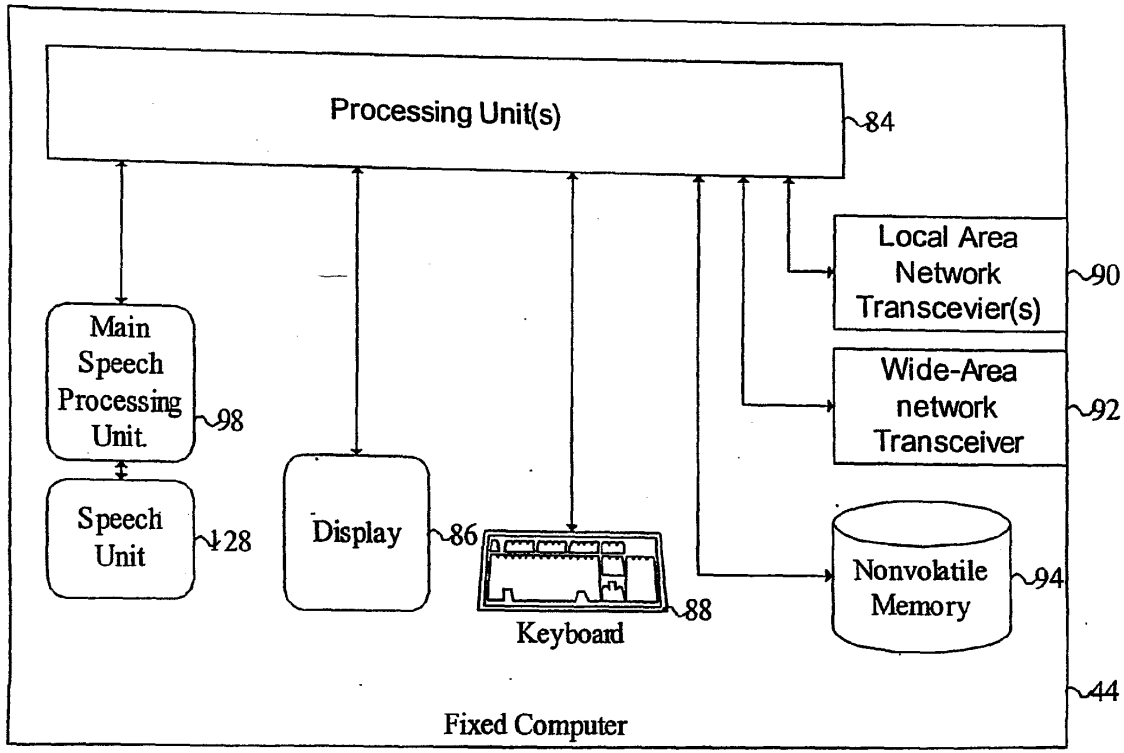

Processing Unit(s)
84
Local Area Network Transcevier(s)
90
Wide-Area network Transceiver
92
Main Speech Processing Unit
98
Speech Unit
128
Display
86
Keyboard
88
Nonvolatile Memory
94
44
Fixed Computer

Figure 4

FIG. 5

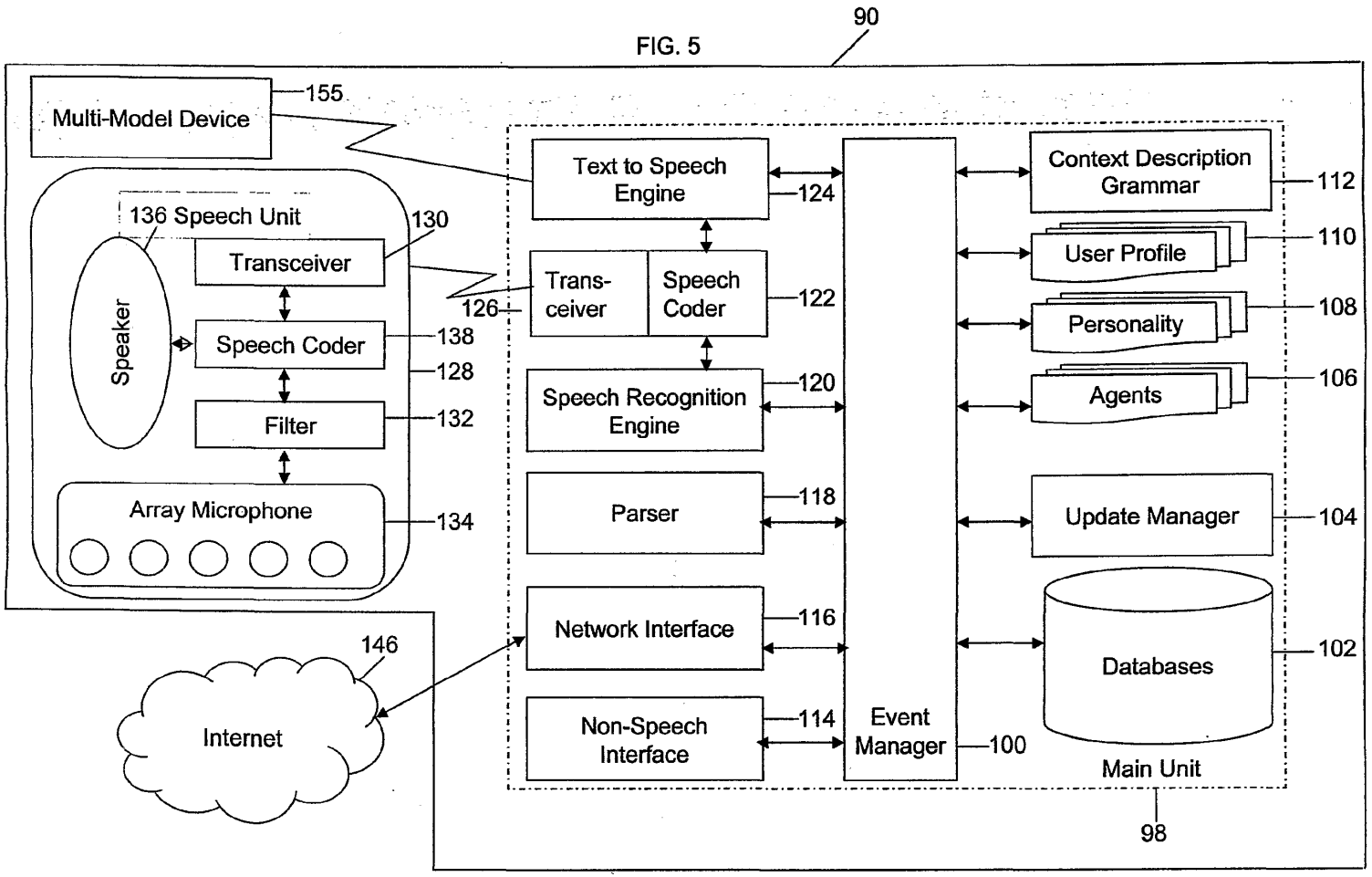

90
Multi-Model Device
155
136 Speech Unit
130
Transceiver
Speaker
Speech Coder
138
128
Filter
132
Array Microphone
134
126
Text to Speech Engine
124
Trans-ceiver
Speech Coder
122
Speech Recognition Engine
120
Parser
118
Network Interface
116
Non-Speech Interface
114
Event Manager
100
Main Unit
98
Context Description Grammar
112
User Profile
110
Personality
108
Agents
106
Update Manager
104
Databases
102
146
Internet

FIG. 6

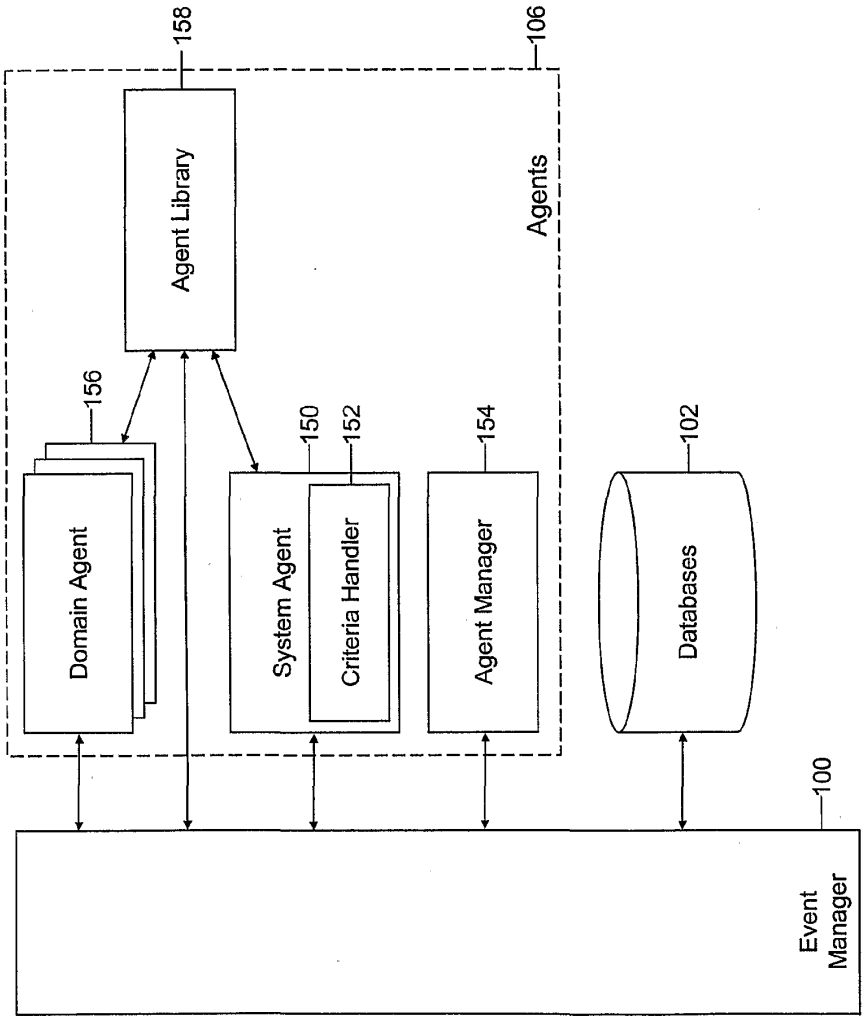
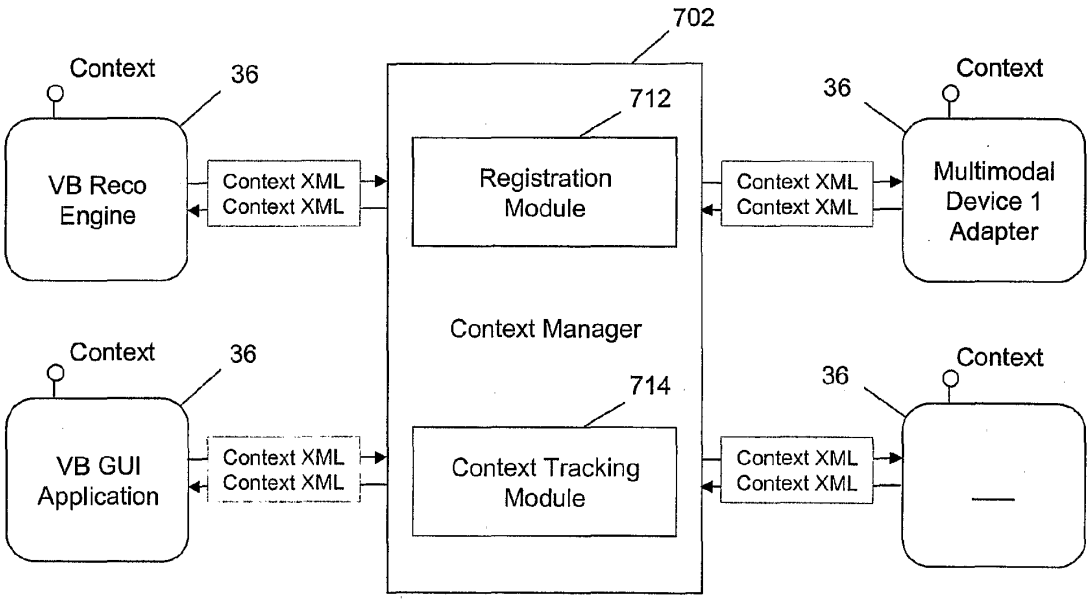

158
106
Agent Library
Agents
156
150
152
154
102
Domain Agent
System Agent
Criteria Handler
Agent Manager
Databases
100
Event Manager

FIG. 7


700
702
Context
36
VB Reco Engine
Context XML
Context XML
712
Registration Module
Context XML
Context XML
36
Context
Multimodal Device 1 Adapter
Context Manager
Context
36
VB GUI Application
Context XML
Context XML
714
Context Tracking Module
Context XML
Context XML
36
Context

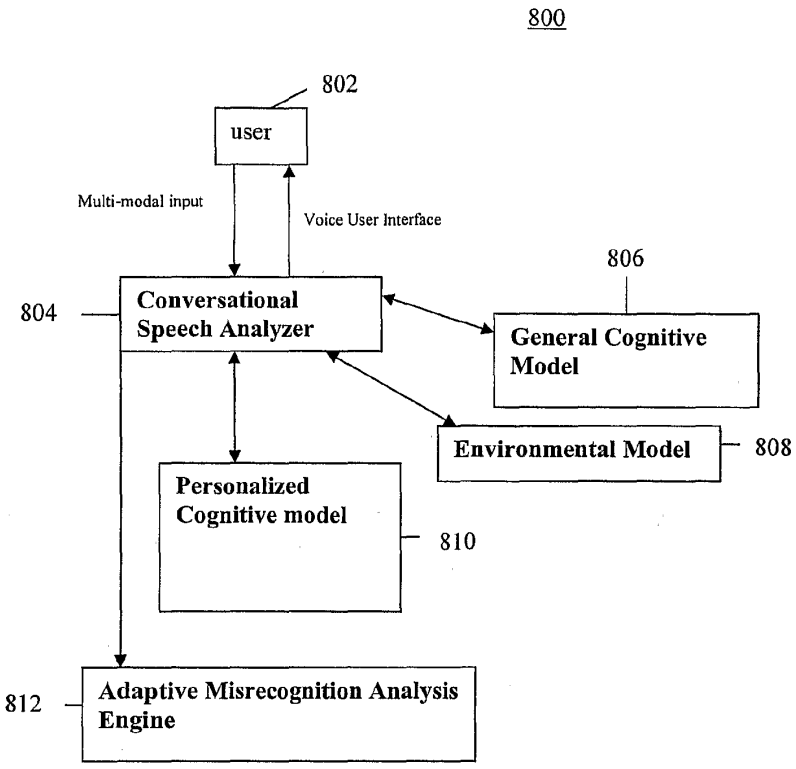

800
802
user
Multi-modal input
Voice User Interface
806
804
Conversational Speech Analyzer
General Cognitive Model
Environmental Model
808
Personalized Cognitive model
810
812
Adaptive Misrecognition Analysis Engine

FIGURE 8

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 March 2007 (08.03.2007)

PCT

(10) International Publication Number
WO 2007/027989 A2

(51) International Patent Classification:
G10L 15/18 (2006.01)

(21) International Application Number:
PCT/US2006/034184

(22) International Filing Date: 31 August 2006 (31.08.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/712,412 31 August 2005 (31.08.2005) US

(71) Applicant *(for all designated States except US)*: **VOICEBOX TECHNOLOGIES, INC.** [US/US]; 10505 NE 38th Place, Building 9, Kirkland, WA 98033 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **DI CRISTO, Philippe** [FR/US]; 177 107th Avenue, N.E., #1007, Bellevue, WA 98004 (US). **KE, Min** [US/US]; 3221 201st Place, S.E., Bothell, WA 98012 (US). **KENNEWICK, Robert, A.** [US/US]; 2204 S. Eddy St., Seattle, WA 98108 (US). **TJALVE, Michael** [DK/US]; 14904 SE 47th Court, Bellevue, WA 98006 (US).

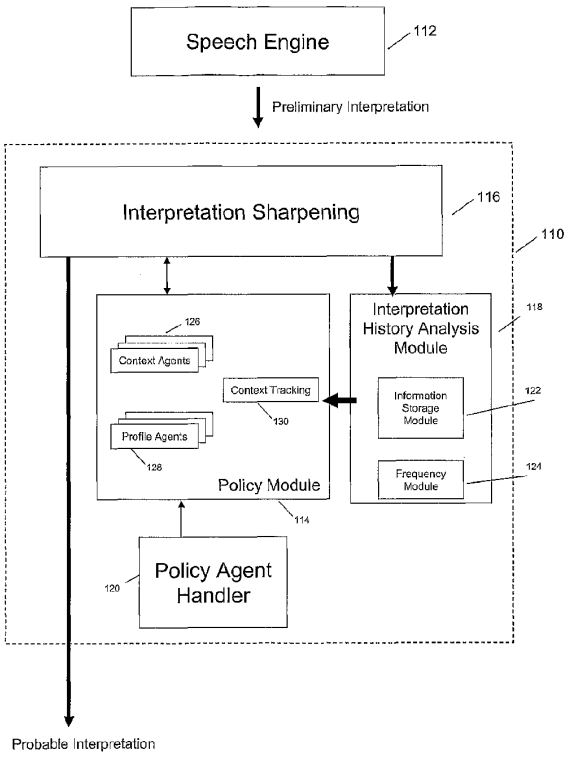
(74) Agents: **GATTO, James, G.** et al.; PILLSBURY WINTHROP SHAW PITTMAN LLP, P.O. Box 10500, Mclean, VA 22102 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,

[Continued on next page]

(54) Title: DYNAMIC SPEECH SHARPENING



(57) Abstract: An enhanced system for speech interpretation is provided. The system may include receiving a user verbalization and generating one or more preliminary interpretations of the verbalization by identifying one or more phonemes in the verbalization. An acoustic grammar may be used to map the phonemes to syllables or words, and the acoustic grammar may include one or more linking elements to reduce a search space associated with the grammar. The preliminary interpretations may be subject to various post-processing techniques to sharpen accuracy of the preliminary interpretation. A heuristic model may assign weights to various parameters based on a context, a user profile, or other domain knowledge. A probable interpretation may be identified based on a confidence score for each of a set of candidate interpretations generated by the heuristic model. The model may be augmented or updated based on various information associated with the interpretation of the verbalization.

WO 2007/027989 A2

FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

DYNAMIC SPEECH SHARPENING

CROSS REFERENCE TO RELATED APPLICATIONS

[001] This application claims the benefit of provisional U.S. Patent Application Serial No. 60/712,412, entitled "Dynamic Speech Sharpening," filed August 31, 2005, and which is hereby incorporated by reference in its entirety.

[002] This application is related to the following co-pending applications, which are hereby incorporated by reference in their entirety: "Systems and Methods for Responding to Natural Language Speech Utterance," Ser. No. 10/452,147, filed June 3, 2003, and published as US Patent Application Publication No. US 2004/0044516; "Mobile Systems and Methods for Responding to Natural Language Speech Utterance," Ser. No. 10/618,633, filed July 15, 2003, and published as US Patent Application Publication No. US 2004/0193420; "Systems and Methods for Responding to Natural Language Speech Utterance," Ser. No. 11/197,504, filed August 5, 2005; "System and Method of Supporting Adaptive Misrecognition in Conversational Speech," Ser. No. 11/200,164, filed August 10, 2005; and "Mobile Systems and Methods of Supporting Natural Language Human-Machine Interactions," Ser. No. 11/212,693, filed August 29, 2005.

FIELD OF THE INVENTION

[003] The invention is related generally to automated speech interpretation, and in particular, to enhancing the accuracy and performance of speech interpretation engines.

BACKGROUND OF THE INVENTION

[004] The field of automated speech interpretation is in increasingly higher demand. One use of automated speech interpretation is to provide voice requests to electronic devices. This may enable a user to simply speak to an electronic device rather than manually inputting requests, or other information, through pressing buttons, uploading information, or by other request input methods. Controlling various electronic devices through speech may enable the user to use the electronic devices more efficiently.

[005] However, existing technology in the field of automated speech interpretation, such as standard speech engines, automatic speech recognition (ASR), and other systems for interpreting speech, are unable to process a speech signal in an efficient manner, often constructing large grammars that include a large number of items,

Amazon/VB Assets
Exhibit 1003
Page 528

nodes, and transitions, which is a concern particularly for large-list recognition for embedded applications. If the grammar for an embedded application grows too much, it may not fit within the constrained space of an embedded application. With limited CPU power, response time and performance is easily affected due to the significant time needed to compile and load the grammar. Response time is further degraded because the speech engine has to parse through a large number of transition states to come up with a recognition result. Even when the speech engine is able recognize a word, the results are often unreliable because large grammars introduce greater risk of confusion between items as the size of the grammar increases. Existing techniques focus on reducing the size of a grammar tree by removing command variants or criteria items, but this approach strips functionality from the application.

[006] In addition to the performance problems associated with speech recognition engines that employ large word grammars, existing speech processing engines are unable to interpret natural human speech with a suitable accuracy to sufficiently control some electronic devices. In particular, speech interpretation engines still have substantial problems with accuracy and interpreting words that are not defined in a predetermined vocabulary or grammar context. Poor quality microphones, extraneous noises, unclear or grammatically incorrect speech by the user, or an accent of the user may also cause shortcomings in accuracy, such as when a particular sound cannot be mapped to a word in the grammar.

[007] In light of these and other problems, there is a need for enhanced automated speech interpretation that may interpret natural human speech with an augmented accuracy.

SUMMARY OF THE INVENTION

[008] According to one aspect of the invention, a system for enhancing automated speech interpretation is provided. The system may include a set of techniques for use in a speech-to-text engine to enhance accuracy and performance, for example, by reducing the search space of the speech engine. The problems with large-list recognition for embedded applications may also be improved by using phonetic dictation, which may recognize a phoneme string by disregarding the notion of words. The system may also use one or more post-processing techniques to sharpen an output of a preliminary speech interpretation made by a speech engine. The system may be modeled at least partially after one or more speech pattern

Amazon/VB Assets
Exhibit 1003
Page 529

recognition techniques used by humans, such as interpreting speech using words, word sequences, word combinations, word positions, context, phonetic similarities between two or more words, parts of speech, or other techniques.

[009] In one implementation of the invention, the system may receive a verbalization made by a user, where a speech engine may receive the verbalization. The speech engine may output information relating to a plurality of preliminary interpretations of the verbalization, where the plurality of preliminary interpretations represent a set of best guesses at the user verbalization. According to one aspect of the invention, the performance of the speech engine may be improved by using phoneme recognition. Phoneme recognition may disregard the notion of words, instead interpreting a verbalization as a series of phonemes, which may provide out-of-vocabulary (OOV) capabilities, such as when a user misspeaks or an electronic capture devices drops part of a speech signal, or for large-list applications, such as city and street names or song titles, for example. Phoneme recognition may be based on any suitable acoustic grammar that maps a speech signal into a phonemic representation. For example, the English language may be broken down into a detailed grammar of the phonotactic rules of the English language. Portions of a word may be represented by a syllable, which may be further broken down into core components of an onset, a nucleus, and a coda, which may be further broken down into sub-categories. Various different acoustic grammars may be formed as trees with various branches representing many different syllables forming a speech signal.

[010] According to another aspect of the invention, the performance of the speech engine and the phonemic recognition may be improved by pruning the search space used by the speech engine using a common phonetic marker. In one implementation, the acoustic grammar may be represented entirely by a loop of phonemes. In another implementation, the speech engine may reduce the search space by reducing the number of transitions in a grammar tree, thereby speeding up the process of compiling, loading, and executing the speech engine. For example, the phoneme loop may include a linking element between transitions. This may reduce the number of grammar transitions, such that grammar paths merge after a first transition and diverge after the linking element. In one implementation of the invention, a common acoustic element that is part of a speech signal may be used as the linking element. In one implementation of the invention, the acoustic element may be one that is very likely to be triggered even if it is unpronounced. For

Amazon/VB Assets
Exhibit 1003
Page 530

example, a *schwa* in the English language may be used as the linking element because *schwa* represents an unstressed, central vowel that is likely to be spoken even if unintended. Those skilled in the art will appreciate that acoustic models for different languages may use other frequently elided phonemes as linking elements to reduce the search space used by the speech engine.

[011] The speech engine may generate a plurality of preliminary interpretations representing a set of best guesses at the user verbalization. The preliminary interpretations may be stored in a matrix, array, or another form, and may be provided to an interpretation sharpening module to determine a probable interpretation of a verbalization made by a user by applying heuristic policies against the preliminary interpretation to identify dominant words and/or phrases. According to various aspects of the invention, the interpretation sharpening module may include a policy module that may manage and/or provide one or more policies that enable the sharpening module to generate a plurality of probable interpretations of the verbalization made by the user. For example, according to one aspect of the invention, the plurality of preliminary interpretations may be applied against one or more policies to generate a set of hypotheses as to a candidate recognition. Each hypothesis may be reanalyzed to generate an interpretation score that may relate to a likelihood of the probable interpretation being a correct interpretation of the verbalization, and the preliminary interpretation corresponding to the highest (or lowest) interpretation score may then be designated as a probable interpretation of the verbalization. The designated probable interpretation may be stored and used for augmenting the policies to improve accuracy.

[012] According to one aspect of the invention, the policy module may include one or more agents that represent domains of knowledge. The agents may compete using a weighted model to revise a preliminary interpretation by determining context and intent. Relevant substitution of suspect words and phrases may be based on phonetic similarities or domain appropriateness. A domain agent may include one or more domain parameters for determining a probable interpretation from a preliminary interpretation. For example, domain parameters may include a policy vocabulary, a word position in the verbalization, a word combination, a sentence structure, or other parameters. A domain agent may include a parameter weighting scheme that may weight individual parameters according to one or more weighting factors, such as, a frequency of use, a difficulty to understand, or other factors.

Amazon/VB Assets
Exhibit 1003
Page 531

[013] According to one aspect of the invention, the domain agents may revise a preliminary interpretation into a probable interpretation using phonetic fuzzy matching (PFM). In one implementation of the invention, the speech engine may output a phoneme stream that is applied against a model of phoneme feature similarities, drawn from domain agents, to identify a closest phonetic match using a multi-pass method. Domain agents may be loaded and prioritized into an M-Tree, which accounts for the possibility of the speech engine dropping or adding phonemes. An M-Tree may be an index structure that resolves similarity queries between phonemes using a closest-distance metric based on relative weightings of phoneme misrecognition, phoneme addition, and phoneme deletion. The M-Tree may be updated using an adaptive misrecognition model. For example, information about a verbalization and its components, as well as a probability that the probable interpretation was correct, may be stored and used for adapting the policy module for the user.

[014] In one implementation of the invention, the domain agents in the policy module may include one or more profile agents that may manage and/or provide one or more profile policies for revising a preliminary interpretation of a phoneme stream. For example, a profile agent may correspond to a user and may include one or more profile parameters tailored to the user. The profile agent may be used as a base policy to interpret any verbalizations made by the user. In other implementations, a profile agent may correspond to a particular language, a regional accent, or other profiles for interpreting a user verbalization. The profile agents may be augmented to enable the system to provide more accurate interpretations of verbalizations made by the user. The augmentation may include a user augmentation, such as providing additional vocabulary (e.g., names in an address book), one or more personalized pronunciations or other pronunciation information, or other user provided augmentations. The augmentation may also include a non-user provided augmentation, such as updates generated by a third party (e.g., a commercial administration and/or maintenance entity), or other non-user provided augmentations. The augmentation may be automated, such as adjusting a profile parameter-weighting scheme through an adaptive misrecognition model, as discussed above.

[015] In another implementation of the invention, the domain agents in the policy module may include one or more context agents that may manage and/or provide

Amazon/VB Assets
Exhibit 1003
Page 532

one or more context policies for revising a preliminary interpretation of a phoneme stream. For example, a context agent may correspond to a context, such as song titles, city and street names, movie titles, finance, or other contexts. A context agent may include one or more context parameters that may be tailored to a verbalization context. The context policy may enhance an ability of the system related to interpreting verbalizations made by the user in the verbalization context corresponding to the context agent. The context agents may be augmented to enable the system to provide more accurate interpretations of verbalizations made in a verbalization context corresponding to the context agent. The augmentation may include a user provided augmentation, a non-user provided augmentation, an automated augmentation, or other augmentations. The augmentation may be automated, such as adjusting a profile parameter-weighting scheme through an adaptive misrecognition model, as discussed above.

[016] According to various implementations of the invention, the policy module may determine which profile agents and/or which context agents to use through a set of heuristics provided in a context-tracking module. In one implementation, the context-tracking module may use phonetic fuzzy matching to track a series of verbalizations by the user to identify a verbalization context. The context-tracking module may utilize one or more M-Trees to track the series of verbalizations and determine a closest phonetic match. The context-tracking module may track one or more past verbalization contexts for the series of verbalizations, one or more current verbalization contexts for the series of verbalizations, and/or make predictions regarding one or more future verbalization contexts for the series of verbalizations. The policy module may utilize information about the verbalization context of the series of verbalizations generated by the context tracking module to manage and/or provide one or more profile and/or context agents.

[017] According to one aspect of the invention, the system may include an interpretation history analysis module that may enable the system to augment one or more domain agents based on an analysis of past interpretations related to previously interpreted verbalizations. The augmentations enabled by the interpretation history analysis module may include a user augmentation, a third-party augmentation, an automated augmentation, or other augmentations. The interpretation history analysis module may include an information storage module that may store interpretation information related to past verbalizations, such as one

Amazon/VB Assets
Exhibit 1003
Page 533

or more preliminary interpretations associated with a past verbalization, one or more interpretation scores associated with a past verbalization, one or more probable interpretations associated with a past verbalization, whether or not a past verbalization was interpreted correctly, or other information. A frequency module may be included in the interpretation history module, and the frequency module may use some or all of the information stored in the information storage module to generate one or more frequencies related to one or more past verbalizations. For example, the frequency module may calculate a frequency of word usage, word combinations, phonetic homonyms, interpretation errors for a particular verbalization, or other frequencies.

[018] The Information generated and/or stored by the interpretation history analysis module may be used to augment the profile and/or context agents in order to enhance the accuracy of subsequent interpretations. In some implementations, an adaptive misrecognition model may use one or more generated frequencies to augment one or more agents. For example, one or more parameters and/or weighting schemes of an agent or phonetic model may be augmented based on a frequency generated by the interpretation history analysis module. Other augmentations using information stored and/or generated by the interpretation history analysis module may be made, and the system may include a policy agent handler that may augment, update, remove, and/or provide one or more domain agents to the system. A domain agent may comprise a profile or context agent, and the policy agent handler may be controlled, directly or indirectly by a third party (e.g. a commercial entity, etc.). The policy agent handler may augment, update, remove, and/or provide domain agents to the system as part of a commercial agreement, such as a licensing agreement, a subscription agreement, a maintenance agreement, or other agreements.

[019] Other objects and advantages of the invention will be apparent to those skilled in the art based on the following detailed description and accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[020] Fig. 1 illustrates an exemplary system for enhancing automated speech interpretation according to one implementation of the invention.

[021] Fig. 2 illustrates an exemplary grammar tree for enhancing the performance of a speech engine according to one implementation of the invention.

Amazon/VB Assets
Exhibit 1003
Page 534

[022] Fig. 3 illustrates an exemplary flow chart of a method for enhancing automated speech interpretation according to one implementation of the invention.

DETAILED DESCRIPTION

[023] Fig. 1 illustrates an exemplary system for enhancing automated speech interpretation according to one implementation of the invention. A speech-to-text processing engine 112 may receive a user verbalization, and speech engine 112 may generate one or more preliminary interpretations of the user verbalization. The preliminary interpretations may represent a set of best guesses as to the user verbalization arranged in any predetermined form or data structure, such as an array, a matrix, or other forms. In one implementation of the invention, speech engine 112 may generate the preliminary interpretations by performing phonetic dictation to recognize a stream of phonemes, instead of a stream of words. Phonemic recognition provides several benefits, particularly in the embedded space, such as offering out-of-vocabulary (OOV) capabilities, improving processing performance by reducing the size of a grammar, and eliminating the need to train Statistic Language Models (SLMs). Those skilled in the art will recognize other advantages of phonemic recognition.

[024] Speech engine 112 may apply the phoneme stream against one or more acoustic grammars that reliably map a speech signal to a phonemic representation in order to generate the plurality of preliminary interpretations. Characteristics of a speech signal may be mapped to a phonemic representation to construct a suitable acoustic grammar, and various acoustic grammars may be included in speech engine 112 to generate one or more preliminary interpretations according to the various acoustic grammars. For example, the English language may be mapped into a detailed acoustic grammar representing the phonotactic rules of English, where words may be divided into syllables, which may further be divided into core components of an onset, a nucleus, and a coda, which may be further broken down into one or more sub-categories.

[025] Once the phonotactic rules of a speech signal have been identified, a detailed acoustic grammar tree may be constructed that accounts for the nuances of the speech signal. The acoustic grammar may include a loop of phonemes, or the phoneme loop may include a linking element to reduce a size of a search space associated with the grammar. Using the English language as an example, the grammar tree may include various branches representing English language

Amazon/VB Assets
Exhibit 1003
Page 535

syllables. The speech engine may traverse one or more grammar trees to generate one or more preliminary interpretations of a phoneme stream as a series of syllables that map to a word or phrase. By using phonemic recognition rather than word recognition, the size of the grammar can be reduced, which reduces the amount of time required to compile, load, and execute speech interpretation. Moreover, because the grammar maintains a high level of phonotactic constraints and therefore a large number of syllables, speech engine 112 may be very precise in generating phonemic representations of human verbalizations.

[026] An acoustic grammar used by speech engine 112 may be further optimized to reduce compile time, load time, and execution time by reducing the size of a search space associated with the acoustic grammar. Referring now to Fig. 2, a traditional grammar tree 120 is compared to an exemplary grammar tree according to one aspect of the invention to demonstrate the performance enhancements of speech engine 112. In traditional speech processing engines, nodes in a grammar tree 210 tend to represent words, or large-list applications may be supported provided through a grammar tree 210 where the nodes represent items in the large-list. This requires the speech engine to parse through a large number of transition states to come up with a recognition result, which degrades response time. An example of this is seen in the following grammar structure:

[027] "<street name> <city name>" → e.g., "NE 24th Street Bellevue"

[028] In the above example, a large list of street names is followed by a large list of city names. Assuming three elements in the list of street names, and three elements in the list of city names, this results in twenty-one transitions, which may be represented by traditional grammar tree 210. Every end-node of the first list is followed by all entries in the second list, potentially leading to very large grammars because most real-world large-list applications are likely to include much more than three list items. For example, a city may have hundreds or thousands of street names, and there may be hundreds or thousands of city names. Moreover, every element in the second segment of traditional grammar tree 210 is repeated, once for each first segment, which introduces redundancy.

[029] According to an aspect of the invention, the problems with traditional grammar trees may be resolved by using phonemic acoustic grammars instead of large-lists. The grammar may further be improved by including linking elements to reduce the number of transition states in the grammar. Thus, a grammar tree with a linking

Amazon/VB Assets
Exhibit 1003
Page 536

element 220 will merge after a first segment and then spread out again at a second segment, where the segments may represent a phoneme in an acoustic grammar, as discussed above. For example, assume a two-syllable word in an acoustic grammar consisting of three phonemes, which is able to reduce the number of transitions from twenty-one in a traditional grammar tree 210 to twelve in a grammar tree with a linking element 220. Two syllables and three phonemes are chosen to show the reduction in search space in a grammar tree with a linking element 220 as opposed to a corresponding traditional grammar tree 210, although a real-world acoustic grammar modeled after a language is likely to have a maximum of roughly fifty phonemes. Moreover, the search space may be further reduced by restricting available transitions based on phonotactic constraints for an acoustic model.

[030] Using the approach described in Fig. 2, adding a linking element to an acoustic grammar may reduce both grammar size and response time. Part of a speech signal may be mapped to the linking element in order to maintain the phonotactic rules of the acoustic grammar. The linking element may be an acoustic element that is likely to be triggered even if unpronounced. For example, a *schwa* represents an unstressed, central vowel in the English language (e.g., the first and last sound in the word "arena" is *schwa*). The phoneme *schwa* is an ideal linking element because of how it is represented in a frequency spectrum. That is, *schwa* is a brief sound and when a person opens their mouth to speak, there is a strong likelihood of passing through the frequencies of *schwa* even if unintended. Those skilled in the art will recognize that this approach may be extended to acoustic models of speech signals for other languages by using frequently elided phonemes as linking elements to reduce the search space of an acoustic grammar.

[031] Referring again to Fig. 1, speech engine 112 may receive a user verbalization and process the verbalization into a plurality of preliminary interpretations using the techniques described above. That is, the verbalization may be interpreted as a series of phonemes, and the series of phonemes may be mapped to one or more preliminary interpretations by traversing one or more acoustic grammars that are modeled after grammar 220 of Fig. 2. The plurality of preliminary interpretations may take the form of words, parts of words, phrases, utterances, or a combination thereof, and the plurality of preliminary interpretations may be arranged as a matrix, an array, or in another form. The plurality of preliminary interpretations are then

Amazon/VB Assets
Exhibit 1003
Page 537

passed to a speech sharpening engine 110 for deducing a most probable interpretation of the verbalization.

[032] According to various aspects of the invention, speech sharpening engine 110 may include an interpretation sharpening module 116, a policy module 114, an interpretation history analysis module 118, and a policy agent handler 120. The plurality of preliminary interpretations may be received by interpretation sharpening module 116, which forwards the preliminary interpretations to policy module 114 for further processing. Policy module 114 may include one or more context agents 126, one or more profile agents 128, and a context tracking module 130 that collectively revise the plurality of preliminary interpretations into a set of hypotheses that represent candidate recognitions of the verbalization. Policy module 114 may assign each hypothesis an interpretation score, and interpretation sharpening module 116 may designate the hypothesis with the highest (or lowest) interpretation score as a probable interpretation.

[033] According to one aspect of the invention, policy module 114 may include one or more context agents 126. Context agents 126 may represent domains of knowledge corresponding to a given context, such as song titles, city and street names, finance, movies, or other contexts. Context agents 126 may use context objects and associated dynamic languages to represent a corresponding context. Policy module 114 may also include one or more profile agents 128. Profile agents 128 may represent domains of knowledge corresponding to a given profile, such as a specific user, language, accent, or other profiles. Profile agents 128 may use profile objects and dynamic languages to represent a corresponding profile. Dynamic languages for context agents 126 or profile agents 128 may specify vocabularies, word combinations, phrases, sentence structures, criteria, and priority weightings for any given context or profile, respectively. The priority weightings may weight individual parameters according to one or more weighting factors, such as assigning a weight according to a frequency of use, a difficulty to understand, or other factors. Policy module 114 may also include a context-tracking module 130. Context tracking module 130 may track a verbalization context of a consecutive series of verbalizations. Context tracking module 130 may utilize one or more conversation trees to track the series of verbalizations. Context tracking sub-module 214 may track one or more past or current verbalization contexts of the series of verbalizations, and/or may make predictions regarding one or more future

Amazon/VB Assets
Exhibit 1003
Page 538

verbalization contexts of the series of verbalizations. Policy module 114 may utilize information about the verbalization context, generated by context tracking module 130, to generate one or more sharpened interpretations and corresponding interpretation scores.

[034] In some implementations, policy module 114 may use context tracking module 130 to apply objects from one or more context agents 126 and/or profile agents 128 to the preliminary interpretations provided by speech engine 112. The various agents may compete with each other using a set of heuristics in a phonetic fuzzy matcher, where an intent or context of the user may be identified based on the set of heuristics about how a request may be phrased in a given domain. A closest phonetic match may be identified for suspect words and/or phrases among the plurality of preliminary interpretations.

[035] The phonetic fuzzy matcher may include an M-Tree that is populated with context objects, profile objects, and/or dynamic language data from one or more of context agents 126 and/or profile agents 128. M-Trees are known to those skilled in the art. The M-Tree may assign relative priority weights to the context objects, profile objects, and/or dynamic language data in order to account for the possibility of misrecognized phonemes, extraneous phonemes, or erroneously deleted phonemes. A closest distance metric associated with the M-Tree may be used given the relative weightings of phoneme misrecognition, phoneme addition, and phoneme deletion for various contexts and/or profiles.

[036] According to one aspect of the invention, one or more passes may be taken over the plurality of preliminary interpretations to identify dominant words and/or phrases among the plurality of preliminary interpretations. Using the M-Tree weighted model, one or more candidate interpretations may be made based on relevant substitution of suspect words and/or phrases based on phonetic similarities and/or domain appropriateness. For example, if a set of dominant words appear to be a movie name, a candidate interpretation will substitute the relevant words and/or phrases to generate a candidate interpretation about movies. After a set of candidate interpretations have been generated, the candidate interpretations are analyzed using the M-Tree weighted model. With the relevant domains constrained by the candidate interpretations, a confidence or interpretation score may be assigned to each candidate interpretation, with the interpretation score representing a likelihood that a particular candidate interpretation is a correct interpretation of the

Amazon/VB Assets
Exhibit 1003
Page 539

verbalization. The candidate interpretations may then be returned to interpretation sharpening module 116, and interpretation sharpening module 116 may select a candidate interpretation with a highest (or lowest) interpretation score as a probable interpretation of the verbalization.

[037] According to various implementations of the invention, speech sharpening engine 110 may include an interpretation history analysis module 118. Interpretation history analysis module 118 may include an information storage module 122 a frequency module 124. Information storage module 122 may store information related to verbalizations, including components of verbalizations, preliminary interpretations, dominant words and/or phrases, candidate interpretations, probable interpretations, and/or interpretation scores associated with verbalizations, as well as whether or not a verbalization was interpreted correctly, or other information. Interpretation history analysis module 118 may also include a frequency module 124. Frequency module 124 may use some or all of the information stored in information storage module 122 to generate one or more frequencies related to one or more past verbalizations. For example, frequency module 124 may calculate a word usage frequency, a word combination frequency, a frequency related to a set of verbalizations that are phonetically similar but have distinct meanings, an interpretation error frequency for a particular verbalization, or other frequencies.

[038] Information stored and/or generated by interpretation history analysis module 118 may be used to augment speech sharpening engine 110. In some implementations, the information may be used to adjust various weights used in phonetic models, such as context agents 126 or profile agents 128, as well as adapting the relative weights in the M-Tree in context tracking module 130 to enhance accuracy for subsequent verbalizations. In another implementation, the stored information may be sent to a third-party or commercial entity for analyzing the data and developing new domain agents or further improving the accuracy of speech sharpening engine 110. For example, one or more parameters and/or weighting schemes of an agent may be augmented based on a frequency generated by interpretation history analysis module 118. Other augmentations related to information stored on and/or generated by interpretation history analysis module 118 may be made. Speech sharpening engine 110 may also include a policy agent handler 120 that may augment, update, remove, and/or provide one or more domain agents to policy module 114. A domain agent may include one or more new,

Amazon/VB Assets
Exhibit 1003
Page 540

modified, or updated context agents 126 and/or profile agents 128. Policy agent handler 120 may also augment or update the M-Tree in context tracking module 130 to adjustments in priority weighting schemes or phonetic models. Policy agent handler 120 may be controlled, directly or indirectly, by a third party, such as a commercial entity, and domain agents may be augmented, updated, removed, and/or provided by policy agent handler 120 as part of a commercial agreement, licensing agreement, subscription agreement, maintenance agreement, or other agreement.

[039] Referring to Fig. 3, a flow chart demonstrating an exemplary method for enhancing the performance and accuracy of speech interpretation is provided. The method may begin by receiving a user verbalization at an operation 312. The received user verbalization may be electronically captured at operation 312, such as by a microphone or other electronic audio capture device. The electronically captured verbalization may be provided to a speech interpretation engine, such as speech engine 112 in Fig. 1.

[040] The speech interpretation may then generate one or more preliminary interpretations of the received verbalization at an operation 314. According to one implementation of the invention, the plurality of preliminary interpretations may be generated using phonetic dictation, grammar trees with linking elements, or any combination thereof to improve performance and enhance accuracy. Phonetic dictation and reducing a search space of a grammar tree by including linking elements is discussed in greater detail above. The preliminary interpretations may be arranged in any predetermined form, such as an array, a matrix, or other forms.

[041] In an operation 320, the preliminary interpretations may be provided to a speech sharpening engine. The speech sharpening engine may take one or more passes over the plurality of preliminary interpretations to identify dominant words and/or phrases in operation 320. This information may then be used to generate one or more candidate interpretations. The candidate interpretations may be based on various domain agents, such as context agents and/or profile agents, which may be organized as a weighted domain model, such as an M-Tree. For example, if a set of dominant words sound like a movie name, apply policies operation 320 may generate a candidate interpretation that substitutes relevant words and/or phrases based on a domain agent populated with movie titles. Additional passes may be made over the candidate interpretations, which may be constrained by domain

Amazon/VB Assets
Exhibit 1003
Page 541

information associated with the candidate interpretations, to thereby generate a confidence score or interpretation score for each candidate interpretation. The interpretation score may represent a likelihood that a particular candidate interpretation is a correct interpretation of the verbalization received in operation 312. The operation of apply policies 320 is described in greater detail above in reference to Fig. 1.

