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.
14/149,296	01/12/2021	10891968	Q203469	4038

23373 7590 12/22/2020
SUGHRUE MION, PLLC
2000 PENNSYLVANIA AVENUE, N.W.
SUITE 900
WASHINGTON, DC 20006

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 0 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):

SAMSUNG ELECTRONICS CO., LTD., Suwon-si, KOREA, REPUBLIC OF;
Ji-hye CHUNG, Seoul, KOREA, REPUBLIC OF;
Cheong-jae LEE, Suwon-si, KOREA, REPUBLIC OF;
Hye-jeong LEE, Seoul, KOREA, REPUBLIC OF;
Yong-wook SHIN, Osan-si, KOREA, REPUBLIC OF;

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.

IR103 (Rev. 10/09)

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.
14/149,296	01/07/2014	Ji-hye CHUNG	Q203469	4038

23373 7590 09/17/2020
SUGHRUE MION, PLLC
2000 PENNSYLVANIA AVENUE, N.W.
SUITE 900
WASHINGTON, DC 20006

EXAMINER
OPSASNICK, MICHAEL N

ART UNIT	PAPER NUMBER
2658	

NOTIFICATION DATE	DELIVERY MODE
09/17/2020	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):

PPROCESSING@SUGHRUE.COM
USPTO@sughrue.com
sughrue@sughrue.com

PTOL-90A (Rev. 04/07)

Supplemental ***Notice of Allowability***	**Application No.** 14/149,296	**Applicant(s)** CHUNG et al.	
	Examiner Michael N Opsasnick	**Art Unit** 2658	**AIA (FITF) 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 IDS filed 8/31/2020.
   ☐ 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 See Continuation Sheet . 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 _____.
3. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material _____.
4. ☐ Interview Summary (PTO-413), Paper No./Mail Date. _____.
5. ☐ Examiner's Amendment/Comment
6. ☐ Examiner's Statement of Reasons for Allowance
7. ☐ Other _____.

/Michael N Opsasnick/
Primary Examiner, Art Unit 2658

Continuation of 3. The allowed claim(s) is/are: 23-38,40-43,45-57,59-67 and 69

Notice of References Cited	Application/Control No. 14/149,296	Applicant(s)/Patent Under Reexamination CHUNG et al.	
	Examiner Michael N Opsasnick	Art Unit 2658	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	CPC Classification	US Classification
*	A	US-20130321340-A1	12-2013	Seo; Joon-kyu	G06F3/04883	345/174
	B					
	C					
	D					
	E					
	F					
	G					
	H					
	I					
	J					
	K					
	L					
	M					

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	CPC 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 STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	14/149,296
	Confirmation Number	4038
	Filing Date	January 7, 2014
	First Named Inventor	Ji-hye CHUNG
	Art Unit	2658
	Examiner Name	OPSASNICK, MICHAEL N
	Attorney Docket Number	Q203469

U.S. PATENTS

Examiner Initials*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines, where Relevant Passages or Relevant Figures Appear

U.S. PATENT APPLICATION PUBLICATIONS

Examiner Initials*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines, where Relevant Passages or Relevant Figures Appear

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No	Foreign Document Number[3]	Country Code[2]	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines, where Relevant Passages or Relevant Figures Appear	T[5]

NON-PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[5]
	1.	Communication dated September 26, 2019 issued by the State Intellectual Property Office of P.R. China in counterpart Chinese Application No. 201410006748.2	English translation

EXAMINER SIGNATURE

Examiner Signature	/Michael N Opsasnick/	Date Considered	12/16/2019

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /m.n.o./

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

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	14/149,296
	Confirmation Number	4038
	Filing Date	January 7, 2014
	First Named Inventor	Ji-hye CHUNG
	Art Unit	2658
	Examiner Name	OPSASNICK, MICHAEL N
	Attorney Docket Number	Q203469

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

☑ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☐ That no item of information contained in the 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 the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

☐ That each item of information contained in the Information Disclosure Statement filed concurrently herewith: (i) was first cited in any communication from a patent office in a counterpart foreign or international application or from the Office, and this communication was not received by any individual designated in § 1.56(c) more than thirty days prior to the filing of the Information Disclosure Statement; or (ii) is a communication that was issued by a patent office in a counterpart foreign or international application or by the Office, and this communication was not received by any individual designated in § 1.56(c) more than thirty days prior to the filing of the Information Disclosure Statement. See 37 CFR 1.704(d).