[042] The candidate interpretations and corresponding interpretation scores may then be analyzed to determine a probable interpretation in an operation 322. In one implementation of the invention, a candidate interpretation with a highest (or lowest) score may be designated as a probable interpretation. The probable interpretation may then be output in an operation 324, such as for use in a voice-activated vehicular navigation system, a voice-controlled server or desktop computer, or other electronic device that can be controlled using voice commands.

[043] Information relating to the verbalization and the interpretations of the verbalization may be provided in a store interpretation operation 325. Store interpretation operation 324 may store interpretation information related to verbalizations, such as components of verbalizations, preliminary interpretations, dominant words and/or phrases, candidate interpretations, probable interpretations, and/or interpretation scores associated with verbalizations, as well as whether or not a verbalization was interpreted correctly, or other information. In some implementations of the invention, some or all of the interpretation information stored at store interpretation operation 324 may be used to determine one or more frequencies at a determine frequencies operation 326. The frequencies calculated at determine frequencies operation 326 may include one or more frequencies related to past verbalizations, such as, a word usage frequency, a word combination frequency, a frequency related to a set of verbalizations that are phonetically similar but have distinct meanings, an interpretation error frequency for a particular verbalization, or other frequencies. Determine frequencies operation 326 may be performed by interpretation history analysis module 118.

[044] In various implementations, a decision may be made whether to augment a speech sharpening engine in an augmentation decision operation 328. The decision concerning system augmentation may be based at least in part on information generated at determine frequencies block 326, such as one or more frequencies, or other information. If it is decided that no augmentation is needed, no further action is

Amazon/VB Assets
Exhibit 1003
Page 542

taken until another verbalization is captured, and the method ends. In some instances, decision operation 328 may determine that augmentation may be made and control passes to an augment system operation 330. Augment system operation 330 may include making an augmentation to a speech sharpening engine. For example, one or more domain agents may be augmented to reflect probabilities of an interpretation being a correct interpretation of a verbalization, to update a user profile, or other augmentation. Dynamic languages associated with context agents and/or profile agents may be augmented, or parameters weights may be augmented to enhance accuracy when interpreting subsequent verbalizations. For example, an adaptive misrecognition technique may adjust the various weights in a phonetic model or update similarity weights for regional accents, or other augmentations may be made. In parallel to augment system operation 330, new agent policies may be received in an operation 332. For example, a third party or commercial entity may redesign or modify various domain agents, new domain agents may be developed and installed as plug-ins, domain agents that are unreliable may be removed, or other augmentations or modifications may be made. Thus, the method continually refines the domain agents and the weighting of various parameters in order to refine the accuracy of the speech sharpening engine for subsequent verbalizations.

[045] The above disclosure has been described in terms of specific exemplary aspects, implementations, and embodiments of the invention. However, those skilled in the art will recognize various changes and modifications that may be made without departing from the scope and spirit of the invention. For example, references throughout the specification to "one implementation," "one aspect," "an implementation," or "an aspect" may indicate that a particular feature, structure, or characteristic is included in at least one implementation. However, the particular features, structures, or characteristics may be combined in any suitable manner in one or more implementations. Therefore, the specification and drawings are to be regarded as exemplary only, and the scope of the invention is to be determined solely by the appended claims.

Amazon/VB Assets
Exhibit 1003
Page 543

CLAIMS

What is claimed is:

1. A computer-implemented method for interpreting speech, the method comprising:

receiving a user verbalization;

parsing the verbalization into a plurality of phonemes; and

generating at least one interpretation of the verbalization based on the plurality of phonemes.

2. The method of claim 1, wherein generating the at least one preliminary interpretation includes traversing an acoustic grammar based on the plurality of phonemes.

3. The method of claim 2, wherein the acoustic grammar is a phonemic representation of a speech signal.

4. The method of claim 3, wherein the acoustic grammar includes a linking element between segments, the linking element being an acoustic element in the speech signal.

5. The method of claim 4, wherein the acoustic element is a phoneme representing an unstressed vowel.

6. The method of claim 1, further comprising:

revising the at least one interpretation, wherein the revising includes generating a plurality of candidate interpretations of the verbalization;

assigning an interpretation score to each candidate interpretation; and

selecting a candidate interpretation with a highest interpretation score as a probable interpretation of the verbalization.

7. The method of claim 6, wherein revising the at least one interpretation includes:

applying the at least one interpretation against at least one domain agent;

identifying suspect words and/or phrases in the at least one interpretation;

identifying a closest phonetic match for the suspect words and/or phrases based on a relative weighting of parameters in at least one domain associated with the at least one domain agent; and

substituting the closest phonetic match for the suspect words and/or phrases to generate a candidate interpretation.

Amazon/VB Assets
Exhibit 1003
Page 544

8. The method of claim 7, wherein the closest phonetic match is stored in an M-Tree, and wherein identifying the closest phonetic match includes using a closest-distance metric associated with the M-Tree.

9. The method of claim 7, wherein the parameters include at least one of a context of the verbalization, a verbalization criteria, a history of user verbalizations, a user profile, or domain specific information.

10. The method of claim 7, further comprising adjusting the relative weighting of parameters in response to the selection of a probable interpretation of the verbalization.

11. A method for sharpening a preliminary interpretation of a user verbalization, the preliminary interpretation generated by a speech interpretation engine, the method comprising:

revising the preliminary interpretation, wherein the revising includes generating a plurality of candidate interpretations of the verbalization;

assigning an interpretation score to each candidate interpretation; and

selecting a candidate interpretation with a highest interpretation score as a probable interpretation of the verbalization.

12. The method of claim 11, wherein revising the preliminary interpretation includes:

applying the preliminary interpretation against at least one domain agent;

identifying suspect words and/or phrases in the preliminary interpretation;

identifying a closest phonetic match for the suspect words and/or phrases based on a relative weighting of parameters in at least one domain associated with the at least one domain agent; and

substituting the closest phonetic match for the suspect words and/or phrases to generate a candidate interpretation.

13. The method of claim 12, wherein the closest phonetic match is stored in an M-Tree, and wherein identifying the closest phonetic match includes using a closest-distance metric associated with the M-Tree.

14. The method of claim 12, wherein the parameters include at least one of a context of the verbalization, a verbalization criteria, a history of user verbalizations, a user profile, or domain specific information.

Amazon/VB Assets
Exhibit 1003
Page 545

15. The method of claim 12, further comprising adjusting the relative weighting of parameters in response to the selection of a probable interpretation of the verbalization.

16. A system for interpreting speech, the system comprising:

a device that receives a user verbalization and generates an electronic signal corresponding to the user verbalization; and

a speech interpretation engine that receives the electronic signal corresponding to the user verbalization, wherein the speech interpretation engine parses the signal into a plurality of phonemes and generates at least one interpretation of the verbalization based on the plurality of phonemes.

17. The system of claim 16, wherein generating the at least one preliminary interpretation includes traversing an acoustic grammar based on the plurality of phonemes.

18. The system of claim 17, wherein the acoustic grammar is a phonemic representation of a speech signal.

19. The system of claim 18, wherein the acoustic grammar includes a linking element between segments, the linking element being an acoustic element in the speech signal.

20. The system of claim 19, wherein the acoustic element is a phoneme representing an unstressed vowel.

21. The system of claim 16, further comprising a speech sharpening engine configured to:

revise the at least one interpretation by generating a plurality of candidate interpretations of the verbalization;

assign an interpretation score to each candidate interpretation; and

select a candidate interpretation with a highest interpretation score as a probable interpretation of the verbalization.

22. The system of claim 21, wherein the speech sharpening engine is configured to revise the at least one interpretation by:

applying the at least one interpretation against at least one domain agent;

identifying suspect words and/or phrases in the at least one interpretation;

identifying a closest phonetic match for the suspect words and/or phrases based on a relative weighting of parameters in at least one domain associated with the at least one domain agent; and

Amazon/VB Assets
Exhibit 1003
Page 546

substituting the closest phonetic match for the suspect words and/or phrases to generate a candidate interpretation.

23. The system of claim 22, wherein the closest phonetic match is stored in an M-Tree, and wherein identifying the closest phonetic match includes using a closest-distance metric associated with the M-Tree.

24. The system of claim 22, wherein the parameters include at least one of a context of the verbalization, a verbalization criteria, a history of user verbalizations, a user profile, or domain specific information.

25. The system of claim 22, wherein the speech sharpening engine is further configured to adjust the relative weighting of parameters in response to the selection of a probable interpretation of the verbalization.

26. A system for sharpening a preliminary interpretation of a user verbalization received by an electronic capture device, the preliminary interpretation generated by a speech interpretation engine, the system comprising a speech sharpening engine containing at least one processor configured to:

revise the preliminary interpretation by generating a plurality of candidate interpretations of the verbalization;

assign an interpretation score to each candidate interpretation; and

select a candidate interpretation with a highest interpretation score as a probable interpretation of the verbalization.

27. The system of claim 26, wherein the speech sharpening engine is further configured to revise the preliminary interpretation by:

applying the preliminary interpretation against at least one domain agent;

identifying suspect words and/or phrases in the preliminary interpretation;

identifying a closest phonetic match for the suspect words and/or phrases based on a relative weighting of parameters in at least one domain associated with the at least one domain agent; and

substituting the closest phonetic match for the suspect words and/or phrases to generate a candidate interpretation.

28. The system of claim 27, wherein the closest phonetic match is stored in an M-Tree, and wherein identifying the closest phonetic match includes using a closest-distance metric associated with the M-Tree.

Amazon/VB Assets
Exhibit 1003
Page 547

29. The system of claim 27, wherein the parameters include at least one of a context of the verbalization, a verbalization criteria, a history of user verbalizations, a user profile, or domain specific information.

30. The system of claim 27, wherein the speech sharpening engine is further configured to adjust the relative weighting of parameters in response to the selection of a probable interpretation of the verbalization.

31. A computer-readable medium containing computer-executable instructions for interpreting speech, the computer-executable instructions configured to:

receive a user verbalization;

parse the verbalization into a plurality of phonemes; and

generate at least one interpretation of the verbalization based on the plurality of phonemes.

32. The computer-readable medium of claim 31, wherein the computer-executable instructions are configured to generate the at least one preliminary interpretation by traversing an acoustic grammar based on the plurality of phonemes.

33. The computer-readable medium of claim 32, wherein the acoustic grammar is a phonemic representation of a speech signal.

34. The computer-readable medium of claim 33, wherein the acoustic grammar includes a linking element between segments, the linking element being an acoustic element in the speech signal.

35. The computer-readable medium of claim 34, wherein the acoustic element is a phoneme representing an unstressed vowel.

36. The computer-readable medium of claim 31, wherein the computer-executable instructions are further configured to:

revise the at least one interpretation by generating a plurality of candidate interpretations of the verbalization;

assign an interpretation score to each candidate interpretation; and

select a candidate interpretation with a highest interpretation score as a probable interpretation of the verbalization.

37. The computer-readable medium of claim 36, wherein the computer-executable instructions are configured to revise the at least one interpretation by:

applying the at least one interpretation against at least one domain agent;

identifying suspect words and/or phrases in the at least one interpretation;

Amazon/VB Assets
Exhibit 1003
Page 548

identifying a closest phonetic match for the suspect words and/or phrases based on a relative weighting of parameters in at least one domain associated with the at least one domain agent; and

substituting the closest phonetic match for the suspect words and/or phrases to generate a candidate interpretation.

38. The computer-readable medium of claim 37, wherein the closest phonetic match is stored in an M-Tree, and wherein identifying the closest phonetic match includes using a closest-distance metric associated with the M-Tree.

39. The computer-readable medium of claim 37, wherein the parameters include at least one of a context of the verbalization, a verbalization criteria, a history of user verbalizations, a user profile, or domain specific information.

40. The computer-readable medium of claim 37, wherein the computer-executable instructions are further configured to adjust the relative weighting of parameters in response to the selection of a probable interpretation of the verbalization.

41. A computer-readable medium containing computer-executable instructions for sharpening a preliminary interpretation of a user verbalization, the preliminary interpretation generated by a speech interpretation engine, the computer-executable instructions configured to:

revise the preliminary interpretation by generating a plurality of candidate interpretations of the verbalization;

assign an interpretation score to each candidate interpretation; and

select a candidate interpretation with a highest interpretation score as a probable interpretation of the verbalization.

42. The computer-readable medium of claim 41, wherein the computer-executable instructions are configured to revise the preliminary interpretation by:

applying the preliminary interpretation against at least one domain agent;

identifying suspect words and/or phrases in the preliminary interpretation;

identifying a closest phonetic match for the suspect words and/or phrases based on a relative weighting of parameters in at least one domain associated with the at least one domain agent; and

substituting the closest phonetic match for the suspect words and/or phrases to generate a candidate interpretation.

Amazon/VB Assets
Exhibit 1003
Page 549

43. The computer-readable medium of claim 42, wherein the closest phonetic match is stored in an M-Tree, and wherein identifying the closest phonetic match includes using a closest-distance metric associated with the M-Tree.

44. The computer-readable medium of claim 42, wherein the parameters include at least one of a context of the verbalization, a verbalization criteria, a history of user verbalizations, a user profile, or domain specific information.

45. The computer-readable medium of claim 42, wherein the computer-executable instructions are further configured to adjust the relative weighting of parameters in response to the selection of a probable interpretation of the verbalization.

Amazon/VB Assets
Exhibit 1003
Page 550

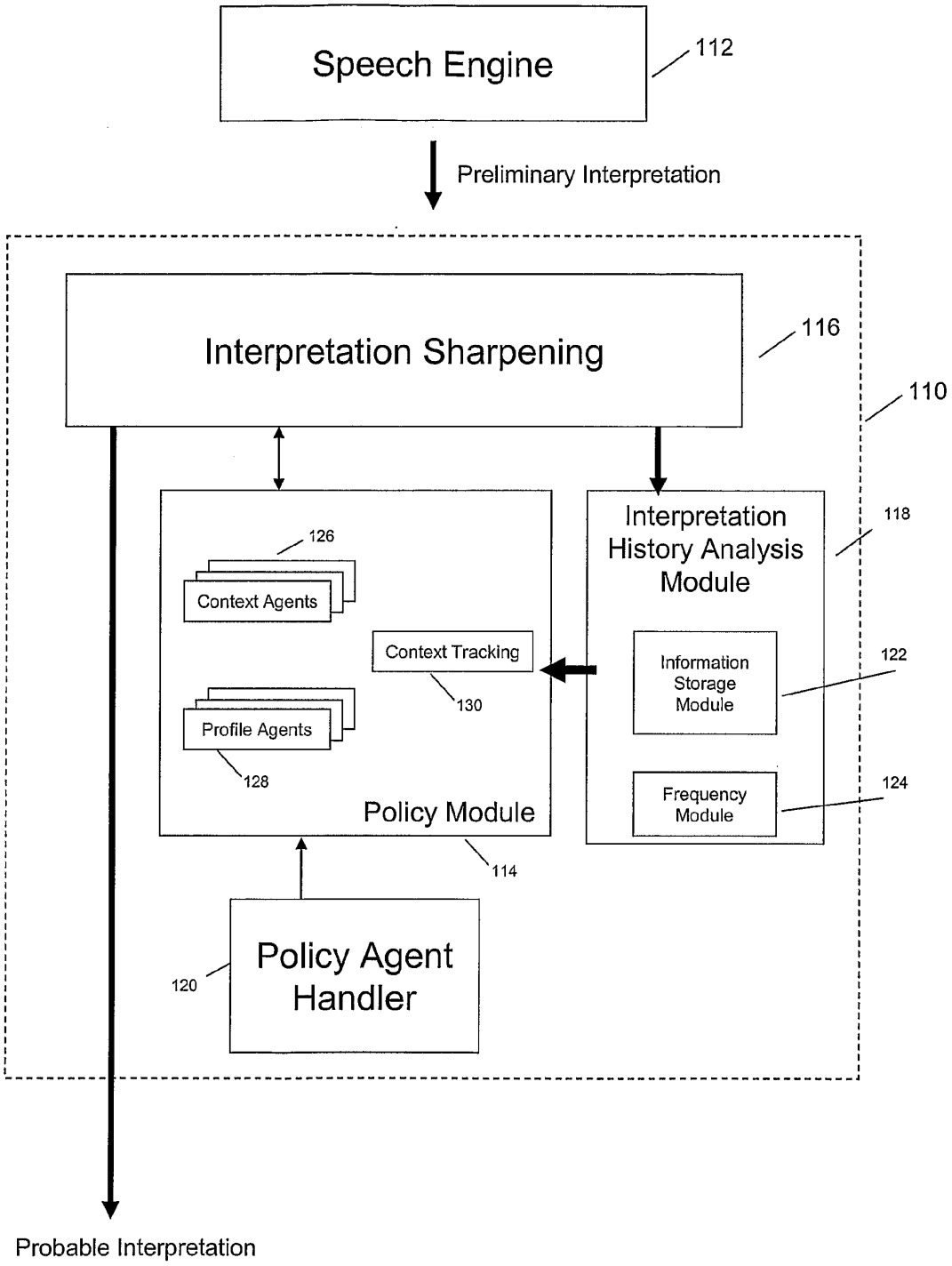
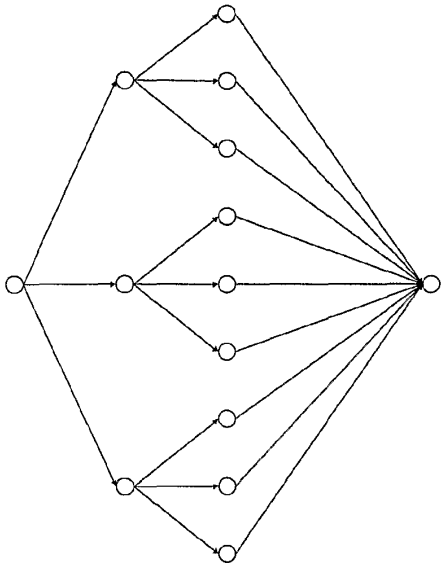

Speech Engine
112
Preliminary Interpretation
Interpretation Sharpening
116
110
Interpretation History Analysis Module
118
126
Context Agents
Context Tracking
130
Information Storage Module
122
Profile Agents
128
Frequency Module
124
Policy Module
114
Policy Agent Handler
120
Probable Interpretation

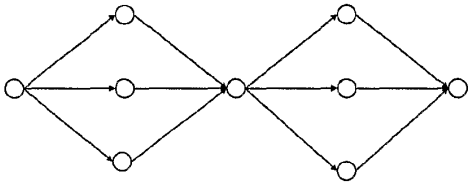
Fig. 1

Traditional Grammar Tree



210

Grammar Tree with Linking Element



220

Figure 2

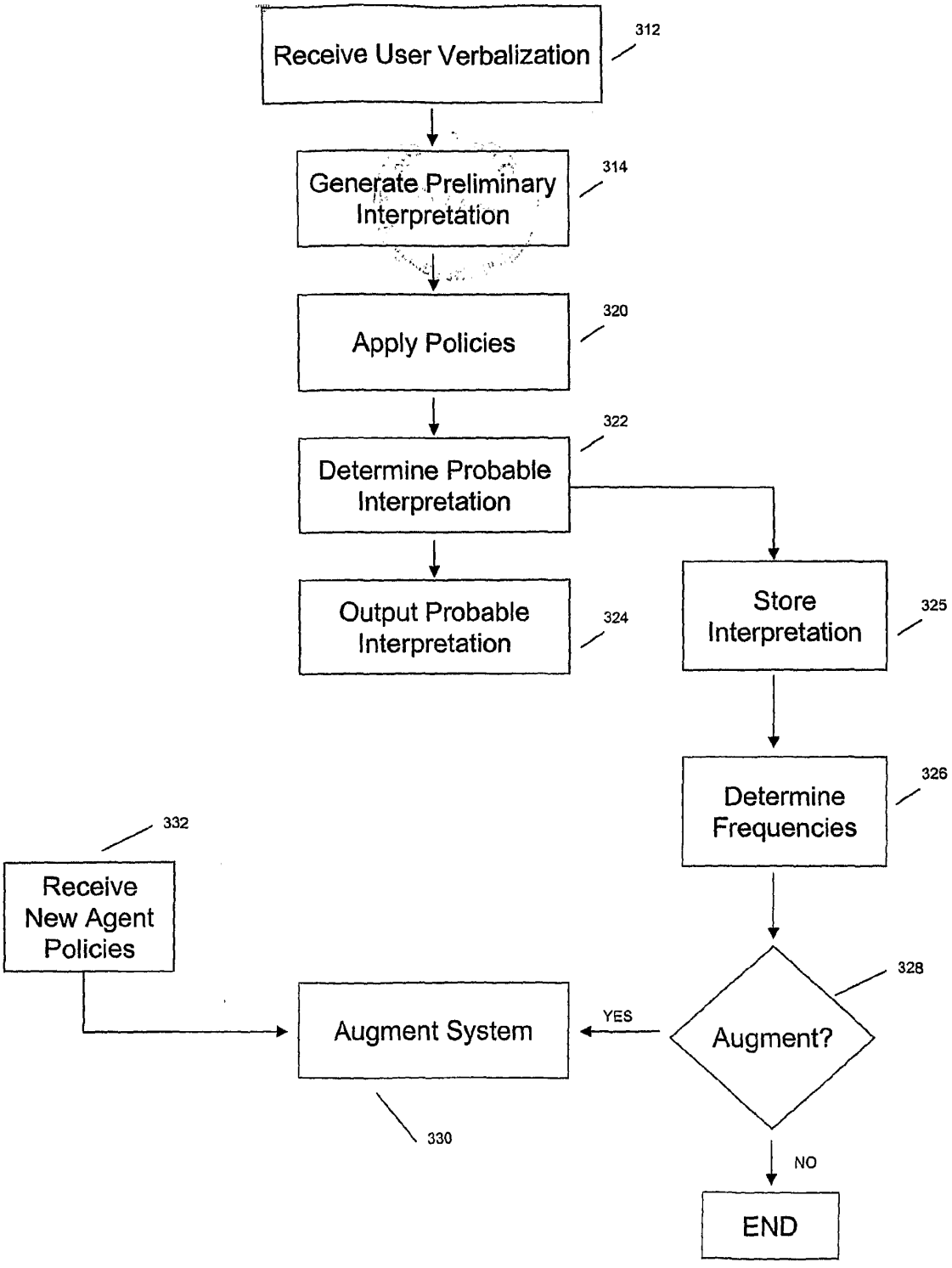

Receive User Verbalization
312
Generate Preliminary Interpretation
314
Apply Policies
320
Determine Probable Interpretation
322
Output Probable Interpretation
324
Store Interpretation
325
Determine Frequencies
326
Augment?
328
YES
NO
Augment System
330
Receive New Agent Policies
332
END

Fig. 3

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 August 2008 (14.08.2008)

PCT

(10) International Publication Number
WO 2008/098039 A1

(51) International Patent Classification:
G06Q 30/00 (2006.01)

(21) International Application Number:
PCT/US2008/053157

(22) International Filing Date: 6 February 2008 (06.02.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/671,526 6 February 2007 (06.02.2007) US

(71) Applicant *(for all designated States except US)*: **VOICEBOX TECHNOLOGIES, INC.** [US/US]; 11980 NE 24th Street, Bellevue, WA 98005 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **FREEMAN, Tom** [US/US]; 6735 83rd Avenue, S.E., Mercer Island, WA 98040 (US). **KENNEWICK, Mike** [US/US]; 13407 Ne 37th Place, Bellevue, WA 98005 (US).

(74) Agents: **GATTO, James, G.** et al.; Pillsbury Winthrop Shaw Pittman LLP, P.O. Box 10500, McLean, VA 22102 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(54) **Title:** SYSTEM AND METHOD FOR SELECTING AND PRESENTING ADVERTISEMENTS BASED ON NATURAL LANGUAGE PROCESSING OF VOICE-BASED INPUT

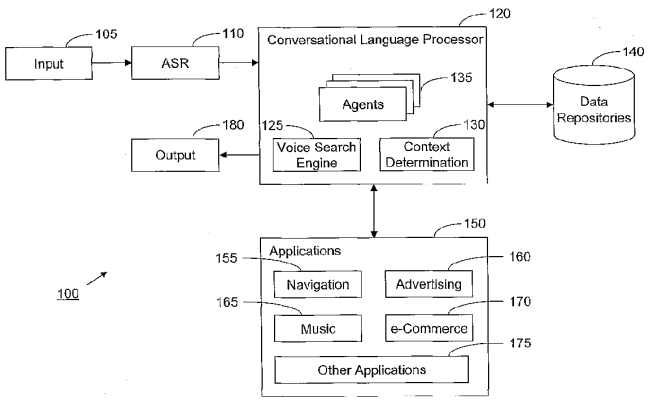


Fig. 1

(57) **Abstract:** A system and method for selecting and presenting advertisements based on natural language processing of voice-based inputs is provided. A user utterance may be received at an input device, and a conversational, natural language processor may identify a request from the utterance. At least one advertisement may be selected and presented to the user based on the identified request. The advertisement may be presented as a natural language response, thereby creating a conversational feel to the presentation of advertisements. The request and the user's subsequent interaction with the advertisement may be tracked to build user statistical profiles, thus enhancing subsequent selection and presentation of advertisements.

WO 2008/098039 A1

System and Method for Selecting and Presenting Advertisements Based on Natural Language Processing of Voice-Based Input

Field of the Invention

[001] The present invention relates to selecting and presenting advertisements based on natural language processing of voice-based input.

Background of the Invention

[002] As technology advances, consumer electronics devices tend to play larger roles due to increased functionality and mobility. For example, mobile phones, navigation devices, embedded devices, and other such devices provide a wealth of functionality beyond core applications. However, increased functionality adds difficulty to the learning curve associated with using electronic devices, and increased mobility intensifies the demand for simple mechanisms to interact with devices on the go. For example, existing systems tend to have complex human to machine interfaces, which may inhibit mass-market adoption for various technologies. For example, when a user wishes to perform a relatively simple task on a mobile phone, such as purchasing a ring tone, the user often is forced to navigate through a series of menus and press a series of buttons. In some instances, this may result in the transaction not necessarily occurring, as the user may prefer to avoid the hassles altogether. As such, there is ever-growing demand for ways to exploit technology in intuitive ways.

[003] Voice recognition software may enable a user to exploit applications and features of a device that may otherwise be unfamiliar, unknown, or difficult to use. However, many existing voice user interfaces (when they actually work) still require significant learning on the part of the user. For example, users often cannot directly issue a request for a system to retrieve information or perform an action without having to memorize specific syntaxes, words, phrases, concepts, semantic indicators, or other keywords/qualifiers. Similarly, when users are uncertain of particular needs, many existing systems do not engage the user in a productive, cooperative dialogue to resolve requests and advance a conversation. Instead, many existing speech interfaces force users to use a fixed set commands or

keywords to communicate requests in ways that systems can understand. Using existing voice user interfaces, there is virtually no option for dialogue between the user and the system to satisfy mutual goals.

[004] The lack of adequate voice user interfaces results in missed opportunities for providing valuable and relevant information to users. Not only does this potentially leave user requests unresolved, in certain instances, providers of goods and services may lose out on potential business. In an increasingly global marketplace, where marketers are continually looking for new and effective ways to reach consumers, the problems with existing voice user interfaces leaves a large segment of consumer demand unfulfilled. Furthermore, existing techniques for marketing, advertising, or otherwise calling consumers to action fail to effectively utilize voice-based information, which is one of the most natural, intuitive methods of human interaction.

[005] Existing systems suffer from these and other problems.

SUMMARY OF THE INVENTION

[006] According to various aspects of the invention, a system and method for selecting and presenting advertisements based on natural language processing of voice-based inputs is provided. A natural language voice-based input may be received by a voice user interface. The voice-based input may include a user utterance, and a request may be identified from the utterance. Appropriate action may be taken to service the request, while one or more advertisements may be selected and presented to the user. Advertisements may be selected based on various criteria, including content of the input (e.g., concepts, semantic indicators, etc.), an activity related to the input (e.g., a relation to a request, a requested application, etc.), user profiles (e.g., demographics, preferences, location, etc.), or in other ways. A user may subsequently interact with the advertisement (e.g., via a voice-based input), and action may be taken in response to the interaction. Furthermore, the interaction may be tracked to build statistical profiles of user behavior based on affinities or clusters among advertisements, user profiles, contexts, topics, semantic indicators, concepts, or other criteria.

[007] According to various aspects of the invention, advertisers may create advertisements, which may be stored in an advertisement repository. For example,

Amazon/VB Assets
Exhibit 1003
Page 556

advertisements may include sponsored messages, calls to action, purchase opportunities, trial downloads, or any other marketing communication, as would be apparent to those skilled in the art. Advertisers may specify various parameters to associate with the advertisements, such as various contexts or topic concepts (e.g., semantic indicators for a "music" concept may include words such as "music," "tunes," "songs," etc.), target demographics (e.g., a preferred audience), marketing criteria or prices for insertion (e.g., dynamic or static pricing based on various marketing criteria), or other information, as would be apparent. The advertisement repository may be associated with a server, where in response to a voice-based input from a user (e.g., at a voice-enabled device), a communications link may be established with the server. Information may be extracted from the voice-based input (e.g., words in the input, applications requested by the input, etc.), and the extracted information may be correlated with user profiles, advertisement parameters, or other information to determine which advertisements to select in relation to the voice-based input. The server may subsequently communicate the selected advertisements to the user, and the server may track the user's subsequent interaction with the selected advertisements.

[008] Other objects and advantages of the invention will be apparent based on the following drawings and detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

[009] Fig. 1 illustrates a block diagram of an exemplary system for implementing a voice user interface according to various aspects of the invention.

[010] Fig. 2 illustrates a block diagram of an exemplary advertising system according to various aspects of the invention.

[011] Fig. 3 illustrates a flow diagram of an exemplary method for selecting and presenting advertisements based on voice-based inputs according to various aspects of the invention.

DETAILED DESCRIPTION

[012] Referring to Fig. 1, an exemplary system 100 for implementing a voice user interface is illustrated according to various aspects of the invention. System 100 may enable users to perform various tasks on a voice-enabled device. For example, users may control navigation

Amazon/VB Assets
Exhibit 1003
Page 557

devices, media devices, personal computers, personal digital assistants, or any other device supporting voice-based inputs. System 100 may enable users to request voice-enabled devices to retrieve information or perform various tasks, among other things, using natural language voice-based inputs. For example, system 100 may interpret natural language voice-based inputs and generate responses using, among other things, techniques described in co-pending U.S. Patent Application Ser. No. 10/452,147, entitled "Systems and Methods for Responding to Natural Language Speech Utterance," filed June 3, 2003; and U.S. Patent Application Ser. No. 10,618,633, entitled "Mobile Systems and Methods for Responding to Natural Language Speech Utterance," filed June 15, 2003, both of which are hereby incorporated by reference in their entirety.

[013] According to various aspects of the invention, system 100 may receive a user input, including at least a voice-based user utterance, at an input device 105. Input device 105 may include any suitable device, or combination of devices, for receiving a voice-based input (e.g., a microphone). In various implementations, input device 105 may include a multi-modal input, such as a touch-screen interface, keypad, or other input. The received utterance may be processed by an Automatic Speech Recognizer 110. Automatic Speech Recognizer 110 may generate one or more preliminary interpretations of the utterance using various techniques. For example, Automatic Speech Recognizer 110 may interpret the utterance using techniques of phonetic dictation to recognize a stream of phonemes. Further, Automatic Speech Recognizer 110 may perform post-processing to enhance the preliminary interpretations. For example, Automatic Speech Recognizer 110 may vary interpretations of an utterance, or components of an utterance, from one context to another. Other techniques for enhancing an interpretation of a user utterance may be used, such as those described in co-pending U.S. Patent Application Ser. No. 11/513,269, entitled "Dynamic Speech Sharpening," filed August 31, which is hereby incorporated by reference in its entirety.

[014] The one or more preliminary interpretations may be provided to a conversational language processor 120. Conversational language processor 120 may include a voice search engine 125, a context determination module 130, and one or more agents 135, among other things, to enable cooperative, conversational interaction between the user and system 100. Conversational language processor 120 may be communicatively coupled to one or more data repositories 140 and one or more applications 150. Conversational language processor

120 may generate a domain-specific conversational response, which may be returned to the user as an output 180. Output 180 may include a multi-modal output (e.g., by simultaneously returning a voice-based response and displaying information on a display device).

[015] System 100 may further include an interaction with one or more applications 150 to service one or more requests in the utterance. For example, the utterance may include one or more requests for performing an action, retrieving information, or various combinations thereof. Output 180 may include a conversational response to advance a conversation to service requests by invoking one or more applications 150, as appropriate. For example, applications 150 may include a navigation application 155, an advertising application 160, a music application, an electronic commerce application 170, and/or other applications 175. Furthermore, Automatic Speech Recognizer 110, conversational language processor 120, data repositories 140, and/or applications 150 may reside locally (e.g., on a user device), remotely (e.g., on a server), and/or hybrid local/remote processing models may be used (e.g., lightweight applications may be processed locally while computationally intensive applications may be processed remotely).

[016] Conversational language processor 120 may build long-term and/or short-term shared knowledge in one or more knowledge source. For example, shared knowledge sources may include information about previous utterances, requests, and other user interactions to inform generating an appropriate response to a current utterance. The shared knowledge may include public/non-private (i.e., environmental) knowledge, as well as personal/private (i.e., historical) knowledge. For example, conversational language processor 120 may use context determination module 130 to establish a context for a current utterance by having domain agents 135 competitively generate a context-based interpretation of the utterance (e.g., by scoring possible interpretations and selecting a highest scoring interpretation). As such, agents 135 may model various domains (e.g., navigation, music, a specific user, global users, advertising, e-commerce, etc.), and conversational language processor 120 may interpret and/or respond to a voice-based input accordingly. For example, context-based interpretations and responses to a voice-based input may be generated using techniques described in co-pending U.S. Patent Application Ser. No. 11/197,504, entitled "Systems and Methods for Responding to Natural Language Speech Utterance," filed August 5, 2005; and U.S. Patent Application Ser. No. 11/212,693,

Amazon/VB Assets
Exhibit 1003
Page 559

entitled "Mobile Systems and Methods of Supporting Natural Language Human-Machine Interactions," filed August 29, 2005, both of which are hereby incorporated by reference in their entirety.

[017] Furthermore, conversational language processor 120 may support adaptive misrecognition to reinterpret a current utterance and/or one or more previous utterances. For example, information contained in a current utterance may indicate that interpretations for one or more previous utterances were incorrect, and therefore, the previous utterances may be reinterpreted to improve subsequent interpretations. Accordingly, conversational language processor 120 may use the techniques described herein, along with various other techniques, to interpret and respond to conversational, natural language utterances. Conversational language processor 120 may use various other techniques as will be apparent, such as those described in co-pending U.S. Patent Application Ser. No. 11/200,164, entitled "System and Method of Supporting Adaptive Misrecognition in Conversational Speech," filed August 10, 2005; and U.S. Patent Application Ser. No. 11/580,926, entitled "System and Method for a Cooperative Conversational Voice User Interface," filed October 16, 2006, both of which are hereby incorporated by reference in their entirety.

[018] Referring to Fig. 2, an exemplary advertising system 200 is illustrated according to various aspects of the invention. System 200 may include a server 230 for receiving one or more advertisements from an advertiser 220, wherein the advertisements may be stored in a data repository 260 associated with server 230. For example, advertisements may include sponsored messages or marketing communications, calls to action, purchase opportunities, trial downloads, coupons, or any other suitable marketing, advertising, campaign, or other information, as would be apparent to those skilled in the art. A voice-enabled device 210 may receive a voice-based input and establish communications with advertising server 230. Subsequently, advertising server 230 may select one or more advertisements from among the advertisements stored in data repository 260, and the selected advertisements may be provided to the voice-enabled device for presentation to a user.

[019] Advertiser 220 may access advertising server 230 via an advertiser interface 245. Advertisers 220 may upload targeted advertisements to server 230 via advertiser interface 245, and server 230 may store the advertisements in data repository 260. The advertisements may include graphically-based advertisements that include banners, images,

Amazon/VB Assets
Exhibit 1003
Page 560

audio, video, or any suitable combination thereof. Furthermore, the advertisements may include interactive or embedded information, such as links, metadata, or computer-executable instructions, or any suitable combination thereof. Advertisers may specify criteria for a campaign or targeting information for an advertisement (e.g., a start date, an end date, budget information, geo-targeting information, conceptual or contextual information, or any other suitable criteria), which may be used to facilitate selecting an advertisement in relation to a particular voice-based input.

[020] In addition to providing interface 245 for advertisers, server 230 may include a content/action identification module 235, a user profile module 240, an advertisement selection module 250, and a tracking module 255. Users may submit voice-based requests to voice-enabled device 210, and voice-enabled device 210 may communicate information about the voice-based input to server 230. Server 230 may invoke advertisement selection module 250 to extract relevant information from about the voice-based input, where advertisement selection module 250 may select one or more advertisements relevant to the voice-based input based on information extracted using content/action identification module 235 and/or user profile module 240.

[021] For example, content/action identification module 235 may identify content of the voice-based input (e.g., words in the input), requested information (e.g., search results, a web page, music, video, graphics, or other information), requested actions (e.g., calculating a navigation route, placing a telephone call, playing a song, etc.), a category or topic related to the input (e.g., music, business, stocks, sports, navigation, movies, etc.), or other criteria to use in selecting an advertisement. Further, user profile module 240 may identify characteristics of a specific user (e.g., demographics, personal preferences, location-based information, etc.), global user profiles (e.g., demographic profiles, click-through rates, etc.), or other criteria to use in selecting an advertisement. Moreover, advertisement selection module 250 may account for where a request originates from. For example, advertisements may be selected based on a default user location (e.g., identified from a user profile), current geolocation information (e.g., identified from a navigation device), whether an affiliate or partner of server 230 initiated the request, or other criteria.

[022] For instance, a user may request airline reservations via voice-enabled device 210, and content/action identification module 235 may identify specific words used in the request, a category related to the request (e.g., travel, airlines, hotels, etc.), or other

Amazon/VB Assets
Exhibit 1003
Page 561

information. Furthermore, user profile module 240 may identify relevant characteristics of the user (e.g., user-specific demographics, location information, preferred airlines or hotels, etc.), as well as global user characteristics (e.g., most popular airlines). In various implementations, advertisements may be selected by assigning a score to each advertisement (e.g., based on click-through rates, relevance metrics, target audiences, etc.). As such, advertisement selection module 250 may correlate the information about the request to select advertisements stored in data repository 260, and server 230 may communicate the selected advertisements to voice-enabled device 210. Furthermore, selected advertisements may be presented according to a predetermined ordering or ranking (e.g., based on a ranking of relevance to an advertisement).

[023] In various implementations, advertisement selection module 250 may retrieve a predetermined number of advertisements for any given request. Furthermore, the selected advertisements may depend upon a presentation format. For example, advertisements may be selected based on an amount of available space on a display of voice-enabled device 210 and/or a size/shape of the selected advertisements. In another example, voice-based advertisements may be selected and presented to the user audibly (e.g., a "hands-free" advertisement may be preferred when voice-enabled device 210 is a telematics device).

[024] Furthermore, the user's subsequent interaction with an advertisement may be tracked using tracking module 255. For example, tracking module 255 may determine whether a conversion or click-through occurs for each advertisement presented to users. Further, tracking module 255 may maintain accounting and/or billing information associated with advertisers 220. For example, advertisers 220 may specify a maximum insertion cost, a cost-per-click-through, an average insertion cost, or other criteria specifying a budget constraint for an advertisement. As such, tracking module 255 may track which advertisements are selected and/or presented, which advertisements result in a conversion or click-through, whether a click-through or conversion results in a transaction or sale, associations between advertisements and users, requests, concepts, semantic indicators, and/or other criteria. For example, tracking user interaction with advertisements may be used to build user-specific and/or global statistical profiles that map or cluster advertisements to topics, semantic indicators, contexts, concepts, etc. based on user behavior, demographics, targeting constraints, content of advertisements, content of requests, actions associated with requests, or other statistically relevant information.

Amazon/VB Assets
Exhibit 1003
Page 562

Accordingly, the tracking information may be used to bill or invoice advertisers 220, as well as to improve subsequent performance and relevance of advertisements selected using advertisement selection module 250. Other techniques and features of selecting and presenting advertisements based on voice-based inputs may suitably be employed, as would be apparent.

[025] Referring to Fig. 3, an exemplary method for selecting and presenting advertisements based on a voice-based input is illustrated according to various aspects of the invention. The method may begin in an operation 305, where a voice-based input, including at least a user utterance, may be received at a voice user interface. The voice user interface may include any suitable mechanism for receiving the utterance (e.g., a microphone), and may interface with any suitable voice-enabled device, as would be apparent, including personal navigation devices, personal digital assistants, media devices, telematics devices, personal computers, mobile phones, or others.

[026] Subsequently, one or more requests included in the voice-based input may be identified in an operation 310. For example, the requests may include requests to retrieve information, perform tasks, explore or gather information, or otherwise interact with a system or device. For example, a voice-based input to a navigation device may include a request to calculate a route or retrieve location-based information. In another example, a voice-based input to a mobile phone may include a request to place a telephone call, purchase a ringtone, or record a voice-memo. Furthermore, in various implementations, voice-based inputs may include multiple requests, multi-modal requests, cross-device requests, cross-application requests, or other types of requests. For example, an utterance received in operation 305 may be: "Get me a route to Chang's Restaurant, and call them so I can make a reservation." The utterance may thus include multiple requests, including cross-device requests (e.g., calculate a route using a navigation device, and make a telephone call using a mobile phone), as well as cross-application requests (e.g., search for an address and/or phone number using a voice search engine, and calculate a route using a navigation application).

[027] The requests may be part of a conversational interaction between a user and a system or device, whereby an interpretation of requests in a current utterance may be based upon previous utterances in a current conversation, utterances in previous conversations, context-based information, local and/or global user profiles, or other

Amazon/VB Assets
Exhibit 1003
Page 563

information. For example, a previous request may be reinterpreted based on information included in subsequent requests, a current request may be interpreted based on information included in previous requests, etc. Furthermore, the conversational interaction may take various forms, including query-based conversations, didactic conversations, exploratory conversations, or other types of conversations. For example, the conversational language processor may identify a type of conversation, and information may be extracted from the utterance accordingly to identify the one or more requests in operation 310. Moreover, the conversational language processor may determine whether any of the requests are incomplete or ambiguous, and action may be taken accordingly (e.g., a system response may prompt a user to clarify an incomplete and/or ambiguous request). The conversational language processor may therefore use various techniques to identify a conversation type, interpret utterances, identify requests, or perform other tasks, such as those described in the aforementioned co-pending patent applications, which are hereby incorporated by reference in their entirety.

[028] Upon identifying the one or more requests, action may be taken based on the identified requests in an operation 315, while one or more advertisements may be selected in an operation 320 (described in greater detail below). For example, one or more context-appropriate applications may be invoked to service the requests in operation 315 (e.g., a voice search engine, a navigation application, an electronic commerce application, or other application may be invoked depending upon the request). Furthermore, in operation 320, information may be communicated to an advertising server to select one or more advertisements related to the request. Thus, as shown in Fig. 3, taking action in operation 315 and selecting advertisements in operation 320 may be related operations (e.g., advertisements may be selected to help in interpreting incomplete and/or ambiguous requests).

[029] Upon taking action in operation 315 (e.g., to service the request) and selecting one or more advertisements in operation 320 (e.g., in relation to the request), an output may be presented to the user in operation 325. The output may indicate a result of the action associated with operation 315. For example, the output may include requested information, an indication of whether a requested task was successfully completed, whether additional information is needed to service the request (e.g., including a prompt for the information), or other information relating to an action based on the request. Furthermore,

Amazon/VB Assets
Exhibit 1003
Page 564

the output may include advertisements, as selected in operation 320. For example, the output may include text-based, graphic-based, video-based, audio-based, or other types of advertisements, as would be apparent to those skilled in the art. Further, the output may include other types of advertisements, including calls to action (e.g., a location-based coupon or purchase opportunity, trial downloads, or other actionable advertising or marketing).

[030] Advertisements may be selected in relation to a request based on various criteria. For example, an advertisement may be selected based on words or other content of the request, relevant words or content related to the words or content of the request, etc. In another example, the advertisement may be selected based on requested tasks/information (e.g., a request for movie showtimes may result in an advertisement being selected for a particular theater). In yet another example, the advertisement may be selected based on a topic or category associated with the requested tasks/information (e.g., a request to purchase airline tickets may result in an advertisement being selected for a hotel in a destination associated with a reserved flight). In still other examples, the advertisement may be selected based on location information, (e.g., advertisements may be selected based on a proximity to a user geolocation identified using a navigation device), user-specific and/or global user profiles (e.g., advertisements may be selected based on user-specific and/or global preferences, advertiser campaign criteria, etc.).