☐ The fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☑ The USPTO is directed and authorized to charge all required fees, except for the Issue Fee and the Publication Fee, to Deposit Account No. 19-4880. Please also credit any overpayments to said Deposit Account.

☐ A certification statement is not submitted herewith.

WASHINGTON OFFICE
23373
CUSTOMER NUMBER

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/Peter A. McKenna/	Date (YYYY-MM-DD)	2019-12-11
Name/Print	Peter A. McKenna	Registration Number	38,551

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /m.n.o./

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

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

Notice of References Cited	Application/Control No. 14/149,296	Applicant(s)/Patent Under Reexamination CHUNG ET AL.	
	Examiner Michael N. Opsasnick	Art Unit 2658	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	CPC Classification	US Classification
*	A	US-2012/0121077 A1	05-2012	Gabay; Gil	H04L65/1063	379/211.02
*	B	US-2008/0313039 A1	12-2008	ALTBERG; Ebbe	G06Q20/10	705/14.54
*	C	US-2008/0183852 A1	07-2008	Pramer; David M.	G06Q30/02	709/223
*	D	US-2007/0121596 A1	05-2007	Kurapati; Krishna	H04L29/06027	370/356
*	E	US-2009/0018832 A1	01-2009	Mukaigaito; Takeya	G10L15/1815	704/251
*	F	US-2003/0235183 A1	12-2003	Skelton, Jeffrey S.	H04M3/562	370/352
	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	CPC 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 STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)	Application Number	14/149,296
	Confirmation Number	4038
	Filing Date	January 7, 2014
	First Named Inventor	Ji-hye CHUNG
	Art Unit	2658
	Examiner Name	OPSASNICK, MICHAEL N
	Attorney Docket Number	Q203469

U.S. PATENTS

Examiner Initials*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines, where Relevant Passages or Relevant Figures Appear
	1.	8073681	B2	2011-12-06	Baldwin et al.	
	2.	8515765	B2	2013-08-20	Baldwin et al.	
	3.	9015049	B2	2015-04-21	Baldwin et al.	

U.S. PATENT APPLICATION PUBLICATIONS

Examiner Initials*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines, where Relevant Passages or Relevant Figures Appear
	4.	20120022857	A1	2012-01-26	BALDWIN et al.	
	5.	20130339022	A1	2013-12-19	Baldwin et al.	
	6.	20150228276	A1	2015-08-13	BALDWIN et al.	

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No	Foreign Document Number[3]	Country Code[2]	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages, Columns, Lines, where Relevant Passages or Relevant Figures Appear	T[5]
	7.	101535983	CN	A	2009-09-16	VOICEBOX TECHNOLOGIES INC		Abstract; Corresponds to US 20120022857 A1, US 20130339022 A1, US 20150228276 A1, US 8073681 B2, US 8515765 B2, & US 9015049 B2; Cited in CN Communication issued 11/6/17 in CN App No. 201410006748.2
	8.	1 033 701	EP	A2	2000-09-06	MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.		

NON-PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city, and/or country where published.	T[5]
	9.	Communication issued November 6, 2017 by the State Intellectual Property Office of P.R. China in counterpart Chinese Patent Application No. 201410006748.2	Y

EXAMINER SIGNATURE

Examiner Signature		Date Considered	

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

1 See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. 2 Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). 3 For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. 4 Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. 5 Applicant is to place a check mark here if English language translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)	Application Number	14/149,296
	Confirmation Number	4038
	Filing Date	January 7, 2014
	First Named Inventor	Ji-hye CHUNG
	Art Unit	2658
	Examiner Name	OPSASNICK, MICHAEL N
	Attorney Docket Number	Q203469

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

☑ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☐ That no item of information contained in the 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 the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

☐ That each item of information contained in the Information Disclosure Statement filed concurrently herewith: (i) was first cited in any communication from a patent office in a counterpart foreign or international application or from the Office, and this communication was not received by any individual designated in § 1.56(c) more than thirty days prior to the filing of the Information Disclosure Statement; or (ii) is a communication that was issued by a patent office in a counterpart foreign or international application or by the Office, and this communication was not received by any individual designated in § 1.56(c) more than thirty days prior to the filing of the Information Disclosure Statement. See 37 CFR 1.704(d).