[031] Content of a voice-based input may be determined based on various criteria, including contextual or conceptual information (e.g., semantic indicators, qualifiers, or other information). For example, a given concept may include various semantically equivalent indicators having an identical meaning. Thus, for instance, a voice-based input may be "Play some tunes!" or "Play some music!" or other variants thereof, each of which may be interpreted as relating to a specific idea (or concept) of "Music." Thus, concept or content information in a request may be used to select an advertisement. For example, a user may request to calculate a route in Seattle, Washington (e.g., "How do I get to the Space Needle?"). Based on a context of the requested task (e.g., "Navigation," "Seattle," etc.), a voice search engine may retrieve an address of the Space Needle and a navigation application may calculate the route. Furthermore, user profile information may indicate that the user is visiting Seattle from out-of-town (e.g., the profile may indicate that the user's home is Sacramento), and therefore, an advertisement for popular points-of-interest

Amazon/VB Assets
Exhibit 1003
Page 565

in Seattle may be selected. In another example, the user may request information about a sporting event (e.g., "Get me the kickoff time for the Eagles game on Sunday"). Based on a context of the requested information (e.g., "Search," "Sports," "Philadelphia," etc.), the requested information may be retrieved, while an advertisement for Eagles apparel or memorabilia may be selected.

[032] In various instances, concepts, semantic indicators, qualifiers, or other information included in, or inferred from, a request may indicate an exploratory nature for the request. In other words, the exploratory request may identify a goal for a conversation, instead of a particular task to perform or information to retrieve. As such, in various implementations, an advertisement may be selected in operation 320 in an effort to advance the conversation towards the goal. For example, an exploratory request may include a request for a navigation route (e.g., "I feel like going to a museum, find me something interesting"). Based on a context of the requested task (e.g., "Navigation," "Points of Interest," etc.), the goal of the conversation may be identified, and the request may be serviced in operation 315 (e.g., a voice search engine may locate nearby points of interest based on user preferred topics). Further, the advertising application may select an appropriate advertisement in operation 320, where the advertisement may be selected in an attempt to advance the conversation towards the goal. For example, statistical profiles (e.g., user profiles, global profiles, topic-based profiles, etc.) may reflect an affinity between an advertisement for a particular museum and other users sharing similar demographics or other characteristics with the requesting user. Thus, in addition to retrieving information about museums in operation 315, an advertisement for a museum likely to be of interest to the user may be selected in operation 320.

[033] In various instances, a request may include incomplete, ambiguous, unrecognized, or otherwise insufficient semantic indicators, context, qualifiers, or other information needed to identify the request. In other words, the request may include inadequate information to identify or infer a task to perform, information to retrieve, or a goal for a conversation. Thus, as much information as possible may be extracted and/or inferred from the request based on shared knowledge such as context, user or global profile information, previous utterances, previous conversations, etc. As such, servicing the request may include generating a response and/or communicating with an advertising application to advance a conversation toward a serviceable request. For example, servicing the request in operation

Amazon/VB Assets
Exhibit 1003
Page 566

315 and selecting an advertisement in operation 320 may include generating a response and/or selecting an advertisement to frame a subsequent user input, thereby advancing the conversation.

[034] For example, the request may include incomplete, ambiguous, or unrecognized information (e.g., "Do you know [*mumbled words*] Seattle?"). A context of the requested task may be identified (e.g., "Seattle"), yet the identified context may be insufficient to adequately take action to service the request. Additional information may be inferred based on previous utterances in the conversation, profile information, or other information. However, when the additional information fails to provide adequate information to infer a reasonable hypothesis, servicing the request in operation 315 may include generating a response to frame a subsequent user input and advance the conversation (e.g., information about various topics may be retrieved based on a user's preferred topics). Further, the advertising application may select an advertisement in operation 320 to advance the conversation (e.g., advertisements may be selected based on user and/or global profiles reflecting an affinity between certain advertisements associated with Seattle and user preferences, profiles, etc.). Thus, by selecting an advertisement, indicating dissatisfaction with an advertisement, or otherwise interacting with an advertisement, the interaction may be used to build context and shared knowledge for a subsequent course of the conversation. For example, a user may select an advertisement, and an interpretation of a subsequent voice-based input (e.g., "Call them," "What's the price range?" etc.) may be interpreted with shared knowledge of the advertisement that the voice-based input relates to. Thus, advertisements may be used in a way that enables advertisers to market to consumers, while also improving the consumers' interaction with a device. Other advantages will be apparent to those skilled in the art.

[035] It will be apparent that operation 320 may use various techniques to select advertisements based on voice-based inputs and/or requests included therein. For example, an advertiser may specify a target audience, marketing criteria, campaign strategies, budget constraints, concepts, semantic indicators, related topics, categories, and/or any other suitable information to associate with an advertisement. For instance, advertisers may pay a premium to prioritize an advertisement in relation to similar advertisements (e.g., advertisements associated with competitors). In another example, various statistical profiles may define affinities between advertisements, topics, users, etc.

(e.g., based on click-through or conversion rates, or other tracking information, as described in more detail below). Thus, advertisements may be selected in operation 320 using various techniques, including content of the request, an activity/action associated with the request, user profiles, user preferences, statistical metrics, advertiser-specified criteria, to advance a conversation, to resolve ambiguous requests, or in various other ways, as will be apparent.

[036] The output presented to the user in operation 325 may be provided to the user in various ways. For example, in various implementations, the output may include a voice-based or otherwise audible response. In another example, when an associated device includes a display mechanism, the output may be displayed on the display device. It will be apparent that many combinations or variants thereof may be used, such as augmenting a voice-based response with information on a display device. For example, a user may request information about restaurants, and an advertisement may be selected based on a user preference indicating a favorite type of restaurant (e.g., a Chinese restaurant may be selected based on a user profile indicating a preference for Chinese). Therefore, in one example, the output presented in operation 325 may display information about various restaurants matching the requested information, while a voice-based advertisement for the Chinese restaurant may be played to the user (e.g., via a speaker or other suitable mechanism for playing voice back to the user). Many other variations will be apparent (e.g., a graphical advertisement may be displayed on a display device, while a corresponding or different voice-based advertisement may be played audibly).

[037] Subsequent interaction between the user and the presented advertisements may be monitored in a decisional operation 330. For instance, when the user elects to interact with the advertisement, action may be taken based on the interaction in an operation 335. The interaction may take various forms, including additional voice-based inputs or other suitable mechanisms for interacting with advertisements (e.g., clicking on an advertisement displayed on a personal digital assistant using an associated stylus). For example, a user may initially request information from a voice-enabled media device (e.g., a satellite radio player) about a song currently playing (e.g., "What is this song?"). In addition to outputting the requested information about the song (e.g., "This song is Double Barrel by Dave and Ansel Collins."), a selected advertisement may enable the user to purchase a ringtone for a mobile phone that corresponds to the song. In this example, the interaction may include a request to purchase the ringtone (e.g., "Yeah, I'll buy that"), and action taken in operation

Amazon/VB Assets
Exhibit 1003
Page 568

335 may include completing a transaction for the ringtone and/or downloading the ringtone to the mobile phone. Furthermore, additional advertisements may be selected in an operation 340 based on the interaction, using similar techniques as described in connection with operation 320 (e.g., advertisements for additional ringtones, similar musicians, etc. may be selected). Processing may subsequently return to operation 325 to present output resulting from the interaction.

[038] User advertisement interaction may be tracked in an operation 345. For example, operation 345 may track historical data about users, conversations, topics, contexts, or other criteria to associate information with the selected advertisement. The tracking information may therefore be used to build statistical profiles defining affinities, click-through or conversion rates, or other information about various advertisements, topics, or other criteria on a user-specific and/or a global-user level. Thus, clusters or mappings may be created between advertisements, topics, concepts, demographics, or other criteria based on user behavior with the advertisements (e.g., whether a user interacts with the advertisement in operation 330).

[039] For instance, certain advertisements may experience high click-through rates in relation to a first context and/or topic, but low click-through rates in relation to a second context and/or topic, and therefore, when requests relate to the first context and/or topic, the advertisement may be more likely to be selected in subsequent operations 320/340. In another example, global statistical profiles may indicate that an advertisement experiences more click-throughs by users of a particular demographic, and therefore, the advertisement may be more likely to be selected for users falling within the demographic. Many different techniques for tracking and building statistical profiles will be apparent.

[040] Implementations of the invention may be made in hardware, firmware, software, or any combination thereof. The invention may also be implemented as instructions stored on a machine-readable medium, which may be read and executed by one or more processors. A machine-readable medium may include any mechanism for storing or transmitting information in a form readable by a machine (e.g., a computing device). For example, a machine-readable storage medium may include read only memory, random access memory, magnetic disk storage media, optical storage media, flash memory devices, and others, and a machine-readable transmission media may include forms of propagated signals, such as carrier waves, infrared signals, digital signals, and others. Further, firmware, software,

Amazon/VB Assets
Exhibit 1003
Page 569

routines, or instructions may be described in the above disclosure in terms of specific exemplary aspects and implementations of the invention, and performing certain actions. However, it will be apparent that such descriptions are merely for convenience and that such actions in fact result from computing devices, processors, controllers, or other devices executing the firmware, software, routines, or instructions.

[041] Aspects and implementations may be described as including a particular feature, structure, or characteristic, but every aspect or implementation may not necessarily include the particular feature, structure, or characteristic. Further, when a particular feature, structure, or characteristic is described in connection with an aspect or implementation, it will be apparent to effect such feature, structure, or characteristic in connection with other aspects or implementations whether or not explicitly described. Thus, various changes and modifications may be made, without departing from the scope and spirit of the invention. The specification and drawings are to be regarded as exemplary only, and the scope of the invention is to be determined solely by the appended claims.

Amazon/VB Assets
Exhibit 1003
Page 570

CLAIMS

What is claimed is:

1. A method for selecting and presenting advertisements in response to voice-based inputs, the method comprising:

receiving a voice-based input, the received input including a request;

retrieving information from at least one knowledge source based on the received request;

selecting an advertisement based on the received request and the retrieved information; and

presenting the selected advertisement to a user.

2. The method of claim 1, wherein the retrieved information is related to one or more of content of the voice-based input, an action associated with the request, a location of the user, a user-specific profile, or a global user profile.

3. The method of claim 1, wherein selecting the advertisement includes identifying an advertisement related to resolving the request when the received input includes insufficient information to resolve the request.

4. The method of claim 1, further comprising tracking an interaction between the user and the presented advertisement.

5. The method of claim 4, wherein the tracking is used to build statistical profiles of one or more users.

6. The method of claim 5, wherein the statistical profiles identify affinities among one or more advertisements, concepts, categories, and/or user profiles.

7. The method of claim 4, wherein the tracking is used to build a context for interpreting subsequent voice-based inputs.

Amazon/VB Assets
Exhibit 1003
Page 571

8. The method of claim 4, further comprising taking an action in response to the interaction.

9. The method of claim 8, wherein the action includes performing a task and/or retrieving information based on the interaction.

10. The method of claim 1, wherein the at least one knowledge source includes public information generic to the user, and private information specific to the user.

11. A system for selecting and presenting advertisements in response to voice-based inputs, the system comprising:

at least one input device that receives a voice-based input, the received input including a request; and

at least one processor operable to:

retrieve information from at least one knowledge source based on the received request;

select an advertisement based on the received request and the retrieved information; and

present the selected advertisement to a user.

12. The system of claim 11, wherein the retrieved information is related to one or more of content of the voice-based input, an action associated with the request, a location of the user, a user-specific profile, or a global user profile.

13. The system of claim 11, wherein selecting the advertisement includes identifying an advertisement related to resolving the request when the received input includes insufficient information to resolve the request.

14. The system of claim 11, further comprising tracking an interaction between the user and the presented advertisement.

Amazon/VB Assets
Exhibit 1003
Page 572

15. The system of claim 14, wherein the tracking is used to build statistical profiles of one or more users.

16. The system of claim 15, wherein the statistical profiles identify affinities among one or more advertisements, concepts, categories, and/or user profiles.

17. The system of claim 14, wherein the tracking is used to build a context for interpreting subsequent voice-based inputs.

18. The system of claim 14, further comprising taking an action in response to the interaction.

19. The system of claim 18, wherein the action includes performing a task and/or retrieving information based on the interaction.

20. The system of claim 11, wherein the at least one knowledge source includes public information generic to the user, and private information specific to the user.

Amazon/VB Assets
Exhibit 1003
Page 573

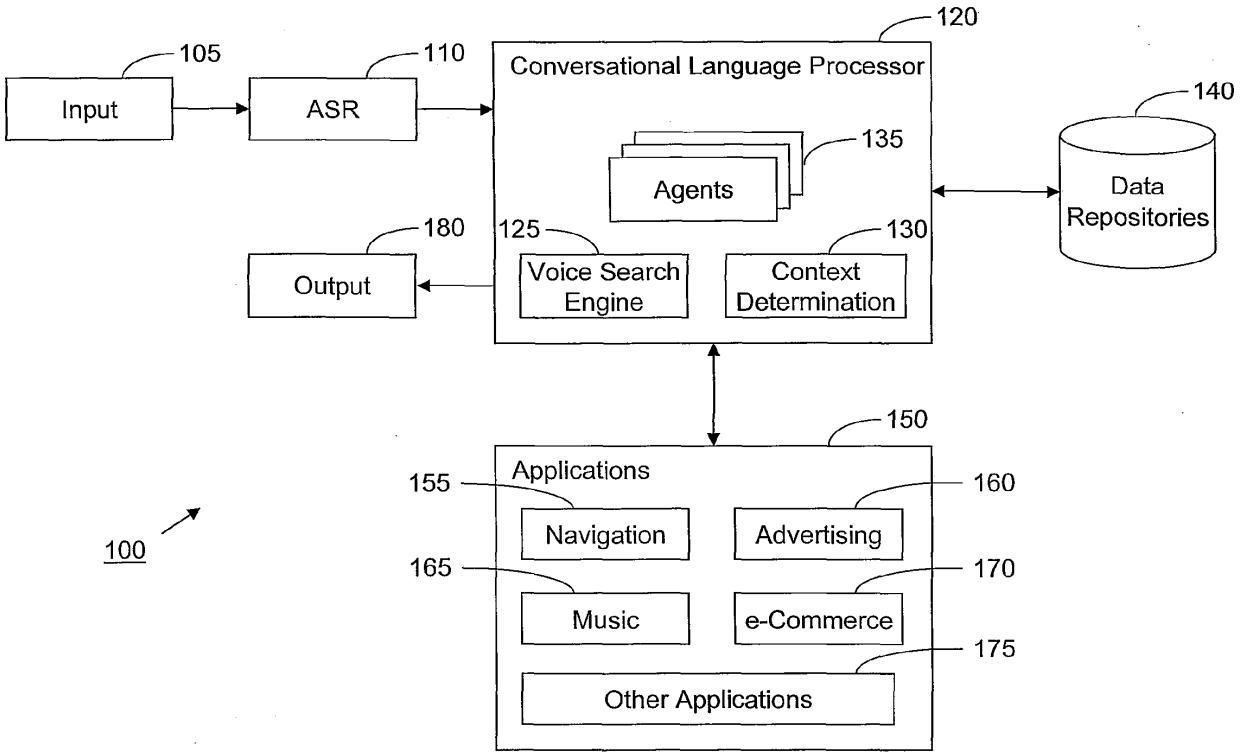

120
105
110
Input
ASR
Conversational Language Processor
140
135
Agents
Data Repositories
180
125
130
Output
Voice Search Engine
Context Determination
150
Applications
155
160
Navigation
Advertising
100
165
170
Music
e-Commerce
175
Other Applications

Fig. 1

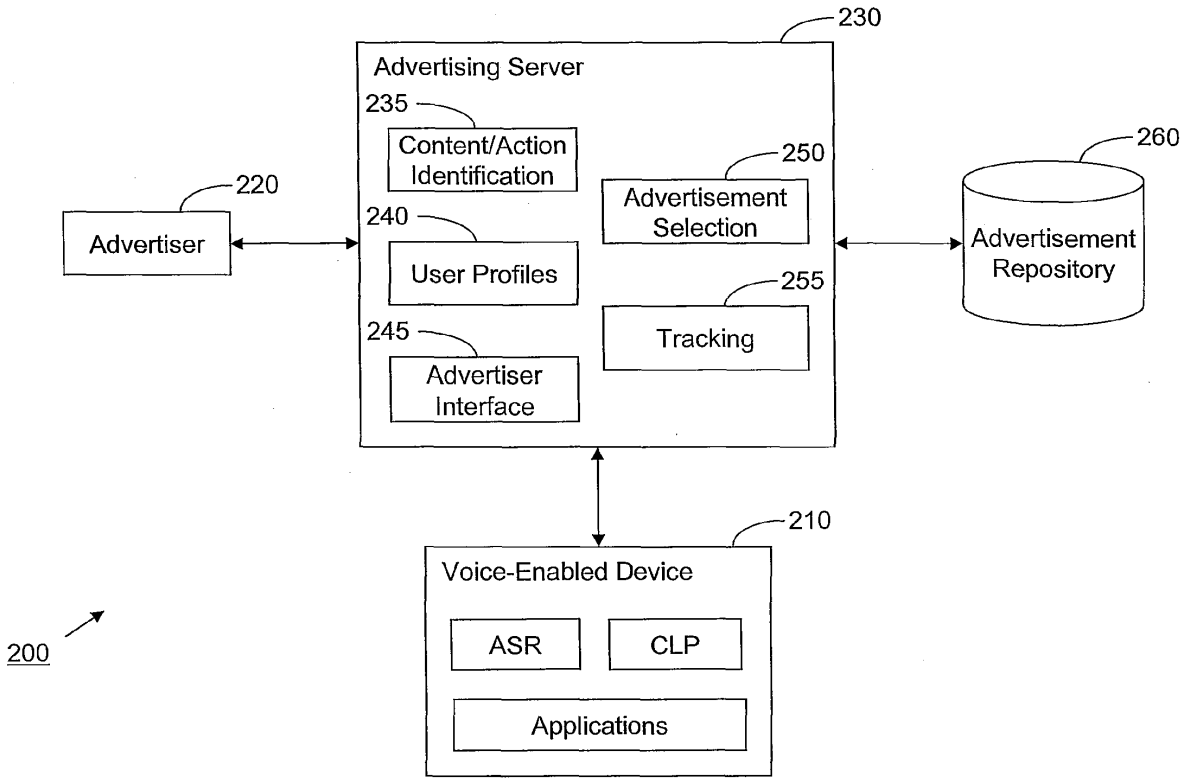


Fig. 2

Amazon/VB Assets
Exhibit 1003
Page 575

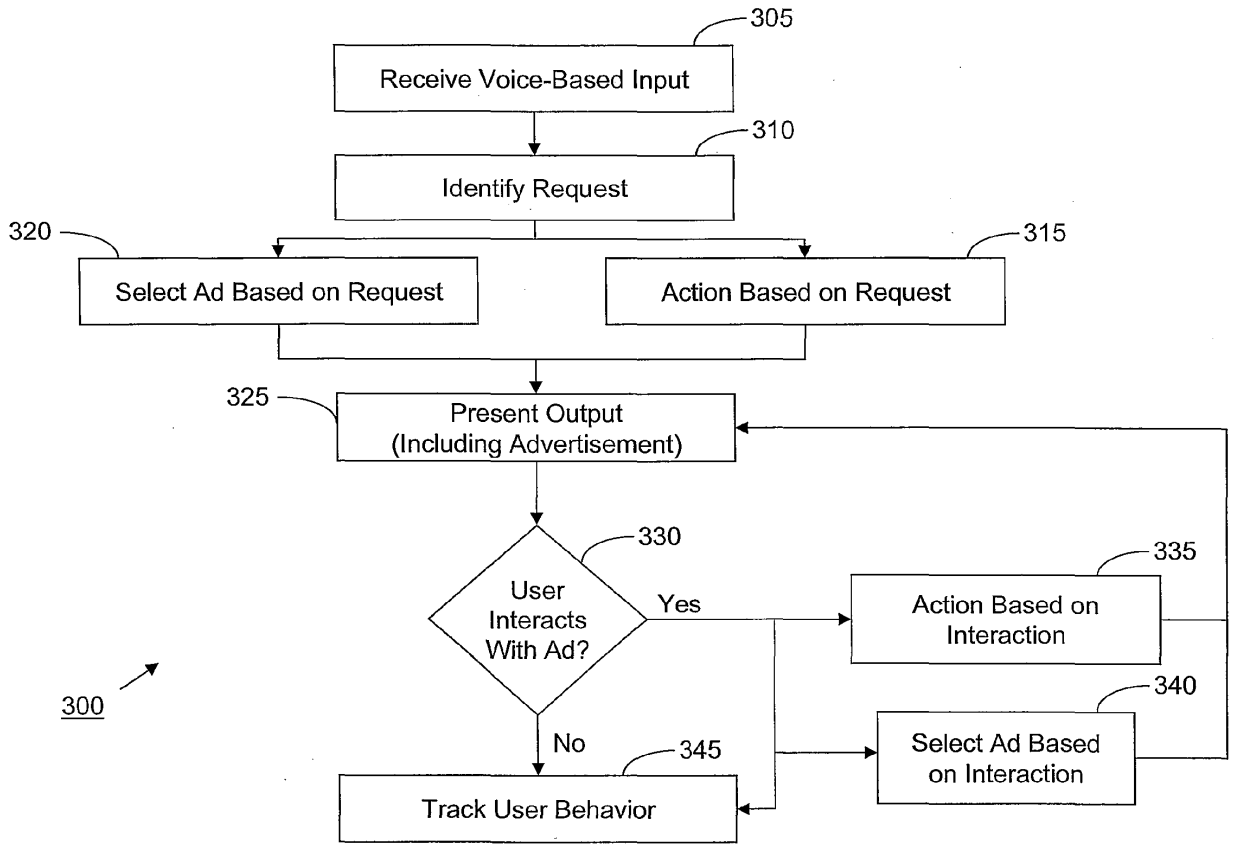

305
Receive Voice-Based Input
310
Identify Request
320
Select Ad Based on Request
315
Action Based on Request
325
Present Output
(Including Advertisement)
330
User
Interacts
With Ad?
Yes
335
Action Based on
Interaction
340
Select Ad Based
on Interaction
No
345
Track User Behavior
300

Fig. 3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 08/53157

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - G06Q 30/00
USPC - 705/14
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8) - G06Q 30/00 (2008.04)
USPC - 705/14

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
705/1, 7, 8; 725/37, 38, 42

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubWEST(USPT,PGPB,EPAB,JPAB); Google Scholar; Google
Search Terms Used: method to provide/select/display advertisement, advertisement, voice input, knowledge source, database, track, profile, user profile, advertisement database/source

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,696,965 A (DEDRICK) 09 December 1997 (09.12.1997) entire document, especially col 4, ln 60-66, col 5, ln 10-20, col 6, ln 12-30 and col 8, ln 24 - col 9, ln 10.	1-20

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
24 April 2008 (24.04.2008)	22 MAY 2008
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (April 2007)

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 October 2008 (02.10.2008)

PCT

(10) International Publication Number
WO 2008/118195 A2

(51) International Patent Classification:
G06F 17/28 (2006.01)

(21) International Application Number:
PCT/US2007/081481

(22) International Filing Date: 16 October 2007 (16.10.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/580,926 16 October 2006 (16.10.2006) US

(71) Applicant *(for all designated States except US)*: **VOICEBOX TECHNOLOGIES, INC.** [US/US]; 11980 Ne 24th Street, Bellevue, WA 98005 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **BALDWIN, Larry** [US/US]; 25498 S.e. 274th Place, Maple Valley, WA 98038 (US). **FREEMAN, Tom** [US/US]; 6735 83rd Avenue S.e., Mercer Island, WA 98040 (US). **TJALVE, Michael** [DK/US]; 14904 S.e. 47th Court, Bellevue, WA 98006 (US). **EBERSOLD, Blane** [US/US]; 4721 Admiral Way, S.w., Seattle, WA 98116 (US). **WEIDER, Chris** [US/US]; 2618 Grand Avenue, #b-508, Everett, WA 98201 (US).

(74) Agents: **BARUFKA, Jack, S.** et al.; Pillsbury Winthrop Shaw Pittman Llp, P.O. Box 10500, Mclean, VA 22102 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

(54) Title: SYSTEM AND METHOD FOR A COOPERATIVE CONVERSATIONAL VOICE USER INTERFACE

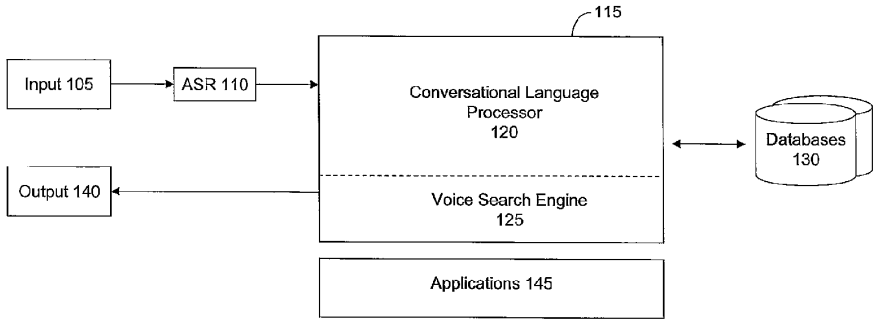


FIG. 1

(57) Abstract: A cooperative conversational voice user interface is provided. The cooperative conversational voice user interface may build upon short-term and long-term shared knowledge to generate one or more explicit and/or implicit hypotheses about an intent of a user utterance. The hypotheses may be ranked based on varying degrees of certainty, and an adaptive response may be generated for the user. Responses may be worded based on the degrees of certainty and to frame an appropriate domain for a subsequent utterance. In one implementation, misrecognitions may be tolerated, and conversational course may be corrected based on subsequent utterances and/or responses.

WO 2008/118195 A2

SYSTEM AND METHOD FOR A COOPERATIVE CONVERSATIONAL VOICE USER INTERFACE

FIELD OF THE INVENTION

[001] The invention relates to a cooperative conversational model for a human to machine voice user interface.

BACKGROUND OF THE INVENTION

[002] Advances in technology, particularly within the convergence space, have resulted in an increase in demand for voice recognition software that can exploit technology in ways that are intuitive to humans. While communication between human beings is most often "cooperative," in that information and/or context is shared to advance mutual conversational goals, existing Human-to-Machine interfaces fail to provide the same level of intuitive interaction. For example, each human participant in a conversation can contribute to an exchange for the benefit of the exchange. This is done through shared assumptions and expectations regarding various aspects of the conversation, such as the topic, participant knowledge about the topic, expectations of the other participant's knowledge about the topic, appropriate word usage for the topic and/or participants, conversational development based on previous utterances, the participants' tone or inflection, the quality and quantity of contribution expected from each participant, and many other factors. Participating in conversations that continually build and draw upon shared information is a natural and intuitive way for humans to converse.

[003] In contrast, complex Human-to-Machine interfaces do not allow users to exploit technology in an intuitive way, which inhibits mass-market adoption for various technologies. Incorporating a speech interface helps to alleviate this burden by making interaction easier and faster, but existing speech interfaces (when they actually work) still require significant learning on the part of the user. That is, existing speech interfaces are unable to bridge the gap between archaic Human-to-Machine interfaces and conversational speech that would make interaction with systems feel normal. Users should be able to directly request what they want from a system in a normal, conversational fashion, without having to memorize exact words or phrases. Alternatively, when users are uncertain of particular needs, they should be able to engage the system in a productive, cooperative dialogue to resolve their

requests. Instead, existing speech interfaces force users to dumb down their requests to match simple sets of instructions in simple languages in order to communicate requests in ways that systems can understand. Using existing speech interfaces, there is virtually no option for dialogue between the user and the system to satisfy mutual goals.

[004] Therefore, existing systems lack a conversational speech model that can provide users with the ability to interact with systems in ways that are inherently intuitive to human beings. Existing systems suffer from these and other problems.

SUMMARY OF THE INVENTION

[005] According to various embodiments and aspects of the invention, a cooperative conversational voice user interface may understand free form human utterances, freeing users from being restricted to a fixed set of commands and/or requests. Rather, users can engage in cooperative conversations with a machine to complete a request or series of requests using a natural, intuitive, free form manner of expression.

[006] According to an aspect of the invention, an exemplary system architecture for implementing a cooperative conversational voice user interface is provided. The system may receive an input, which may include a human utterance received by an input device, where the utterance may include one or more requests. As used herein, an "utterance" may be words, syllables, phonemes, or any other audible sound made by a human being. As used herein, a "request" may be a command, directive, or other instruction for a device, computer, or other machine to retrieve information, perform a task, or take some other action. In one implementation, the input may be a multi-modal input, where at least part of the multi-modal input is an utterance. The utterance component of the input may be processed by an Automatic Speech Recognizer to generate one or more preliminary interpretations of the utterance. The one or more preliminary interpretations may then be provided to a conversational speech engine for further processing, where the conversational speech engine may communicate with one or more databases to generate an adaptive conversational response, which may be returned to the user as an output. In one implementation, the output may be a multi-modal output. For example, the utterance may include a request to perform an action, and the output may include a conversational response reporting success or failure, as well as an execution of the action.

[007] According to another aspect of the invention, an exemplary conversational

Amazon/VB Assets
Exhibit 1003
Page 580

speech engine may generate an adaptive conversational response to a request or series of requests. The conversational speech engine may include a free form voice search module that may understand an utterance made using typical, day-to-day language (i.e., in free form), and may account for variations in how humans normally speak, the vocabulary they use, and the conditions in which they speak. To account for intangible variables of human speech, the free form search module may include models of casual human speech. For example, in one implementation, the free form search module may understand specialized jargon and/or slang, tolerate variations in word order, and tolerate verbalized pauses or stuttered speech. For example, formalized English requests, where a verb precedes a noun, may be treated in an equivalent manner to requests where the noun precedes the verb. In another implementation, compound requests and/or compound tasks with multiple variables may be identified in a single utterance. By identifying all relevant information for completing one or more tasks from a single utterance, advantages may be provided over existing voice user interfaces, such as Command and Control systems that use verbal menus to restrict information that a person can provide at a given point. In another implementation, inferring intended requests from incomplete or ambiguous requests may provide a conversational feel. By modeling what contextual signifiers, qualifiers, or other information may be required to perform a task in an identified context, an adaptive response may be generated, such as prompting a user for missing contextual signifiers, qualifiers, or other information. In one implementation, the response may ask for missing information in a way that most restricts possible interpretations, and the response may be framed to establish a domain for a subsequent user utterance. In another implementation, common alternatives for nouns and verbs may be recognized to reflect variations in usage patterns according to various criteria. Thus, variations in expression may be supported because word order is unimportant or unanticipated, and nouns and/or verbs may be represented in different ways to give simplistic, yet representative, examples. In another implementation, requests may be inferred from contradictory or otherwise inaccurate information, such as when an utterance includes starts and stops, restarts, stutters, run-on sentences, or other imperfect speech. For example, a user may sometimes change their mind, and thus alter the request in mid-utterance, and the imperfect speech feature may nonetheless be able to infer a request based on models of human speech. For example, various models may indicate that a last criterion is most likely to be correct, or intonation, emphasis, stress, use of the word "not," or other

Amazon/VB Assets
Exhibit 1003
Page 581

models may indicate which criterion is most likely to be correct.

[008] According to another aspect of the invention, the conversational speech engine may include a noise tolerance module that may discard words or noise which has no meaning in a given context to reduce a likelihood of confusion. Moreover, the noise tolerance module may filter out environmental and non-human noise to further reduce a likelihood of confusion. In one implementation, the noise tolerance module may cooperate with other modules and features to filter out words that do not fit into an identified context. For example, the noise tolerance module may filter other human conversations and/or utterances within a range of one or more microphones. For example, a single device may include multiple microphones, or multiple devices may each include one or more microphones, and the noise tolerance module may collate inputs and cooperatively filter out sound by comparing a speech signal from the various microphones. The noise tolerance module may also filter out non-human environmental noise within range of the microphones, out-of-vocabulary words caused by speaker ambiguity or malapropisms, or other noise that may be unrelated to a target request. Performance benchmarks for the noise tolerance module may be defined by noise models based on human criteria. For example, if a driver of a car is 92% likely to be understood by a passenger when traveling at 65 miles-per-hour with windows cracked, then performance benchmarks for the noise tolerance module may have a similar performance under such conditions.

[009] According to another aspect of the invention, the conversational speech engine may include a context determination process that determines one or more contexts for a request to establish meaning within a conversation. The one or more contexts may be determined by having one or more context domain agents compete to determine a most appropriate domain for a given utterance. Once a given domain "wins" the competition, the winning domain may be responsible for establishing or inferring further contexts and updating short-term and long-term shared knowledge. If there is a deadlock between context domain agents, an adaptive conversational response may prompt the user to assist in disambiguating between the deadlocked agents. Moreover, the context determination process may infer intended operations and/or context based on previous utterances and/or requests, whereas existing systems consider each utterance independently, potentially making the same errors over and over again. For example, if a given interpretation turns

Amazon/VB Assets
Exhibit 1003
Page 582

out to be incorrect, the incorrect interpretation may be removed as a potential interpretation from one or more Automatic Speech Recognizer grammars and/or from possible interpretations determined by the conversational speech engine, thereby assuring that a mistake will not be repeated for an identical utterance.

[010] The context determination process may provide advantages over existing voice user interfaces by continually updating one or more models of an existing context and establishing context as a by-product of a conversation, which cannot be established a priori. Rather, the context determination process may track conversation topics and attempt to fit a current utterance into recent contexts, including switching between contexts as tasks are completed, partially completed, requested, etc. The context determination process may identify one or more context domains for an utterance by defining a collection of related functions that may be useful for users in various context domains. Moreover, each context domain may have relevant vocabularies and thought collections to model word groupings, which when evaluated together, may disambiguate one context domain from another. Thus, eliminating out-of-context words and noise words when searching for relevant combinations may enhance accuracy of inferences. This provides advantages over existing systems that attempt to assign meaning to every component of an utterance (i.e., including out-of-context words and noise words), which results in nearly infinite possible combinations and greater likelihood of confusion. The context determination process may also be self-aware, assigning degrees of certainty to one or more generated hypotheses, where a hypothesis may be developed to account for variations in environmental conditions, speaker ambiguity, accents, or other factors. By identifying a context, capabilities within the context, vocabularies within the context, what tasks are done most often historically in the context, what task was just completed, etc., the context determination process may establish intent from rather meager phonetic clues. Moreover, just as in human-to-human conversation, users may switch contexts at any time without confusion, enabling various context domains to be rapidly selected, without menu-driven dead ends, when an utterance is unambiguous.

[011] According to another aspect of the invention, an exemplary cooperative conversational model may build upon free form voice search, noise tolerance, and context determination to implement a conversational Human-to-Machine interface that reflects human interaction and normal conversational behavior. That is, the cooperative

Amazon/VB Assets
Exhibit 1003
Page 583

conversational model enables humans and machines to participant in a conversation with an accepted purpose or direction, with each participant contributing to the conversation for the benefit of the conversation. By taking advantage of human presumptions about utterances that humans rely upon, both as speakers and listeners, a Human-to-Machine interface may be analogous to everyday human-to-human conversation. In one implementation, the exemplary cooperative conversation model may take incoming data (shared knowledge) to inform a decision (intelligent hypothesis building), and then may refine the decision and generate a response (adaptive response building).

[012] According to another aspect of the invention, shared knowledge may include both short-term and long-term knowledge. Short-term knowledge may accumulate during a single conversation, where input received during a single conversation may be retained. The shared knowledge may include cross-modality awareness, where in addition to accumulating input relating to user utterances, requests, locations, etc., the shared knowledge may accumulate a current user interface state relating to other modal inputs to further build shared knowledge models. The shared knowledge may be used to build one or more intelligent hypotheses using current and relevant information, build long-term shared knowledge by identifying information with long-term significance, and generate adaptive responses with relevant state and word usage information. Moreover, because cooperative conversations model human conversations, short-term session data may be expired after a psychologically appropriate amount of time, thereby humanizing system behavior, reducing a likelihood of contextual confusion based on stale data, while also adding relevant information from an expired session context to long-term knowledge models. Long-term shared knowledge may generally be user-centric, rather than session-based, where inputs may be accumulated over time to build user, environmental, cognitive, historical, or other long-term knowledge models. Long-term and short-term shared knowledge may be used simultaneously anytime a user engages in a cooperative conversation. Long-term shared knowledge may include explicit and/or implicit user preferences, a history of recent contexts, requests, tasks, etc., user-specific jargon related to vocabularies and/or capabilities of a context, most often used word choices, or other information. The long-term shared knowledge may be used to build one or more intelligent hypotheses using current and relevant information, generate adaptive responses with appropriate word choices when unavailable via short-term shared knowledge, refine long-term shared knowledge models, identify a frequency of

Amazon/VB Assets
Exhibit 1003
Page 584

specific tasks, identify tasks a user frequently has difficulty with, or provide other information and/or analysis to generate more accurate conversational responses. Shared knowledge may also be used to adapt a level of unprompted support (e.g., for novices versus experienced users, users who are frequently misrecognized, etc.) Thus, shared knowledge may enable a user and a voice user interface to share assumptions and expectations such as topic knowledge, conversation history, word usage, jargon, tone, or other assumptions and/or expectations that facilitate a cooperative conversation between human users and a system.

[013] According to another aspect of the invention, a conversation type may be identified for any given utterance. Categorizing and developing conceptual models for various types of exchanges may consistently align user expectations and domain capabilities. One or more intelligent hypotheses may be generated as to a conversation type by considering conversational goals, participant roles, and/or an allocation of information among the participants. Based on the conversational goals, participant roles, and allocation of information, the intelligent hypotheses may consider various factors to classify a conversation (or utterance) into general types of conversations that can interact with one another to form many more variations and permutations of conversation types (e.g., a conversation type may change dynamically as information is reallocated from one participant to another, or as conversational goals change based on the reallocation of information).

[014] According to another aspect of the invention, the intelligent hypotheses may include one or more hypotheses of a user's intent in an utterance. In addition, the intelligent hypotheses may use short-term and/or long-term shared knowledge to proactively build and evaluate interaction with a user as a conversation progresses or over time. The hypotheses may model human-to-human interaction to include a varying degree of certainty for each hypothesis. That is, just as humans rely on knowledge shared by participants to examine how much and what kind of information was available, the intelligent hypotheses may leverage the identified conversation type and shared knowledge to generate a degree of certainty for each hypothesis.

[015] According to another aspect of the invention, syntactically, grammatically, and contextually sensitive "intelligent responses" may be generated from the intelligent hypotheses that can be used to generate a conversational experience for a user, while

Amazon/VB Assets
Exhibit 1003
Page 585

also guiding the user to reply in a manner favorable for recognition. The intelligent responses may create a conversational feel by adapting to a user's manner of speaking, framing responses appropriately, and having natural variation and/or personality (e.g., by varying tone, pace, timing, inflection, word use, jargon, and other variables in a verbal or audible response).

[016] According to another aspect of the invention, the intelligent responses may adapt to a user's manner of speaking by using contextual signifiers and grammatical rules to generate one or more sentences that may cooperate with the user. By taking advantage of shared knowledge about how a user utters a request, the responses may be modeled using similar techniques used to recognize requests. The intelligent responses may rate possible responses statistically and/or randomize responses, which creates an opportunity to build an exchange with natural variation and conversational feel. This provides advantages over existing voice user interfaces where input and output is incongruous, as the input is "conversational" and the output is "computerese."

[017] According to another aspect of the invention, the intelligent responses may frame responses to influence a user reply utterance for easy recognition. For example, the responses may be modeled to illicit utterances from the user that may be more likely to result in a completed request. Thus, the responses may conform to a cooperative nature of human dialog and a natural human tendency to "parrot" what was just heard as part of a next utterance. Moreover, knowledge of current context may enhance responses to generate more meaningful conversational responses. Framing the responses may also deal with misrecognitions according to human models. For example, humans frequently remember a number of recent utterances, especially when one or more previous utterances were misrecognized or unrecognized. Another participant in the conversation may limit correction to a part of the utterance that was misrecognized or unrecognized, or over subsequent utterances and/or other interactions, clues may be provided to indicate the initial interpretation was incorrect. Thus, by storing and analyzing multiple utterances, utterances from earlier in a conversation may be corrected as the conversation progresses.

[018] According to another aspect of the invention, the intelligent responses may include multi-modal, or cross-modal, responses to a user. In one implementation, responses may be aware of and control one or more devices and/or interfaces, and users may respond by using whichever input method, or combination of input methods, is most

Amazon/VB Assets
Exhibit 1003
Page 586

convenient.

[019] According to another aspect of the invention, the intelligent responses may correct a course of a conversation without interrupting conversational flow. That is, even though the intelligent responses may be reasonably "sure," the intelligent responses may nonetheless sometimes be incorrect. While existing voice user interfaces tend to fail on average conversational missteps, normal human interactions may expect missteps and deal with them appropriately. Thus, responses after misrecognitions may be modeled after clarifications, rather than errors, and words may chosen in subsequent responses to move conversation forward and establish an appropriate domain to be explored with the user.

[020] Other objects and advantages of the invention will be apparent to those skilled in the art based on the following drawings and detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

[021] Fig. 1 is an exemplary block diagram of a system architecture according to one aspect of the invention.

[022] Fig. 2 is an exemplary block diagram of a conversational speech engine according to one aspect of the invention.

[023] Fig. 3 is an exemplary block diagram of a cooperative conversational model according to one aspect of the invention.

DETAILED DESCRIPTION

[024] Referring to Fig. 1, an exemplary system architecture for implementing a cooperative conversational voice user interface is illustrated according to one aspect of the invention. The system may receive an input 105 from a user, where in one implementation, input 105 may be an utterance received by an input device (e.g., a microphone), where the utterance may include one or more requests. Input 105 may also be a multi-modal input, where at least part of the multi-modal input is an utterance. For example, the input device may include a combination of a microphone and a touch-screen device, and input 105 may include an utterance that includes a request relating to a portion of a display on the touch-screen device that the user is touching. For instance, the touch-screen device may be a navigation device, and input 105 may include an utterance

Amazon/VB Assets
Exhibit 1003
Page 587

of "Give me directions to here," where the user may be requesting directions to a desired destination on the display of the navigation device.

[025] The utterance component of input 105 may be processed by an Automatic Speech Recognizer 110 to generate one or more preliminary interpretations of the utterance. Automatic Speech Recognizer 110 may process the utterance using any suitable technique known in the art. For example, in one implementation, Automatic Speech Recognizer 110 may interpret the utterance using techniques of phonetic dictation to recognize a phoneme stream, as described in copending U.S. Patent Application Ser. No. 11/513,269, entitled "Dynamic Speech Sharpening," which is hereby incorporated by reference in its entirety. The one or more preliminary interpretations generated by Automatic Speech Recognizer 110 may then be provided to a conversational speech engine 115 for further processing. Conversational speech engine 115 may include a conversational language processor 120 and/or a voice search engine 125, described in greater detail in Fig. 2 below. Conversational speech engine 115 may communicate with one or more databases 130 to generate an adaptive conversational response, which may be returned to the user as an output 140. In one implementation, output 140 may be a multi-modal output and/or an interaction with one or more applications 145 to complete the request. For example, output 140 may include a combination of an audible response and a display of a route on a navigation device. For example, the utterance may include a request to perform an action, and output 140 may include a conversational response reporting success or failure, as well as an execution of the action. In addition, in various implementations, Automatic Speech Recognizer 110, conversational speech engine 115, and/or databases 130 may reside locally (e.g., on a user device), remotely (e.g., on a server), or a hybrid model of local and remote processing may be used (e.g., lightweight applications may be processed locally while computationally intensive applications may be processed remotely).

[026] Referring to Fig. 2, an exemplary block diagram is provided illustrating a conversational speech engine 215 according to one aspect of the invention. Conversational speech engine 215 may include a conversational language processor 220 that generates an adaptive conversational response to a request or series of requests using a free form voice search module 245, a noise tolerance module 250, and/or a context determination process 255. According to one aspect of the invention, modules 245-255 may communicate with a voice search engine 225 that includes one or more context

Amazon/VB Assets
Exhibit 1003
Page 588

domain agents 230 and/or one or more vocabularies 235 to aid in interpreting utterances and generating responses, as described in "Enhancing the VUE™ (Voce-User-Experience) Through Conversational Speech," by Tom Freeman and Larry Baldwin, which is herein incorporated by reference in its entirety. Conversational speech engine 215 may generate an adaptive conversational response to one or more requests, where the requests may depend on unspoken assumptions, incomplete information, context established by previous utterances, user profiles, historical profiles, environmental profiles, or other information. Moreover, conversational speech engine 215 may track which requests have been completed, which requests are being processed, and/or which requests cannot be processed due to incomplete or inaccurate information, and the response may be generated accordingly.

[027] According to one aspect of the invention, free form voice search module 245 may understand an utterance made using typical, day-to-day language (i.e., in free form), and may account for variations in how humans normally speak, the vocabulary they use, and the conditions in which they speak. Because variables such as stress, distraction, and serendipity are always different and infinitely varied, free form search module 245 may be designed with a goal of understanding that no human will come to the same Human-to-Machine interface situation in the same way twice. Thus, free form search module 245 may implement one or more features that model casual human speech. In various implementations, free form search module 245 may include, among other things, a free form utterance feature, a one-step access feature, an inferencing intended operations feature, an alternative expression feature, and/or an imperfect speech feature.

[028] The free form utterance feature may understand specialized jargon and/or slang, tolerate variations in word order (e.g., whether a subject of a request comes before or after a verb may be irrelevant), and tolerate verbalized pauses (e.g., "um," "ah," "eh," and other utterances without meaning). For example, the free form utterance feature may treat formalized English verb-before-noun requests in an equivalent manner to free form requests where a noun may precede a verb. For example, user utterances of "Change it to the Squizz" and "You know, um, that Squizz channel, ah, switch it there" may be treated equivalently (where Squizz is a channel on XM Satellite Radio). In either case, the free form utterance feature is able to identify "Squizz" as a subject of the utterance and "Change it" or "switch it" as a verb or request for the utterance (e.g., by

Amazon/VB Assets
Exhibit 1003
Page 589

cooperating with context determination process 255, or other features, and identifying a relevant context domain agent 230 and/or vocabulary 235 to interpret the utterance).

[029] The one-step access feature may understand utterances that include compound requests with multiple variables. For example, a user utterance may be "What is the forecast for Boston this weekend?" The one-step access feature may identify "weather" as a context (e.g., by cooperating with context determination process 255, or other features, and identifying "forecast" as a synonym of "weather"), and search for a city equal to "Boston" and a time equal to "weekend." By identifying all relevant information for completing a task from a single utterance, the one-step access feature may overcome drawbacks of existing voice user interfaces, such as Command and Control systems that use verbal menus to restrict information that a person can provide at a given point (e.g., a Command and Control system for a phone directory service may say: "State please," . . . "City please," . . . "What listing," etc.). Moreover, some utterances may include compound requests, and the one-step access feature may decompose the compound requests into sub-tasks. For example, a user utterance of "I need to be at a meeting tomorrow in San Francisco at 8:00 am" may be decomposed into a set of sub-tasks such as (1) checking availability and reserving a flight on an evening before the meeting, (2) checking availability and reserving a hotel, (3) checking availability and reserving a car, etc., where users may further designate preferences for various tasks (e.g., first check availability on an airline for which the user is a frequent flyer). Depending on a level of shared knowledge about a user's preferences and/or historical patterns, the one-step access feature may infer additional tasks from a request. For example, in the above example, the one-step access feature may also check a weather forecast, and if the weather is "nice" (as defined by the user preferences and/or as inferred from historical patterns), the one-step access feature may schedule a tee-time at a preferred golf course in San Francisco.

[030] The inferencing intended operations feature may identify an intended request from incomplete or ambiguous requests. For example, when a user utters "Route <indecipherable> Chicago <indecipherable> here," where the user intended to say "Route calculation to Chicago from here," the inferencing intended operations feature may model what is required to calculate a route (an origination point and a destination point). Because the utterance includes the origination point and the destination point, a request to calculate a route from the user's present location to Chicago may be inferred.

Amazon/VB Assets
Exhibit 1003
Page 590

Similarly, when the inferencing intended operations feature does not have sufficient information to infer a complete request, an adaptive conversational response may be generated to prompt the user for missing information. For example, when an utterance includes a request for a stock quote but not a company name (e.g., "Get me the stock price for <indecipherable>"), the response may be "What company's stock quote do you want?" The user may then provide an utterance including the company name, and the request may be completed. In one implementation, the response may ask for missing information in a way that most restricts possible interpretations (e.g., in a request for a task that requires both a city and a state, the state may be asked for first because there are fewer states than cities). Moreover, the inferencing intended operations feature may model compound tasks and/or requests by maintaining context and identifying relevant and/or missing information at both a composite and sub-task level.

[031] The alternative expression feature may recognize common alternatives for nouns and verbs to reflect variations in usage patterns according to various criteria. For example, users may vary expression based on age, socio-economics, ethnicity, user whims, or other factors. Thus, the alternative expression feature may support variations in expression where word order is unimportant or unanticipated. Alternatives in expression based on various criteria or demographics may be loaded into context domain agents 230 and/or vocabularies 235, and the alternative expression feature may update context domain agents 230 and/or vocabularies 235 based on inferred or newly discovered variations. In one implementation, conversational speech engine 215 may include a subscription interface to update changes to context domain agents 230 and/or vocabularies 235 (e.g., a repository may aggregate various user utterances and deploy updates system wide). In operation, the alternative expression feature may allow nouns and/or verbs to be represented in different ways to give simplistic, yet representative, examples. For example, a user interested in a weather forecast for Washington, D.C. may provide any of the following utterances, each of which are interpreted equivalently: "What's the weather like in DC," "Is it raining inside the Beltway," Gimme the forecast for the capital," etc. Similarly, utterances of "Go to my home," "Go home," "Show route to home," and "I would like to know my way home" may all be interpreted equivalently, where a user profile may include the user's home address and a navigation route to the home address may be calculated.

[032] The imperfect speech feature may be able to infer requests from contradictory or

Amazon/VB Assets
Exhibit 1003
Page 591

otherwise inaccurate information, such as when an utterance includes starts and stops, restarts, stutters, run-on sentences, or other imperfect speech. For example, a user may sometimes change their mind, and thus alter the request in mid-utterance, and the imperfect speech feature may nonetheless be able to infer a request based on models of human speech. For example, for an utterance of "Well, I wanna . . . Mexi . . . no, steak restaurant please, I'm hungry," existing voice user interfaces make no assumptions regarding models of human speech and would be unable to infer whether the user wanted a Mexican or steak restaurant. The imperfect speech feature overcomes these drawbacks by using various models of human understanding that may indicate that a last criterion is most likely to be correct, or intonation, emphasis, stress, use of the word "not," or other models may indicate which criterion is most likely to be correct. Thus, in the above example, the imperfect speech feature may infer that the user wants a steak restaurant.

[033] According to one aspect of the invention, noise tolerance module 250 may be closely related to the imperfect speech feature, and may operate to discard words or noise that has no meaning in a given context so as not to create confusion. Moreover, noise tolerance module 250 may filter out environmental and non-human noise to further reduce a likelihood of confusion. In one implementation, noise tolerance module 250 may cooperate with other modules and features to filter out words that do not fit into a context. For example, one or more contexts may be identified, and words that have no meaning with respect to system capabilities, random human utterances without meaning and other noise may be filtered out. Thus, noise tolerance module 250 may model real-world conditions to identify meaningful requests. For example, noise tolerance module 250 may filter other human conversations and/or utterances within a range of one or more microphones, For example, a single device may include multiple microphones, or multiple devices may each include one or more microphones, and the noise tolerance module may collate inputs and cooperatively filter out sound by comparing a speech signal from the various microphones. Noise tolerance module 250 may also filter out non-human environmental noise within the range of the microphones, out-of-vocabulary words, which could be a result of speaker ambiguity or malapropisms, or other noise that may be unrelated to a target request. Noise models in noise tolerance module 250 may define performance benchmarks based on human criteria. For example, if a driver of a car, traveling at 65 miles-per-hour, with windows cracked is 92% likely to be understood by a passenger, then noise tolerance module 250 may have a similar performance under

Amazon/VB Assets
Exhibit 1003
Page 592

those conditions.

[034] According to one aspect of the invention, conversational speech engine 215 may include a context determination process 255 that determines one or more contexts for a request to establish meaning within a conversation. The one or more contexts may be determined by having one or more context domain agents compete to determine a most appropriate domain for a given utterance, as described in copending U.S. Patent Application Ser. No. 11/197,504, entitled "Systems and Methods for Responding to Natural Language Speech Utterance," and copending U.S. Patent Application Ser. No. 11/212,693, entitled "Mobile Systems and Methods of Supporting Natural Language Human-Machine Interactions," both of which are hereby incorporated by reference in their entirety. Once a given context domain agent "wins" the competition, the winning agent may be responsible for establishing or inferring further contexts and updating short-term and long-term shared knowledge. If there is a deadlock between context domain agents, an adaptive conversational response may prompt the user to assist in disambiguating between the deadlocked agents. For example, a user utterance of "What about traffic?" may have a distinct meaning in various contexts. That is, "traffic" may have a first meaning when the user is querying a system's media player (i.e., "traffic" would be a Rock and Roll band led by singer/songwriter Steve Winwood), a second meaning when the user is querying a search interface regarding Michael Douglas films (i.e., "traffic" would be a film directed by Steven Soderbergh), a third meaning when the user is querying a navigation device for directions to an airport (i.e., "traffic" would be related to conditions on roads along a route to the airport).

[035] Moreover, context determination process 255 may infer intended operations and/or context based on previous utterances and/or requests, whereas existing systems consider each utterance independently, potentially making the same errors over and over again. For example, if a given interpretation turns out to be incorrect, the incorrect interpretation may be removed as a potential interpretation from one or more Automatic Speech Recognizer grammars and/or from possible subsequent interpretations determined by context determination process 255, thereby assuring that a mistake will not be repeated for an identical utterance.

[036] Context determination process 255 may overcome drawbacks of existing systems by continually updating one or more models of an existing context, where establishing

Amazon/VB Assets
Exhibit 1003
Page 593

context may be a by-product of a conversation, which cannot be established a priori. Context determination process 255 may establish a first context domain, change to a second context domain, change back to the first context domain, and so on, as tasks are completed, partially completed, requested, etc, and a context stack may track conversation topics and attempt to fit a current utterance into a most-recent context, next-most-recent topic, etc., traversing the context stack until a most likely intent can be established. For example, a user may utter "What's the traffic report," and context determination process 255 may establish Traffic as a context, and return an output including a traffic report, which does not happen to mention traffic on Interstate-5. The user may then utter "What about I-5?" and context determination process 255 may know that the current context is Traffic, a traffic report including information about Interstate-5 may be searched for, and the traffic report indicating that Interstate-5 is crowded may be returned as an output. The user may then utter "Is there a faster way?" and context determination module 255 may know that the current context is still Traffic, and may search for routes to a specified destination with light traffic and avoiding Interstate-5. Moreover, context determination process 255 may build context based on user profiles, environmental profiles, historical profiles, or other information to further refine the context. For example, the profiles may indicate that Interstate-5 is a typical route taken Monday through Friday.

[037] The profiles may be particularly meaningful when attempting to disambiguate between contexts where a word has different meanings in different contexts. For example, a user may utter "What's the weather in Seattle?" and context determination process 255 may establish Weather as a context, as well as establishing Seattle as an environmental context. The user may then utter "and Portland?" and context determination process 255 may return a weather report for Portland, Oregon based on the Weather and an environmental proximity between Portland, Oregon and Seattle, Washington. The user may then ask "What time does the game start?"" and a search for sports events with teams from Seattle and/or Portland may occur, with results presented conversationally according to methods described in greater detail below in Fig. 3. Correlatively, had user originally uttered "What's the weather in Portsmouth, New Hampshire," in the second utterance, context determination process 255 may instead retrieve a weather report for Portland, Maine based on an environmental proximity to New Hampshire. Moreover, when environmental profiles, contextual shared knowledge, and/or other short-term and/or long-term shared knowledge does not provide enough information to

Amazon/VB Assets
Exhibit 1003
Page 594

disambiguate between possibilities, responses may prompt the user with a request for further information (e.g., "Did you mean Portland, Maine, or Portland, Oregon?").

[038] Context determination process 255 may cooperate with context domain agents 230, where each context domain agent 230 may define a collection of related functions that may be useful for users. Moreover, each context domain agent 230 may include a relevant vocabulary 235 and thought collections that model word groupings, which when evaluated together, may disambiguate one context domain from another (e.g., a Music context domain agent 230 may include a vocabulary 235 for songs, artists, albums, etc., whereas a Stock context domain agent 230 may include a vocabulary 235 for company names, ticker symbols, financial metrics, etc.). Thus, accuracy in identifying meaning may be enhanced by eliminating out-of-context words and noise words when searching for relevant combinations. In contrast, existing systems attempt to assign meaning to every component of an utterance (e.g., including out-of-context words and noise words), which results in nearly infinite possible combinations and greater likelihood of confusion. Moreover, context domain agents 230 may include metadata for each criteria to further assist in interpreting utterances, inferring intent, completing incomplete requests, etc. (e.g., a Space Needle vocabulary word may include metadata for Seattle, landmark, tourism, Sky City restaurant, etc.). Given a disambiguated criterion, context determination process 255 may thus be able to automatically determine other information needed to complete a request, discard importance of word order, and perform other enhancements for conversational speech.

[039] Context domain agents 230 may also be self-aware, assigning degrees of certainty to one or more generated hypotheses, where a hypothesis may be developed to account for variations in environmental conditions, speaker ambiguity, accents, or other factors. Conceptually, context domain agents 230 may be designed to model utterances as a hard-of-hearing person would at a noisy party. By identifying a context, capabilities within the context, vocabularies within the context, what tasks are done most often historically in the context, what task was just completed, etc., a context domain agent 230 may establish intent from rather meager phonetic clues. Moreover, the context stack may be one of a plurality of components for establishing context, and thus not a constraint upon the user. All context domains may be accessible, allowing the user to switch contexts at any time without confusion. Thus, just as in human- to-human conversation, context domains

Amazon/VB Assets
Exhibit 1003
Page 595

may be rapidly selected, without menu-driven dead ends, when an utterance is unambiguous. For example, a user may utter, "Please call Rich Kennewick on his cell phone," and a system response of "Do you wish me to call Rich Kennewick on his cell?" may be generated. The user may decide at that point to call Rich Kennewick later, and instead, listen to some music. Thus, the user may then utter, "No, play the Louis Armstrong version of Body and Soul from my iPod," and a system response of "Playing Body and Soul by Louis Armstrong" may be generated as Body and Soul is played through a media player. In this example, the later utterance has no contextual connection to the first utterance, yet because request criteria in the utterances are unambiguous, contexts can be switched easily without relying on the context stack.

[040] Referring to Fig. 3, an exemplary cooperative conversational model 300 is illustrated according to an aspect of the invention. Cooperative conversational model 300 may build upon free form voice search 245, noise tolerance 250, and context determination 255 to implement a conversational Human-to-Machine interface that reflects how humans interact with each other and their normal behavior in conversation. Simply put, cooperative conversational model 300 enables humans and machines to participate in a conversation with an accepted purpose or direction, with each participant contributing to the conversation for the benefit of the conversation. That is, cooperative conversational model 300 incorporates technology and process-flow that takes advantage of human presumptions about utterances that humans rely upon, both as speakers and listeners, thereby creating a Human-to-Machine interface that is analogous to everyday human-to-human conversation. In one implementation, a cooperative conversation may take incoming data (shared knowledge) 305 to inform a decision (intelligent hypothesis building) 310, and then may refine the decision and generate a response (adaptive response building) 315.

[041] According to one aspect of the invention, shared knowledge 305 includes both short-term and long-term knowledge about incoming data. Short-term knowledge may accumulate during a single conversation, while long-term knowledge may accumulate over time to build user profiles, environmental profiles, historical profiles, cognitive profiles, etc.

[042] Input received during a single conversation may be retained in a Session Input Accumulator. The Session Input Accumulator may include cross-modality

Amazon/VB Assets
Exhibit 1003
Page 596

awareness, where in addition to accumulating input relating to user utterances, requests, locations, etc., the Session Input Accumulator may accumulate a current user interface state relating to other modal inputs to further build shared knowledge models and more accurate adaptive responses (e.g., when a user utters a request relating to a portion of a touch-screen device, as described above). For example, the Session Input Accumulator may accumulate inputs including recognition text for each utterance, a recorded speech file for each utterance, a list-item selection history, a graphical user interface manipulation history, or other input data. Thus, the Session Input Accumulator may populate Intelligent Hypothesis Builder 310 with current and relevant information, build long-term shared knowledge by identifying information with long-term significance, provide Adaptive Response Builder 315 with relevant state and word usage information, retain recent contexts for use with Intelligent Hypothesis Builder 310, and/or retain utterances for reprocessing during multi-pass evaluations. Moreover, because cooperative conversations 300 model human conversations, short-term session data may be expired after a psychologically appropriate amount of time, thereby humanizing system behavior. For example, a human is unlikely to recall a context of a conversation from two years ago, but because the context would be identifiable by a machine, session context is expired after a predetermined amount of time to reduce a likelihood of contextual confusion based on stale data. However, relevant information from an expired session context may nonetheless be added to user, historical, environmental, cognitive, or other long-term knowledge models.

[043] Long-term shared knowledge may generally be user-centric, rather than session-based. That is, inputs may be accumulated over time to build user, environmental, cognitive, historical, or other long-term knowledge models. Long-term and short-term shared knowledge (collectively, shared knowledge 305) may be used simultaneously anytime a user engages in a cooperative conversation 300. Long-term shared knowledge may include explicit and/or implicit user preferences, a history of most recently used agents, contexts, requests, tasks, etc., user-specific jargon related to vocabularies and/or capabilities of an agent and/or context, most often used word choices, or other information. The long-term shared knowledge may be used to populate Intelligent Hypothesis Builder 310 with current and relevant information, provide Adaptive Response Builder 315 with appropriate word choices when the appropriate word choices are unavailable via the Session Input Accumulator, refine long-term shared

Amazon/VB Assets
Exhibit 1003
Page 597

knowledge models, identify a frequency of specific tasks, identify tasks a user frequently has difficulty with, or provide other information and/or analysis to generate more accurate conversational responses.

[044] As described above, shared knowledge 305 may be used to populate Intelligent Hypothesis Builder 310, such that a user and a voice user interface may share assumptions and expectations such as topic knowledge, conversation history, word usage, jargon, tone (e.g., formal, humorous, terse, etc.), or other assumptions and/or expectations that facilitate interaction at a Human-to-Machine interface.

[045] According to an aspect of the invention, one component of a successful cooperative conversation may be identifying a type of conversation from an utterance. By categorizing and developing conceptual models for various types of exchanges, user expectations and domain capabilities may be consistently aligned. Intelligent Hypothesis Builder 310 may generate a hypothesis as to a conversation type by considering conversational goals, participant roles, and/or an allocation of information among the participants. Conversational goals may broadly include: (1) getting a discrete piece of information or performing a discrete task, (2) gathering related pieces of information to make a decision, and/or (3) disseminating or gathering large amounts of information to build expertise. Participant roles may broadly include: (1) a leader that controls a conversation, (2) a supporter that follows the leader and provides input as requested, and/or (3) a consumer that uses information. Information may be held by one or more of the participants at the outset of a conversation, where a participant may hold most (or all) of the information, little (or none) of the information, or the information may be allocated roughly equally amongst the participants. Based on the conversational goals, participant roles, and allocation of information, Intelligent Hypothesis Builder 310 may consider various factors to classify a conversation (or utterance) into general types of conversations that can interact with one another to form many more variations and permutations of conversation types (e.g., a conversation type may change dynamically as information is reallocated from one participant to another, or as conversational goals change based on the reallocation of information).

[046] For example, in one implementation, a query conversation may include a conversational goal of getting a discrete piece of information or performing a particular task, where a leader of the query conversation may have a specific goal in

Amazon/VB Assets
Exhibit 1003
Page 598

mind and may lead the conversation toward achieving the goal. The other participant may hold the information and may support the leader by providing the information. In a didactic conversation, a leader of the conversation may control information desired by a supporter of the conversation. The supporter's role may be limited to regulating an overall progression of the conversation and interjecting queries for clarification. In an exploratory conversation, both participants share leader and supporter roles, and the conversation may have no specific goal, or the goal may be improvised as the conversation progresses. Based on this model, Intelligent Hypothesis Builder 310 may broadly categorize a conversation (or utterance) according to the following diagram:

QUERY		
	Participant A	*Participant B*
	User	Voice User Interface
GOAL	Get information/action	Provide information/action
ROLE	Leader/Consumer	Supporter/Dispenser
INFORMATION ALLOCATION	Less	More

Amazon/VB Assets
Exhibit 1003
Page 599

DIDACTIC		
	Participant A	*Participant B*
	User	Voice User Interface
GOAL	Get information	Provide information
ROLE	Follower/Consumer	Leader/Dispenser
INFORMATION ALLOCATION	Less	More

EXPLORATORY		
	Participant A	*Participant B*
	User	Voice User Interface
GOAL	Gather/share information	Gather/share information
ROLE	Follower/Consumer and Leader/Dispenser	Follower/Consumer and Leader/Dispenser
INFORMATION ALLOCATION	Equal or alternating	Equal or alternating

[047] Intelligent Hypothesis Builder 310 may use an identified conversation type to assist in generating a set of hypotheses as to a user's intent in an utterance. In addition, Intelligent Hypothesis Builder 310 may use short-term shared knowledge from the Session Input Accumulator to proactively build and evaluate interaction with a user as a conversation progresses, as well as long-term shared knowledge to proactively build and evaluate interaction with the user over time. Intelligent Hypothesis Builder 310 may thus adaptively arrive at a set of n-best hypotheses about user intent, and the n-best hypotheses may be provided to an Adaptive Response Builder 315. In addition, Intelligent Hypothesis Builder 310 may model human-to-human interaction by calculating a degree of certainty for each of the hypotheses. That is, just as humans rely on knowledge shared by participants to examine how much and what kind of information was available, Intelligent Hypothesis Builder 310 may leverage the identified conversation type and short-term

Amazon/VB Assets
Exhibit 1003
Page 600

and long-term shared knowledge to generate a degree of certainty for each hypothesis.

[048] According to another aspect of the invention, Intelligent Hypothesis Builder 310 may generate one or more explicit hypotheses of a user's intent when an utterance contains all information (including qualifiers) needed to complete a request or task. Each hypothesis may have a corresponding degree of certainty, which may be used to determine a level of unprompted support to provide in a response. For example, a response may include a confirmation to ensure the utterance was not misunderstood. or the response may adaptively prompt a user to provide missing information.

[049] According to another aspect of the invention, Intelligent Hypothesis Builder 310 may use short-term knowledge to generate one or more implicit hypotheses of a user's intent when an utterance may be missing required qualifiers or other information needed to complete a request or task. Each hypothesis may have a corresponding degree of certainty. For instance, when a conversation begins, short-term knowledge stored in the Session Input Accumulator may be empty, and as the conversation progresses, the Session Input Accumulator may build a history of the conversation. Intelligent Hypothesis Builder 310 may use data in the Session Input Accumulator to supplement or infer additional information about a current utterance. For example, Intelligent Hypothesis Builder 310 may evaluate a degree of certainty based on a number of previous requests relevant to the current utterance. In another example, when the current utterance contains insufficient information to complete a request or task, data in the Session Input Accumulator may be used to infer missing information so that a hypothesis can be generated. In still another example, Intelligent Hypothesis Builder 310 may identify syntax and/or grammar to be used by Adaptive Response Builder 315 to formulate personalized and conversational response. In yet another example, when the current utterance contains a threshold amount of information needed to complete a request or task, data in the Session Input Accumulator may be relied upon to tune a degree of certainty.

[050] According to another aspect of the invention, Intelligent Hypothesis Builder 310 may use long-term shared knowledge to generate one or more implicit hypotheses of a user's intent when an utterance is missing qualifiers or other information needed to complete a request or task. Each hypothesis may have a corresponding degree of certainty. Using long-term knowledge may be substantially similar to using short-term shared knowledge, except that information may be unconstrained by a current session,

Amazon/VB Assets
Exhibit 1003
Page 601

and an input mechanism may include information from additional sources other than conversational sessions. For example, Intelligent Hypothesis Builder 310 may use information from long-term shared knowledge at any time, even when a new conversation is initiated, whereas short-term shared knowledge may be limited to an existing conversation (where no short-term shared knowledge would be available when a new conversation is initiated). Long-term shared knowledge may come from several sources, including user preferences or a plug-in data source (e.g., a subscription interface to a remote database), expertise of a user (e.g., based on a frequency of errors, types of tasks requested, etc., the user may be identified as a novice, intermediate, experienced, or other type of user), agent-specific information and/or language that may also apply to other agents (e.g., by decoupling information from an agent to incorporate the information into other agents), frequently used topics passed in from the Session Input Accumulator, frequently used verbs, nouns, or other parts of speech, and/or other syntax information passed in from the Session Input Accumulator, or other sources of long-term shared knowledge may be used.

[051] According to another aspect of the invention, knowledge-enabled utterances, as generated by Intelligent Hypothesis Builder 310, may include one or more explicit (supplied by a user), and one or more implicit (supplied by Intelligent Hypothesis Builder 310) contextual signifiers, qualifiers, criteria, and other information that can be used to identify and evaluate relevant tasks. At that point, Intelligent Hypothesis Builder 310 may provide an input to Adaptive Response Builder 315. The input received by Adaptive Response Builder 315 may include at least a ranked list of hypotheses, including explicit and/or implicit hypotheses, each of which may have a corresponding degree of certainty. A hypothesis may be assigned one of four degrees of certainty: (1) "sure," where contextual signifiers and qualifiers relate to one task, context and qualifiers relate to one task, and an ASR confidence level exceeds a predetermined threshold; (2) "pretty sure," where contextual signifiers and qualifiers relate to more than one task (select top-ranked task) and criteria relates to one request, and/or the ASR confidence level is below the predetermined threshold; (3) "not sure," where additional contextual signifiers or qualifiers are needed to indicate or rank a task; and (4) "no hypothesis," where little or no information can be deciphered. Each degree of certainty may further be classified as explicit or implicit, which may be used to adjust a response. The input received by Adaptive Response Builder 310 may also include a context, user syntax and/or

Amazon/VB Assets
Exhibit 1003
Page 602

grammar, context domain agent specific information and/or preferences (e.g., a travel context domain agent may know a user frequently requests information about France, which may be shared with a movie context domain agent so that responses may occasionally include French movies).

[052] According to another aspect of the invention, Adaptive Response Builder 315 may build syntactically, grammatically, and contextually sensitive "intelligent responses" that can be used with one or more agents to generate a conversational experience for a user, while also guiding the user to reply in a manner favorable for recognition. In one implementation, the intelligent responses may include a verbal or audible reply played through an output device (e.g., a speaker), and/or an action performed by a device, computer, or machine (e.g., downloading a web page, showing a list, executing an application, etc.). In one implementation, an appropriate response may not require conversational adaptation, and default replies and/or randomly selected response sets for a given task may be used.

[053] According to another aspect of the invention, Adaptive Response Builder 310 may draw on information maintained by Intelligence Hypothesis Builder 310 to generate responses that may be sensitive to context, task recognition of a current utterance, what a user already knows about a topic, what an application already knows about the topic, shared knowledge regarding user preferences and/or related topics, appropriate contextual word usage (e.g., jargon), words uttered by the user in recent utterances, conversational development and/or course correction, conversational tone, type of conversation, natural variation in wording of responses, or other information. As a result, Adaptive Response Builder 315 may generate intelligent responses that create conversational feel, adapt to information that accumulates over a duration of a conversation, maintain cross-modal awareness, and keep the conversation on course.

[054] According to another aspect of the invention, Adaptive Response Builder 315 may create a conversational feel by adapting to a user's manner of speaking, framing responses appropriately, and having natural variation and/or personality (e.g., by varying tone, pace, timing, inflection, word use, jargon, and other variables in a verbal or audible response). Adapting to a user's manner of speaking may include using contextual signifiers and grammatical rules to generate one or more sentences for use as response sets that may cooperate with the user. By taking advantage of short-term (from the Session Input

Amazon/VB Assets
Exhibit 1003
Page 603

Accumulator) and long-term (from one or more profiles) shared knowledge about how a user utters a request, the responses may be modeled using techniques used to recognize requests. Adaptive Response Builder 315 may rate possible responses statistically and/or randomize responses, which creates an opportunity to build an exchange with natural variation and conversational feel. This may be a significant advantage over existing voice user interfaces with incongruous input and output, where the input is "conversational" and the output is "computerese." The following examples may demonstrate how a response may adapt to a user's input word choices and manner of speaking:

User	**Do you know** [*mumbled words*] Seattle [*more mumbled words*]?
Voice User Interface	Did you want Seattle sports scores, weather, traffic, or news?

User	**Find me** [*mumbled words*] Seattle [*more mumbled words*]?
Voice User Interface	I *found* Seattle, did you want sports scores, weather, traffic, or news?

User	**Get me** [*mumbled words*] Seattle [*more mumbled words*]?
Voice User Interface	I've *got* Seattle, did you want me to *get* sports scores, weather, traffic, or news?

[055] According to another aspect of the invention, Adaptive Response Builder 315 may frame responses to influence a user to reply with an utterance that may be easily recognized. For example, a user may utter, "Get me the news" and a voice user interface response may be "Which of these categories? Top news stories, international news, political news, or sports news?" The response may be likely to illicit utterances from the user, such as "Top news stories" or "International news," which are more likely to result in a completed request. Thus, the responses may conform to a cooperative nature of human dialog, and a natural human tendency to "parrot" what was just heard as part of a next utterance. Moreover, knowledge of current context may enhance responses to generate more meaningful conversational responses, such as in the following exchange:

User	What's the weather like in Dallas?

Amazon/VB Assets
Exhibit 1003
Page 604

Voice User Interface	In Dallas, it's sunny and 90 degrees.
User	What theaters are showing the movie "The Fantastic Four" there?
Voice User Interface	10 theaters in Dallas are showing "The Fantastic Four." Do you want show times for a particular theater?

[056] Framing the responses may also deal with misrecognitions according to human models. For example, humans frequently remember a number of recent utterances, especially when one or more previous utterances were misrecognized or unrecognized. Another participant in the conversation may limit correction to a part of the utterance that was misrecognized or unrecognized, or over subsequent utterances and/or other interactions, clues may be provided to indicate the initial interpretation was incorrect. Thus, by storing and analyzing multiple utterances, utterances from earlier in a conversation may be corrected as the conversation progresses.

[057] According to another aspect of the invention, Adaptive Response Builder 315 may generate multi-modal, or cross-modal, responses to a user. In one implementation, responses may be aware of and control one or more devices and/or interfaces, and users may respond by using whichever input method, or combination of input methods, is most convenient. For example, a response asking the user to direct an utterance with a "Yes" or "No" in a multi-modal environment may also display alternatives visually.

[058] According to another aspect of the invention, Adaptive Response Builder 315 may correct a course of a conversation without interrupting conversational flow. Adaptive Response Builder 315 may generate intelligent responses based on the ranked list of hypotheses and corresponding degrees of certainty, which may be used to correct a course of a conversation without interrupting conversational flow. That is, even though the intelligent responses may be reasonably "sure," the intelligent responses may nonetheless sometimes be incorrect. While existing voice user interfaces tend to fail on average conversational missteps, normal human interactions may expect missteps and deal with them appropriately. Thus, responses after a misrecognition may be modeled after clarifications, rather than errors, and words may chosen in subsequent responses that move conversation forward and establish an appropriate domain to be explored with the user. For example, course correction may result in the following exchange:

Amazon/VB Assets
Exhibit 1003
Page 605

User	Can you **get** [*mumbled words here*] Mariners [*more mumbled words*]?
Voice User Interface	*I've got* the score for the Mariners game? 4-2 Mariners.
User	No, the start time for tomorrow's game.
Voice User Interface	Oh, 7:05.

[059] The above disclosure has been described in terms of specific exemplary aspects, implementations, and embodiments of the invention. However, those skilled in the art will recognize various changes and modifications that may be made without departing from the scope and spirit of the invention. Therefore, the specification and drawings are to be regarded as exemplary only, and the scope of the invention is to be determined solely by the appended claims.

Amazon/VB Assets
Exhibit 1003
Page 606

CLAIMS

What is claimed is:

1. A method for providing a voice user interface that generates cooperative conversational responses to user utterances, the responses being based on shared knowledge about a current conversation and/or one or more past conversations, the method comprising:

receiving an input, the input including at least one utterance;

updating the shared knowledge to include knowledge about the at least one utterance;

determining an intended meaning for the at least one utterance based on the updated shared knowledge; and

generating a response based on the determined intended meaning.

2. The method of claim 1, wherein updating the shared knowledge includes populating a short-term context stack with information about the at least one utterance.

3. The method of claim 2, further comprising expiring the information about the at least one utterance in the short-term context stack after a predetermined amount of time.

4. The method of claim 1, wherein updating the shared knowledge includes updating one or more long-term profiles based on information about the at least one utterance.

5. The method of claim 1, wherein determining the intended meaning includes identifying a conversation type for the at least one utterance, the conversation type based on one or more of a goal for the utterance, one or more participant roles for the utterance, and/or an allocation of information for the utterance.

6. The method of claim 1, wherein determining the intended meaning includes generating at least one hypothesis of an intended meaning for the at least one utterance, the at least one hypothesis including a degree of certainty.

7. The method of claim 6, wherein generating the at least one hypothesis includes identifying contextual signifiers and/or qualifiers in the at least one

Amazon/VB Assets
Exhibit 1003
Page 607

utterance.

8. The method of claim 7, further comprising inferring missing contextual signifiers and/or qualifiers for the at least one utterance based on the shared knowledge.

9. The method of claim 8, wherein the degree of certainty for the at least one hypothesis is based on one or more of the identified contextual signifiers and/or qualifiers, the inferred contextual signifiers and/or qualifiers, and/or contextual signifiers and/or qualifiers necessary for generating a response.

10. The method of claim 1, wherein generating the response includes adapting contextual signifiers and/or grammar in the response to reflect contextual signifiers and/or grammar in the at least one utterance.

11. The method of claim 1, wherein generating the response includes framing the response to influence a subsequent utterance.

12. The method of claim 1, further comprising:

learning that the determined intended meaning is incorrect;

further updating the shared knowledge to include knowledge that the determined intended meaning is incorrect; and

determining a second intended meaning for the at least one utterance based on the further updated shared knowledge, wherein the incorrect determined intended meaning is excluded as a possible second intended meaning.

13. The method of claim 1, wherein the input is a multi-modal input.

14. The method of claim 1, wherein the response is a multi-modal response.

15. A computer readable medium containing computer-executable instructions for providing a voice user interface that generates cooperative conversational responses to user utterances, the responses being based on shared knowledge about a current conversation and/or one or more past conversations, the computer-executable

Amazon/VB Assets
Exhibit 1003
Page 608

instructions operable when executed to:

receive an input, the input including at least one utterance;

update the shared knowledge to include knowledge about the at least one utterance;

determine an intended meaning for the at least one utterance based on the updated shared knowledge; and

generate a response based on the determined intended meaning.

16. The computer readable medium of claim 15, wherein the computer-executable instructions are operable to update the shared knowledge by populating a short-term context stack with information about the at least one utterance.

17. The computer readable medium of claim 16, wherein the computer-executable instructions are operable to expire the information about the at least one utterance in the short-term context stack after a predetermined amount of time.

18. The computer readable medium of claim 15, wherein the computer-executable instructions are operable to update the shared knowledge by updating one or more long-term profiles based on information about the at least one utterance.

19. The computer readable medium of claim 15, wherein the computer-executable instructions are operable to determine the intended meaning by identifying a conversation type for the at least one utterance, the conversation type based on one or more of a goal for the utterance, one or more participant roles for the utterance, and/or an allocation of information for the utterance.

20. The computer readable medium of claim 15, wherein the computer-executable instructions are operable to determine the intended meaning by generating at least one hypothesis of an intended meaning for the at least one utterance, the at least one hypothesis including a degree of certainty.

21. The computer-readable medium of claim 20, wherein the computer-executable instructions are operable to generate the at least one hypothesis by identifying contextual signifiers and/or qualifiers in the at least one utterance.

Amazon/VB Assets
Exhibit 1003
Page 609

22. The computer readable medium of claim 21, wherein the computer-executable instructions are operable to infer missing contextual signifiers and/or qualifiers for the at least one utterance based on the shared knowledge.

23. The computer readable medium of claim 22, wherein the degree of certainty for the at least one hypothesis is based on one or more of the identified contextual signifiers and/or qualifiers, the inferred contextual signifiers and/or qualifiers, and/or contextual signifiers and/or qualifiers necessary for generating a response.

24. The computer readable medium of claim 15, wherein the computer-executable instructions are operable to generate the response by adapting contextual signifiers and/or grammar in the response to reflect contextual signifiers and/or grammar in the at least one utterance.

25. The computer readable medium of claim 15, wherein the computer-executable instructions are operable to generate the response by framing the response to influence a subsequent utterance.

26. The computer readable medium of claim 15, wherein the computer-executable instructions are further operable when executed to:

learn that the determined intended meaning is incorrect;

further update the shared knowledge to include knowledge that the determined intended meaning is incorrect; and

determine a second intended meaning for the at least one utterance based on the further updated shared knowledge, wherein the incorrect determined intended meaning is excluded as a possible second intended meaning.

27. The computer readable medium of claim 15, wherein the input is a multi-modal input.

28. The computer readable medium of claim 15, wherein the response is a multi-modal response.

Amazon/VB Assets
Exhibit 1003
Page 610

29. A system for providing a voice user interface that generates cooperative conversational responses to user utterances, the system including at least one memory for storing shared knowledge about a current conversation and/or one or more past conversations, the system comprising:

at least one input device that receives an input, the input including at least one utterance; and

one or more processors collectively operable to:

update the shared knowledge to include knowledge about the at least one utterance;

determine an intended meaning for the at least one utterance based on the updated shared knowledge; and

generate a response based on the determined intended meaning.

30. The system of claim 29, wherein the one or more processors are operable to update the shared knowledge by populating a short-term context stack with information about the at least one utterance.

31. The system of claim 30, wherein the one or more processors are operable to expire the information about the at least one utterance in the short-term context stack after a predetermined amount of time.

32. The system of claim 29, wherein the one or more processors are operable to update the shared knowledge by updating one or more long-term profiles based on information about the at least one utterance.

33. The system of claim 29, wherein the one or more processors are operable to determine the intended meaning by identifying a conversation type for the at least one utterance, the conversation type based on one or more of a goal for the utterance, one or more participant roles for the utterance, and/or an allocation of information for the utterance.

34. The system of claim 29, wherein the one or more processors are operable to

Amazon/VB Assets
Exhibit 1003
Page 611

determine the intended meaning by generating at least one hypothesis of an intended meaning for the at least one utterance, the at least one hypothesis including a degree of certainty.

35. The system of claim 34, wherein the one or more processors are operable to generate the at least one hypothesis by identifying contextual signifiers and/or qualifiers in the at least one utterance.

36. The system of claim 35, wherein the one or more processors are operable to infer missing contextual signifiers and/or qualifiers for the at least one utterance based on the shared knowledge.

37. The system of claim 36, wherein the degree of certainty for the at least one hypothesis is based on one or more of the identified contextual signifiers and/or qualifiers, the inferred contextual signifiers and/or qualifiers, and/or contextual signifiers and/or qualifiers necessary for generating a response.

38. The system of claim 29, wherein the one or more processors are operable to generate the response by adapting contextual signifiers and/or grammar in the response to reflect contextual signifiers and/or grammar in the at least one utterance.

39. The system of claim 29, wherein the one or more processors are operable to generate the response by framing the response to influence a subsequent utterance.

40. The system of claim 29, wherein the one or more processors are further operable to:

learn that the determined intended meaning is incorrect;

further update the shared knowledge to include knowledge that the determined intended meaning is incorrect; and

determine a second intended meaning for the at least one utterance based on the further updated shared knowledge, wherein the incorrect determined intended meaning is excluded as a possible second intended meaning.

Amazon/VB Assets
Exhibit 1003
Page 612

41. The system of claim 29, wherein the input is a multi-modal input.

42. The system of claim 29, wherein the response is a multi-modal response.

Amazon/VB Assets
Exhibit 1003
Page 613

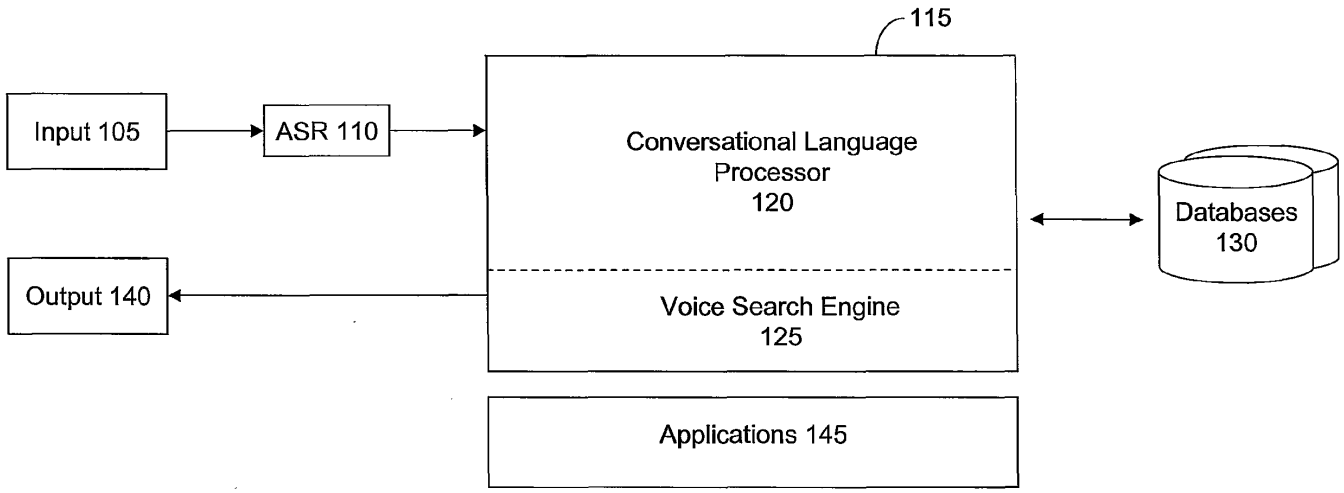

115
Input 105
ASR 110
Conversational Language Processor 120
Databases 130
Output 140
Voice Search Engine 125
Applications 145

FIG. 1

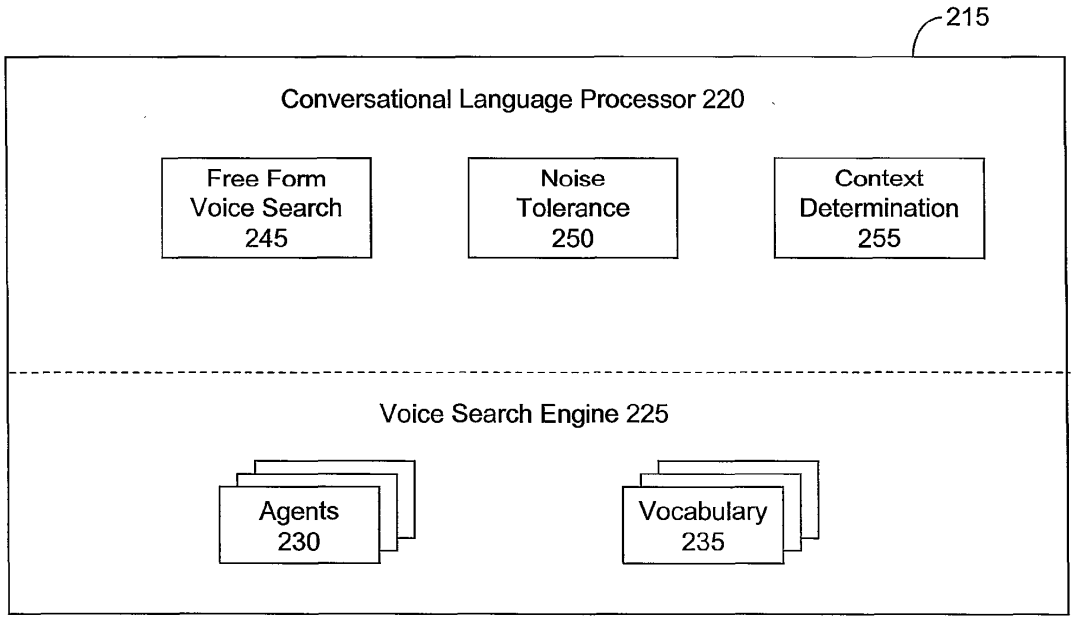

215
Conversational Language Processor 220
Free Form Voice Search 245
Noise Tolerance 250
Context Determination 255
Voice Search Engine 225
Agents 230
Vocabulary 235

FIG. 2

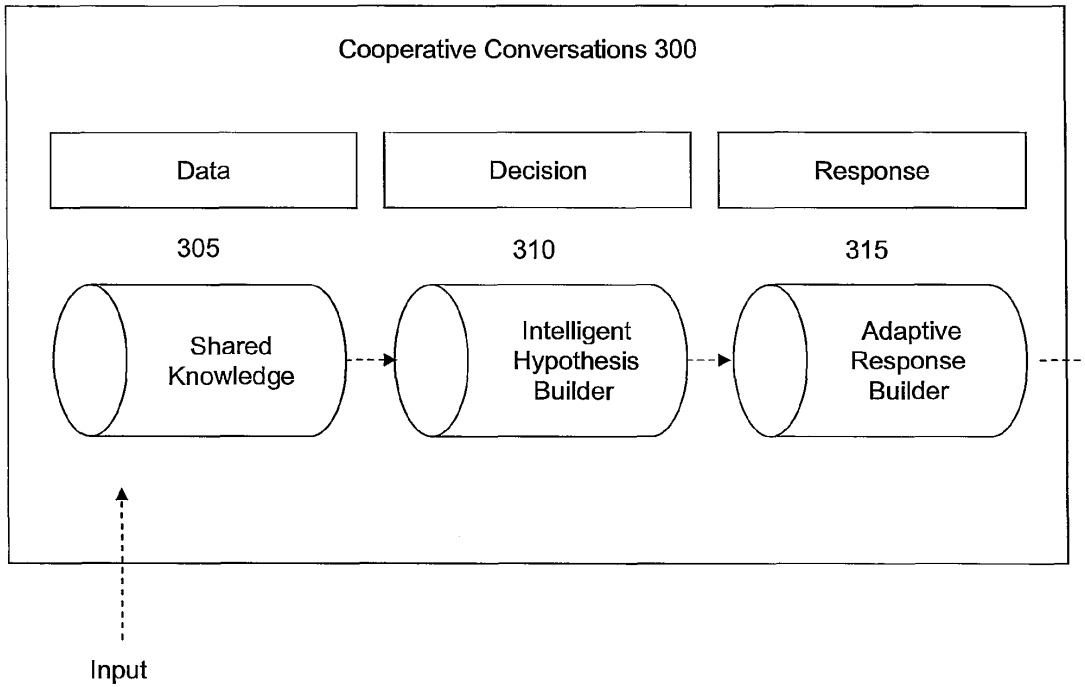


FIG. 3

Amazon/VB Assets
Exhibit 1003
Page 616

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 June 2009 (18.06.2009)

PCT

(10) International Publication Number
WO 2009/075912 A1

(51) International Patent Classification:
G06F 17/27 (2006.01)

(21) International Application Number:
PCT/US2008/069508

(22) International Filing Date: 9 July 2008 (09.07.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/954,064 11 December 2007 (11.12.2007) US

(71) Applicant *(for all designated States except US)*: **VOICEBOX TECHNOLOGIES, INC.** [US/US]; 11980 NE 24th Street, Bellevue, Washington 98005 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **KENNEWICK, Michael** [US/US]; 12196 NE 24th Street, Bellevue, Washington 98005 (US). **CHEUNG, Catherine** [US/US]; 1639 207th Avenue NE, Sammamish, Washington 98074 (US). **BALDWIN, Larry** [US/US]; 23125 SE 243rd Place, Maple Valley, Washington 98038 (US). **SALOMON, Ari** [US/US]; 4133 96th Avenue SE, Mercer Island, Washington 98040 (US). **TJALVE, Michael** [DK/US]; 14904 SE 47th Court, Bellevue, Washington 98006 (US). **GUTTIGOLI, Sheetal** [IN/US]; 11314 Ohde Circle, Kirkland, Washington 98033 (US). **ARMSTRONG, Lynn** [US/US]; 19807 183rd Place NE, Woodinville, Washington 98077 (US). **DI CRISTO, Philippe** [FR/US]; 177 107th Avenue NE, #1007, Bellevue, Washington 98004 (US). **ZIMMERMAN, Bernie** [US/US]; 135 S. 107th Street, Seattle, Washington 98168 (US). **MENAKER, Sam** [US/US]; 480 150th Place NE, Bellevue, Washington 98007 (US).