☐ The fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☑ The USPTO is directed and authorized to charge all required fees, except for the Issue Fee and the Publication Fee, to Deposit Account No. 19-4880. Please also credit any overpayments to said Deposit Account.

☐ None

WASHINGTON OFFICE
23373
CUSTOMER NUMBER

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/Peter A. McKenna/	Date (YYYY-MM-DD)	2018-01-31
Name/Print	Peter A. McKenna	Registration Number	38,551

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

1 See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. 2 Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). 3 For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. 4 Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. 5 Applicant is to place a check mark here if English language translation is attached.

Espacenet

Bibliographic data: CN101535983 (A) — 2009-09-16

System and method for a cooperative conversational voice user interface

Inventor(s): LARRY BALDWIN [US]; TOM FREEMAN [US]; MICHAEL TJALVE [US]; BLANE EBERSOLD [US]; CHRIS WEIDER [US] ± (BALDWIN LARRY, ; FREEMAN TOM, ; TJALVE MICHAEL, ; EBERSOLD BLANE, ; WEIDER CHRIS)

Applicant(s): VOICEBOX TECHNOLOGIES INC [US] ± (VOICEBOX TECHNOLOGIES INC)

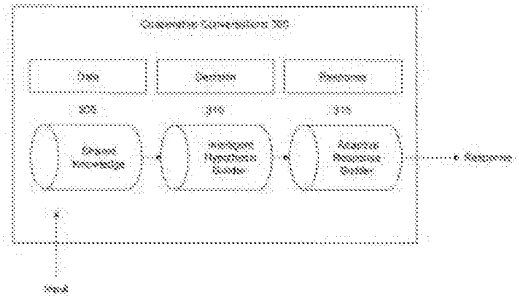
Classification: - international: *G06F17/28*
- cooperative: G06F3/167; G10L15/18; G10L15/22; G10L17/22; G10L15/183; G10L2015/0631

Application number: CN2007842315 20071016

Priority number(s): WO2007US81481 20071016 ; US20060580926 20061016

Also published as: CN101535983 (B) EP2082335 (A2) EP2082335 (A4) US2008091406 (A1) US2012022857 (A1) US2013339022 (A1) US2015228276 (A1) US8073681 (B2) US8515765 (B2) US9015049 (B2) WO2008118195 (A2) WO2008118195 (A3) less

Abstract of CN101535983 (A)



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.

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
G06F 17/28 (2006.01)

[12] 发明专利申请公布说明书

[21] 申请号 200780042315.0

[43] 公开日 2009 年 9 月 16 日

[11] 公开号 CN 101535983A

[22] 申请日 2007.10.16
[21] 申请号 200780042315.0
[30] 优先权
[32] 2006.10.16 [33] US [31] 11/580，926
[86] 国际申请 PCT/US2007/081481 2007.10.16
[87] 国际公布 WO2008/118195 英 2008.10.2
[85] 进入国家阶段日期 2009.5.14
[71] 申请人 沃伊斯博克斯科技公司
地址 美国华盛顿
[72] 发明人 L·贝尔德文 T·弗莱曼
M·特加尔弗 B·艾弗尔索德
C·威德尔
[74] 专利代理机构 中国国际贸易促进委员会专利商标事务所
代理人 李向英