(74) Agents: **GATTO, James G.** et al.; Pillsbury Winthrop Shaw Pittman LLP, P.O. Box 10500, McLean, Virginia 22102 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),

[Continued on next page]

(54) Title: PROVIDING A NATURAL LANGUAGE VOICE USER INTERFACE IN AN INTEGRATED VOICE NAVIGATION SERVICES ENVIRONMENT

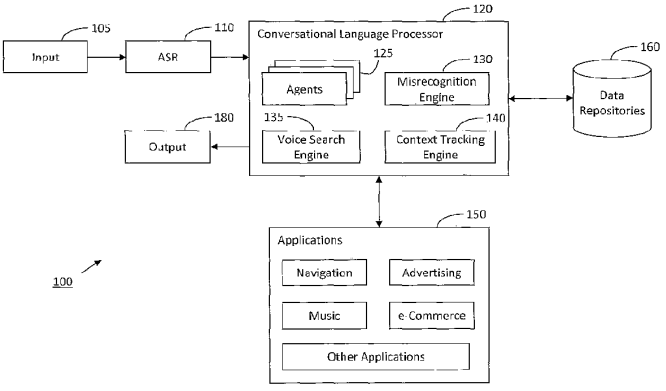


Figure 1

(57) Abstract: A conversational, natural language voice user interface may provide an integrated voice navigation services environment. The voice user interface may enable a user to make natural language requests relating to various navigation services, and further, may interact with the user in a cooperative, conversational dialogue to resolve the requests. Through dynamic awareness of context, available sources of information, domain knowledge, user behavior and preferences, and external systems and devices, among other things, the voice user interface may provide an integrated environment in which the user can speak conversationally, using natural language, to issue queries, commands, or other requests relating to the navigation services provided in the environment.

WO 2009/075912 A1

European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

PROVIDING A NATURAL LANGUAGE VOICE USER INTERFACE IN AN INTEGRATED VOICE NAVIGATION SERVICES ENVIRONMENT

FIELD OF THE INVENTION

[001] The present invention relates to a natural language voice user interface that facilitates cooperative, conversational interactions in an integrated voice navigation services environment, and in particular, to a natural language voice user interface in which users can request navigation services using conversational, natural language queries or commands.

BACKGROUND OF THE INVENTION

[002] As technology advances, consumer electronics tend to play increasingly significant roles in everyday life. As a result, users tend to expect greater functionality, mobility, and convenience from their electronic devices, as exemplified by modern mobile phones, navigation devices, personal digital assistants, portable media players, and other devices often providing a wealth of functionality beyond core applications. However, the greater functionality often tends to be accompanied by significant learning curves and other barriers that prevent users from fully exploiting device capabilities (e.g., features may often be buried within difficult to navigate menus or interfaces). Moreover, although increasing demand for mobility magnifies the need for simple on-the-go device interaction mechanisms, existing systems often have complex human to machine interfaces. For example, existing human to machine interfaces tend to primarily utilize various combinations of keyboards, keypads, point and click techniques, touch screen displays, or other interface mechanisms. However, these interfaces may often be unsuitable for mobile or vehicular devices (e.g., navigation devices), as they tend to be cumbersome in environments where speed of interaction and dangers of distraction pose significant issues. As such, existing systems often fall short in providing simple and intuitive interaction mechanisms, potentially inhibiting mass-market adoption for certain technologies. As such, there is an ever-growing demand for ways to exploit technology in intuitive ways.

[003] In response to these and other problems, various existing systems have turned to voice recognition software to simplify human to machine interactions. For example, voice recognition software can enable a user to exploit applications and features of a device that may otherwise be unfamiliar, unknown, or difficult to use. However, existing voice user interfaces, when they

Amazon/VB Assets
Exhibit 1003
Page 619

actually work, still require significant learning on the part of the user. For example, existing voice user interfaces (e.g., command and control systems) often require users to memorize syntaxes, words, phrases, or other keywords or qualifiers in order to issue queries or commands. Similarly, when users may be uncertain of exactly what to request, or what a device may be capable of, existing systems cannot engage with the user in a productive, cooperative, natural language dialogue to resolve requests and advance conversations. Instead, many existing speech interfaces force users to use predetermined commands or keywords to communicate requests in ways that systems can understand. By contrast, cognitive research on human interaction demonstrates that a person asking a question or giving a command typically relies heavily on context and shared knowledge of an answering person. Similarly, the answering person also tends to rely on the context and shared knowledge to inform what may be an appropriate response. However, existing voice user interfaces do not adequately utilize context, shared knowledge, or other similar information to provide an environment in which users and devices can cooperate to satisfy mutual goals through conversational, natural language interaction.

[004] Furthermore, demand for global positional systems and other navigation-enabled devices has grown significantly in recent years. Navigation devices often tend to be used while a user may be driving, on-the-go, or in other environments where having a hands-free interface provides critical advantages. For example, a user may want to avoid being distracted by looking away from the road, yet the user may also want to interact with a navigation device, for example, to calculate a route to a destination, recalculate the route in response to traffic, find a local restaurant, gas station, or other point of interest, or perform another navigation related task. In these and other instances, efficiently processing a natural language voice-based input could enable the user to interact with the navigation device in a safer, simpler, and more effective way. However, existing systems often fall short in providing an integrated, conversational, natural language voice user interface that can provide such advantages in navigation and other mobile environments.

[005] Existing systems suffer from these and other problems.

SUMMARY OF THE INVENTION

Amazon/VB Assets
Exhibit 1003
Page 620

[006] According to various aspects of the invention, various problems associated with existing systems may be addressed by a conversational, natural language voice user interface that provides an integrated voice navigation services environment.

[007] According to various aspects of the invention, the natural language voice user interface may resolve voice requests relating to navigation (e.g., calculating routes, identifying locations, displaying maps, etc.). The navigation application can provide a user with interactive, data-driven directions to a destination or waypoint, wherein the user can specify the destination or waypoint using free-form natural language (e.g., the user can identify full or partial destinations, including a specific address, a general vicinity, a city, a name or type of a place, a name or type of a business, a name of a person, etc.). As free form voice destination inputs may be provided in many different forms, post-processing may be performed on full or partial voice destination inputs to identify a suitable destination address for calculating a route (e.g., a closest address that makes "sense"). For example, an utterance containing a full or partial destination may be analyzed to identify one or more probable destinations (e.g., an N-best list of destinations). The N-best list may be post-processed to assign weights or rankings to the probable destinations (e.g., based on a degree of certainty that a given probable destination corresponds to an intended destination). Thus, a route may be calculated from a user's current location to a most heavily weighted one of the probable destinations in the N-best list. Further, when a voice destination entry includes a partial destination, a final destination may be successively refined over one or more subsequent voice destination entries. The navigation application may also provide dynamic, data-driven directions or routing to a destination. For instance, the navigation application may access data associated with various user-specific and environmental data sources to provide personalized data-driven directions along a route, which can be recalculated or modified based on information taken from the data sources. As such, data may be obtained dynamically to identify alternate routes, recalculate routes, or otherwise provide routing services. Further, possible answers or responses to a given utterance may be filtered according to a current route.

[008] According to various aspects of the invention, the natural language voice user interface may dynamically generate and/or load recognition grammars for interpreting what was said in an utterance (e.g., content of the utterance). Information contained in the dynamic recognition grammars may be used by a navigation agent, an Automatic Speech Recognizer, a context stack,

Amazon/VB Assets
Exhibit 1003
Page 621

or various other components in the voice user interface that use grammar information. By efficiently generating, updating, loading, extending, or otherwise building dynamic grammars based on various factors, processing bottlenecks can be avoided, conflicts can be reduced, and other aspects of interpreting an utterance using a recognition grammar can be optimized. For example, a size of a generated grammar may be constrained by an amount of resources available in a system (e.g., in embedded devices or other devices having low amounts of dynamic memory, the constrained grammar size may limit a quantity of resources to be occupied). In another example, the size of the dynamic grammar can be reduced by eliminating redundant keywords, criteria, or other information available in the context stack, the shared knowledge, or other local sources. Thus, favorability of correct interpretations may be improved by reducing perplexity in the grammar (e.g., when two or more elements may likely be confused, one or more of the elements may be eliminated to reduce confusion).

[009] According to various aspects of the invention, the natural language voice user interface may generate dynamic recognition grammars using techniques of geographical chunking. A user's location can be determined at a given moment to determine one or more geographic proximities, which can be used to form an appropriate topological domain for the grammar. For example, the topological domains may reflect physical proximities (e.g., a distance from a current location), civil organization proximities (e.g., regions, states, cities, neighborhoods, subdivisions, localities, etc.), temporal proximities (e.g., amounts of travel time from the current location), directional proximities (e.g., based on directional travel vectors), or various combinations thereof. As a result, by mapping the user's geographic proximities to one or more topological domains, dynamic grammars may be pruned, extended, swapped in or out of memory, or otherwise generated and/or loaded to provide optimal recognition based on location, time, travel, or other factors (e.g., information may be swapped in and out of a grammar as a user moves from one area to another, ensuring that system resources utilize information presently relevant to a given location).

[010] According to various aspects of the invention, the natural language voice user interface may include dynamic grammars formed from one or more topological domains, which may be subdivided into a plurality of tiles, which may be further subdivided into a plurality of subtiles. Thus, information used to build the dynamic grammar can be subdivided or weighed in various ways to determine what information should be included in the grammar. Moreover,

Amazon/VB Assets
Exhibit 1003
Page 622

geographical chunks based on physical, civil organization, temporal, directional, or other proximities may be extended into other domains in which a topological taxonomy can be placed. As a result, in addition to having relevance in navigation or other location-dependent systems, the geographical chunking techniques can be applied in other contexts or domains in which geography or location may be relevant. Further, a server operably coupled to the voice user interface may analyze various forms of information to build or refine a source of grammar information. For example, when various devices communicate with the server, information may be communicated to the server may be used to update proximities, topological domains, tiles, subtiles, peer-to-peer affinities, or other grammar information.

[011] According to various aspects of the invention, the natural language voice user interface may calculate routes, provide dynamic data-driven directions to a destination, provide dynamic routing to a destination, perform post-processing of full or partial destination entries, or otherwise provide various voice navigation services. Further, destinations and/or routes may be identified using techniques of successive refinement for voice destination entries, wherein context, agent adaptation, and shared knowledge, among other things, can help a user to narrow down a final destination using voice commands, multi-modal commands, or various combinations thereof. However, it will be apparent that the successive refinement techniques can be applied to various tasks in which generalized approximations can be successively refined through voice or multi-modal commands to narrow down information sought by a user, including various other domains, contexts, applications, devices, or other components that employ the techniques described herein.

[012] According to various aspects of the invention, the natural language voice user interface may enable successive refinement of a final destination by progressively narrowing the final destination. For example, successively refining the destination may be modeled after patterns of human interaction in which a route or a destination may be narrowed down or otherwise refined over a course of interaction. For example, a user may generally approximate a destination, which may result in a route being calculated along a preferred route to the approximated destination. While en route to the approximated destination, the user and the voice user interface may cooperatively refine the final destination through one or more subsequent interactions. Thus, a user may provide a full or partial destination input using free form natural language, for example, including voice commands and/or multi-modal commands. One or more

Amazon/VB Assets
Exhibit 1003
Page 623

interpretations of a possible destination corresponding to the voice destination input may be organized in an N-best list. The list of possible destinations may be post-processed to assign weights or ranks to one or more of the entries therein, thus determining a most likely intended destination from a full or partial voice destination input. Thus, the post-processing operation may rank or weigh possible destinations according to shared knowledge about the user, domain-specific knowledge, dialogue history, or other factors. As a result, the full or partial destination input may be analyzed to identify an address to which a route can be calculated, for example, by resolving a closest address that makes "sense" relative to the input destination. Subsequent inputs may provide additional information relating to the destination, and the weighted N-best list may be iteratively refined until the final destination can be identified through successive refinement. As a result, when a suitable final destination has been identified, the route to the final destination may be completed.

[013] According to various aspects of the invention, the natural language voice user interface may include one or more advertising models for generating and/or detecting events relating to location dependent advertisements for navigation systems (e.g., as generated by a local or remote advertising engine, or via a data channel, or in other ways). For example, navigation systems typically include various mechanisms for determining a current location (e.g., a global positioning system, a radio frequency identification system, a system that determines location based on a distance to an identifiable wireless tower or access point, etc.). The location detection system may thus detect information associated with a radio frequency identifier over a data channel used by a marketer to provide advertisements. The marketer may broadcast the advertisement via the data channel, such that the navigation system triggers an event when within a suitable proximity of the RFIDs. Thus, information associated with the event may be filtered according to the current routing information or other contextual parameters to determine what action should be taken in response thereto. In other instances, advertisements may be uploaded to a server by one or more advertising partners, wherein the uploaded advertisements may be associated with metadata or other descriptive information that identifies a target audience, location-dependent information, or other criteria. In another example, a plurality of advertisements may be stored locally at the voice user interface, and an inferencing engine may determine appropriate circumstances in which an event should be generated to deliver one or more of the advertisements to a user. As a result, it will be apparent that advertising events may

Amazon/VB Assets
Exhibit 1003
Page 624

be generated in a number of ways, and may be generated and/or detected locally, remotely, by detecting RFIDs, or in other ways.

[014] According to various aspects of the invention, the natural language voice user interface may track user interactions with delivered advertisements. In this way, affinity based models may be generated, for example, to ensure that promotions or advertisements will be delivered to a likely target audience. Thus, an event relating to a given advertisement may be generated and/or detected when shared knowledge about a user's behavior, preferences, or other characteristics match one or more criteria associated with peer-to-peer affinities associated with the advertisement. In other examples, an advertising model may include mobile pay-per-use systems, peer-to-peer local guides or recommendations, or other forms of advertising. Additionally, various aspects of the advertising model, such as the local guides and recommendations, may be generated according to a mapping applied to various topological domains. For example, certain types of advertisements may be dependent on geographic or topological characteristics, and such advertisements may be associated with a topological taxonomy based on geographical chunks. As a result, various advertising events may be generated and/or detected according to physical proximities, temporal proximities, directional proximities, civil organization proximities, or various combinations thereof.

[015] According to various aspects of the invention, the natural language voice user interface may enable a user to provide requests (e.g., queries, commands, or other requests) to the navigation device using natural language. As such, the user and the navigation device may engage in a cooperative, conversational dialogue to resolve the request. For example, the voice user interface may use prior context, dialogue histories, domain knowledge, short and long-term shared knowledge relating to user behavior and preferences, noise tolerance, and cognitive models, among various other things, to provide an integrated environment in which users can speak conversationally, using natural language, to issue queries, commands, or other requests that can be understood and processed by a machine. Accordingly, the voice user interface may understand free form human utterances, freeing the user from restrictions relating to how commands, queries, or other types of requests should be formulated. Instead, the user can use a natural or casual manner of speaking to request various voice services in an integrated environment, in which various devices can be controlled in a conversational manner, using natural language. For example, the voice user interface may be aware of data and services

Amazon/VB Assets
Exhibit 1003
Page 625

associated with the navigation device, a media device, a personal computer, a personal digital assistant, a mobile phone, or various other computing devices or systems available in the environment.

[016] According to various aspects of the invention, the natural language voice user interface may include an input mechanism that receives a voice-based input, which includes at least an utterance or verbalization spoken by a user. The input mechanism may include a suitable device or combination of devices that can receive voice-based inputs (e.g., a directional microphone, an array of microphones, or other devices that encode speech). The input mechanism can be optimized to maximize gain in a direction of a user, cancel echoes, null point noise sources, perform variable rate sampling, filter out background conversations or environmental noise, or otherwise optimize fidelity of encoded speech. As such, the input mechanism may generate encoded speech generated in a manner that tolerates noise or other factors that could otherwise interfere with interpreting speech. Further, in various implementations, the input mechanism may include one or more other (non-voice) input modalities, which can be processed and/or correlated with one or more previous, current, or subsequent utterances or other voice-based inputs. As such, a user can provide other forms of input using a touch-screen interface, a stylus/tablet interface, a keypad or keyboard, or other input interfaces, for example, to clarify utterances or provide additional information about the utterances using other input modalities. For instance, the user could touch a stylus or other pointing device to a portion of a map displayed on a touch-screen interface, while also providing an utterance relating to the touched portion (e.g., "Show me restaurants around here."). In this example, the inputs can be correlated to interpret "around here" as likely referring to the touched portion of the map, as distinct from the user's current location or some other meaning.

[017] According to various aspects of the invention, the natural language voice user interface may include an Automatic Speech Recognizer that processes encoded speech to generate one or more preliminary interpretations of what was said in an utterance (e.g., content of the utterance). For example, the Automatic Speech Recognizer may generate the preliminary interpretations using phonetic dictation to recognize a stream of phonemes based on a dynamically adaptable recognition grammar. The dynamically adaptable recognition grammar may be based on dictionaries or phrases from various input domains (e.g., domains for languages, navigation, music, movies, weather, various temporal or geographic proximities, or various other domains).

Amazon/VB Assets
Exhibit 1003
Page 626

Thus, the Automatic Speech Recognizer may generate one or more interpretations of an utterance, which may be represented as a series of phonemes or syllables. The one or more interpretations can be analyzed (e.g., utilizing phonotactic rules or models of human speech) to generate an N-best list of preliminary interpretations as to what was spoken by the user. The preliminary interpretations may then be provided to a conversational language processor, which utilizes shared knowledge, contextual information, and various other sources of information to generate an intelligent hypothesis as to an actual meaning, the user's intent, or other aspects of the utterance. By formulating the hypothesis using various features and components that model everyday human-to-human conversations, the conversational language processor may generate a hypothesis as to the meaning or intent of the utterance, which can inform a process of resolving one or more requests contained in the utterance.

[018] According to various aspects of the invention, the natural language voice user interface may include, among other things, a context tracking engine that establishes meaning for a given utterance. For example, the context tracking engine can manage competitions among one or more context-specific domain agents that establish the meaning (e.g., redistributable, adaptable engines or modules that provide functionality for a given context, domain, system, or application). The domain agents may analyze preliminary interpretations of an utterance to generate a domain-specific probable interpretation. For example, one or more of the agents may include adaptable vocabularies, concepts, available tasks, or other forms of information specific to the respective domain or context. In addition, the agents can use a voice search engine to search a network for information that may not be available within the system. The probable interpretation can be assigned a weighted ranking or score, which can be used to select a "winning" one of the agents. Thus, the winning one of the agents may be designated responsible for establishing or inferring further information (e.g., based on domain or context-specific information), updating the shared knowledge, or resolving requests in the utterance, among other things. The context tracking engine may also maintain a context stack to track conversation topics, track previously invoked agents, evaluate criteria, weigh parameters, or otherwise maintain information relating to a conversational context (e.g., the context stack may be traversed in light of recent contexts, frequently used contexts, or other information included therein to determine a most likely intent of the user). By identifying a context, or correlatively, by identifying capabilities, tasks, vocabularies, or other information within the context, the

Amazon/VB Assets
Exhibit 1003
Page 627

context tracking engine can provide relevant information for establishing intent in addition to phonetic clues associated with the utterance (e.g., a word with multiple possible meanings may be disambiguated based on its meaning in a given context, previous usage in a dialogue, etc.).

[019] According to various aspects of the invention, the natural language voice user interface may utilize various forms of information to enable sharing of assumptions and expectations relating to a given utterance, conversation, or other human to machine interaction. For example, to inform decision making in the voice user interface, the voice user interface may include information sources containing short-term and long-term shared knowledge relating to user behavior, preferences, or other characteristics (e.g., short-term and long-term profiles of a specific user, global users, peer users, etc.). The short-term shared knowledge may accumulate information during a current conversation to dynamically establish awareness of a state of the voice user interface (e.g., recognition text for previous utterances, a cross-modal user interface manipulation history, a list of previously selected tasks or invoked queries, or other information). Storage of the short-term knowledge may be modeled after human interaction, such that certain information may be expired after a psychologically appropriate amount of time (e.g., to expunge stale data), whereas information with long-term significance can be added to the long-term shared knowledge (e.g., to establish persistent awareness of data likely to remain static over time). As such, the long-term shared knowledge may profile or otherwise model various characteristics, preferences, behavior, or other information relating to a user based on information accumulated over time (e.g., user-specific jargon, demographics, cognitive patterns, frequently requested tasks, favorite topics or concepts, etc.). Accordingly, the voice user interface may utilize various forms of information available via a context tracking engine, domain agents, a voice search engine, shared knowledge, internal or external databases, data associated with other devices, or other knowledge sources. As a result, a conversational type or goal associated with utterance can be identified. Based on the available information and the type or goal of the conversation, the voice user interface may attempt to resolve the utterance (e.g., by invoking an agent, which utilizes one or more applications to perform a requested task, retrieve requested information, etc.).

[020] According to various aspects of the invention, the natural language voice user interface may generate cross-modal intelligent responses sensitive to syntax, grammar, and context, which may provide a conversational feel to system-generated responses. When available, a generated

Amazon/VB Assets
Exhibit 1003
Page 628

intelligent response may present results of a resolved request (e.g., feedback relating to a task, information retrieved as a result of a query, etc.). Further, the intelligent responses may be provided across modalities, for example, using various combinations of verbal and/or non-verbal outputs (e.g., information may be presented on a display device, via an audible alert, a verbalized output, another output mechanism, or various combinations thereof). Further, verbalized components of the cross-modal response may be adapted to the user's manner of speaking (e.g., by varying tone, pace, timing, or other variables), thus creating verbal responses having natural variation and personality. The intelligent responses may also be formulated to provide an output that guides the user towards a subsequent response (e.g., a subsequent utterance) that may be more likely to be recognized. For example, when an utterance cannot be resolved because of an ambiguous context, unrecognizable words or phrases, or other factors that can result in an ambiguous or unrecognized interpretation, the intelligent response can be framed to disambiguate context or request additional information from the user to clarify a meaning of the utterance. Further, when subsequent information indicates that a given interpretation or hypothesis was incorrect, one or more previous utterances may be reinterpreted to refine context and update short-term or long-term conversational models (e.g., an utterance of "No, I meant . . ." may indicate that a previous utterance was interpreted incorrectly, in which case various previous utterances may be reinterpreted based on a correct interpretation, thus building a more accurate conversational context). As a result, the voice user interface may tolerate full or partial failure through adaptive mechanisms.

[021] According to various aspects of the invention, the natural language voice user interface may provide voice navigation services within an agent-based architecture. The architecture may include a plurality of adaptable agents (e.g., specialized software, data, content, or other information that provide functionality, behavior, services, data, and other information in a plurality of respective contextual domains), at least one of which provides navigation services (e.g., route calculation, map control, location-sensitive information, data-driven directions, etc.). As the agents process requests, the agents may autonomously react, adapt, and otherwise reconfigure to provide optimal voice services in respective domains. For example, by building context over time (e.g., by generating short and long-term profiles of a user, conversations with the user, frequent topics or preferences, etc.), the agents may automatically combine knowledge, adapt preferences, remove conflicts, or perform other adaptations to refine or otherwise optimize

Amazon/VB Assets
Exhibit 1003
Page 629

an operational framework thereof. Adaptation of the agents, which include at least a navigation agent, may occur across the plurality of agents, for example, in response to various ones of the agents resolving voice-based requests. As such, the adaptation may occur autonomously as a by-product of the agents providing voice services, generating inferences, identifying affinities (e.g., among users, peers, communities, etc.), receiving updates from external sources (e.g., using an update manager), or in other ways, as will be apparent.

[022] According to various aspects of the invention, the natural language voice user interface may include an agent-based architecture for providing voice navigation services. For example, the agent-based architecture may include one or more domain or context-specific agents, which include at least a navigation agent. The navigation agent may include, among other things, various navigation-specific content packages (e.g., dictionaries, available queries, tasks, commands, dynamic grammars, etc.), response lists (e.g., appropriate responses to commands, queries, or other requests), personality profiles (e.g., for creating a natural feel to system-generated speech), substitution lists (e.g., for substituting or transforming data into a structured form that can be understood by a target information source), or various other forms of navigation-specific information. Further, the navigation agent may be associated with pointers to local or remote data sources, parameters and operating data provided associated with other services in the architecture, or various other forms of information. For example, the data sources used by the navigation agent may include data relating to navigation, points-of-interest, traffic, events, parking, personal data, peer affinities, or various other sources of information. Further, the data sources may be populated, extended, pruned, or otherwise constructed through adaptation, analysis of various models, communication with a data service, or in other ways, as will be apparent.

[023] According to various aspects of the invention, the natural language voice user interface may include a navigation agent, which may be coupled with various sources of information, and may make use of context, communicating with various other adaptable agents and other system components to provide voice navigation services. For example, the navigation agent may use contextual information relating to a navigation domain, including tracked topics, user locations, routes traveled, previous requests, user interface states, user behaviors, preferences, demographics, or other characteristics, or various other types of contextual information. As a result, the navigation agent may have various sources of knowledge and resources available to

Amazon/VB Assets
Exhibit 1003
Page 630

resolve voice navigation requests. For example, the navigation agent may generate inferences (e.g., using an inference engine) using the available knowledge and resources to apply various rules, policies, or other inferencing techniques to generate interpretations of an utterance (e.g., phonetic fuzzy matching, inductive logic, Bayesian probability analysis, monotonic or non-monotonic reasoning, etc.). As such, the navigation agent can infer keywords or criteria not explicitly provided in the utterance, determine suitable responses to subjective or indeterminate utterances, generate events, identify peer affinities, or otherwise generate inferences for resolving navigation-related requests.

[024] According to various aspects of the invention, the natural language voice user interface may include one or more inference engines, which can generate various inferences through awareness of previous context, short-term or long-term shared knowledge, command histories, states of vehicular systems, user interface states, and various other data sources. In various implementations, one or more of the agents may be associated with a respective inference engine that can generate inferences using domain-specific knowledge, rules, policies, or other criteria. For instance, the inference engines may identify keywords or criteria missing in an utterance, infer intended meanings, autonomously suggest available tasks, or otherwise assist an associated agent in identifying queries, commands, or other requests contained in an utterance. Moreover, when information cannot be suitably resolved using information sources associated with the navigation agent, or through generating inferences, the information may be requested from one or more other agents, other devices, network information sources (e.g., via a voice search engine), or in other ways, as will be apparent. Upon identifying the information through one or more of the other sources, the requesting agent may be adapted to make the information subsequently available. Thus, various devices, applications, systems, and other components of an architecture may cooperatively share available information and services (e.g., context, dialogue histories, shared knowledge, maps, points of interest, contact lists, user or peer affinities, dynamic grammars, available applications, command histories, etc.). Accordingly, the architecture may provide an integrated voice navigation services environment in which users can speak natural language requests relating to various available contexts, domains, applications, devices, information sources, or various combinations thereof.

[025] According to various aspects of the invention, the natural language voice user interface may accept natural language voice-based inputs to control an electronic device that can provide

Amazon/VB Assets
Exhibit 1003
Page 631

navigational information, in addition to various other devices associated with an environment in which the voice user interface operates. Furthermore, various functional aspects of the voice user interface may reside at a client device, at a server, or various combinations thereof.

[026] According to various aspects of the invention, the natural language voice user interface may support multi-modal voice inputs. Thus, a given multi-modal voice input may include at least a voice component (e.g., an utterance) in addition to one or more non-voice input components (e.g., inputs provided via a keypad, a touch-screen, a stylus/tablet combination, a mouse, a keyboard, or other input modalities). As such, the non-voice input component can provide additional information or clarification relating to the utterance, adding to an amount of input information available when processing voice. For example, the user can use other input modalities to clarify a meaning of the utterance, provide additional information about the utterance, reduce a number of device interactions needed to make a given request, or otherwise provide additional in relation to a given utterance.

[027] According to various aspects of the invention, the natural language voice user interface may utilize various of cognitive models, contextual models, user-specific models, or other models to identify queries, commands, or other requests in a voice input. For example, a given input may include information relating to one or more contextual domains, one or more of which may be invoked to interpret and/or infer keywords, concepts, or other information contained in the input. Moreover, short-term and long-term shared knowledge about a user's behavior and preferences may be used in a hybrid recognition model that also considers semantic analysis and contextual inferences. For example, certain syllables, words, phrases, requests, queries, commands, or other information may be more likely to occur in a given context. Thus, the hybrid recognition model may analyze semantic patterns to resolve what was said by an utterance, and may further rely on contextual history or other information to resolve what was meant by the utterance. The hybrid recognition model may be used in conjunction with, or independently of, a peer to peer recognition model. For example, recognition models may include awareness of global usage patterns, preferences, or other characteristics of peer users, where certain keywords, concepts, queries, commands, or other aspects of a contextual framework may employed by peer users within the context.

[028] Other objects and advantages of the invention will be apparent based on the following drawings and detailed description.

Amazon/VB Assets
Exhibit 1003
Page 632

BRIEF DESCRIPTION OF THE DRAWINGS

[029] Figure 1 illustrates a block diagram of an exemplary system for providing a navigation device with a conversational, natural language voice user interface according to various aspects of the invention.

[030] Figure 2 illustrates a block diagram of exemplary agent-based architecture for providing a navigation device with a conversational, natural language voice user interface according to various aspects of the invention.

[031] Figure 3 illustrates a flow diagram of an exemplary method for dynamically generating recognition grammars for a navigation device having a conversational, natural language voice user interface according to various aspects of the invention.

[032] Figure 4 illustrates a flow diagram of an exemplary method for processing multi-modal voice-based inputs provided to a navigation device having a conversational, natural language voice user interface according to various aspects of the invention.

[033] Figure 5 illustrates a flow diagram of an exemplary method for calculating routes based on voice-based inputs provided to a navigation device having a conversational, natural language voice user interface according to various aspects of the invention.

[034] Figure 6 illustrates a flow diagram of an exemplary method for providing voice services based on a current navigation route according to various aspects of the invention.

DETAILED DESCRIPTION

[035] According to various aspects of the invention, Figure 1 illustrates a block diagram of an exemplary system 100 for providing a navigation device with a conversational, natural language voice user interface. As illustrated in Figure 1, the system 100 may enable a user to engage in a natural language, cooperative, conversational dialogue with a voice-enabled navigation device. For example, the system 100 may understand free form human utterances, freeing the user from being restricted in how commands, queries, or other types of requests should be formulated. Instead, the user can use a casual or natural manner of speaking to request various voice navigation services, among various other voice services (e.g., services relating to telematics, communications, media, messaging, external systems, marketing, information retrieval, or various other computational services). As such, the user can use system 100 to control

Amazon/VB Assets
Exhibit 1003
Page 633

navigation devices, media devices, personal computers, personal digital assistants, wireless phones, or other computing devices or systems in a conversational manner, using natural language. By making use of context, prior information, domain knowledge, short and long-term shared knowledge, noise tolerance, and cognitive models, among other things, system 100 can provide an integrated environment in which users can speak conversationally, using natural language, to issue queries, commands, or other requests that can be understood and processed by a machine. For example, in various implementations, system 100 may utilize one or more techniques as described in co-pending U.S. Patent Application 10/452,147, entitled "Systems and Methods for Responding to Natural Language Speech Utterance," filed June 3, 2003, and co-pending U.S. Patent Application 10/618,633, entitled "Mobile Systems and Methods for Responding to Natural Language Speech Utterance," filed June 15, 2003, the disclosure of which are hereby incorporated by reference in their entirety.

[036] According to various aspects of the invention, system 100 may include an input mechanism 105 that receives a voice-based input, including at least one utterance or verbalization spoken by a user. The input mechanism 105 may include a device or combination of devices suitable for receiving a voice-based input (e.g., a directional microphone, an array of microphones, or other devices that encode speech). In various implementations, the input mechanism 105 can be optimized to receive the voice-based input, for example, by maximizing gain in a direction of a user, cancelling echoes, nulling point noise sources, performing variable rate sampling, filtering out background conversations or environmental noise, or performing various other techniques for maximizing fidelity of encoded speech. As such, the encoded speech generated by input mechanism 105 may be highly tolerant of noise or other factors that may potentially interfere with interpreting human speech.

[037] Further, in various implementations, the input mechanism 105 may be coupled to other input modalities, in that various forms of input other than voice can be processed and/or correlated with one or more previous, current, or subsequent voice-based inputs. For example, the input mechanism 105 may be coupled to a touch-screen interface, a stylus/tablet interface, a keypad or keyboard, or other devices or system interfaces, as will be apparent. As a result, an amount of input information potentially available for system 100 to process voice may be maximized, as the user can clarify utterances or provide additional information about utterances using other input modalities. For instance, the user could touch a stylus or other pointing device

Amazon/VB Assets
Exhibit 1003
Page 634

to a portion of a map displayed on a touch-screen interface, while also providing an utterance relating to the touched portion (e.g., "Show me restaurants around here."). In this example, system 100 may correlate the inputs to interpret "around here" as likely referring to the touched portion of the map, as distinct from the user's current location or another meaning.

[038] System 100 may also include an Automatic Speech Recognizer 110 that receives the encoded voice input and generates one or more preliminary interpretations thereof. For example, the Automatic Speech Recognizer 110 may recognize the voice-based input using phonetic dictation to recognize a stream of phonemes based on a dynamically adaptable recognition grammar. As a result, the Automatic Speech Recognizer 110 may provide out-of-vocabulary capabilities, which may be tolerant of a user misspeaking, portions of a speech signal being dropped, or other factors that could interfere with interpreting an utterance. The dynamically adaptable recognition grammar may be based on dictionaries or phrases from various input domains (e.g., domains for different languages, navigation, music, movies, weather, various temporal or geographic proximities, or various other domains). Further, performance of the Automatic Speech Recognizer 110 may be improved, for example, by pruning a search space associated with the recognition grammar (e.g., the grammar can include a linking element, such as *schwa*, to represent an unstressed, central vowel that tends to be spoken frequently, even without the user's conscious awareness). Thus, using these and other techniques, Automatic Speech Recognizer 110 may analyze an incoming encoded utterance to represent portions of the utterance as a series of phonemes or syllables, which can be further broken down into core components of an onset, a nucleus, and a coda, among other sub-categories. The series of phonemes or syllables can then be analyzed (e.g., utilizing phonotactic rules that model human speech) to identify a plurality of preliminary interpretations or best guesses (e.g., an N-best list) as to what was actually spoken by the user. It will be apparent, however, that the Automatic Speech Recognizer 110 may use various techniques to generate the preliminary interpretations of the encoded utterance, including those described, for example, in co-pending U.S. Patent Application Ser. No. 11/513,269, entitled "Dynamic Speech Sharpening," filed August 31, 2006, the disclosure of which is hereby incorporated by reference in its entirety.

[039] The plurality of preliminary interpretations generated by the Automatic Speech Recognizer 110 may be provided to a conversational language processor 120, which utilizes shared knowledge to generate an intelligent hypothesis of an actual meaning, a user's intent, or

Amazon/VB Assets
Exhibit 1003
Page 635

other aspect of the voice-based input. The conversational language processor 120 may formulate the hypothesis using various features and components that collectively operate to model everyday human-to-human conversations.

[040] For example, the conversational language processor 120 may include a context tracking engine 140 that establishes meaning for a given utterance by, among other things, managing a competition among one or more context-specific domain agents 125 (e.g., redistributable, adaptable packages or modules that provide functionality for a given context, domain, system, or application). For example, the preliminary interpretations may be ranked by the Automatic Speech Recognizer 110, and the agents 125 may further analyze the preliminary interpretations to generate a weighted ranking or score, which can be used to select a "winning" one of the agents 125. The winning one of the agents 125 may then be responsible for establishing or inferring further information, updating the shared knowledge, or performing other tasks to aid in generating the intelligent hypothesis. Moreover, the context tracking engine 140 may use a context stack to track conversation topics, evaluate criteria, weigh parameters, or otherwise maintain contextual information for generating the hypothesis based on a conversational context (e.g., a context stack may be traversed in light of recent contexts, frequently used contexts, or other information included therein to determine a most likely intent of the user). By identifying a context, or correlatively, by identifying capabilities, tasks, vocabularies, or other information within the context, the context tracking engine 140 can provide highly relevant information for establishing intent, apart from meager phonetic clues.

[041] Furthermore, the conversational language processor 120 may utilize various other forms of knowledge to inform the generation of the intelligent hypothesis. For example, various agents 125 may adaptively include domain-specific or context-specific vocabularies, concepts, available tasks, or other forms of information relevant to the respective domain or context. In addition, the various components associated with the conversational language processor 120 can invoke a voice search engine 135 (e.g., an engine that searches a network for information) to resolve information that may not be internally available (e.g., when an external knowledge source may be useful in clarifying an intent behind a particular word, command, query subject, or other aspect of an utterance).

[042] System 100 may also enable the user and the system 100 to share assumptions and expectations relating to a given utterance, conversation, or other interaction. For example, the

Amazon/VB Assets
Exhibit 1003
Page 636

conversational language processor 120 may be coupled to one or more data repositories 160 that store short-term and long-term shared knowledge that inform decision making in the conversational language processor 120. The short-term shared knowledge may accumulate information during a current conversation (e.g., recognition text for previous utterances, a cross-modal user interface manipulation history, a list of previously selected tasks or invoked queries, or other information), thus dynamically establishing awareness of a cross-modal state of the voice user interface. Storage of the short-term knowledge may be modeled after human interaction, thus certain data may be expired after a psychologically appropriate amount of time (e.g., to expunge stale data), and information with long-term significance can be added to long-term shared knowledge (e.g., a new address of a long-term contact of the user). As such, the long-term shared knowledge may profile or otherwise model environmental, cognitive, historical, demographic, or other aspects of a user based on information accumulated over time (e.g., user-specific jargon, frequently requested tasks, favorite topics or concepts, etc.).

[043] Accordingly, the conversational language processor 120 includes various features that can be used to generate intelligent hypotheses as to a user's intent in a given utterance. For example, the hypothesis may be based on information provided via the context tracking engine 140, the agents 125, the voice search engine 135, the shared knowledge, or other knowledge sources. As a result, the conversational language processor 120 may attempt to identify a conversational type or goal of the utterance (e.g., a query for retrieving a discrete piece of information, a didactic interaction for clarifying information provided by the voice user interface, or an exploratory interaction in which conversational goals may be improvised as the conversation progresses, etc.). Based on the information available and the type or goal of the conversation, the generated hypothesis can be assigned a degree of certainty, which can inform how conversational language processor 120 resolves the utterance. For example, when the degree of certainty indicates that sufficient information has been identified, the conversational language processor may invoke one or more of the agents 125, which in turn can utilize one or more applications 150 to perform a requested task (e.g., a task relating to a navigation application, an advertising application, a music application, an electronic commerce application, or other suitable applications or tasks). In another example, one or more of the agents 125 may query the data repositories 160 or the voice search engine 135 to retrieve requested information, or otherwise take action to resolve a request for information contained in the utterance.

Amazon/VB Assets
Exhibit 1003
Page 637

[044] Additionally, the conversational language processor 120 may generate cross-modal intelligent responses, which may be syntactically, grammatically, and contextually sensitive, thus providing a conversational feel to system-generated responses. When available, the intelligent response may present results of a performed task or an executed query to the user, and the response may be provided across modalities. For example, verbal and/or non-verbal outputs 180 may be used separately or in concert (e.g., by presenting information using a display device, an audible alert, a verbalized output, another output mechanism, or various combinations thereof). Further, a verbalized component of the cross-modal output 180 may be adapted to the user's manner of speaking (e.g., by varying tone, pace, timing, or other variables), thus creating verbal responses having natural variation and personality.

[045] The intelligent responses may also be formulated to provide an output 180 that guides the user towards a subsequent response likely to be favorable for recognition. For example, when the degree of certainty reflects an ambiguous context (e.g., when the competition results in a deadlock among the agents 125), an adaptive misrecognition engine 130 may identify the utterance as ambiguous or unrecognized, and the intelligent response can be framed to disambiguate context, or request a subsequent response from the user to clarify the meaning of the utterance, for example. Further, when the conversational language processor 120 or the misrecognition engine 130 determines that a given interpretation or hypothesis was incorrect, one or more previous utterances may be reinterpreted to refine context and build more accurate short-term or long-term conversational models (e.g., an utterance of "No, I meant . . ." may indicate that a previous utterance was interpreted incorrectly, in which case various previous utterances may be reinterpreted based on a correct interpretation, thus building a more accurate conversational context). As a result, the conversational language processor 120 may recover from full or partial failure. Additional techniques for adaptively responding to misrecognition or ambiguity may be included, such as those described in co-pending U.S. Patent Application Ser. No. 11/200,164, entitled "System and Method of Supporting Adaptive Misrecognition in Conversational Speech," filed August 10, 2005, the disclosure of which is hereby incorporated by reference in its entirety.

[046] Additional information relating to the various techniques described herein, as well as other techniques that system 100 uses to provide a conversational, natural language interaction may be provided, for example, in co-pending U.S. Patent Application Ser. No. 11/197,504,

Amazon/VB Assets
Exhibit 1003
Page 638

entitled "Systems and Methods for Responding to Natural Language Speech Utterance," filed August 5, 2005, co-pending U.S. Patent Application Ser. No. 11/212,693, entitled "Mobile Systems and Methods of Supporting Natural Language Human-Machine Interactions," filed August 29, 2005, and co-pending U.S. Patent Application Ser. No. 11/580,926, entitled "System and Method for a Cooperative Conversational Voice User Interface," filed October 16, 2006, the disclosures of which are hereby incorporated by reference in their entirety.