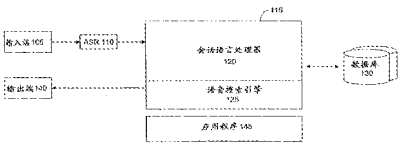
权利要求书 6 页 说明书 26 页 附图 3 页

[54] 发明名称

协作会话语音用户界面的系统和方法

[57] 摘要

提供了协作会话语音用户界面。 协作会话语音用户界面可以基于短期和长期共享知识，以生成有关用户发声的意图的一个或多个显式的和/或隐式的假设。 可以基于变化的确定程度，对假设进行分级，并可以为用户生成适应性响应。 可以基于确定程度，用言辞来表达响应，并使随后发声的适当域成帧。 在一种实现方式中，错误识别可以容忍，可以基于随后的发声和/或响应，纠正会话过程。



1. 一种用于提供生成对用户发声的协作会话响应的语音用户界面的方法，所述响应基于有关当前谈话和/或一个或多个过去的谈话的共享知识，所述方法包括：

接收输入，所述输入至少包括一个发声；

更新所述共享知识以包括关于所述至少一个发声的知识；

基于所述更新的共享知识，确定所述至少一个发声的预定的含义；以及

基于所述确定的预定的含义生成响应。

2. 根据权利要求 1 所述的方法，其中，更新所述共享知识包括用关于所述至少一个发声的信息填充短期上下文堆栈。

3. 根据权利要求 2 所述的方法，进一步包括在预定的时间量之后，使所述短期上下文堆栈中的关于所述至少一个发声的信息过期。

4. 根据权利要求 1 所述的方法，其中更新所述共享知识包括，基于关于所述至少一个发声的信息，更新一个或多个长期配置文件。

5. 根据权利要求 1 所述的方法，其中，确定所述预定的含义包括标识所述至少一个发声的谈话类型，所述谈话类型基于发声的一个或多个目标，发声的一个或多个参与者角色，和/或发声的信息的分配。

6. 根据权利要求 1 所述的方法，其中，确定所述预定的含义包括生成所述至少一个发声的预定的含义的至少一个假设，所述至少一个假设包括确定程度。

7. 根据权利要求 6 所述的方法，其中，生成所述至少一个假设包括标识所述至少一个发声中的上下文记号和/或限定符。

8. 根据权利要求 7 所述的方法，进一步包括基于所述共享知识推断所述至少一个发声的丢失的上下文记号和/或限定符。

9. 根据权利要求 8 所述的方法，其中，所述至少一个假设的所述确定程度基于一个或多个所述标识的上下文记号和/或限定符，所述推断的上下文记号和/或限定符，和/或生成响应所需的上下文记号和/

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 13 of 174

或限定符。

10. 根据权利要求 1 所述的方法，其中，生成所述响应包括改编所述响应中的上下文记号和/或语法，以反映所述至少一个发声中的上下文记号和/或语法。

11. 根据权利要求 1 所述的方法，其中，生成所述响应包括使所述响应成帧，以影响随后的发声。

12. 根据权利要求 1 所述的方法，进一步包括:

了解所述确定的预定的含义是不正确的;

进一步更新所述共享知识，以包括所述确定的预定含义是不正确的知识；以及

基于所述进一步更新的共享知识，确定所述至少一个发声的第二预定的含义，其中，排除所述不正确的确定的预定的含义作为可能的第二预定的含义。

13. 根据权利要求 1 所述的方法，其中，所述输入是多模输入。

14. 根据权利要求 1 所述的方法，其中，所述响应是多模响应。

15. 用于提供生成对用户发声的协作会话响应的语音用户界面的包含计算机可执行的指令的计算机可读取的介质，所述响应基于有关当前谈话和/或一个或多个过去的谈话的共享知识，当执行时，所述计算机可执行的指令用于:

接收输入，所述输入至少包括一个发声;

更新所述共享知识以包括关于所述至少一个发声的知识;

基于所述更新的共享知识，确定所述至少一个发声的预定的含义；以及

基于所述确定的预定的含义生成响应。

16. 根据权利要求 15 所述的计算机可读取的介质，其中，计算机可执行的指令用于，通过用关于所述至少一个发声的信息填充短期上下文堆栈来更新所述共享知识。

17. 根据权利要求 16 所述的计算机可读取的介质，其中，计算机可执行的指令用于在预定的时间量之后，使所述短期上下文堆栈中

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 14 of 174

的关于所述至少一个发声的信息过期。

18. 根据权利要求 **15** 所述的计算机可读取的介质，其中，计算机可执行的指令用于，通过基于关于所述至少一个发声的信息更新一个或多个长期配置文件来更新所述共享知识。