[047] As such, system 100 may provide an environment where conversational, natural language interactions can occur between a user and the system 100. Further, as will be described in greater below, system 100 can be implemented to provide the conversational, natural language interactions, for example, as a voice user interface to one or more electronic devices or applications, including a navigation device. As will also be apparent from the further descriptions provided herein, the system 100 has capabilities for cross-device and cross-modal awareness, such that system 100 provides an environment in which the user can engage in a cooperative, natural language dialogue to request voice services relating to many different devices or applications.

[048] According to various aspects of the invention, Figure 2 illustrates a block diagram of exemplary agent-based architecture 200, which can provide a conversational, natural language voice user interface to a navigation device. The architecture 200 may include a plurality of agents 225a-n, which include specialized software, data, content, or other information that provide functionality, behavior, services, data, and other information in a plurality of respective contextual domains. The architecture 200 may be an integrated and dynamically adaptable environment, in which the agents 225 autonomously react, adapt, and reconfigure to provide optimal service in the environment. For example, by building context over time (e.g., by generating short and long-term profiles of a user, conversations with the user, frequent topics or preferences, etc.), the navigation agent 225a may automatically combine knowledge, adapt preferences, remove conflicts, or perform other adaptations to refine or otherwise optimize an operational framework thereof.

[049] For example, the agents 225, which include at least a navigation agent 225a, may adapt through ongoing use of short-term shared knowledge 270 and long-term shared knowledge 275 about a user's behavior, preferences, or other characteristics. Further, agent adaptation may occur across the plurality of agents 225, for example, in response to various ones of the agents

Amazon/VB Assets
Exhibit 1003
Page 639

225 resolving voice-based requests (e.g., by invoking one or more applications 250, querying one or more data sources 260, cooperating with one or more other agents 225, or learning through a misrecognition engine 230, a voice search engine 235, a context tracking engine 24, etc.). In various implementations, adaptation may occur autonomously as a by-product of the agents 225 providing voice services, generating inferences, identifying affinities (e.g., among users, peers, communities, etc.), receiving updates from external sources (e.g., an update manager can update content, behavior, data, commands, domain knowledge, keywords, concepts, dictionaries, or other information for one or more of the agents 225), or in other ways, as will be apparent.

[050] As illustrated in Figure 2, the architecture 200 provides the conversational, natural language voice user interface to the navigation device by way of at least a navigation agent 225a. The navigation agent 225a can include, among other things, various navigation-specific content packages (e.g., dictionaries, available queries, tasks, commands, dynamic grammars, etc.), response lists (e.g., appropriate responses to commands, queries, or other requests), personality profiles (e.g., for creating a natural feel to system-generated speech), substitution lists (e.g., for substituting or transforming data into a structured form that can be understood by a target information source), or various other forms of navigation-specific information. Further, the navigation agent 225a may be associated with parameters and operating data provided by other services in the architecture 200 (e.g., available devices, applications 250, agents 225, etc.), pointers to local or remote data sources (e.g., short-term shared knowledge 270, long-term shared knowledge 275, data sources 260, voice search engine 235, etc.), among various other forms of information.

[051] The data sources 260 used by the navigation agent 225a may include, among other things, data relating to navigation (e.g., maps, addresses, street names, directories, alternate routes, etc.), points-of-interest (e.g., restaurants, hotels, museums, tourist attractions, gas stations, etc.), traffic (e.g., dynamic road conditions, traffic, detours, or other traffic-related information), events (e.g., sporting events, concerts, protests, etc.), parking (e.g., parking garages, restricted areas or times, street lots, street parking, etc.), personal data (e.g., telephone numbers and addresses associated with a user's contact lists), peer affinities (e.g., recommendations based on similarities among users having similar preferences, demographic profiles, or other characteristics). The data

Amazon/VB Assets
Exhibit 1003
Page 640

sources 260 may be populated in various ways, such as being based on one or models, received via a data service, extended or refined through adaptation, or in other ways, as will be apparent.

[052] Further, contextual information relating to a navigation domain may be maintained via the context tracking engine 240, short-term shared knowledge 270, and long-term shared knowledge 275, among other things. For example, the contextual information may relate to tracked topics, user location, routes traveled, previous requests, user interface state, user behavior, preferences, demographics, or other characteristics, among other types of contextual information. For example, context may be built and utilized using techniques described in greater detail above with reference to Figure 1, or in the above-referenced co-pending U.S. Patent Applications, or various combinations thereof. Moreover, when navigation agent 225a cannot find information locally within architecture 200, and the information cannot be inferred, the information may be requested from one or more other agents 225b-n (e.g., to request domain-specific information), other devices, the voice search engine 235 (e.g., to search a network), external sources (e.g., the update manager may be invoked to download available updates, which may have the information), or other sources, as will be apparent. If the information can be located through one or more of the other sources, the navigation agent 225a may be adapted, extended, or otherwise updated to make the information available subsequently.

[053] Accordingly, the navigation agent 225a may be coupled with various sources of information, and may make use of context, while communicating with various other adaptable agents 225b-n and other system components to provide voice navigation services. Thus, context, shared knowledge, maps, points of interest, contact lists, user or peer affinities, dynamic grammars, available applications, or various other aspects of the navigation architecture 200 can be shared or otherwise made available to various components, devices, or aspects of the architecture 200. As a result, an inference engine 255 included with the navigation agent 225a may various knowledge sources and other resources available to provide an integrated voice navigation services environment. For example, the inference engine 255 may generate inferences from the available knowledge and resources by applying one or more rules, policies, or other inferencing mechanisms to generate probable interpretations of utterances in the environment (e.g., phonetic fuzzy matching, inductive logic, Bayesian probability analysis, monotonic or non-monotonic reasoning, etc.).

Amazon/VB Assets
Exhibit 1003
Page 641

[054] For instance, as described in greater detail above, the navigation agent 225a and one or more other agents 225b-n may compete in analyzing one or more preliminary interpretations of an utterance to generate one or more respective probable interpretations of the utterance. The navigation agent 225a may then be invoked to process the probable interpretation (e.g., by resolving information contained in the utterance) upon being identified as a "winning" one of the competing agents. For example, which of the competing agents wins may depend on context or other information contained in the utterance, whereby the navigation agent 225a may win and recalculate a route in response to an utterance of "This traffic is terrible, maybe back roads," while a music agent may win and change a radio channel in response to an utterance of "Traffic is terrible, maybe some Oldies," while a movie agent may win and search for movie show times in response to an utterance of "Traffic was terrible, maybe a comedy." Thus, when the navigation agent 225a generates a probable interpretation that results in being a winning one of the agents 225, the navigation agent 225a may manage processed for searching, inferring, or otherwise identifying information in light of context to provide voice navigation services using one or more applications 250 (e.g., navigation, grammar generation, location determination, response generation, advertising, device control, or other applications available within architecture 200).

[055] For example, the navigation agent 225a may utilize the inference engine 255 to infer keywords or criteria not explicitly provided in the utterance, determine suitable responses to subjective or indeterminate utterances (e.g., selecting a most likely answer to a query, asking a user for more information, etc.), generate events, identify peer affinities, or otherwise generate inferences for resolving navigation-related requests. The inference engine 255 may generate such inferences through awareness of previous context (e.g., through information provided by the context tracking engine 240), short-term or long-term shared knowledge 270, command histories, states of vehicular systems, user interfaces, or other devices, data sources 260, or other available information. For example, in an exemplary illustration, previous context, shared knowledge, and other information sources may be utilized in conjunction with the inference engine 255 and/or various other components of architecture 200 to enable a human to machine interaction that may occur as follows:

User:	"Show me restaurants around here."

Amazon/VB Assets
Exhibit 1003
Page 642

Voice User Interface:	"All I can find is fast food, which you don't like."
User:	"Okay, how about closer to the hotel?"
Voice User Interface:	"Found an Indian restaurant with good reviews. Want me to call them?"
User:	"Yeah go ahead, that sounds good."

[056] In the above illustrated example, the first utterance establishes a context of "restaurants," and the voice user interface responds to the utterance in light of a shared knowledge inference indicating that the user does not like fast food. Furthermore, while the second utterance does not explicitly reference a restaurant, the context established in the first utterance can be used to infer the unreferenced information, forming an unambiguous request out of what would otherwise be ambiguous. Further still, long-term shared knowledge 275 about the user (e.g., the user may frequently go to Indian restaurants) and short-term knowledge 270 about the user (e.g., a hotel associated with a current route) can narrow a search space for resolving the second utterance. Moreover, the voice user interface may invoke the response generation application to advance the conversation towards a resolution (e.g., by prompting the user to call the restaurant). Because the integrated environment has cross-device awareness, the user's subsequent response can result in the device control application formulating an appropriate command structure to dial the user's mobile phone, which may or may not be distinct from the navigation device. Accordingly, as illustrated in this example, robust awareness of context, shared knowledge, available applications, and available systems or devices, among other things, has enabled the user to use natural language to locate a restaurant along a current route matching the user's preferences, while also enabling the user to dial a phone to call the restaurant, without the user having to manually interact with devices or otherwise interact with cumbersome human to machine interfaces.

[057] It will be understood, however, that the above illustrated dialogue provides but one example of a cooperative conversation that can occur between a user and a system that includes the above-described architecture 200. Because the architecture 200 supports free form interactions, and can tolerate serendipitous variations in human speech, vocabulary usage, or other factors, it will be understood that no two conversational interactions between the user and the system will necessarily be alike. As such, the user can seamlessly switch among contexts,

Amazon/VB Assets
Exhibit 1003
Page 643

allowing subsequent utterances to occur in various available domains (e.g., a navigation-related utterance of "Take me to the Seahawks game" may be followed up with subsequent cross-domain utterances, such as "Are there tickets left?" or "When does it start?"). Thus, although various descriptions illustrate exemplary ways in which the agent-based architecture 200 can be used for voice navigation services, it will be apparent that, in a given utterance, the user can provide natural language voice-based requests relating to various available contexts, domains, applications, devices, information sources, or various combinations thereof.

[058] The navigation application available within architecture 200 may resolve natural language, voice-based requests relating to navigation (e.g., calculating routes, identifying locations, displaying maps, etc.). The navigation application can provide a user with interactive, data-driven directions to a destination or waypoint, wherein the user can specify the destination or waypoint using free-form natural language (e.g., the user can identify full or partial destinations, including a specific address, a general vicinity, a city, a name or type of a place, a name or type of a business, a name of a person, etc.). As a result, in various implementations, the navigation application may perform post-processing on the full or partial destinations to identify a suitable address for calculating a route (e.g., a closest address that makes "sense"). For example, an utterance containing a full or partial destination may be analyzed to identify one or more probable destinations (e.g., an N-best list of destinations), which may be subject to post-processing that weighs the identified probable destinations based on various factors. Thus, the probable destination most likely to provide a suitable (preliminary or complete) route may be identified from the N-best list.

[059] For example, in an illustrative conversation between a user and the voice user interface, the user may utter a voice-based input of "Take me to Seattle." Based on a current location of the user (e.g., as determined using one or more navigation sensors, radio frequency identifiers, local or remote map databases, etc.), the navigation application may select an address in Seattle that provides a route from the current location to Seattle (e.g., a central point in Seattle may be selected for long distance travel, whereas an address in northern Seattle may be selected for a current location being close to north Seattle). Further, when the requested destination does not specify a final destination (e.g., as in the above example), the user may successively refine the final destination in subsequent requests, and the post-processing may continue to select additional addresses based on the current location, a current route, or other factors. For instance,

Amazon/VB Assets
Exhibit 1003
Page 644

continuing the above-provided example, a subsequent input of "Take me to Pike Street" may result in the post-processing analyzing the current route to determine an appropriate address on Pike Street, and possibly recalculating the route, as necessary. Thus, the current route may have the user driving north-bound on Interstate-5, such that an address may be selected on Pike Street on a same side as north-bound lanes of Interstate-5 (e.g., preserving current routing information). The final destination may continue to be successively refined, in combination with the post-processing. As a result, a user knowing only that a destination falls along a 500-block of Pike Street can specify, "501 Pike Street," and even if such an address does not exist, the post-processing may find a closest address that makes sense, thus routing the user to the 500-block of Pike Street.

[060] In addition, the navigation application may provide dynamic, data-driven directions to the destination or dynamically provide routing information, among other things. Furthermore, the response generation application may work together with the navigation application and/or one or more other applications to generate system responses personalized for the user. For instance, the navigation application may access data associated with various user-specific and environmental data sources to provide personalized data-driven directions along a route, which can be recalculated or modified based on information taken from the data sources. Thus, in one example, the response generation application may generate personalized data-driven directions as follows:

Voice User Interface:	"Go to Bob's House."
Voice User Interface:	"Take a left on 14th Street."
Voice User Interface:	"Make a right here to get on the highway, and prepare for heavy traffic."

[061] In the above-illustrated example, the response has been personalized by framing directions in terms of one of the user's contacts (i.e., Bob), while an available source of traffic data has been used to provide data-driven directions. In another example, data that may affect routing can be identified and used to recalculate a route or identify alternate routes (e.g., weather data may be used to avoid snowy roads, or event data may be used to avoid a route along which a protest may be occurring). As such, data may be obtained dynamically to identify alternate

Amazon/VB Assets
Exhibit 1003
Page 645

routes, recalculate routes, or otherwise provide optimal routing service. In various implementations, the source of data may be based on persistent models, dynamically received data (e.g., a given highway may regularly have high volume during rush hour or sporting events), or other information. Routes may be recalculated automatically (e.g., user preferences may dictate that heavy traffic be avoided, or sporting event traffic may not be avoided when the user has tickets to the event), or the user can request recalculation through additional utterances (e.g., "I don't want to sit in traffic, can we go a different way?"). Further, possible answers or responses to a given utterance may be filtered according to a current route (e.g., possible results provided for an utterance of "Find me a gas station" may be filtered according to those gas stations within a given proximity of the route). Moreover, continuing the above-provided example, the user may later return to a point of origin, and personalized responses may include return directions, such as "Go back the way you came on 14th Street." Thus, the personalized responses may also create a natural feel to the directions, and may be based on context that builds over time.

[062] The navigation architecture 200 may be used in conjunction with an electronic device capable of accepting voice-based inputs and providing navigation-related information. For example, architecture 200 may be implemented within a handheld device (e.g., personal navigation device, mobile phone, etc.), a telematics devices (e.g., a vehicular navigation system), or various other suitable devices. Furthermore, various functional aspects of the architecture 200 may reside at a client device, at a server, or various combinations thereof. For example, a voice-based input may be received by the client device (e.g., a personal navigation device), which may communicate with the server to retrieve information, perform a task, process the input, or for other reasons. The server may subsequently return information to the client device in response to the request. In various implementations, however, the voice-based input may be processed within the system (e.g., processing delays may be reduced by minimizing communication with the server). Other suitable arrangements may be used, as will be apparent, and specific arrangements as to how information can be processed, retrieved, or returned to the user, among other things, may be highly dependent on context or specific implementations (e.g., what information can be obtained locally, remotely, or otherwise, what services a user subscribes to, etc.).

Amazon/VB Assets
Exhibit 1003
Page 646

[063] According to various aspects of the invention, Figure 3 illustrates a flow diagram of an exemplary method 300 for dynamically generating and/or loading recognition grammars in a conversational, natural language voice user interface. For example, a grammar generator application (e.g., as illustrated in Figure 2) may dynamically generate and/or load recognition grammars for use by a navigation agent, an Automatic Speech Recognizer, a context stack, or other components in the voice user interface that incorporate a recognition grammar. While grammar generating method 300 may be used by various components of the natural language voice user interface, the method 300 may be particularly useful in reducing processing overhead and reducing perplexity or confusion. For example, by efficiently generating, updating, loading, extending, or otherwise building dynamic grammars based on various factors, as will described in greater detail herein, bottlenecks can be avoided, conflicts can be reduced, or other aspects of interpreting utterances using a grammar can be optimized.

[064] For example, in various implementations, grammar generation may be keyed to an amount of available resources in a system that includes the voice user interface. Thus, at an operation 310, the amount of system available resources may be calculated, and a subsequent operation 320 may constrain a size of the dynamic grammar based on the calculated resources. For instance, in embedded devices or other devices having low amounts of dynamic memory, the constrained grammar size may ensure that only as many resources as necessary will be occupied. Thus, throughput in the low-memory device will not be degraded by a large-list grammar over-occupying processing capacity. In another example, the size of the dynamic grammar can be reduced by eliminating redundant keywords, criteria, or other information available in the context stack, the shared knowledge, or other local sources. However, it will be apparent that various other techniques may be used to constrain grammar size, and that effectively considering available system resources can provide processing improvements in various environments.

[065] In a navigational context, in which a user's location can change from one moment to another, grammar generation may be further optimized using techniques of geographical chunking. For instance, a user's location can be determined at a given moment (e.g., via navigation sensors, global positioning systems, etc.), and an operation 330 can use the determined location to determine one or more geographic proximities associated with the location. Subsequently, at an operation 340, the determined proximities may be used to identify one or more topological domains appropriate for the determined proximities.

Amazon/VB Assets
Exhibit 1003
Page 647

[066] As such, the proximities determined in operation 330 may be weighed or analyzed in various ways to determine how to utilize the topological domains identified in operation 340 in generating the dynamic grammar. For example, the topological domains may reflect physical proximities (e.g., distances from a current location, such as everything within a five-mile radius), civil organization proximities (e.g., regions, states, cities, neighborhoods, subdivisions, localities, etc.), temporal proximities (e.g., amounts of travel time from the current location, where road classes or other information may be used to estimate fastest/safest routes, such as highways and interstates being faster than city or surface roads), directional proximities (e.g., based on directional travel vectors, where a point ten miles ahead may be considered "closer" than a point one mile behind based on a direction of travel, an amount of time to turn around, etc.), or various combinations thereof. As a result, by mapping the user's geographic proximities to one or more topological domains, dynamic grammars may be pruned, extended, swapped in or out of memory, or otherwise generated and/or loaded to provide optimal recognition based on location, time, travel, or other factors (e.g., information may be swapped in and out of a grammar as a user drives from one area to another, optimizing the information along the drive to ensure that the system resources utilize information presently relevant for a given location).

[067] Furthermore, the topological domains may be subdivided into a plurality of tiles, which may in turn be subdivided into a plurality of subtiles. For example, a civil organization topological domain may include a tile representing grammar information for a state, and the tile may include one or more subtiles for counties within the state. Further, the county subtiles may include subtiles for various cities, neighborhoods, or other organizational boundaries within respective counties. Thus, the geographic proximities and topological domains used to build the dynamic grammar may be based on a variety of factors, and can be subdivided or weighed in various ways to determine what information to include in the grammar. Moreover, geographical chunks based on physical, civil organization, temporal, directional, or other proximities may be extended into various other domains in which a topological taxonomy can be placed. As a result, the geographical chunking techniques may have particular relevance in navigation or other location-dependent voice recognition systems, yet the techniques can be suitably applied in various contexts or domains in which geography or location may have relevance.

[068] As various implementations generate the recognition grammars dynamically, for example, to preserve system resources, information available to extend, update, or otherwise

Amazon/VB Assets
Exhibit 1003
Page 648

update the grammar may be stored remotely (e.g., on a server). Thus, various implementations may include a system having a network connection, a data service, or another communication mechanism for establishing a link to the remote source. Based on the proximities and/or the topological domains identified for a given location, context, utterance, or otherwise, one or more grammar tiles and/or subtiles may be downloaded from the remote source at an operation 350. Further, the remote source may be adapted to store, analyze, or otherwise process certain information to refine the grammar information stored therein. For example, when a given device communicates with the remote source at operation 350, the remote source may receive identifying information relating to a user of the device (e.g., user preferences, user characteristics, request characteristics, etc.), requested information, or other information. Based on the received information, the remote source may dynamically update various tiles and/or subtiles, build affinities (e.g., among users, peers, communities, etc.), or perform other actions to refine a process by which relevant information can be identified (e.g., the remote source may determine that various different users, who share certain common demographics, may be interested in particular events, locations, etc.).

[069] Thus, the grammar information downloaded at operation 350 may be based on a variety of factors, including information in a current voice-based input, current geographical proximities of the user, topological domains or geographical chunks related to the current geographical proximities, or various kinds of affinities, among other things. Based on these and other factors, appropriate information may be retrieved from various local or remote data sources (e.g., a local repository may contain a great deal of grammar information, a portion of which may be used by the dynamic grammar to interpret an utterance). The retrieved information may then be used to generate the dynamic grammars at an operation 360. As a result, the dynamic grammars may be adapted to include information favorable for correctly interpreting a given utterance. Furthermore, at an operation 370, the favorability of correct interpretations may further be improved by post-processing the generated grammars using one or more optimization techniques, such as reducing perplexity in the grammar. For example, when the grammar generated at operation 360 includes two or more elements likely to be confused (e.g., two or more streets having a same name), the optimization operation 370 may eliminate one or more of the elements to reduce confusion (e.g., selecting the element closest to a user's current location, or other criteria). In another example, when an interpretation turns out to be incorrect (e.g., a

Amazon/VB Assets
Exhibit 1003
Page 649

subsequent utterance may be "No, 14th Street Northwest, not Northeast"), the dynamic grammar may be pruned to eliminate information associated with the incorrect interpretation (e.g., information associated with Northeast may be removed). As a result, the optimization operation 370 can reduce overhead, assure that mistakes will not be repeated, and/or improve accuracy of interpretations, among other things.

[070] Although various exemplary operations have been described for dynamically extending, pruning, loading, or otherwise generating recognition grammars, it will be apparent that in operation, the natural language voice user interface may use various forms of information or techniques to determine an optimal grammar. For instance, in addition to being based on available system resources, likelihood of perplexity, geographical proximities, or user affinities, the grammar may be generated in response to conversational contexts, shared knowledge about a user, and/or adaptable agent behavior, among other factors. Accordingly, the dynamic grammar can be structured to recognize queries, commands, or other information in an utterance based on information most likely to be relevant for the utterance, with minimal ambiguity. Further, when the utterance cannot be resolved using the information currently in the grammar, various information sources or techniques may be available to update the grammar in a way that will be favorable for recognition. Thus, the above-described aspects and features of the grammar generation method 300 should be regarded as exemplary only, as many variations will be possible without departing from the inventive concepts described.

[071] According to various aspects of the invention, Figure 4 illustrates a flow diagram of an exemplary method 400 for processing multi-modal voice inputs using a conversational, natural language voice user interface. As described in connection with Figure 1 above, method 400 may be performed in a voice user interface having a mechanism for receiving voice inputs (e.g., a directional microphone, an array of microphones, or other devices that encode speech), wherein the voice inputs include at least a voice component (e.g., an utterance). The utterance, which may be received in an operation 410, may be provided using natural language or other free form manner of expression.

[072] Furthermore, to enable processing of multi-modal voice inputs, the voice user interface may have one or more input mechanisms for receiving non-voice input components (e.g., inputs provided via a keypad, a touch-screen, a stylus/tablet combination, a mouse, a keyboard, or other input modalities). As such, in operation 410, the received voice input may include a non-voice

Amazon/VB Assets
Exhibit 1003
Page 650

input component in addition to the utterance, such that an amount of input information may be maximized through usage of multi-modal inputs. For example, the user can use other input modalities to clarify a meaning of the utterance, provide additional information about the utterance, reduce a number of device interactions needed to make a given request, or otherwise maximize an amount of information provided in connection with a given utterance. For example, a multi-modal voice input may be used to efficiently interact with a map displayed on a navigation device (e.g., an utterance of "I only need to see this area" may be coupled with a stylus input that circles a given portion of a map, and the combination of inputs may result in a display of the circled portion of the map).

[073] Thus, for a given voice input, a decisional operation 420 may determine whether a non-voice component accompanies an utterance in the voice input. For example, in various implementations, the various input mechanisms associated with the voice user interface may be coupled to an Alternating Voice and Data (AVD) system. As a result, input data may alternate between voice and data, or the input data may be multiplexed together in a single stream, for example, to allocate processing resources where needed at a given moment.

[074] When decisional operation 420 determines that a non-voice input component has been received, the voice and non-voice components may be correlated in an operation 430. For example, the voice component may be processed using accurate speech recognition techniques, and the voice component may further be parsed and interpreted in light of context, shared knowledge, and available data sources, among other things, as described in greater detail above. As such, correlation operation 430 may generate additional information that can inform context, shared knowledge, or other information for processing the voice input. For instance, as discussed above, short-term shared knowledge may include awareness of various user interface states, which may be affected by non-voice components. Thus, correlation operation 430 may assist the voice user interface in developing one or more preliminary interpretations of the utterance, for example, by relating information in the utterance to information in the non-voice inputs, or vice versa. For example, a multi-modal input could include a user touching a stylus to one of a number of restaurant listings displayed on a touch-screen interface, while also providing an utterance relating to the touched listing (e.g., "Call this one."). In this example, correlation operation 430 can interpret "this one" to refer to the touched telephone listing, providing additional information for processing the voice input.

Amazon/VB Assets
Exhibit 1003
Page 651

[075] However, it will be apparent that voice inputs need not necessarily include non-voice components. As such, when the voice input includes only utterances, verbalizations, or other spoken components, decisional operation 420 may advance method 400 directly to an operation 440. For instance, as will be apparent from the further descriptions provided elsewhere in this specification, the natural language voice user interface may efficiently resolve many requests that include only voice. As such, a received voice input can include stand-alone utterances, such that a navigation device, vehicular system, mobile phone, or other device can be controlled in one step using voice. For example, using existing navigation systems or other navigation interfaces, a user must often take multiple steps to configure a displayed map, even for relatively simple tasks (e.g., multiple button and/or stylus inputs may be required to display a desired portion of a map). By contrast, using the natural language voice user interface, the user can simplify controlling a device using one-step voice control (e.g., a map may be controlled using utterances of "Move the map up," "Move the map Northeast," "Show me downtown Seattle," or other utterances that provide all information necessary to control a device in one-step), which may substantially reduce a number of device interactions needed to control the device.

[076] In various instances, (e.g., whether the input received in operation 410 includes an utterance, a non-voice component, or various combinations thereof), one or more queries and/or commands contained in the voice input may be identified in operation 440. Once the one or more preliminary interpretations of the multi-modal input have been generated, appropriate contextual information, shared knowledge, external knowledge sources, or other information may be used to identify the queries and/or the commands contained in the utterance. The information sources relied upon may be associated with various devices, systems, or other components of the integrated voice user interface environment. Thus, information associated with a navigation device (e.g., maps, points-of-interest, etc.) may be distributed for sharing with other devices in the environment (e.g., a personal digital assistant capable of rendering maps may display a map associated with a navigation device, or a mobile phone may dial a telephone number of a point-of-interest associated with the navigation device, etc.). Moreover, the information may be shared in a bi-directional manner, such that information sources associated various devices can be shared with other devices in the integrated environment.

[077] Information may be associated persistently (e.g., by storing all information on a server, an online data backup system, or other data repositories), or may be built dynamically as a

Amazon/VB Assets
Exhibit 1003
Page 652

function of processing multi-modal voice inputs. For example, a user may maintain an address book of contacts, e-mail addresses, street addresses, telephone numbers, or other information using a commercial service provider (e.g., GMail, Yahoo Mail, or another service provider capable of storing address books). When a user makes a request (e.g., via voice input, non-voice input, or multi-modal input) in which the service provider manages information needed to resolve the request, the information may be pulled and shared with a device that will be used to resolve the request. For example, a mobile phone may be integrated with the navigation voice user interface (e.g., a mobile phone having a GPS receiver), and the mobile phone may initially (e.g., at a first-time use) have no local contacts, addresses, telephone numbers, or other information. Thus, a user utterance of "I'm running late for dinner at Jane's house, call her to let her know I'm on the way" may result in a query being formulated to the service provider in order to retrieve an address and telephone number associated with Jane's house for the mobile phone to process accordingly. Moreover, information relating to being "late" may be processed to generate a route that avoids traffic, uses highways, or otherwise reduces an amount of travel time.

[078] A variety of cognitive models, contextual models, user-specific models, or other models may be used to identify the query and/or command in operation 440. For example, a given input may include information relating to one or more contextual domains, one or more of which may be invoked to interpret and/or infer certain keywords, concepts, or other information contained in the input. Moreover, short-term and long-term shared knowledge about a user's behavior and preferences may be used in a hybrid recognition model, for example, relating to navigation queries and/or commands. The hybrid recognition model may be used to generate probable or likely interpretations of an utterance using a combination of semantic analysis and contextual inferences. For example, certain syllables, words, phrases, requests, queries, commands, or other identifiable aspects of an utterance may be identified as being more likely to occur based on a given contextual history. The contextual history can be based on the short-term or the long-term shared knowledge about the user, previous utterances in a current conversation, common requests in a given environment (e.g., a navigation device may be used most often for displaying maps, calculating routes, and identifying locations), or other information. Thus, the hybrid recognition model for operation 440 may include various processes for analyzing semantic

Amazon/VB Assets
Exhibit 1003
Page 653

patterns to resolve what was said by an utterance, in addition to various processes for relying on contextual history to resolve what was meant by the utterance.

[079] In addition, the hybrid recognition model may be used in conjunction with, or independently of, a peer to peer recognition model. For example, contextual histories may include various preferences or user characteristics, in addition to providing a basis for inferring information about a user based on patterns of usage or behavior, among others. As a result, the recognition models may include additional awareness relating to global usage patterns, preferences, or other characteristics of peer users on a global basis. For example, certain keywords, concepts, queries, commands, or other aspects of a contextual framework may be commonly employed by all users within a context. In another example, users of certain demographics may share common jargon, slang, or other semantic speech patterns. As a result, operation 440 may utilize various recognition models, which consider context and semantics in various dimensions, to identify queries or command. For example, in addition to information generally available within a given environment or context, a voice input may be recognized using context and semantics based on short-term or long-term behavior or preferences for a specific user, global users, peer users, or other meaningful user abstractions.

[080] Accordingly, upon identifying a suitable query and/or command contained in the utterance, a suitable response may be generated, or a task may be performed in an operation 450. As will be apparent from the descriptions already provided, the generated response and/or the performed task may be highly dependent on context and information contained in a given voice input. As a result, when certain voice inputs may be unclear, ambiguous, or otherwise unclear, the generated response may prompt a user for an additional input, and the response may be framed in a way favorable for recognition (e.g., an utterance of "Where is Joe's Pizza" provided while a user is in New York City may return too many results to be meaningful, such that results may be weighed and displayed for a user to select a correct one of the results).

[081] Furthermore, voice inputs may be used to perform compound requests, which could otherwise be impossible to perform using a single manual input. For example, a single voice-based input may include a compound map control request, such as "Show me downtown Seattle." As such, operation 450 may perform tasks of retrieving a map of Seattle and automatically zooming in on a downtown area of the Seattle map. Further, one or more responses or other outputs may be generated, such as suggesting a route to a frequent destination,

Amazon/VB Assets
Exhibit 1003
Page 654

identifying possible points-of-interest to the user, or searching for traffic or event notifications, among other things. Many other variations will be apparent, including characteristics of the received inputs (e.g., utterances, non-voice inputs, etc.), the requested queries or commands, the generated responses, and the performed tasks, among other things. As such, it will be apparent that the method 400 illustrated herein may enable users to request many different navigation related tasks verbally, non-verbally, or various combinations thereof, such that the users can request various kinds of information or tasks available in the environment (e.g., relating to various devices, information sources, etc.). Thus, method 400 may operate in an environment in which the natural language voice user interface has been associated with one or more devices that provide navigation services.

[082] According to various aspects of the invention, Figure 5 illustrates a flow diagram of an exemplary method for calculating navigation routes from voice destination entries provided to a navigation device having a conversational, natural language voice user interface. As previously described, the voice user interface may calculate routes, provide dynamic data-driven directions to a destination, provide dynamic routing to a destination, perform post-processing of full or partial destination entries, or otherwise provide various voice navigation services. Thus, as illustrated in Figure 5, the method for calculating routes may support successive refinement of a voice destination entry, wherein context, agent adaptation, and shared knowledge, among other things, can help a user to narrow down a final destination using voice commands, multi-modal commands, or various combinations thereof. However, though Figure 5 specifically illustrates an exemplary implementation of successive refinement for voice destination entry, it will be apparent that the techniques described herein can be applied in performing various tasks in which generalized approximations can suitably be refined through successive voice or multi-modal commands to narrow down information sought by a user in various domains, contexts, applications, devices, or other components that employ the techniques described herein. As a result, successive refinement can be implemented in various domains that enable a user to "drill down" to a specific piece of information or data through specifying more and more specific criteria about the information sought.

[083] As illustrated in Figure 5, for example, successive refinement of a final destination may includes instances in which a user specifies the destination in a single utterance, over a plurality of utterances (e.g., which sequentially narrow down approximations of the destination), as well

Amazon/VB Assets
Exhibit 1003
Page 655

as instances in which a system identifies the destination (e.g., through generating inferences, using context or shared knowledge, etc.). For example, a given approximation of a final destination may be associated with a finite number of possible destinations, and the system may be able to unambiguously identify the final destination from among the finite possible destinations by generating inferences, by relying on context, shared knowledge, and other information, or by other ways, as will be apparent. Moreover, in another example, successively refining the destination may be modeled after patterns of human interaction between passengers in a taxicab, or other similar situations in which a route or a destination may be narrowed down or otherwise refined over a course of interaction. For example, passengers in taxicabs sometimes specify a general approximation of a destination (e.g., a passenger picked up in Arlington, Virginia may state, "I'm going to the Foggy Bottom area of D.C."), which may result in a cab driver beginning to drive along a preferred route to the approximated destination (e.g., the driver may know to head towards Rock Creek Parkway when driving from Arlington to Foggy Bottom). While en route to the approximated destination, the passenger and/or the driver may cooperate to refine the final destination through one or more subsequent interactions (e.g., the cab driver may ask, "Where in Foggy Bottom are you headed?", and the passenger may specify "McFadden's," which may provide sufficient information for the driver to determine a route to the destination).

[084] However, it will be apparent that no two sequences of interactions will necessarily be alike. For example, the passenger may provide additional information relating to the final destination without being prompted by the driver, or the driver may not begin driving until the final destination has been identified (e.g., when several routes may be available, and the final destination has an impact on which route will be selected). As such, cooperative, conversational interactions may have a general goal of calculating a route to a final destination, yet the interactions may occur in many different ways, where a given interaction will likely be highly dependent on context, shared knowledge, agent adaptation, points of origin and/or presence, user preferences, dynamic data (e.g., relating to traffic, events, external systems, etc.), among many other factors.

[085] Accordingly, as further described herein, the method illustrated in Figure 5 may enable voice destination entry through successive refinement using the aforementioned factors, among other things, to inform route calculation based on full or partial destination inputs. For instance,

Amazon/VB Assets
Exhibit 1003
Page 656

in similar fashion as described in the above cab driver example, the voice user interface (e.g., the cab driver) may have knowledge that certain routes may be preferable to reach a certain destination from a current point of presence (e.g., Arlington to Foggy Bottom via Rock Creek Parkway). Further, as will be described in greater detail in Figure 6, a search space associated with subsequent interactions may be constrained according to a current route, context, or other knowledge (e.g., locations named "McFadden's" that fall outside a certain proximity of Foggy Bottom may be excluded).

[086] Thus, in an operation 510, a user may provide a full or partial destination input using free form natural language, for example, including voice commands and/or multi-modal commands (e.g., an input may include an utterance of "I'm going here," coupled with a touched point on a display). The full or partial destination may be specified in various ways, including by specific address, place name, person's name, business name, neighborhood, city, region, or various other ways (e.g., a voice destination entry may be provided in an exploratory manner, such as when a user wants to visit a museum, but has yet to decide which one to visit). Furthermore, because successive refinement tolerates incomplete or partial addresses, voice destination inputs may be provided in a way that uses the successive refinement in reverse, for example, by narrowly specifying a desired destination in order to calculate a route to a broader vicinity of the desired destination (e.g., a route to South Philadelphia could be identified through an utterance of "Pat's Steaks in Philly," even though Pat's Steaks may not be a user's actual destination).

[087] The voice destination input may be parsed or otherwise analyzed using one or more dynamically adaptable recognition grammars, for example, as described above in reference to Figure 3. For example, recognition grammars may be loaded, generated, extended, pruned, or otherwise adapted based on various factors, including a proximity to a user's point of presence (e.g., as the user moves from one area to another, the recognition grammar may be optimized based on a current location, a direction of travel, temporal constraints, etc.), a contextual history (e.g., as the user interacts with the voice user interface, the grammar may adapt based on dictionaries, keywords, concepts, or other information associated with other contexts, domains, devices, applications, etc.), or other factors, as will be apparent. As such, an operation 520 may include generating one or more interpretations of the voice destination input, which may be analyzed using various data sources in order to generate an N-best list of possible destinations (e.g., a navigation agent may query a directory, a voice search engine, or other components to

Amazon/VB Assets
Exhibit 1003
Page 657

identify one or more destinations that at least partially match criteria contained in the voice destination input).

[088] The generated list of possible destinations may be post-processed in an operation 530 in order to assign weights or ranks to one or more of the entries in the N-best list. The post-processing may include analyzing the destination list generated in operation 520 according to various factors in order to determine a most likely intended destination from a full or partial voice destination input. For example, post-processing operation 530 may rank or weigh possible destinations according to shared knowledge about the user, domain-specific knowledge, dialogue history, or other factors. Furthermore, the post-processing operation 530 may analyze the full or partial destination input in order to identify an address to which a route can be calculated, for example, by resolving a closest address that makes "sense" relative to the input destination. For example, a user may specify a partial destination that identifies a broad and approximated area (e.g., "Take me to Massachusetts"), and depending on a user's current location, direction of travel, preferences, or other information, post-processing operation 530 may select an address makes sense for calculating a route (e.g., an address in Cape Cod may be selected for a user having relatives that live in Cape Cod, whereas an address in Boston may be selected for a user who may be traveling to various popular sightseeing areas, etc.).

[089] As a result, the weighed list of N-best destinations may be evaluated in an operation 540 to determine a suitably identifiable destination exists in the list. For example, a full or partial voice destination entry may be ambiguous, or certain criteria or keywords in a voice destination entry may be unrecognizable, such that a highest ranked destination in the weighted list does not exceed a minimal confidence level needed to identify a destination. For instance, a user located in Oklahoma may utter a partial destination, as in "I'm heading to Washington," and the decisional operation 540 may return a negative indication when the voice user interface cannot disambiguate between Washington state, Washington, D.C., Washington University in Saint Louis, and a town of Washington located slightly south of Oklahoma City. In another example, a user originating in Saint Louis may provide a voice destination entry of "Take me to Springfield," which could result in an unidentifiable destination even though multiple destinations may satisfy the minimal confidence level. In this example, "Springfield" may be unidentifiable as a destination because Springfield, Illinois and Springfield, Missouri both reside within reasonable distances of Saint Louis, yet directions of travel to either destination may be

Amazon/VB Assets
Exhibit 1003
Page 658

entirely opposite (e.g., a route to the Illinois Springfield includes traveling north on Interstate-55, whereas a route to the Missouri Springfield includes traveling southwest on Interstate-44). Thus, to avoid routing the user in a direction opposite from an intended destination, processing may instead branch to an operation 550 to generate a prompt for resolving the destination.

[090] For example, the prompt may ask the user a question or provide information that can clarify an intended meaning of the received voice destination input. For instance, in the Washington example given above, a multi-modal user prompt may display the four possible destinations on a display device (i.e., Washington state, Washington, D.C., Washington University, and Washington, Oklahoma). The multi-modal prompt may further include system-generated speech, for example, stating, "I found several possible Washingtons, did you mean one of these or something else?" As a result, processing then proceeds back to operation 510, in which the user can disambiguate the intended destination through another input (e.g., by touching one of the displayed destinations when the display device includes a touch-screen, or by verbalizing one of the destinations, or by providing additional information indicating that none of the displayed destinations were intended). Thereafter, subsequent N-best lists may be more likely to result in an identifiable destination, as an original voice destination input can be interpreted in view of a dialogue history.

[091] Alternatively, when at least one of the possible destinations can be identified unambiguously, while meeting the minimal confidence level, a positive indication may result in processing proceeding to another decisional operation 560, which controls how a route will be calculated to the identified destination (e.g., a highest ranking entry in the weighted N-best list of destinations). The decisional operation 560 may determine whether the identified destination provides a full or otherwise complete destination. For example, through successive refinement, a partial voice destination entry may result in an identifiable, yet incomplete, destination. As such, an operation 570 may calculate a preliminary route to the identified partial destination. Thereafter, subsequent processing returns to operation 510, where the voice user interface may await an additional voice destination input that refines the partial destination. Further, in various implementations, one or more system prompts may be generated to request the additional voice destination inputs (e.g., as the user approaches the partial destination and additional information will be needed to provide further routing, or as the user approaches a point in the preliminary

Amazon/VB Assets
Exhibit 1003
Page 659

route where distinct routes may be taken to different topographical subtiles or points within the partial destination, etc.).

[092] As such, the processing operations 510-570 may be iteratively repeated until the final destination can be identified through successive refinement using one or more multi-modal voice destination entries. As a result, when a suitable final destination has been identified, operation 560 may be followed by an operation 580 that completes the route to the final destination. For example, the route may be completed by extrapolating the route to the partial destination into a complete route to the final destination. Further, it will be apparent that voice destination entries may be successively refined into a final destination en route (e.g., the final destination may be successively refined as a user proceeds along a preliminary route), in advance (e.g., the user may choose to drive upon a complete route to a final destination being identified), or in other ways.

[093] Moreover, as described in greater detail herein, the calculated route may be dynamically adjusted or rerouted based on subsequent inputs, generated inferences, dynamic data, or various other sources of information (e.g., inferences may be generated, which may result in dynamic routing, in response to dynamic data relating to traffic conditions, detours, events, weather conditions, or user preferences, or other factors). Thus, Figure 6 illustrates a flow diagram of an exemplary method for providing voice services based on a current navigation route according to various aspects of the invention. The voice user interface may use the method illustrated in Figure 6 to provide the voice services in a dynamic, data-driven manner, wherein the voice services may be provided in a customized manner, for example, based on context, shared knowledge about a user, or other sources of information.

[094] Accordingly, providing voice navigation services based on a current route includes identifying the current navigation route in an operation 610. The route may be identified from a preliminary route or a complete route, as calculated according to the techniques described above in reference to Figure 5. Thus, current routing information may be made available to various system components in order to process events or voice inputs, generate responses, formulate queries, resolve commands, or perform other processing services available in a voice navigation services environment. As such, the current routing information may provide further context relevant to interpreting voice inputs, taking action towards resolving the inputs, or responding to events, among other things. Thus, though not illustrated in Figure 6, various other sources of contextual information may be utilized to provide voice navigation services in view of the

Amazon/VB Assets
Exhibit 1003
Page 660

current route. For example, user preferences may limit places to which a vehicle can travel, or for how long the vehicle can travel, among other things (e.g., for keeping a delivery or passenger route on a regular route, or for placing parental controls on teenage drivers, etc.). In another example, the voice user interface may have awareness of one or more external systems or devices, sources of dynamic data, or other information, which can be utilized to determine how to best provide the voice navigation services.

[095] Thus, upon identifying the current route in operation 610, along with other appropriate information relating to context, shared knowledge, or otherwise, a decisional operation 620 may determine whether the identified information results in an event being generated and/or detected. When no event can be generated and/or detected, processing proceeds to a decisional operation 630, which determines whether a voice input has been received from the user. As such, the method illustrated herein may loop through decisional operations 620 and 630 until one or more of an event has been detected and/or a voice input has been received.

[096] One type of event that can be detected and/or generated in decisional operation 620 may relate to location dependent advertisements for navigation systems (e.g., as generated by a local or remote advertising engine, or via a data channel, or in other ways). For example, in various implementations, the advertising events may be generated and/or detected using techniques described in co-pending U.S. Patent Application 11/671,526, entitled "System and Method for Selecting and Presenting Advertisements Based on Natural Language Processing of Voice-Based Input," filed February 6, 2007, the disclosure of which is hereby incorporated by reference in its entirety. For instance, as will be apparent, a navigation system will necessarily include one or more systems or mechanisms for determining the navigation system's current location (e.g., a global positioning system, a radio frequency identification system, a system that determines location based on a distance to an identifiable wireless tower or access point, etc.). Further, the natural language voice user interface may be able to determine location in a number of other ways, in addition to or apart from the location detection system (e.g., an adaptable prompt may affirmatively request the location information from a user, for example, upon failure of the location detection system, or other techniques utilizing shared knowledge or context may be used, as will be apparent).

[097] Thus, in one example, operation 620 may include detecting and/or generating events relating to a location dependent advertisement. A marketer or other entity seeking to reach a

Amazon/VB Assets
Exhibit 1003
Page 661

target audience may simply scatter radio frequency identifiers (RFIDs) along a roadside, or another place in which a system may detect the RFIDs. The marketer or other entity may then broadcast one or more advertisements via a data channel associated with the RFIDs, such that the navigation system may trigger an event when within a suitable proximity of the RFIDs. Thus, processing would proceed to an operation 570, in which information associated with the event may be filtered according to the current routing information or other contextual parameters. As applied to this example, the filtering operation 650 may include monitoring the data channel to analyze information associated with the detected RFID, thus receiving the advertisements broadcasted over the data channel. Thus, an advertisement including speech, for example, may be analyzed in a similar manner to how user speech would be analyzed (e.g., maintaining awareness that the speech relates to an advertisement).

[098] Accordingly, the advertisement event data may be filtered in view of the current route, personal information (e.g., previous queries, preferences, history, etc.), or other factors to determine what action should be taken in an operation 660. For example, the advertisement may be associated with a local restaurant seeking to reach personal navigation devices that may not necessarily have a complete business database. Therefore, in one example, filtering operation 650 could discard the advertisement when shared knowledge indicates that the user prefers to reach destinations as fast as possible, without making unnecessary stops. In this case, operation 660 may effectively be a null operation, or may simply include displaying an image associated with the advertisement to allow the user to make an independent choice as to whether or not to stop. However, in a distinct example, the user may have provided a previous voice input indicating a desire to stop driving and eat lunch at some point along the route. Thus, in this example, operation 660 may include a spoken output that brings the restaurant to the user's attention, possibly asking the user whether the route should be temporarily diverted to stop for lunch at the restaurant.

[099] However, it should be appreciated that the above described RFID-based advertisement provides but one example of a possible advertising event that may be detected in operation 620. For instance, advertisements may be uploaded to a server by one or more advertising partners, wherein the uploaded advertisements may be associated with metadata or other descriptive information that identifies a target audience, location-dependent information, or other criteria. In another example, a plurality of advertisements may be stored locally at the voice user interface,

Amazon/VB Assets
Exhibit 1003
Page 662

and an inferencing engine may determine appropriate circumstances in which an event should be generated to deliver one or more of the advertisements to a user. As a result, it will be apparent that advertising events may be generated in a number of ways, and may be generated and/or detected locally, remotely, by detecting RFIDs, or in other ways.

[0100] Further, as advertisements get delivered to users (e.g., in operation 660), the users' subsequent interaction with the delivered advertisements may be tracked. In this way, affinity based models may be generated, for example, to ensure that promotions or advertisements will be delivered to a likely target audience. Thus, an event relating to a given advertisement may be generated and/or detected in operation 620 when shared knowledge about a user's behavior, preferences, or other characteristics match one or more criteria associated with peer-to-peer affinities associated with the advertisement. For example, a marketer may be offering promotional coupons intended for a specific demographic audience, and the peer-to-peer affinities may be used to ensure that the coupons will only be provided to users likely to be in the intended audience (e.g., to avoid providing an entire market, including unlikely customers, with the coupons). Thus, when the user falls within the specified demographic audience, one or more events may be generated on a server, locally by an advertising engine, or in other ways to deliver one or more of the promotional coupons to the user.

[0101] In other examples, the advertising model may include mobile pay-per-use systems (e.g., network access may be provided through a wireless access point on a pay-per-use basis), peer-to-peer local guides or recommendations (e.g., affinities may indicate that various categories of peer users may be interested in similar points-of-interests, events, local restaurants, stores, activities, tourist attractions, or other aspects of a local area). Additionally, as described in greater detail above in reference to Figure 3, various aspects of the advertising model, such as the local guides and recommendations, may be generated according to a mapping applied to various topological domains. For example, local guides or recommendations may be dependent on topological characteristics (e.g., different guides for various cities, communities, localities, etc.). As such, the local guides, recommendations, or other aspects of the advertising model may be associated with a topological taxonomy based on geographical chunking. As a result, various advertising events may be generated and/or detected according to physical proximities, temporal proximities, directional proximities, civil organization proximities, or various combinations

Amazon/VB Assets
Exhibit 1003
Page 663

thereof (e.g., peer-to-peer recommendations may be varied depending on a user's location, direction of travel, temporal constraints, etc.).

[0102] In addition to the advertising related events that can be generated and/or detected in operation 620, the natural language voice user interface may generate additional events through awareness of context, shared knowledge about a user, external systems and devices, or other information. For example, as discussed above in reference to Figure 2, the voice user interface may implemented within an agent-based architecture, in which one or more of a plurality of agents may include an inference engine. However, it will be apparent that the inference engine may be arranged within the voice user interface in various other ways. For example, the voice user interface may include the inference engine within a conversational language processor, where in such implementations, the inference engine may generate, detect, and distribute events to various components of the voice user interface. In another example, a managing inference engine may coordinate event generation and/or detection among the inference engines associated with the agents. As such, it will be apparent that the voice user interface may include various suitable arrangements of one or more inference engines, such that events and inferences can be detected, generated, and distributed to various system components as they may arise.

[0103] Thus, the one or more inference engines may utilize awareness of context, shared knowledge about a user, dynamic data sources, data and services associated with external or devices, among other information to generate and/or detect the events or other inferences identified in operation 620. When the events or other inferences occur when a current navigation route exists, the events or other inferences may be filtered in operation 650 prior to being a response being generated and/or a task being performed in operation 660.

[0104] For example, the inference engine may utilize personal information, such as a user frequently visiting a particular person in the user's address book, to generate an event that suggests rerouting the current route. Thus, in this example, the inference engine may generate the event (which may be detected in operation 620) when the current route or current location information reflects a certain proximity to an address associated with the person in the address book. Operation 650 could then filter the event by determining possible responses or courses of action for handling the event in a manner consistent with the current route. For example, information associated with the current route may indicate that the user was out running errands and may now be heading home. Further, a calendar in the user's address book may indicate that

Amazon/VB Assets
Exhibit 1003
Page 664

the user has an appointment later in the day, but that no calendar entries exist for a few hours. Thus, an exemplary response generated in operation 660 may include a speech-based output of, "You are close to Jane's house, would you like to stop by?"

[0105] Although the above-provided example illustrates an event generated using personal information, it will be apparent that various available information sources may be utilized to generate and/or detect events. For example, events may be based on transient or dynamic data relating to communities, traffic, weather, and/or many other sources of data (e.g., a generated event may result in an output of "Tonight's soccer game is being played at a field near your house, do you want to go?", in which the soccer game may be a local community event).

[0106] In yet another example, various types of external system awareness may trigger events and resulting voice responses from the voice user interface. For example, the voice user interface may be coupled to various vehicular or telematics systems, providing awareness over a state of such systems. Thus, an event may be generated when the vehicular system indicates that the vehicle will soon run out of gas, resulting in a voice response prompting a user to stop for gas. Further, the external system awareness may be filtered according to the current routing information (e.g., in conjunction with the inference engine), wherein a calculation may be made indicating that a level of gas will be sufficient to get to a next gas station along the route, but that the level of gas will not be sufficient to get to a second gas station along the route. As a result, the route-based filtering operation 650 may result in a response being generated in operation 660 that provides a voice warning that the user must stop for gas at the next gas station.

[0107] Returning to operation 630, user-provided voice inputs may also be processed in view of current routing information. As previously discussed, interpretations of what was said in a voice input (e.g., content of an utterance) may be based on various phonetic models, dictionaries, context histories, dialogue histories, and other information that can form a dynamic recognition grammar. Moreover, context, shared knowledge, and available data and services, among other information, may be used to interpret what meant by the voice input (e.g., an intent of the utterance). As a result, using the various techniques described herein and in the incorporated co-pending U.S. Patent Applications, a query, command, or other request contained in the utterance may be identified in an operation 640. When a current route exists, a domain of possible tasks, commands, query answers, responses, and/or other system actions may be filtered according to the current route. For example, a voice input of "Where's the closest bathroom" may be

Amazon/VB Assets
Exhibit 1003
Page 665

analyzed in view of the current route. Various proximities (e.g., a direction of travel, physical distances, temporal distances, etc.) may then be utilized to determine what bathrooms may be appropriate in view of the current route. Thus, when a route has the user traveling a long distance, for example, a bathroom at a rest area twenty miles ahead may be favored over a bathroom at a restaurant ten miles off a highway (e.g., to avoid diverting the user from the current route).

[0108] Furthermore, in many instances, an agent or other processing component may need to formulate a plurality of queries to various different data sources in order to resolve a request. As such, operation 660 for performing a task may determine what inferences, metadata, or other information may be needed to formulate queries to the data sources in order to resolve the request. As a result, voice requests relating to navigation may include information that may not necessarily be linked to navigation, or that may not necessarily be in a database or other source of navigation information. Even so, by being coupled with various available sources of information, complex or compound queries may be formulated to retrieve information from various data sources in order to resolve a given request. Furthermore, a given utterance may include multiple requests, queries, or commands, such that a given response or task performed in response to the utterance may include interactions with various different system components. Thus, in an illustrative example, an utterance of "Find the 7-11 closest to the Woodland Park Zoo" may include a first query to identify a location of Woodland Park Zoo, a second query to find one or more 7-11's within a certain radius of the Woodland Park Zoo location, a third query to determine a current location of the user, and a fourth query that invokes an inference engine to determine which 7-11 may be "closest" based on various user proximities. Thus, supposing that two 7-11's exist equidistantly from the Woodland Park Zoo, one in a direction east of the Zoo and one in a direction west of the Zoo, the user's current location may be used to determine which 7-11 to select; by contrast, a selection based purely on physical proximity to the Zoo could result in ambiguity.

[0109] Although the above-provided descriptions highlight various specific examples and illustrations of how the natural language voice user interface may provide voice navigation services, it will be understood that the descriptions, examples, and modes of operation provided above should be regarded as exemplary only. The natural language voice user interface may be highly tolerant of variations in how humans interact verbally, and may utilize context, shared

Amazon/VB Assets
Exhibit 1003
Page 666

knowledge, and available data and systems to effectively interact with any given user in a conversational manner that advances mutual goals. As a result, interactions between the user and the voice user interface may be regarded as unique occurrences, as users may speak free form requests relating to various available data sources, applications, or devices in a given interaction. Thus, the voice user interface may dynamically adapt, learn, and otherwise refine an operational framework to provide optimal services in the voice navigation services environment.

[0110] Implementations of the invention may be made in hardware, firmware, software, or various combinations thereof. The invention may also be implemented as instructions stored on a machine-readable medium, which may be read and executed by one or more processors. A machine-readable medium may include a mechanism for storing or transmitting information in a form readable by a machine (e.g., a computing device). For example, a machine-readable storage medium may include read only memory, random access memory, magnetic disk storage media, optical storage media, flash memory devices, and others, and a machine-readable transmission media may include forms of propagated signals, such as carrier waves, infrared signals, digital signals, and others. Further, firmware, software, routines, or instructions may be described in the above disclosure in terms of specific exemplary aspects and implementations of the invention, and performing certain actions. However, those skilled in the art will recognize that such descriptions are merely for convenience and that such actions in fact result from computing devices, processors, controllers, or other devices executing the firmware, software, routines, or instructions.

[0111] Aspects and implementations may be described as including a particular feature, structure, or characteristic, but every aspect or implementation may not necessarily include the particular feature, structure, or characteristic. Further, when a particular feature, structure, or characteristic is described in connection with an aspect or implementation, it will be understood that one skilled in the art may be able to effect such feature, structure, or characteristic in connection with other aspects or implementations, whether or not explicitly described. Thus, various changes and modifications may be made to the provided description without departing from the scope or spirit of the invention. As such, the specification and drawings should be regarded as exemplary only, and the scope of the invention to be determined solely by the appended claims.

Amazon/VB Assets
Exhibit 1003
Page 667

CLAIMS

What is claimed is:

1. A computer-implemented method for providing a natural language voice user interface for a navigation device, comprising:

receiving a navigation-related voice-based input at an input mechanism associated with the navigation device, the voice-based input including at least a natural language utterance provided by a user;

generating one or more preliminary interpretations of the natural language utterance using a dynamic recognition grammar associated with a speech recognition engine;

analyzing the preliminary interpretations using a conversational language processor, the conversational language processor using shared knowledge and information associated with a navigation context to determine a probable interpretation of the natural language utterance in the navigation context; and

communicating the probable interpretation of the natural language utterance to a navigation agent associated with the navigation context, the navigation agent operable to:

identify one or more navigation-related requests contained in the voice-based input based on the probable interpretation of the natural language utterance; and

resolve the requests using information associated with one or more of a plurality of information sources, the plurality of information sources including at least a navigation-specific information source.

2. The method of claim 1, the navigation agent further operable to determine that at least one of the requests includes an approximation of requested information, wherein the conversational language processor manages a dialogue with the user to resolve the approximated information, the managed dialogue including one or more interactions with the user that successively refine the approximated information.

3. The method of claim 2, each of the interactions including one or more of prompting the user to refine the approximated information or receiving a subsequent input from the user that refines the approximated information.

Amazon/VB Assets
Exhibit 1003
Page 668

4. The method of claim 1, the navigation-related requests including a request for a route to a full or partial address, the navigation agent operable to resolve the request for the route by:

selecting a destination for the route, the selected destination having an address that best corresponds to the full or partial address;

calculating a route from a current location of the user to the selected destination; and

generating directions to guide the user on the calculated route to the selected destination, the generated directions dynamically driven by data from a plurality of information sources.

5. The method of claim 4, the dynamic data-driven directions including information about destinations, points of interest, traffic, parking, weather, events, or other information relevant to the calculated route.

6. The method of claim 4, further comprising:

receiving a subsequent voice-based input at the input mechanism associated with the navigation device, the subsequent voice-based input including at least one request provided by the user; and

invoking at least one of a plurality of domain agents to resolve the request included in the subsequent voice-based input, the at least one domain agent operable to filter results of the request according to the calculated route.

7. The method of claim 1, the dynamic recognition grammar including information associated with one or more topological domains.

8. The method of claim 7, the topological domains including physical, temporal, directional, and civil organizational proximities with respect to a current location of the user.

9. The method of claim 1, the conversational language processor coupled to an inferencing engine operable to generate one or more inferences, wherein the conversational language processor uses the generated inferences to determine the probable interpretation of the natural language utterance.

Amazon/VB Assets
Exhibit 1003
Page 669

10. The method of claim 9, wherein the conversational language processor further uses the generated inferences to suggest one or more services available in the navigation context to the user.

11. The method of claim 1, the navigation-related requests including a multi-modal request to control a map display, the navigation agent operable to resolve the request to control the map display by:

associating a non-voice component of the voice-based input with the probable interpretation of the utterance, the non-voice component identifying a portion of the map display; and

issuing a command to control the identified portion of the map display in accordance with the probable interpretation of the utterance.

12. The method of claim 1, wherein the plurality of information sources include the navigation-specific information source, the shared knowledge, the information associated with the navigation context, and one or more information sources relating to maps, destinations, directories, points of interest, traffic, parking, weather, events, user address books, user devices and systems, a search engine, and a plurality of domain agents.

13. The method of claim 12, the shared knowledge including dialogue history information, request history information, user interface state information, short-term user profile information, long-term user profile information, peer user profile information, and user location information.

14. A system for providing a natural language voice user interface for a navigation device, comprising:

an input mechanism that receives a navigation-related voice-based input, the voice-based input including at least a natural language utterance provided by a user;

a speech recognition engine having a dynamic recognition grammar associated therewith, the speech recognition engine operable to generate one or more preliminary interpretations of the natural language utterance using the dynamic recognition grammar;

Amazon/VB Assets
Exhibit 1003
Page 670

a plurality of domain agents associated with respective contexts, the plurality of domain agents including at least a navigation agent associated with a navigation context; and

a conversational language processor operable to:

analyze the preliminary interpretations using shared knowledge and information associated with the navigation context to determine a probable interpretation of the natural language utterance in the navigation context; and

communicate the probable interpretation of the natural language utterance to the navigation agent associated with the navigation context, the navigation agent operable to identify one or more navigation-related requests contained in the voice-based input based on the probable interpretation of the natural language utterance and resolve the requests using information associated with one or more of a plurality of information sources, the plurality of information sources including at least a navigation-specific information source.

15. The system of claim 14, the navigation agent further operable to determine that at least one of the requests includes an approximation of requested information; and

the conversational language processor further operable to manage a dialogue with the user to resolve the approximated information, the managed dialogue including one or more interactions with the user that successively refine the approximated information.

16. The system of claim 15, each of the interactions including one or more of prompting the user to refine the approximated information or receiving a subsequent input from the user that refines the approximated information.

17. The system of claim 14, the navigation-related requests including a request for a route to a full or partial address, the navigation agent operable to resolve the request for the route by:

selecting a destination for the route, the selected destination having an address that best corresponds to the full or partial address;

calculating a route from a current location of the user to the selected destination; and

generating directions to guide the user on the calculated route to the selected destination, the generated directions dynamically driven by data from a plurality of information sources.

Amazon/VB Assets
Exhibit 1003
Page 671

18. The system of claim 17, the dynamic data-driven directions including information about destinations, points of interest, traffic, parking, weather, events, or other information relevant to the calculated route.

19. The system of claim 17, the one or more input mechanisms further receiving a subsequent voice-based input, the subsequent voice-based input including at least one request provided by the user, and the conversational language processor further operable to:

invoking at least one of a plurality of domain agents to resolve the request included in the subsequent voice-based input, the at least one domain agent operable to filter results of the request according to the calculated route.

20. The system of claim 14, the dynamic recognition grammar including information associated with one or more topological domains.

21. The system of claim 20, the topological domains including physical, temporal, directional, and civil organizational proximities with respect to a current location of the user.

22. The system of claim 14, further comprising an inferencing engine coupled to the conversational language processor that generates one or more inferences, wherein the conversational language processor uses the generated inferences to determine the probable interpretation of the natural language utterance.

23. The system of claim 22, further comprising an inferencing engine coupled to the conversational language processor, wherein the conversational language processor further uses the generated inferences to suggest one or more services available in the navigation context to the user.

24. The system of claim 14, the navigation-related requests including a multi-modal request to control a map display, the navigation agent operable to resolve the request to control the map display by:

Amazon/VB Assets
Exhibit 1003
Page 672

associating a non-voice component of the voice-based input with the probable interpretation of the utterance, the non-voice component identifying a portion of the map display; and

issuing a command to control the identified portion of the map display in accordance with the intended meaning of the utterance.

25. The system of claim 14, wherein the plurality of information sources include the navigation-specific information source, the shared knowledge, the information associated with the navigation context, and one or more information sources relating to maps, destinations, directories, points of interest, traffic, parking, weather, events, user address books, user devices and systems, a search engine, and a plurality of domain agents.

26. The system of claim 25, the shared knowledge including dialogue history information, request history information, user interface state information, short-term user profile information, long-term user profile information, peer user profile information, and user location information.

27. A method for providing a natural language voice user interface in a voice navigation services environment, comprising:

receiving a voice-based destination input provided by a user, the voice-based destination input including at least a natural language utterance;

generating one or more preliminary interpretations of the natural language utterance using a dynamic recognition grammar associated with a speech recognition engine;

analyzing the preliminary interpretations using a conversational language processor, the conversational language processor using shared knowledge and information associated with a navigation context to determine an intended destination provided in the natural language utterance, the intended destination including an approximation of a destination; and

communicating the intended destination to a navigation agent, the navigation agent operable to provide a route to the intended destination by:

selecting a preliminary destination for the route, the selected preliminary destination having an address that best corresponds to the approximated destination;

Amazon/VB Assets
Exhibit 1003
Page 673

calculating a route from a current location of the user to the selected preliminary destination; and

participating in a dialogue with the user to resolve a final destination for the calculated route, the dialogue includes one or more interactions with the user that successively refine the approximated destination until the final destination has been resolved.

28. The method of claim 27, each of the interactions including one or more of prompting the user to refine the approximated destination or receiving a subsequent input from the user that refines the approximated information.

29. The method of claim 27, the navigation agent further operable to dynamically recalculate the route in response to the successive refinements of the approximated destination.

30. The method of claim 27, the final destination being resolvable substantially later in time with respect to an initial calculation of the route.

31. The method of claim 27, the navigation agent operable to select the preliminary destination by identifying one or more addresses possibly corresponding to the approximated destination and selecting a highest ranked one of the identified addresses as the preliminary destination.

32. The method of claim 27, the identified addresses ranked according to a proximity to the current location of the user or the approximated destination.

33. The method of claim 32, the proximity to the based on one or more topological domains, including one or more of physical, temporal, directional, and civil organizational proximities with respect to the current location of the user or the approximated destination.

Amazon/VB Assets
Exhibit 1003
Page 674

34. The method of claim 27, the navigation agent further operable to generate directions to guide the user on the calculated route, the generated directions dynamically driven by data from a plurality of information sources.

35. The method of claim 34, the plurality of information sources including information relating to destinations, points of interest, traffic, parking, weather, events, or other information relevant to the calculated route.

36. The method of claim 27, the shared knowledge including dialogue history information, request history information, user interface state information, short-term user profile information, long-term user profile information, peer user profile information, and user location information.

37. A method for providing a natural language voice user interface in a voice navigation services environment, comprising:

receiving a navigation-related voice-based input provided by a user, the voice-based input including at least a natural language utterance;

identifying a current location of a user using a location detection system;

determining one or more topological domains associated with the current location of the user;

generating a dynamic recognition grammar that includes grammar information associated with the determined topological domains; and

using the dynamic recognition grammar to generate one or more interpretations of the natural language utterance.

38. The method of claim 37, the dynamic recognition grammar organizing the grammar information according to geographic chunks associated with the topological domains.

39. The method of claim 38, the geographic chunks including physical proximities defined by a distance from the current location of the user.

Amazon/VB Assets
Exhibit 1003
Page 675

40. The method of claim 38, the geographic chunks including temporal proximities defined by a time of travel from the current location of the user.

41. The method of claim 38, the geographic chunks including directional proximities defined by a directional vector of travel by the user.

42. The method of claim 38, the geographic chunks including civil organizational proximities defined by continents, countries, regions, states, cities, localities, neighborhoods, and communities.

43. The method of claim 38, further comprising dividing one or more of the geographic chunks into a plurality of tiles, the dynamic recognition grammar further organizing the grammar information according to the plurality of tiles.

44. The method of claim 43, further comprising dividing one or more of the plurality of tiles into a plurality of subtiles, the dynamic recognition grammar further organizing the grammar information according to the plurality of subtiles.

45. The method of claim 37, the dynamic recognition grammar having a size constrained according to memory or resource availability in the voice navigation services environment.

46. The method of claim 37, further comprising modifying the grammar information included in the dynamic recognition grammar in response to changes in the topological domains associated with the current location of the user.

47. The method of claim 37, further comprising determining affinities between the user and one or more peers of the user, the dynamic recognition grammar further organizing the grammar information according to the determined affinities.

48. A method for providing advertisements in a voice navigation services environment, comprising:

Amazon/VB Assets
Exhibit 1003
Page 676

identifying a current location of a user using a location detection system;

retrieving shared knowledge and information associated with a navigation context, the retrieved information used to determine probable interpretations of natural language utterances received in the voice navigation services environment;

identifying one or more advertisements to provide to the user, the one or more advertisements identified based on an affinity to one or more of the current location of the user, the shared knowledge, or the information associated with the navigation context; and

generating a multi-modal output to provide the one or more identified advertisements to the user.

49. The method of claim 48, the multi-modal output including a system-generated speech utterance.

50. The method of claim 48, the location detection system receiving information relating to the one or more advertisements over a data channel.

51. The method of claim 50, the data channel associated with a radio frequency identifier detected by the location detection system.

52. The method of claim 48, the shared knowledge including dialogue history information, request history information, user interface state information, short-term user profile information, long-term user profile information, peer user profile information, and user location information.

53. The method of claim 48, further comprising:

receiving a voice-based input that includes at least a natural language utterance;

using the shared knowledge and the information associated with the navigation context to determine a probable interpretation of the received natural language utterance, the advertisements identified based on the probable interpretation of the natural language utterance.

54. The method of claim 48, the advertisement including a guide to a community local to the current location of the user.

Amazon/VB Assets
Exhibit 1003
Page 677

55. The method of claim 54, the local community guide including recommendations for one or more points of interest, events, restaurants, stores, activities, or tourist attractions.

56. The method of claim 54, the local community guide including information associated with one or more of maps, destinations, directories, traffic, parking, or weather.

57. The method of claim 54, the local community guide based on affinities between the user and one or more of the current location of the user, one or more peers of the user, or one or more communities associated with the user.

58. The method of claim 48, the advertisement including one or more of an image, a banner, an audio message, a video message, a promotional offer, a coupon, and a data stream.

Amazon/VB Assets
Exhibit 1003
Page 678

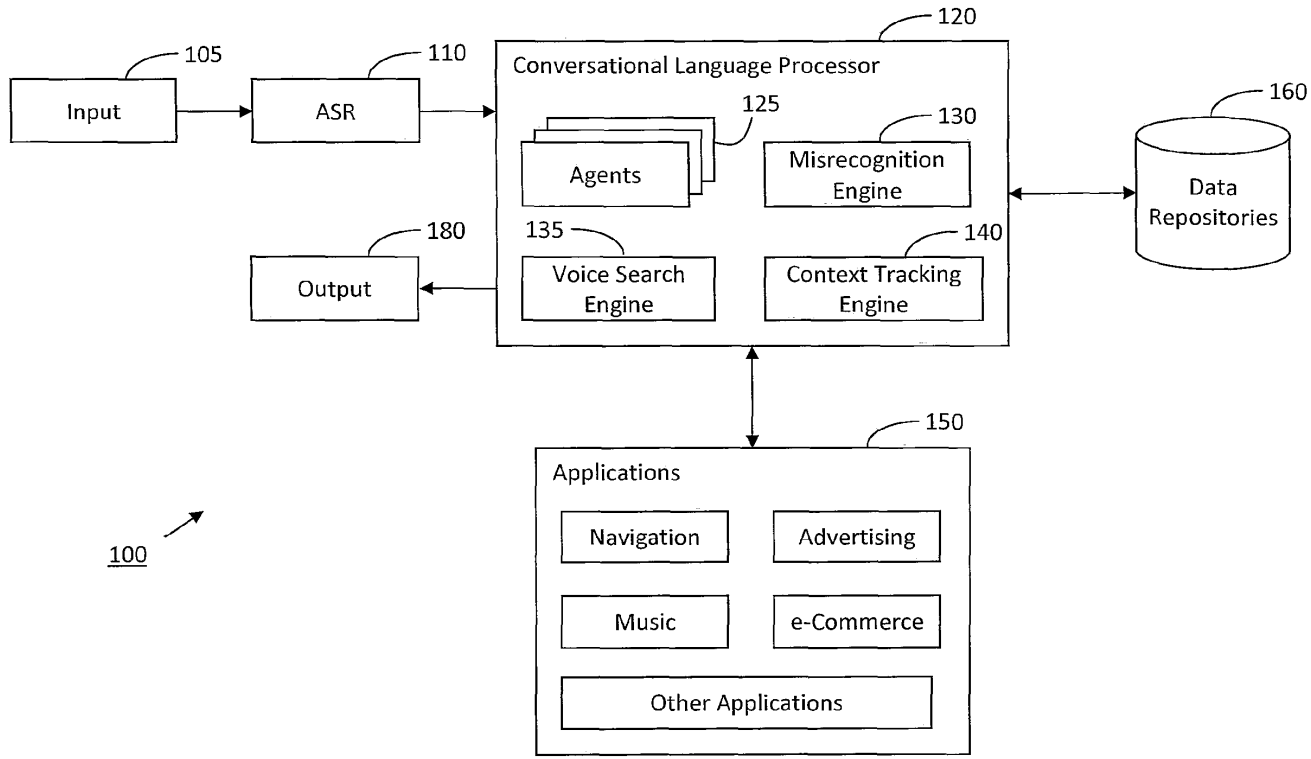

120
105
110
Input
ASR
Conversational Language Processor
125
130
160
Agents
Misrecognition Engine
Data Repositories
180
135
140
Output
Voice Search Engine
Context Tracking Engine
150
Applications
Navigation
Advertising
Music
e-Commerce
Other Applications
100

Figure 1

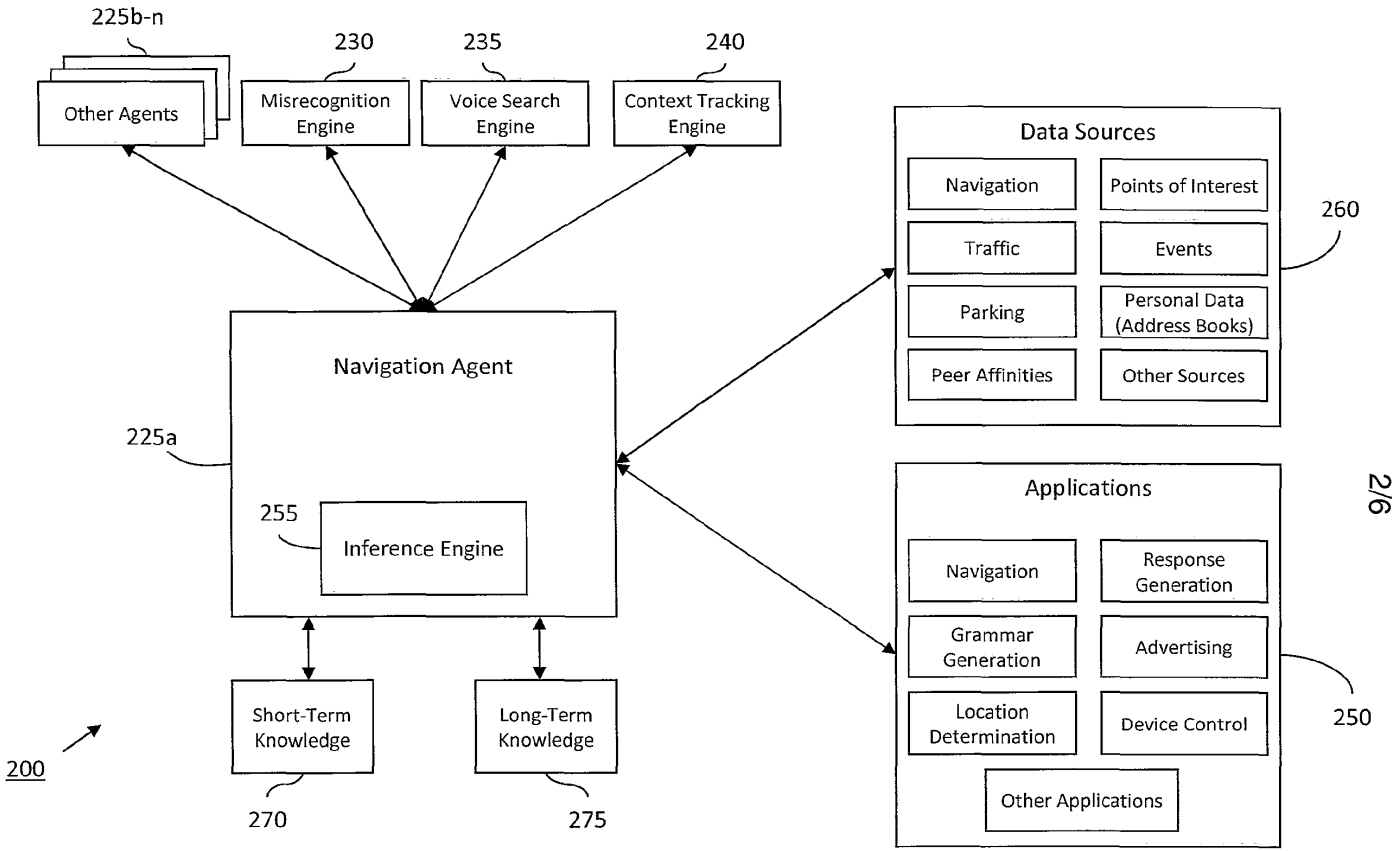


Figure 2

Amazon/VB Assets
Exhibit 1003
Page 680

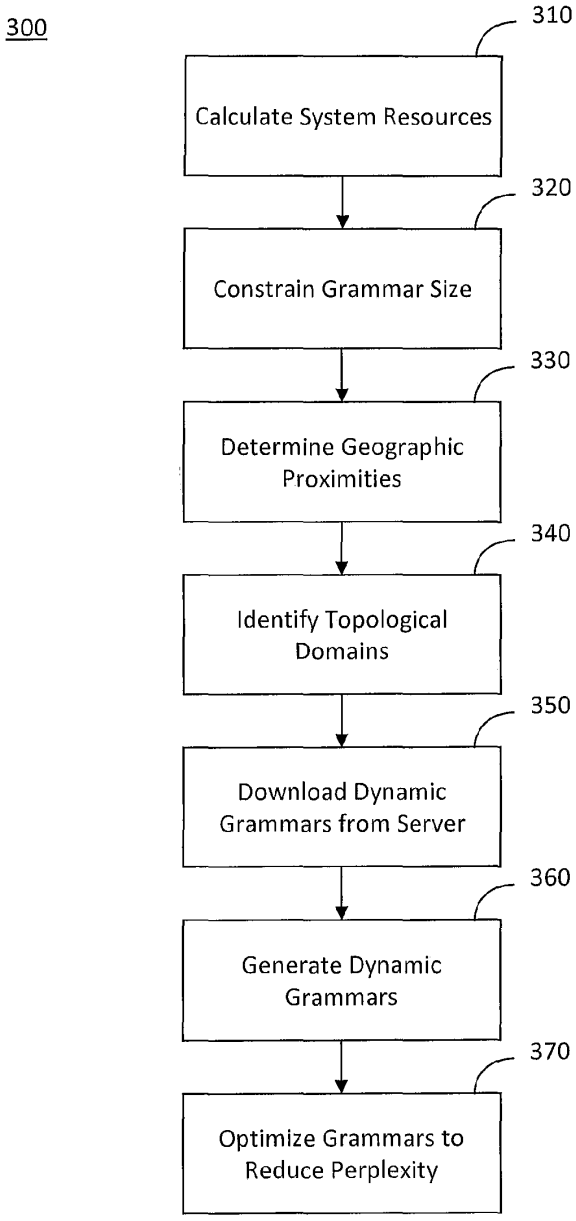

300
310
Calculate System Resources
320
Constrain Grammar Size
330
Determine Geographic Proximities
340
Identify Topological Domains
350
Download Dynamic Grammars from Server
360
Generate Dynamic Grammars
370
Optimize Grammars to Reduce Perplexity

Figure 3

400

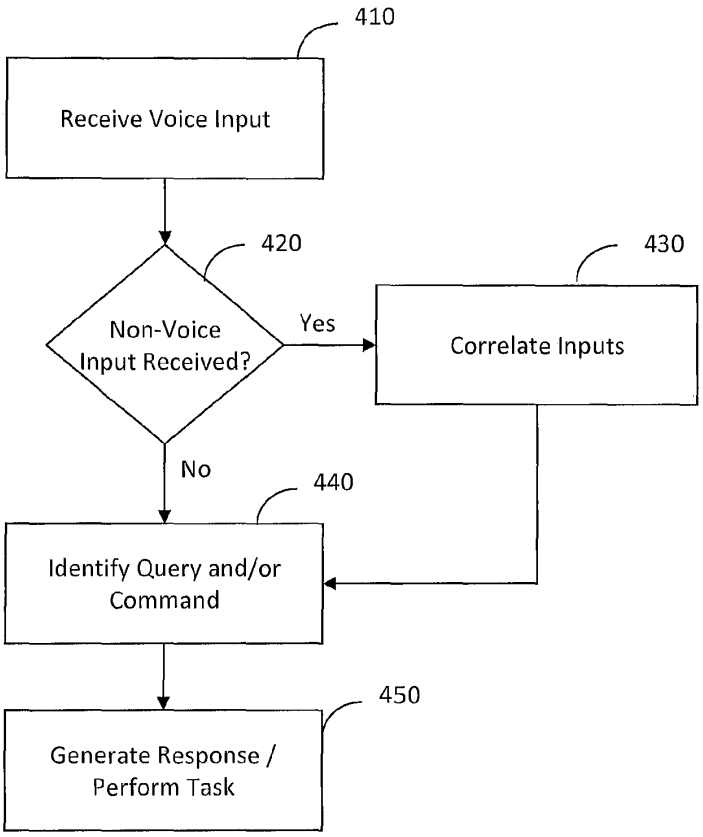

410
Receive Voice Input
420
Non-Voice Input Received?
Yes
430
Correlate Inputs
No
440
Identify Query and/or Command
450
Generate Response / Perform Task

Figure 4

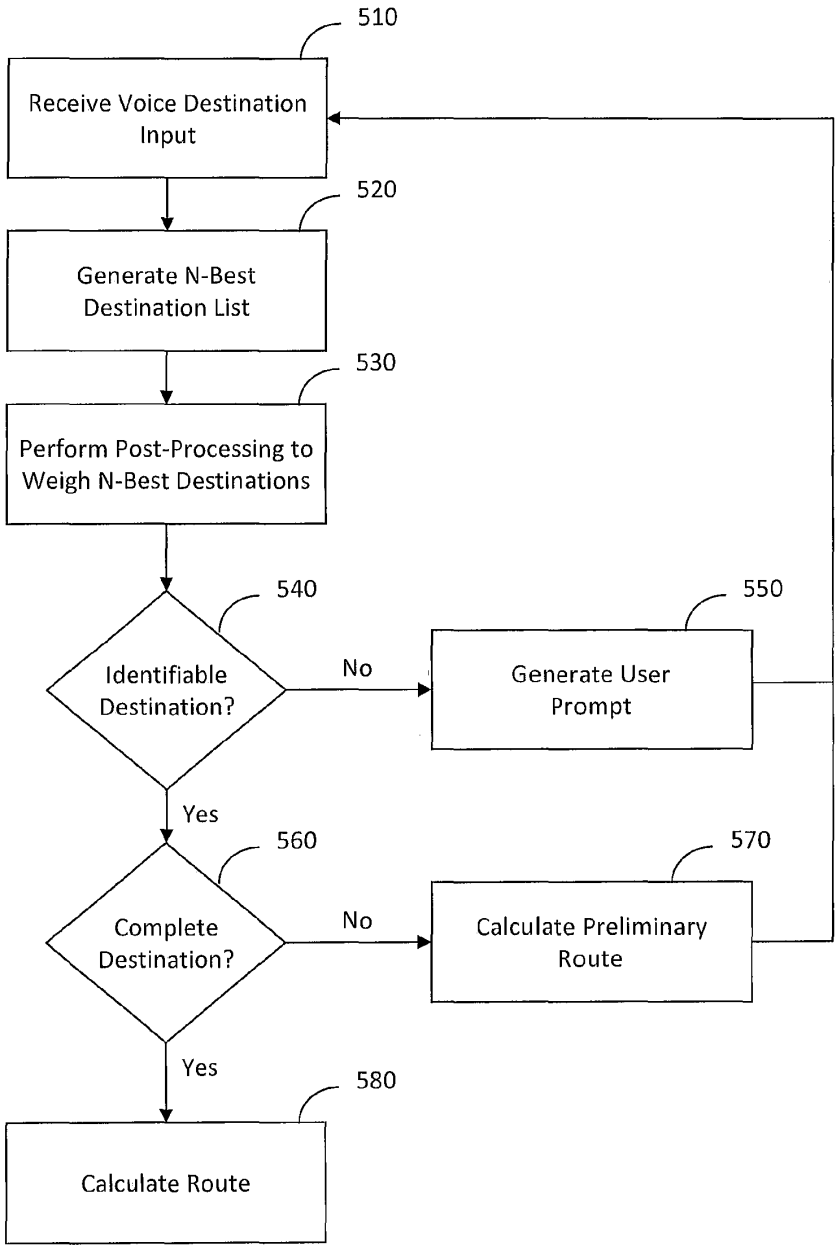

510
Receive Voice Destination Input
520
Generate N-Best Destination List
530
Perform Post-Processing to Weigh N-Best Destinations
540
Identifiable Destination?
No
550
Generate User Prompt
Yes
560
Complete Destination?
No
570
Calculate Preliminary Route
Yes
580
Calculate Route

Figure 5

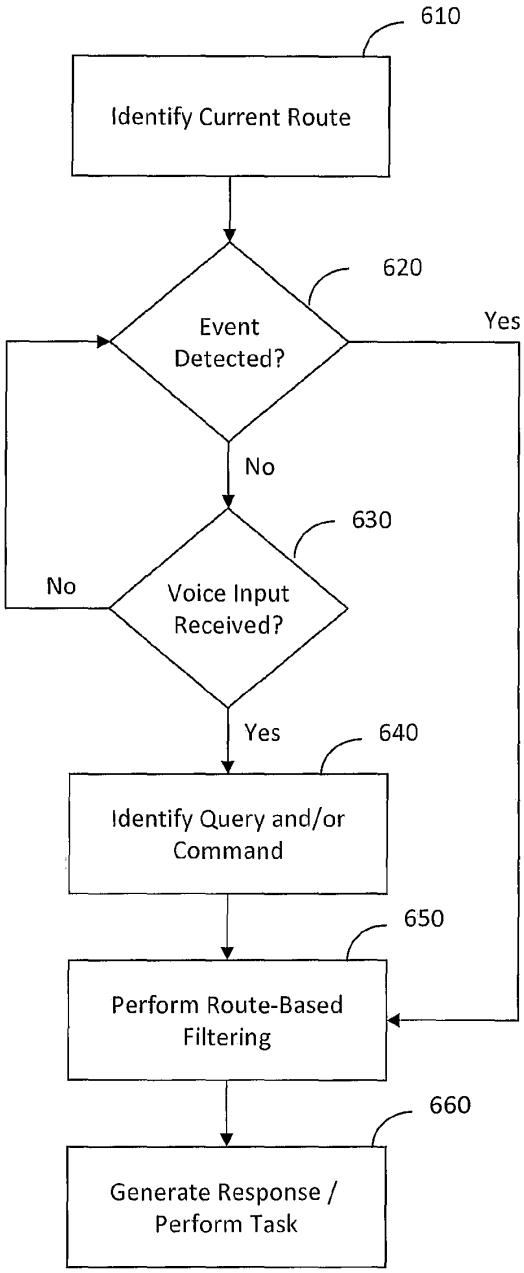

610
Identify Current Route
620
Event Detected?
Yes
No
630
Voice Input Received?
No
Yes
640
Identify Query and/or Command
650
Perform Route-Based Filtering
660
Generate Response / Perform Task

Figure 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 08/69508

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - G06F 17/27 (2008.04)
USPC - 704/9
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
USPC - 704/9

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC - 704/1, 2, 8, 9, 200, 206, 221, E15.045

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubWEST (PGPB, USPT, EPAB, JPAB); Google Scholar
Search Terms Used: map, gps, speech, voice, talk, navigation, natural, conversational, grammar, interpret, software, computer, input, interface, user interface, refine, zoom, route, directions

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	US 2007/0050191 A1 (WEIDER et al.) 01 March 2007 (01.03.2007), entire document, especially paras. [0061]-[0073], [0100], [0121], [0123], [0126], [0133], [0140], [0142], [0147], [0153], [0167], [0189]-[0194], [0210]-[0211], [0215]	1-42, 45-58 ---------- 43-44
Y	US 5,902,347 A (BACKMAN et al.) 11 May 1999 (11.05.1999), entire document, especially col 1 ln 32-34, col 4 ln 33-44	43-44
A	US 5,274,560 A (LARUE) 28 December 1993 (28.12.1993), entire document	1-58
A	US 6,598,018 B1 (JUNQUA) 22 July 2003 (22.07.2003), entire document	1-58

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 September 2008 (25.09.2008)	02 OCT 2008
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (April 2007)

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 December 2009 (03.12.2009)

PCT

(10) International Publication Number
WO 2009/145796 A1

(51) International Patent Classification:
G06F 13/12 (2006.01)

(21) International Application Number:
PCT/US2008/069524

(22) International Filing Date:
9 July 2008 (09.07.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
12/127,343 27 May 2008 (27.05.2008) US

(71) Applicant *(for all designated States except US)*: **VOICEBOX TECHNOLOGIES, INC.** [US/US]; 11980 NE 24th Street, Bellevue, Washington 98005 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **KENNEWICK, Robert** [US/US]; 13660 NE 42nd St., Bellevue, WA 98005 (US). **WEIDER, Chris** [US/US]; 10757 Greenwood Avenue N. #220, Seattle, Washington 98133 (US).