19. 根据权利要求 **15** 所述的计算机可读取的介质，其中，所述计算机可执行的指令用于，通过标识所述至少一个发声的谈话类型，确定所述预定的含义，所述谈话类型基于所述发声的一个或多个目标，发声的一个或多个参与者角色，和/或发声的信息的分配。

20. 根据权利要求 **15** 所述的计算机可读取的介质，其中，所述计算机可执行的指令用于，通过生成所述至少一个发声的预定的含义的至少一个假设来确定所述预定的含义，所述至少一个假设包括确定程度。

21. 根据权利要求 **20** 所述的计算机可读取的介质，其中，所述计算机可执行的指令用于，通过标识所述至少一个发声中的上下文记号和/或限定符来生成所述至少一个假设。

22. 根据权利要求 **21** 所述的计算机可读取的介质，其中，所述计算机可执行的指令用于，基于所述共享知识推断所述至少一个发声的丢失的上下文记号和/或限定符。

23. 根据权利要求 **22** 所述的计算机可读取的介质，其中，所述至少一个假设的所述确定程度基于一个或多个所述标识的上下文记号和/或限定符，所述推断的上下文记号和/或限定符，和/或生成响应所需的上下文记号和/或限定符。

24. 根据权利要求 **15** 所述的计算机可读取的介质，其中，所述计算机可执行的指令用于，通过改编所述响应中的上下文记号和/或语法以反映所述至少一个发声中的上下文记号和/或语法来生成所述响应。

25. 根据权利要求 **15** 所述的计算机可读取的介质，其中，所述计算机可执行的指令用于，通过使所述响应成帧以影响随后的发声来生成所述响应。

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 15 of 174

26. 根据权利要求 15 所述的计算机可读取的介质，其中，当执行时，所述计算机可执行的指令进一步用于:

了解所述确定的预定的含义是不正确的;

进一步更新所述共享知识，以包括所述确定的预定含义是不正确的知识; 以及

基于所述进一步更新的共享知识，确定所述至少一个发声的第二预定的含义，其中，排除所述不正确的确定的预定的含义作为可能的第二预定的含义。

27. 根据权利要求 15 所述的计算机可读取的介质，其中，所述输入是多模输入。

28. 根据权利要求 15 所述的计算机可读取的介质，其中，所述响应是多模响应。

29. 一种用于提供生成对用户发声的协作会话响应的语音用户界面的系统，所述系统至少包括一个用于存储有关当前谈话和/或一个或多个过去的谈话的共享知识的存储器，所述系统包括:

接收输入的至少一个输入设备，所述输入至少包括一个发声; 以及

一个或多个处理器，用于:

更新所述共享知识以包括关于所述至少一个发声的知识;

基于所述更新的共享知识，确定所述至少一个发声的预定的含义; 以及

基于所述确定的预定的含义生成响应。

30. 根据权利要求 29 所述的系统，其中，所述一个或多个处理器用于，通过用关于所述至少一个发声的信息填充短期上下文堆栈来更新所述共享知识。

31. 根据权利要求 30 所述的系统，其中，所述一个或多个处理器用于在预定的时间量之后，使所述短期上下文堆栈中的关于所述至少一个发声的信息过期。

32. 根据权利要求 29 所述的系统，其中，所述一个或多个处理

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 16 of 174

器用于，通过基于关于所述至少一个发声的信息更新一个或多个长期配置文件来更新所述共享知识。

33. 根据权利要求 **29** 所述的系统，其中，所述一个或多个处理器用于，通过标识所述至少一个发声的谈话类型，确定所述预定的含义，所述谈话类型基于所述发声的一个或多个目标，发声的一个或多个参与者角色，和/或发声的信息的分配。

34. 根据权利要求 **29** 所述的系统，其中，所述一个或多个处理器用于，通过生成所述至少一个发声的预定的含义的至少一个假设来确定所述预定的含义，所述至少一个假设包括确定程度。

35. 根据权利要求 **34** 所述的系统，其中，所述一个或多个处理器用于，通过标识所述至少一个发声中的上下文记号和/或限定符来生成所述至少一个假设。

36. 根据权利要求 **35** 所述的系统，其中，所述一个或多个处理器用于，基于所述共享知识推断所述至少一个发声的丢失的上下文记号和/或限定符。

37. 根据权利要求 **36** 所述的系统，其中，所述至少一个假设的所述确定程度基于一个或多个所述标识的上下文记号和/或限定符，所述推断的上下文记号和/或限定符，和/或生成响应所需的上下文记号和/或限定符。

38. 根据权利要求 **29** 所述的系统，其中，所述一个或多个处理器用于，通过改编所述响应中的上下文记号和/或语法以反映所述至少一个发声中的上下文记号和/或语法来生成所述响应。

39. 根据权利要求 **29** 所述的系统，其中，所述一个或多个处理器用于，通过使所述响应成帧以影响随后的发声来生成所述响应。

40. 根据权利要求 **29** 所述的系统，其中，所述一个或多个处理器进一步用于：

了解所述确定的预定的含义是不正确的；

进一步更新所述共享知识，以包括所述确定的预定含义是不正确的知识；以及

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 17 of 174

基于所述进一步更新的共享知识，确定所述至少一个发声的第二预定的含义，其中，排除所述不正确的确定的预定的含义作为可能的第二预定的含义。

41. 根据权利要求 **29** 所述的系统，其中，所述输入是多模输入。

42. 根据权利要求 **29** 所述的系统，其中，所述响应是多模响应。

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 18 of 174

协作会话语音用户界面的系统和方法

技术领域

[001] 本发明涉及人机语音用户界面的协作性会话模型。

背景技术

[002] 技术的进步，特别是在收敛（convergence）空间内，导致了对语音识别软件的需求的增加，该软件可以以对人类直观的方式利用该技术。尽管由于共享信息和/或上下文以推进相互的谈话目标，人类之间的通信常常是“协作性的”，但是，现有的人机界面却不能提供相同级别的直观的交互。例如，谈话中的每一个人类参与者都可以为了受益于交换而对交换给出贡献。这是通过关于谈话的各个方面的共享的假设和预期完成的，如主题、参与者的有关主题的知识，对另一个参与者的有关主题的知识的预期，对于主题和/或参与者的适当的措词，基于以前的发声的谈话的发展，参与者的语气或语气变化，从每一个参与者预期的贡献的质量和数量，以及许多其他因素。参与连续地形成的并依赖共享的信息的谈话是人类交谈的天然而直观的方式。

[003] 相比之下，复杂的人机界面不允许用户直观地利用技术，这阻碍了各种技术的大量的采用。利用语音界面，通过使交互更加容易，更快，有助于减轻此负担，但是，现有的语音界面（当他们实际工作时）仍需要用户进行很多的学习。即，现有的语音界面不能在古老的人机界面和谈话语音之间架起桥梁，使得与系统的交互感觉起来很普通。用户应该能够以普通的谈话方式直接从系统请求他们所需要的东西，无需记住准确的词语。或者，当用户不能确定特定需求时，他们应该能够使系统参加有成果的对话，以解决他们的请求。相反，现有的语音界面迫使用户为了配合简单语言的简单指令集以便以系统可以理解的方式发出请求而不说出他们的请求。通过使用现有的语

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 19 of 174

音界面，用户和系统之间的对话几乎没有办法满足相互的目标。

[004] 因此，现有的系统缺乏一种谈话语音模型，谈话语音模型可以给用户提供以对人类固有地直观的方式与系统进行交互的能力。现有的系统存在这些及其他问题。

发明内容

[005] 根据本发明的各个实施例和方面，协作性的谈话语音用户界面可以理解自由形式的人类发声，把用户从被限制于固定的命令集和/或请求中解放出来。相反地，用户可以使用天然的、直观的，自由形式的表达方式，参与与机器的协作性的谈话，以完成请求或一系列请求。