(74) Agents: **GATTO, James, G.** et al.; Pillsbury Winthrop Shaw Pittman LLP, P.O. Box 10500, McLean, Virginia 22102 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

(54) **Title**: SYSTEM AND METHOD FOR AN INTEGRATED, MULTI-MODAL, MULTI-DEVICE NATURAL LANGUAGE VOICE SERVICES ENVIRONMENT

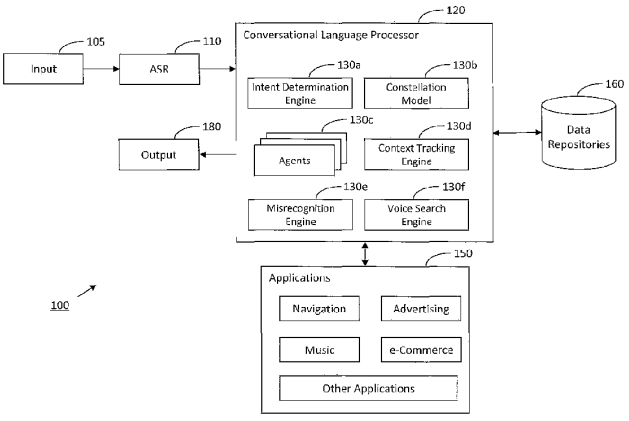


Figure 1

(57) **Abstract**: A system and method for an integrated, multi-modal, multi-device natural language voice services environment may be provided. In particular, the environment may include a plurality of voice-enabled devices each having intent determination capabilities for processing multi-modal natural language inputs in addition to knowledge of the intent determination capabilities of other devices in the environment. Further, the environment may be arranged in a centralized manner, a distributed peer-to-peer manner, or various combinations thereof. As such, the various devices may cooperate to determine intent of multi-modal natural language inputs, and commands, queries, or other requests may be routed to one or more of the devices best suited to take action in response thereto.

WO 2009/145796 A1

SYSTEM AND METHOD FOR AN INTEGRATED, MULTI-MODAL, MULTI-DEVICE NATURAL LANGUAGE VOICE SERVICES ENVIRONMENT

FIELD OF THE INVENTION

[001] The invention relates to an integrated voice services environment in which a plurality of devices can provide various voice services by cooperatively processing free form, multi-modal, natural language inputs, thereby facilitating conversational interactions between a user and one or more of the devices in the integrated environment.

BACKGROUND OF THE INVENTION

[002] As technology has progressed in recent years, consumer electronic devices have emerged to become nearly ubiquitous in the everyday lives of many people. To meet the increasing demand that has resulted from growth in the functionality and mobility of mobile phones, navigation devices, embedded devices, and other such devices, a wealth of features and functions are often provided therein in addition to core applications. Greater functionality also introduces the trade-offs, however, including learning curves that often inhibit users from fully exploiting all of the capabilities of their electronic devices. For example, many existing electronic devices include complex human to machine interfaces that may not be particularly user-friendly, which inhibits mass-market adoption for many technologies. Moreover, cumbersome interfaces often result in otherwise desirable features being buried (e.g., within menus that may be tedious to navigate), which has the tendency of causing many users to not use, or even know about, the potential capabilities of their devices.

[003] As such, the increased functionality provided by many electronic devices often tends to be wasted, as market research suggests that many users only use only a fraction of the features or applications available on a given device. Moreover, in a society where wireless networking and broadband access are increasingly prevalent, consumers tend to naturally desire seamless mobile capabilities from their electronic devices. Thus, as consumer demand intensifies for simpler mechanisms to interact with electronic devices, cumbersome interfaces that prevent quick and focused interaction can become an important concern. Accordingly, the ever-growing demand for mechanisms to use technology in intuitive ways remains largely unfulfilled.

[004] One approach towards simplifying human to machine interactions in electronic devices includes the use of voice recognition software, which can enable users to exploit features that

Amazon/VB Assets
Exhibit 1003
Page 687

could otherwise be unfamiliar, unknown, or difficult to use. For example, a recent survey conducted by the Navteq Corporation, which provides data used in a variety of applications such as automotive navigation and web-based applications, demonstrates that voice recognition often ranks among the features most desired by consumers of electronic devices. Even so, existing voice user interfaces, when they actually work, still tend to require significant learning on the part of the user.

[005] For example, many existing voice user interface only support requests formulated according to specific command-and-control sequences or syntaxes. Furthermore, many existing voice user interfaces cause user frustration or dissatisfaction because of inaccurate speech recognition. Similarly, by forcing a user to provide pre-established commands or keywords to communicate requests in ways that a system can understand, existing voice user interfaces do not effectively engage the user in a productive, cooperative dialogue to resolve requests and advance a conversation towards a mutually satisfactory goal (e.g., when users may be uncertain of particular needs, available information, or device capabilities, among other things). As such, existing voice user interfaces tend to suffer from various drawbacks, including significant limitations on engaging users in a dialogue in a cooperative and conversational manner.

[006] Additionally, many existing voice user interfaces fall short in utilizing information distributed across various different domains or devices in order to resolve natural language voice-based inputs. Thus, existing voice user interfaces suffer from being constrained to a finite set of applications for which they have been designed, or to devices on which they reside. Although technological advancement has resulted in users often having several devices to suit their various needs, existing voice user interfaces do not adequately free users from device constraints. For example, users may be interested in services associated with different applications and devices, but existing voice user interfaces tend to restrict users from accessing the applications and devices as they see fit. Moreover, users typically can only practicably carry a finite number of devices at any given time, yet content or services associated with users' other devices that currently being used may be desired in various circumstances. Accordingly, although users tend to have varying needs, where content or services associated with different devices may be desired in various contexts or environments, existing voice technologies tend to fall short in providing an integrated environment in which users can request content or services associated with virtually any device or network. As such, constraints on information availability

Amazon/VB Assets
Exhibit 1003
Page 688

and device interaction mechanisms in existing voice services environments tend to prevent users from experiencing technology in an intuitive, natural, and efficient way.

[007] Existing systems suffer from these and other problems.

SUMMARY OF THE INVENTION

[008] According to various aspects of the invention, a system and method for an integrated, multi-modal, multi-device natural language voice services environment may include a plurality of voice-enabled devices each having intent determination capabilities for processing multi-modal natural language inputs in addition to knowledge of the intent determination capabilities of other devices in the environment. Further, the environment may be arranged in a centralized manner, a distributed peer-to-peer manner, or various combinations thereof. As such, the various devices may cooperate to determine intent of multi-modal natural language inputs, and commands, queries, or other requests may be routed to one or more of the devices best suited to take action in response thereto.

[009] According to various aspects of the invention, the integrated natural language voice services environment arranged in the centralized manner includes an input device that receives a multi-modal natural language input, a central device communicatively coupled to the input device, and one or more secondary devices communicatively coupled to the central device. Each of the input device, the central device, and the one or more secondary devices may have intent determination capabilities for processing multi-modal natural language inputs. As such, an intent of a given multi-modal natural language input may be determined in the centralized manner by communicating the multi-modal natural language input from the input device to the central device. Thereafter, the central device may aggregate the intent determination capabilities of the input device and the one or more secondary devices and determine an intent of the multi-modal natural language input using the aggregated intent determination capabilities. The input device may then receive the determined intent from the central device and invoke at least one action at one or more of the input device, the central device, or the secondary devices based on the determined intent.

[010] According to various aspects of the invention, the integrated natural language voice services environment arranged in the distributed manner includes an input device that receives a multi-modal natural language input, a central device communicatively coupled to the input

Amazon/VB Assets
Exhibit 1003
Page 689

device and one or more secondary devices communicatively coupled to the input device, wherein each of the input device and the one or more secondary devices may have intent determination capabilities for processing multi-modal natural language inputs, as in the centralized implementation. However, the distributed implementation may be distinct from the centralized implementation in that a preliminary intent of the multi-modal natural language input may be determined at the input device using local intent determination capabilities. The multi-modal natural language input may then be communicated to one or more of the secondary devices (e.g., when a confidence level of the intent determination at the input device falls below a given threshold). In such cases, each of the secondary devices determine an intent of the multi-modal natural language input using local intent determination capabilities. The input device collates the preliminary intent determination and the intent determinations of the secondary devices, and may arbitrate among the collated intent determinations to determine an actionable intent of the multi-modal natural input.

[011] According to various aspects of the invention, the integrated natural language voice services environment arranged in a manner that dynamically selects between a centralized model and a distributed model. For example, the environment includes an input device that receives a multi-modal natural language input one or more secondary devices communicatively coupled to the input device, each of which have intent determination capabilities for processing multi-modal natural language inputs. A constellation model may be accessible to each of the input device and the one or more secondary devices, wherein the constellation model describes the intent determination capabilities of the input device and the one or more secondary devices. The multi-modal natural language input can be routed for processing at one or more of the input device or the secondary devices to determine an intent thereof based on the intent determination capabilities described in the constellation model. For example, when the constellation model arranges the input device and the secondary devices in the centralized manner, one of the secondary devices may be designated the central device and the natural language input may be processed as described above. However, when the multi-modal natural language cannot be communicated to the central device, the constellation model may be dynamically rearranged in the distributed manner, whereby the input device and the secondary devices share knowledge relating to respective local intent determination capabilities and operate as cooperative nodes to

Amazon/VB Assets
Exhibit 1003
Page 690

determine the intent of the multi-modal natural language input using the shared knowledge relating to local intent determination capabilities.

[012] Other objects and advantages of the invention will be apparent based on the following drawings and detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

[013] Figure 1 illustrates a block diagram of an exemplary multi-modal electronic device that may be provided in an integrated, multi-device natural language voice services environment, according to various aspects of the invention.

[014] Figure 2 illustrates a block diagram of an exemplary centralized implementation of the integrated, multi-modal, multi-device natural language voice service environment, according to various aspects of the invention.

[015] Figure 3 illustrates a flow diagram of an exemplary method for processing multi-modal, natural language inputs at an input device in the centralized implementation of the integrated, multi-modal, multi-device natural language voice service environment, according to various aspects of the invention.

[016] Figure 4 illustrates a flow diagram of an exemplary method for processing multi-modal, natural language inputs at a central device in the centralized implementation of the integrated, multi-modal, multi-device natural language voice service environment, according to various aspects of the invention.

[017] Figure 5 illustrates a flow diagram of an exemplary method for processing multi-modal, natural language inputs at a secondary device in the centralized implementation of the integrated, multi-modal, multi-device natural language voice service environment, according to various aspects of the invention.

[018] Figure 6 illustrates a block diagram of an exemplary distributed implementation of the integrated, multi-modal, multi-device natural language voice service environment, according to various aspects of the invention.

[019] Figure 7 illustrates a flow diagram of an exemplary method for processing multi-modal, natural language inputs at an input device in the distributed implementation of the integrated, multi-modal, multi-device natural language voice service environment, according to various aspects of the invention.

Amazon/VB Assets
Exhibit 1003
Page 691

DETAILED DESCRIPTION

[020] According to various aspects of the invention, Figure 1 illustrates a block diagram of an exemplary multi-modal electronic device 100 that may be provided in a natural language voice services environment that includes one or more additional multi-modal devices (e.g., as illustrated in Figures 2 and 6). As will be apparent, the electronic device 100 illustrated in Figure 1 may be any suitable voice-enabled electronic device (e.g., a telematics device, a personal navigation device, a mobile phone, a VoIP node, a personal computer, a media device, an embedded device, a server, or another electronic device). The device 100 may include various components that collectively provide a capability to process conversational, multi-modal natural language inputs. As such, a user of the device 100 may engage in multi-modal conversational dialogues with the voice-enabled electronic device 100 to resolve requests in a free form, cooperative manner.

[021] For example, the natural language processing components may support free form natural language utterances to liberate the user from restrictions relating to how commands, queries, or other requests should be formulated. Rather, the user may employ any manner of speaking that feels natural in order to request content or services available through the device 100 (e.g., content or services relating to telematics, communications, media, messaging, navigation, marketing, information retrieval, etc.). For instance, in various implementations, the device 100 may process natural language utterances utilizing techniques described in co-pending U.S. Patent Application 10/452,147, entitled "Systems and Methods for Responding to Natural Language Speech Utterance," filed June 3, 2003, and U.S. Patent Application 10/618,633, entitled "Mobile Systems and Methods for Responding to Natural Language Speech Utterance," filed June 15, 2003, the disclosures of which are hereby incorporated by reference in their entirety.

[022] Moreover, because the device 100 may be deployed in an integrated multi-device environment, the user may further request content or services available through other devices deployed in the environment. In particular, the integrated voice services environment may include a plurality of multi-modal devices, each of which include natural language components generally similar to those illustrated in Figure 1. The various devices in the environment may serve distinct purposes, however, such that available content, services, applications, or other capabilities may vary among the devices in the environment (e.g., core functions of a media

Amazon/VB Assets
Exhibit 1003
Page 692

device may vary from those of a personal navigation device). Thus, each device in the environment, including device 100, may have knowledge of content, services, applications, intent determination capabilities, and other features available through the other devices by way of a constellation model 130b. Accordingly, as will be described in greater detail below, the electronic device 100 may cooperate with other devices in the integrated environment to resolve requests by sharing context, prior information, domain knowledge, short-term knowledge, long-term knowledge, and cognitive models, among other things.

[023] According to various aspects of the invention, the electronic device 100 may include an input mechanism 105 that can receive multi-modal natural language inputs, which include at least an utterance spoken by the user. As will be apparent, the input mechanism 105 may include any appropriate device or combination of devices capable of receiving a spoken input (e.g., a directional microphone, an array of microphones, or any other device that can generate encoded speech). Further, in various implementations, the input mechanism 105 can be configured to maximize fidelity of encoded speech, for example, by maximizing gain in a direction of the user, cancelling echoes, nulling point noise sources, performing variable rate sampling, or filtering environmental noise (e.g., background conversations). As such, the input mechanism 105 may generate encoded speech in a manner that can tolerate noise or other factors that could otherwise interfere with accurate interpretation of the utterance.

[024] Furthermore, in various implementations, the input mechanism 105 may include various other input modalities (i.e., the input mechanism 105 may be arranged in a multi-modal environment), in that non-voice inputs can be correlated and/or processed in connection with one or more previous, contemporaneous, or subsequent multi-modal natural language inputs. For example, the input mechanism 105 may be coupled to a touch-screen interface, a stylus and tablet interface, a keypad or keyboard, or any other suitable input mechanism, as will be apparent. As a result, an amount of information potentially available when processing the multi-modal inputs may be maximized, as the user can clarify utterances or otherwise provide additional information in a given multi-modal natural language input using various input modalities. For instance, in an exemplary illustration, the user could touch a stylus or other pointing device to a portion of a touch-screen interface of the device 100, while also providing an utterance relating to the touched portion of the interface (e.g., "Show me restaurants around here"). In this example, the natural language utterance may be correlated with the input received

Amazon/VB Assets
Exhibit 1003
Page 693

via the touch-screen interface, resulting in "around here" being interpreted in relation to the touched portion of the interface (e.g., as opposed to the user's current location or some other meaning).

[025] According to various aspects of the invention, the device 100 may include an Automatic Speech Recognizer 110 that generates one or more preliminary interpretations of the encoded speech, which may be received from the input mechanism 105. For example, the Automatic Speech Recognizer 110 may recognize syllables, words, or phrases contained in an utterance using one or more dynamically adaptable recognition grammars. The dynamic recognition grammars may be used to recognize a stream of phonemes through phonetic dictation based on one or more acoustic models. Furthermore, as described in co-pending U.S. Patent Application 11/197,504, entitled "Systems and Methods for Responding to Natural Language Speech Utterance," filed August 5, 2005, the disclosure of which is hereby incorporated by reference in its entirety, the Automatic Speech Recognizer 110 may be capable of multi-pass analysis, where a primary speech recognition engine may generate a primary interpretation of an utterance (e.g., using a large list dictation grammar) and request secondary transcription from one or more secondary speech recognition engines (e.g., using a virtual dictation grammar having decoy words for out-of-vocabulary words).

[026] Thus, the Automatic Speech Recognizer 110 may generate preliminary interpretations of an utterance in various ways, including exclusive use of a dictation grammar or virtual dictation grammar, or use of various combinations thereof (e.g., when the device 100 supports multi-pass analysis). In any event, the Automatic Speech Recognizer 110 may provide out-of-vocabulary capabilities and may tolerate portions of a speech signal being dropped, the user misspeaking, or other variables that may occur in natural language speech (e.g., stops and starts, stutters, etc.). Furthermore, the recognition grammars employed by the Automatic Speech Recognizer 110 may include vocabularies, dictionaries, syllables, words, phrases, or other information optimized according to various contextual or application-specific domains (e.g., navigation, music, movies, weather, shopping, news, languages, temporal or geographic proximities, or other suitable domains). Moreover, environmental knowledge (e.g., peer-to-peer affinities, capabilities of devices in the environment, etc.), historical knowledge (e.g., frequent requests, prior context, etc.), or other types of knowledge can be used to continually optimize the information contained in the recognition grammars on a dynamic basis.

Amazon/VB Assets
Exhibit 1003
Page 694

[027] For example, information contained in the recognition grammars may be dynamically optimized to improve a likelihood of a given utterance being recognized accurately (e.g., following an incorrect interpretation of a word, the incorrect interpretation may be removed from the grammar to reduce a likelihood of the incorrect interpretation being repeated). Accordingly, the Automatic Speech Recognizer 110 may use a number of techniques to generate preliminary interpretations of natural language utterances, including those described, for example, in co-pending U.S. Patent Application 11/513,269, entitled "Dynamic Speech Sharpening," filed August 31, 2006, the disclosure of which is hereby incorporated by reference in its entirety. Furthermore, the techniques used by the Automatic Speech Recognizer 110 associated with the device 100 may be considered in defining intent determination capabilities of the device 100, and such capabilities may be shared with other devices in the environment to enable convergence of speech recognition throughout the environment (e.g., because various devices may employ distinct speech recognition techniques or have distinct grammars or vocabularies, the devices may share vocabulary translation mechanisms to enhance system-wide recognition).

[028] According to various aspects of the invention, the Automatic Speech Recognizer 110 may provide one or more preliminary interpretations of a multi-modal input, including an utterance contained therein, to a conversational language processor 120. The conversational language processor 120 may include various components that collectively operate to model everyday human-to-human conversations in order to engage in cooperative conversations with the user to resolve requests based on the user's intent. For example, the conversational language processor 120 may include, among other things, an intent determination engine 130a, a constellation model 130b, one or more domain agents 130c, a context tracking engine 130d, a misrecognition engine 130e, and a voice search engine 130f. Furthermore, the conversational language processor 120 may be coupled to one or more data repositories 160 and applications associated with one or more domains. Thus, the intent determination capabilities of the device 100 may be defined based on the data and processing capabilities of the Automatic Speech Recognizer 110 and the conversational language processor 120.

[029] More particularly, the intent determination engine 130a may establish meaning for a given multi-modal natural language input based on a consideration of the intent determination capabilities of the device 100 as well as the intent determination capabilities of other devices in the integrated voice services environment. For example, the intent determination capabilities of

Amazon/VB Assets
Exhibit 1003
Page 695

the device 100 may be defined as a function of processing resources, storage for grammars, context, agents, or other data, and content or services associated with the device 100 (e.g., a media device with a small amount of memory may have a smaller list of recognizable songs than a device with a large amount of memory). Thus, the intent determination engine 130a may determine whether to process a given input locally (e.g., when the device 100 has intent determination capabilities that suggest favorable conditions for recognition), or whether to route information associated with the input to other devices, which may assist in determining the intent of the input.

[030] As such, to determine which device or combination of devices should process an input, the intent determination engine 130a may evaluate the constellation model 130b, which provides a model of the intent determination capabilities for each of the devices in the integrated voice services environment. For instance, the constellation model 130b may contain, among other things, knowledge of processing and storage resources available to each of the devices in the environment, as well as the nature and scope of domain agents, context, content, services, and other information available to each of the devices in the environment. As such, using the constellation model 130b, the intent determination engine 130a may be able to determine whether any of the other devices have intent determination capabilities that can be invoked to augment or otherwise enhance the intent determination capabilities of the device 100 (e.g., by routing information associated with a multi-modal natural language input to the device or devices that appear best suited to analyze the information and therefore determine an intent of the input). Accordingly, the intent determination engine 130a may establish the meaning of a given utterance by utilizing the comprehensive constellation model 130b that describes capabilities within the device 100 and across the integrated environment. The intent determination engine 130a may therefore optimize processing of a given natural language input based on capabilities throughout the environment (e.g., utterances may be processed locally to the device 100, routed to a specific device based on information in the constellation model 130b, or flooded to all of the devices in the environment in which case an arbitration may occur to select a best guess at an intent determination).

[031] Although the following discussion will generally focus on various techniques that can be used to determine the intent of multi-modal natural language inputs in the integrated multi-device environment, it will be apparent that the natural language processing capabilities of any

Amazon/VB Assets
Exhibit 1003
Page 696

one of the devices may extend beyond the specific discussion that has been provided herein. As such, in addition to the co-pending U.S. Patent Applications referenced above, further natural language processing capabilities that may be employed include those described in co-pending U.S. Patent Application 11/197,504, entitled "Systems and Methods for Responding to Natural Language Speech Utterance," filed August 5, 2005, U.S. Patent Application 11/200,164, entitled "System and Method of Supporting Adaptive Misrecognition in Conversational Speech," filed August 10, 2005, U.S. Patent Application 11/212,693, entitled "Mobile Systems and Methods of Supporting Natural Language Human-Machine Interactions," filed August 29, 2005, U.S. Patent Application 11/580,926, entitled "System and Method for a Cooperative Conversational Voice User Interface," filed October 16, 2006, U.S. Patent Application 11/671,526, entitled "System and Method for Selecting and Presenting Advertisements Based on Natural Language Processing of Voice-Based Input," filed February 6, 2007, and U.S. Patent Application 11/954,064, entitled "System and Method for Providing a Natural Language Voice User Interface in an Integrated Voice Navigation Services Environment," filed December 11, 2007, the disclosures of which are hereby incorporated by reference in their entirety.

[032] According to various aspects of the invention, Figure 2 illustrates a block diagram of an exemplary centralized implementation of the integrated, multi-modal, multi-device natural language voice service environment. As will be apparent from the further description to be provided herein, the centralized implementation of the integrated, multi-device voice services environment may enable a user to engage in conversational, multi-modal natural language interactions with any one of voice-enabled devices 210a-n or central voice-enabled device 220. As such, the multi-device voice services environment may collectively determine intent for any given multi-modal natural language input, whereby the user may request content or voice services relating to any device or application in the environment, without restraint.

[033] As illustrated in Figure 2, the centralized implementation of the multi-device voice service environment may include a plurality of voice-enabled devices 210a-n, each of which include various components capable of determining intent of natural language utterances, as described above in reference to Figure 1. Furthermore, as will be apparent, the centralized implementation includes a central device 220, which contains information relating to intent determination capabilities for each of the other voice-enabled devices 210a-n. For example, in various exemplary implementations, the central device 220 may be designated as such by virtue

Amazon/VB Assets
Exhibit 1003
Page 697

of being a device most capable of determining the intent of an utterance (e.g., a server, home data center, or other device having significant processing power, memory resources, and communication capabilities making the device suitable to manage intent determination across the environment). In another exemplary implementation, the central device 220 may be dynamically selected based on one or more characteristics of a given multi-modal natural language input, dialogue, or interaction (e.g., a device may be designated as the central device 220 when a current utterance relates to a specific domain).

[034] In the centralized implementation illustrated in Figure 2, a multi-modal natural language input may be received at one of the voice-enabled devices 210a-n. Therefore, the receiving one of the devices 210a-n may be designated as an input device for that input, while the remaining devices 210a-n may be designated as secondary devices for that input. In other words, for any given multi-modal natural language input, the multi-device environment may include an input device that collects the input, a central device 220 that aggregates intent determination, inferencing, and processing capabilities for all of the devices 210a-n in the environment, and one or more secondary devices that may also be used in the intent determination process. As such, each device 210 in the environment may be provided with a constellation model that identifies all of the devices 210 having incoming and outgoing communication capabilities, thus indicating an extent to which other devices may be capable of determining intent for a given multi-modal natural language input. The constellation model may further define a location of the central device 220, which aggregates context, vocabularies, content, recognition grammars, misrecognitions, shared knowledge, intent determination capabilities, inferencing capabilities, and other information from the various devices 210a-n in the environment.

[035] Accordingly, as communication and processing capabilities permit, the central device 220 may be used as a recognizer of first or last resort. For example, because the central device 220 converges intent determination capabilities across the environment (e.g., by aggregating context, vocabularies, device capabilities, and other information from the devices 210a-n in the environment), inputs may be automatically routed to the central device 220 when used as a recognizer of first resort, or as a recognizer of last resort when local processing at the input device 210 cannot determine the intent of the input with a satisfactory level of confidence. However, it will also be apparent that in certain instances the input device 210 may be unable to make contact with the central device 220 for various reasons (e.g., a network connection may be

Amazon/VB Assets
Exhibit 1003
Page 698

unavailable, or a processing bottleneck at the central device 220 may cause communication delays). In such cases, the input device 210 that has initiated contact with the central device 220 may shift into decentralized processing (e.g., as described in reference to Figure 6) and communicate capabilities with one or more of the other devices 210a-n in the constellation model. Thus, when the central device 220 cannot be invoked for various reasons, the remaining devices 210a-n may operate as cooperative nodes to determine intent in a decentralized manner.

[036] Additionally, in the multi-device voice services environment, the central device 220 and the various other devices 210a-n may cooperate to create a converged model of capabilities throughout the environment. For example, as indicated above, in addition to having intent determination capabilities based on processing resources, memory resources, and device capabilities, each of the devices 210a-n and the central device 220 may include various other natural language processing components. The voice services environment may therefore operate in an integrated manner by maintaining not only a complete model of data, content, and services associated with the various devices 210a-n, but also of other natural language processing capabilities and dynamic states associated with the various devices 210a-n. As such, the various devices 210a-n may operate with a goal of converging capabilities, data, states, and other information across the device, either on one device (e.g., the central device 220) or distributed among the various devices 210a-n (e.g., as in the decentralized implementation to be described in Figure 6).

[037] For example, as discussed above, each device 210 includes an Automatic Speech Recognizer, one or more dynamically adaptable recognition grammars, and vocabulary lists used to generate phonemic interpretations of natural language utterances. Moreover, each device 210 includes locally established context, which can range from information contained in a context stack, context and namespace variables, vocabulary translation mechanisms, short-term shared knowledge relating to a current dialogue or conversational interaction, long-term shared knowledge relating to a user's learned preferences over time, or other contextual information. Furthermore, each device 210 may have various services or applications associated therewith, and may perform various aspects of natural language processing locally. Thus, additional information to be converged throughout the environment may include partial or preliminary utterance recognitions, misrecognitions or ambiguous recognitions, inferencing capabilities, and

Amazon/VB Assets
Exhibit 1003
Page 699

overall device state information (e.g., songs playing in the environment, alarms set in the environment, etc.).

[038] Thus, various data synchronization and referential integrity algorithms may be employed in concert by the various devices 210a-n and the central device 220 to provide a consistent worldview of the environment. For example, information may be described and transmitted throughout the environment for synchronization and convergence purposes using the Universal Plug and Play protocol designed for computer ancillary devices, although the environment can also operate in a peer-to-peer disconnected mode (e.g., when the central device 220 cannot be reached). However, in various implementations, the environment may also operate in a peer-to-peer mode regardless of the disconnected status, as illustrated in Figure 6, for example, when the devices 210a-n have sufficient commensurate resources and capabilities for natural language processing.

[039] In general, the algorithms for convergence in the environment can be executed at various intervals, although it may be desirable to limit data transmission so as to avoid processing bottlenecks. For example, because the convergence and synchronization techniques relate to natural language processing, in which any given utterance will typically be expressed over a course of several seconds, information relating to context and vocabulary need not be updated on a time frame of less than a few seconds. However, as communication capabilities permit, context and vocabulary could be updated more frequently to provide real-time recognition or the appearance of real-time recognition. In another implementation, the convergence and synchronization may be permitted to run until completion (e.g., when no requests are currently pending), or the convergence and synchronization may be suspended or terminated when a predetermined time or resource consumption limit has been reached (e.g., when the convergence relates to a pending request, an intent determination having a highest confidence level at the time of cut-off may be used).

[040] By establishing a consistent view of capabilities, data, states, and other information throughout the environment, an input device 210 may cooperate with the central device 220 and one or more secondary devices (i.e., one or more of devices 210a-n, other than the input device) in processing any given multi-modal natural language input. Furthermore, by providing each device 210 and the central device 220 with a constellation model that describes a synchronized state of the environment, the environment may be tolerant of failure by one or more of the

Amazon/VB Assets
Exhibit 1003
Page 700

devices 210a-n, or of the central device 220. For example, if the input device 210 cannot communicate with the central device 220 (e.g., because of a server crash), the input device 210 may enter a disconnected peer-to-peer mode, whereby capabilities can be exchanged with one or more devices 210a-n with which communications remain available. To that end, when a device 210 establishes new information relating to vocabulary, context, misrecognitions, agent adaptation, intent determination capabilities, inferencing capabilities, or otherwise, the device 210 may transmit the information to the central device 220 for convergence purposes, as discussed above, in addition to consulting the constellation model to determine whether the information should be transmitted to one or more of the other devices 210a-n.

[041] For example, suppose the environment includes a voice-enabled mobile phone that has nominal functionality relating to playing music or other media, and which further has a limited amount of local storage space, while the environment further includes a voice-enabled home media system that includes a mass storage medium that provides dedicated media functionality. If the mobile phone were to establish new vocabulary, context, or other information relating to a song (e.g., a user downloads the song or a ringtone to the mobile phone while on the road), the mobile phone may transmit the newly established information to the home media system in addition to the central device 220. As such, by having a model of all of the devices 210a-n in the environment and transmitting new information to the devices where it will most likely be useful, the various devices may handle disconnected modes of operation when the central device 220 may be unavailable for any reason, while resources may be allocated efficiently throughout the environment.

[042] Thus, based on the foregoing discussion, it will be apparent that a centralized implementation of an integrated multi-device voice services environment may generally include a central device 220 operable to aggregate or converge knowledge relating to content, services, capabilities, and other information associated with various voice-enabled devices 210a-n deployed within the environment. In such centralized implementations, the central device 220 may be invoked as a recognizer of first or last resort, as will be described in greater detail with reference to Figures 3-5, and furthermore, the other devices 210a-n in the environment may be configured to automatically enter a disconnected or peer-to-peer mode of operation when the central device 220 cannot be invoked for any reason (i.e., devices may enter a decentralized or distributed mode, as will be described in greater detail with reference to Figures 6-7).

Amazon/VB Assets
Exhibit 1003
Page 701

Knowledge and capabilities of each of the devices 210a-n may therefore be made available throughout the voice services environment in a centralized manner, a distributed manner, or various combinations thereof, thus optimizing an amount of natural language processing resources used to determine an intent of any given multi-modal natural language input.

[043] According to various aspects of the invention, Figure 3 illustrates a flow diagram of an exemplary method for processing multi-modal, natural language inputs at an input device in the centralized implementation of the integrated, multi-modal, multi-device natural language voice service environment. Similarly, Figures 4 and 5 illustrate corresponding methods associated with a central device and one or more secondary devices, respectively, in the centralized voice service environment. Furthermore, it will be apparent that the processing techniques described in relation to Figures 3-5 may generally be based on the centralized implementation illustrated in Figure 2 and described above, whereby the input device may be assumed to be distinct from the central device, and the one or more secondary devices may be assumed to be distinct from the central device and the input device. However, it will be apparent that various instances may involve a natural language input being received at the central device or at another device, in which case the techniques described in Figures 3-5 may be vary depending on circumstances of the environment (e.g., decisions relating to routing utterances to a specific device or devices may be made locally, collaboratively, or in other ways depending on various factors, such as overall system state, communication capabilities, intent determination capabilities, or otherwise).

[044] As illustrated in Figure 3, a multi-modal natural language input may be received at an input device in an operation 310. The multi-modal input may include at least a natural language utterance provided by a user, and may further include other input modalities such as audio, text, button presses, gestures, or other non-voice inputs. It will also be apparent that prior to receiving the natural language input in operation 310, the input device may be configured to establish natural language processing capabilities. For example, establishing natural language processing capabilities may include, among other things, loading an Automatic Speech Recognizer and any associated recognition grammars, launching a conversational language processor to handle dialogues with the user, and installing one or more domain agents that provide functionality for respective application domains or contextual domains (e.g., navigation, music, movies, weather, information retrieval, device control, etc.).

Amazon/VB Assets
Exhibit 1003
Page 702

[045] The input device may also be configured to coordinate synchronization of intent determination capabilities, shared knowledge, and other information with the central device and the secondary devices in the environment prior to receiving the input at operation 310. For example, when the input device installs a domain agent, the installed domain agent may bootstrap context variables, semantics, namespace variables, criteria values, and other context related to that agent from other devices in the system. Similarly, misrecognitions may be received from the central device and the secondary devices in order to enable correction of agents that use information relevant to the received misrecognitions, and vocabularies and associated translation mechanisms may be synchronized among the devices to account for potential variations between the Automatic Speech Recognizers used by the various devices (e.g., each device in the environment cannot be guaranteed to use the same Automatic Speech Recognizer or recognition grammars, necessitating vocabulary and translation mechanisms to be shared among the devices that share intent determination capabilities).

[046] Upon establishing and synchronizing natural language processing capabilities and subsequently receiving a multi-modal natural language input in operation 310, the input device may determine whether the environment has been set up to automatically transmit the input to the central device in a decisional operation 320. In such a case, processing proceeds to an operation 360 for transmitting the input to the central device, which may then process the input according to techniques to be described in relation to Figure 4. If the environment has not been set up to automatically communicate the input to the central device, however, processing proceeds to an operation 330, where the input device performs transcription of the natural language utterance contained in the multi-modal input. For example, the input device may transcribe the utterance using the Automatic Speech Recognizer and recognition grammars associated therewith according to techniques described above and in the above-referenced U.S. Patent Applications.

[047] Subsequently, in an operation 340, an intent of the multi-modal natural language input may be determined at the input device using local natural language processing capabilities and resources. For example, any non-voice input modalities included in the input may be merged with the utterance transcription and a conversational language processor associated with the input device may utilize local information relating to context, domain knowledge, shared knowledge, context variables, criteria values, or other information useful in natural language

Amazon/VB Assets
Exhibit 1003
Page 703

processing. As such, the input device may attempt to determine a best guess as to an intent of the user that provided the input, such as identifying a conversation type (e.g., query, didactic, or exploratory) or request that may be contained in the input (e.g., a command or query relating to one or more domain agents or application domains).

[048] The intent determination of the input device may be assigned a confidence level (e.g., a device having an Automatic Speech Recognizer that implements multi-pass analysis may assign comparatively higher confidence levels to utterance transcriptions created thereby, which may result in a higher confidence level for the intent determination). The confidence level may be assigned based on various factors, as described in the above-referenced U.S. Patent Applications. As such, a decisional operation 350 may include determining whether the intent determination of the input device meets an acceptable level of confidence. When the intent determination meets the acceptable level confidence, processing may proceed directly to an operation 380 where action may be taken in response thereto. For example, when the intent determination indicates that the user has requested certain information, one or more queries may be formulated to retrieve the information from appropriate information sources, which may include one or more of the other devices. In another example, when the intent determination indicates that the user has requested a given command (e.g., to control a specific device), the command may be routed to the appropriate device for execution.

[049] Thus, in cases where the input device can determine the intent of a natural language input without assistance from the central device or the secondary devices, communications and processing resources may be conserved by taking immediate action as may be appropriate. On the other hand, when the intent determination of the input device does not meet the acceptable level of confidence, decisional operation 350 may result in the input device requesting assistance from the central device in operation 360. In such a case, the multi-modal natural language input may be communicated to the central device in its entirety, whereby the central device processes the input according to techniques described in Figure 4. However, should transmission to the central device fail for some reason, the input device may shift into a disconnected peer-to-peer mode where one or more secondary devices may be utilized, as will be described below in relation to Figure 7. When transmission to the central device occurs without incident, however, the input device may receive an intent determination from the central device in an operation 370, and may further receive results of one or more requests that the central device was able to

Amazon/VB Assets
Exhibit 1003
Page 704

resolve, or requests that the central device has formulated for further processing on the input device. As such, the input device may take action in operation 380 based on the information received from the central device in operation 370. For example, the input device may route queries or commands to local or remote information sources or devices based on the intent determination, or may present results of the requests processed by the central device to the user.

[050] Referring to Figure 4, the central device may receive the multi-modal natural language input from the input device in an operation 410. The central device, having aggregated context and other knowledge from throughout the environment, may thus transcribe the utterance in an operation 420 and determine an intent of the input from the transcribed utterance in an operation 430. As such, the central device may consider information relating to context, domain agents, applications, and device capabilities throughout the environment in determining the intent of the utterance, including identification of one or more domains relevant to the input. However, it will be apparent that utilizing information aggregated from throughout the environment may cause ambiguity or uncertainty in various instances (e.g., an utterance containing the word "traffic" may have a different intent in domains relating to movies, music, and navigation).

[051] As such, once the central device has attempted to determine the intent of the natural language input, a determination may be made in an operation 440 as to whether one or more secondary devices (i.e., other devices in the constellation besides the input device) may also be capable of intent determination in the identified domain or domains. When no such secondary devices can be identified, decisional operation 440 may branch directly to an operation 480 to send to the input device the determined intent and any commands, queries, or other requests identified from the determined intent.

[052] On the other hand, when one or more secondary devices in the environment have intent determination capabilities in the identified domain or domains, the natural language input may be sent to such secondary devices in an operation 450. The secondary devices may then determine an intent as illustrated in Figure 5, which may include techniques generally similar to those described above in relation to the input device and central device (i.e., the natural language input may be received in an operation 510, an utterance contained therein may be transcribed in an operation 520, and an intent determination made in an operation 530 may be returned to the central device in an operation 540).

Amazon/VB Assets
Exhibit 1003
Page 705

[053] Returning to Figure 4, the central device may collate intent determination responses received from the secondary devices in an operation 460. For example, as indicated above, the central device may identify one or more secondary devices capable of determining intent in a domain that the central device has identified as being relevant to the natural language utterance. As will be apparent, the secondary devices invoked in operation 450 may often include a plurality of devices, and intent determination responses may be received from the secondary devices in an interleaved manner, depending on processing resources, communications throughput, or other factors (e.g., the secondary devices may include a telematics device having a large amount of processing power and a broadband network connection and an embedded mobile phone having less processing power and only a cellular connection, in which case the telematics device may be highly likely to provide results to the central device before the embedded mobile phone). Thus, based on potential variations in response time of secondary devices, the central device may be configured to place constraints on collating operation 460. For example, the collating operation 460 may be terminated as soon as an intent determination has been received from one of the secondary devices that meets an acceptable level of confidence, or the operation 460 may be cut off when a predetermined amount of time has lapsed or a predetermined amount of resources have been consumed. In other implementations, however, it will be apparent that collating operation 460 may be configured to run to completion, regardless of whether delays have occurred or suitable intent determinations have been received. Further, it will be apparent that various criteria may be used to determine whether or when to end the collating operation 460, including the nature of a given natural language input, dialogue, or other interaction, or system or user preferences, among other criteria, as will be apparent.

[054] In any event, when the collating operation 460 has completed, a subsequent operation 470 may include the central device arbitrating among the intent determination responses received from one or more of the secondary devices previously invoked in operation 450. For example, each of the invoked secondary devices that generate an intent determination may also assign a confidence level to that intent determination, and the central device may consider the confidence levels in arbitrating among the responses. Moreover, the central device may associate other criteria with the secondary devices or the intent determinations received from the secondary devices to further enhance a likelihood that the best intent determination will be used. For example, various ones of the secondary devices may only be invoked for partial recognition in

Amazon/VB Assets
Exhibit 1003
Page 706

distinct domains, and the central device may aggregate and arbitrate the partial recognitions to create a complete transcription. In another example, a plurality of secondary devices may be invoked to perform overlapping intent determination, and the central device may consider capabilities of the secondary devices to weigh the respective confidence levels (e.g., when one of two otherwise identical secondary devices employs multi-pass speech recognition analysis, the secondary device employing the multi-pass speech recognition analysis may be weighed as having a higher likelihood of success). It will be apparent that the central device may be configured to arbitrate and select one intent determination from among all of the intent hypotheses, which may include the intent determination hypothesis generated by the central device in operation 430. Upon selecting the best intent determination hypothesis, the central device may then provide that intent determination to the input device in operation 480, as well as any commands, queries, or other requests that may be relevant thereto. The input device may then take appropriate action as described above in relation to Figure 3.

[055] According to various aspects of the invention, Figure 6 illustrates a block diagram of an exemplary distributed implementation of the integrated, multi-modal, multi-device natural language voice service environment. As described above, the distributed implementation may also be categorized as a disconnected or peer-to-peer mode that may be employed when a central device in a centralized implementation cannot be reached or otherwise does not meet the needs of the environment. The distributed implementation illustrated in Figure 6 may be generally operate with similar purposes as described above in relation to the centralized implementation (i.e., to ensure that the environment includes a comprehensive model of aggregate knowledge and capabilities of a plurality of devices 610a-n in the environment). Nonetheless, the distributed implementation may operate in a somewhat different manner, in that one or more of the devices 610a-n may be provided with the entire constellation model, or various aspects of the model may be distributed among the plurality of devices 610a-n, or various combinations thereof.

[056] Generally speaking, the plurality of voice-enabled devices 610a- may be coupled to one another by a voice services interface 630, which may include any suitable real or virtual interface (e.g., a common message bus or network interface, a service-oriented abstraction layer, etc.). The various devices 610a-n may therefore operate as cooperative nodes in determining intent for multi-modal natural language utterances received by any one of the devices 610. Furthermore,

Amazon/VB Assets
Exhibit 1003
Page 707

the devices 610a-n may share knowledge of vocabularies, context, capabilities, and other information, while certain forms of data may be synchronized to ensure consistent processing among the devices 610a-n. For example, because natural language processing components used in the devices 610a-n may vary (e.g., different recognition grammars or speech recognition techniques may exist), vocabulary translation mechanisms, misrecognitions, context variables, criteria values, criteria handlers, and other such information used in the intent determination process should be synchronized to the extent that communication capabilities permit.

[057] By sharing intent determination capabilities, device capabilities, inferencing capabilities, domain knowledge, and other information, decisions as to routing an utterance to a specific one of the devices 610a-n may be made locally (e.g., at an input device), collaboratively (e.g., a device having particular capabilities relevant to the utterance may communicate a request to process the utterance), or various combinations thereof (e.g., the input device may consider routing to secondary devices only when an intent of the utterance cannot be determined). Similarly, partial recognition performed at one or more of the devices 610a-n may be used to determine routing strategies for further intent determination of the utterance. For example, an utterance that contains a plurality of requests relating to a plurality of different domains may be received at an input device that can only determine intent in one of the domains. In this example, the input device may perform partial recognition for the domain associated with the input device, and the partial recognition may also identify the other domains relevant to the utterance for which the input device does not have sufficient recognition information. Thus, the partial recognition performed by the input device may result in identification of other potentially relevant domains and a strategy may be formulated to route the utterance to other devices in the environment that include recognition information for those domains.

[058] As a result, multi-modal natural language inputs, including natural language utterances, may be routed among the various devices 610a-n in order to perform intent determination in a distributed manner. However, as the capabilities and knowledge held by any one of the devices 610a-n may vary, each of the devices 610a-n may be associated with a reliability factor for intent determinations generated by the respective devices 610a-n. As such, to ensure that final intent determinations can be relied upon with a sufficient level of confidence, knowledge may be distributed among the devices 610a-n to ensure that reliability metrics for intent determinations provided by each of the devices 610a-n are commensurable throughout the environment. For

Amazon/VB Assets
Exhibit 1003
Page 708

example, additional knowledge may be provided to a device having a low intent determination reliability, even when such knowledge results in redundancy in the environment, to ensure commensurate reliability of intent determination environment-wide.

[059] Therefore, in distributed implementations of the integrated voice services environment, utterances may be processed in various ways, which may depend on circumstances at a given time (e.g., system states, system or user preferences, etc.). For example, an utterance may be processed locally at an input device and only routed to secondary devices when an intent determination confidence level falls below a given threshold. In another example, utterances may be routed to a specific device based on the modeling of knowledge and capabilities discussed above. In yet another example, utterances may be flooded among all of the devices in the environment, and arbitration may occur whereby intent determinations may be collated and arbitrated to determine a best guess at intent determination.

[060] Thus, utterances may be processed in various ways, including through local techniques, centralized techniques, distributed techniques, and various combinations thereof. Although many variations will be apparent, Figure 7 illustrates an exemplary method for combined local and distributed processing of multi-modal, natural language inputs in a distributed implementation of the voice service environment, according to various aspects of the invention. In particular, the distributed processing may begin in an operation 710, where a multi-modal natural language input may be received at an input device. The input device may then utilize various natural language processing capabilities associated therewith in an operation 720 to transcribe an utterance contained in the multi-modal input (e.g., using an Automatic Speech Recognizer and associated recognition grammars), and may subsequently determine a preliminary intent of the multi-modal natural language input in an operation 730. It will be apparent that operations 710 through 730 may generally be performed using local intent determination capabilities associated with the input device.

[061] Thereafter, the input device may invoke intent determination capabilities of one or more secondary devices in an operation 740. More particularly, the input device may provide information associated with the multi-modal natural language input to one or more of the secondary devices, which may utilize local intent determination capabilities to attempt to determine intent of the input using techniques as described in relation to Figure 5. It will also be apparent that, in various implementations, the secondary devices invoked in operation 740 may

Amazon/VB Assets
Exhibit 1003
Page 709