[006] 根据本发明的一个方面，提供了用于实现协作会话语音用户界面的示范性系统体系结构。系统可以接收输入，输入可以包括由输入设备接收到的人的发声，发声可以包括一个或多个请求。如这里所使用的，“发声”可以是字、音节、音素或由人发出的任何其他可听的声音。如这里所使用的，“请求”可以是设备、计算机或其他机器检索信息、执行任务或执行某种其他操作的命令、指令或其他指令。在一种实现方式中，输入可以是多模输入，多模输入的至少一部分是发声。输入的发声分量可以由自动语音识别器进行处理，以生成发声的一个或多个初步解释。然后，可以将一个或多个初步解释提供到会话语音引擎，以便进一步地进行处理，其中，会话语音引擎可以与一个或多个数据库进行通信，以生成自适应会话响应，响应可以作为输出返回到用户。在一种实现方式中，输出可以是多模输出。例如，发声可以包括执行操作的请求，而输出可以包括报告成功或失败的会话响应，以及操作的执行。

[007] 根据本发明的另一个方面，示范性会话语音引擎可以生成对一个请求或一系列请求的自适应会话响应。会话语音引擎可以包括自由形式的语音搜索模块，该模块可以理解使用典型的日常的语言（即，自由形式的）的发声，并可以解释人通常如何说话的变化，它们使用的词汇量，以及他们说话时所处的条件。为解决人类语音的不

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 20 of 174

可捉模的变化，自由形式的搜索模块可以包括偶然的人类语音的模型。例如，在一种实现方式中，自由形式的搜索模块可以理解特殊的行话和/或俚语，容忍词语顺序的变化，并容忍唠叨的暂停或口吃的语音。例如，动词在名词前面的形式化的英语请求可以以与名词在动词前面的请求等效的方式加以对待。在另一种实现方式中，在单一发声中可以标识带有多种变化的复合请求和/或复合任务。通过从单一发声中标识用于完成一个或多个任务的所有相关信息，可以提供优于诸如命令与控制系统之类的现有的语音用户界面（它们使用口头菜单将人可以提供的信息限制于给定点）的优点。在另一种实现方式中，通过从不完整的或模糊的请求推断预定的请求，可以提供谈话的感觉。通过模型化执行标识的上下文中的任务可能需要的上下文记号、限定符或其他信息，可以生成适应性响应，如提示用户提供丢失的上下文记号、限定符或其他信息。在一种实现方式中，响应可以以最限制可能的解释的方式要求遗漏信息，并可以使响应成帧，以建立随后的用户发声的域。在另一种实现方式中，可以识别名词和动词的常见的备选方案，以根据各种条件反映使用模式的变化。如此，表达中的变化可以得到支持，因为词序是不重要的或不曾预料到的，名词和/或动词可以以不同的方式表示，以给出简单化的，但典型的示例。在另一种实现方式中，可以从矛盾的或由于别的原因不准确的信息推断请求，如当发声包括开始和停止、重新开始、口吃、不停顿（run-on）的句子或其他有缺点的语音时。例如，用户可能有时改变主意，如此，在发声的过程中改变请求，尽管如此，有缺点的语音功能能够基于人语音的模型来推断请求。例如，各种模型可以指出最后一个条件最有可能正确、或语调、着重点、强调，词语“not”的使用等各种模型，或其他模型可以表示哪一个条件最有可能是正确的。

[008] 根据本发明的另一个方面，会话语音引擎可以包括噪声耐量模块，该噪声耐量模块可以丢弃给定上下文中的没有含义的词或噪声，以降低混淆的可能性。此外，噪声耐量模块 **250** 还可以过滤掉环境和非人类的噪声，以进一步降低混淆的可能性。在一种实现方式

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 21 of 174

中，噪声耐量模块可以与其他模块和功能协作，以过滤掉不适合标识的上下文的词。例如，噪声耐量模块可以过滤掉一系列一个或多个麦克风内的其他人类对话和/或发声。例如，单一设备可以包括多个麦克风，或者多个设备可以各自包括一个或多个麦克风，噪声耐量模块可以校对输入，并通过比较来自各个麦克风的语音信号来协作地过滤声音。噪声耐量模块也可以过滤掉多个麦克风内的非人类的环境噪音，由说话者的歧义或词语误用所引起的词汇表之外的词，或可能与目标请求不相关的其他噪声。噪声耐量模块的性能基准可以由基于人类的条件的噪声模型进行定义。例如，当汽车以 65 英里每小时行驶时窗户毁损，汽车的驾驶员有 92% 的可能被乘客所理解，那么，噪声耐量模块的性能基准在这样的条件下可以具有类似的性能。

[009] 根据本发明的另一个方面，会话语音引擎可以包括上下文判断过程，用于确定一个请求的一个或多个上下文，以确定谈话内的含义。可以通过让一个或多个上下文域代理竞争来确定给定发声的最适当的域，确定一个或多个上下文。一旦一个给定的域“赢得”了竞争，获胜的域可以负责确定或推断进一步的上下文，并更新短期和长期共享知识。如果在上下文域代理之间存在僵持，则自适应会话响应可以提示用户消除僵持的代理之间的歧义。此外，上下文判断过程可以基于以前的发声和/或请求，推断预定的操作和/或上下文，而现有的系统独立地考虑每一个发声，潜在地一次又一次地产生相同的错误。例如，如果给定解释不正确，则可以从一个或多个自动语音识别器语法和/或从由会话语音引擎确定的可能的解释中删除不正确的解释作为潜在的解释，从而确保，对于完全相同的发声，同一个错误不会重复。

[010] 上下文判断过程可以通过连续地更新现有上下文的一个或多个模型并确定不能先验地确定的上下文作为谈话的副产品，提供优于现有的语音用户界面的优点。相反，上下文判断过程可以跟踪谈话主题，并试图将当前发声匹配到最近的上下文中，包括在完成、部分地完成、请求了任务时在上下文之间进行切换。上下文判断过程可以通过定义在各种上下文域中可能对用户有用的相关的功能的集合，

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 22 of 174

标识发声的一个或多个上下文域。此外，每一个上下文域都可以具有相关的词汇表和思想集合，以模型化词语分组，当一起评估时，可以消除一个上下文域与另一个上下文域的歧义。如此，当搜索相关组合时删去不符上下文的词语和噪声字可以增强推理的准确性。这提供了优于试图向发声的每个组件分配含义（即，包括不符上下文的词语和噪声字）的现有的系统的优点，导致几乎无限的可能的组合和混淆的可能性更大。上下文判断过程也可以是自我意识的，向一个或多个生成的假设分配确定程度，其中，假设可以发展，以反映环境条件的变化，说话者的歧义，口音或其他因素。通过标识上下文，上下文内的功能、上下文内的词汇表，从历史来看在上下文中最常常执行什么任务，什么任务刚刚完成等等，上下文判断过程可以根据相当弱的语音线索确定意图。此外，正如在人与人的谈话中，用户可以在任何时间切换上下文，而不会产生混淆，当发声明确时，可以快速地选择各种上下文域，而不出现菜单驱动的死路。

[011] 根据本发明的另一个方面，示范性协作性会话模型可以基于自由形式的语音搜索、噪声耐量，以及上下文判断，以实现会话人机界面，该会话人机界面反映了人的交互和普通会话行为。即，协作性会话模型允许人和机器参与带有接受的目的或方向的谈话中，每一个参与者都对从中受益的谈话有贡献。通过利用有关人所依赖的发声的人的假定，无论是作为说话者和听者，人机界面可以类似于日常的人与人的谈话。在一种实现方式中，示范性协作性谈话模型可以获取传入的数据（共享知识），以通知决定（智能假设构建），然后，可以精炼判断，并生成响应（适应性响应构建）。

[012] 根据本发明的另一个方面，共享知识可以包括短期和长期知识。短期知识可以在单一谈话过程中累积，其中，在单一谈话过程中接收到的输入可以保留。共享知识可以包括跨模态意识，其中，除累积涉及用户发声、请求、位置等等的输入之外，共享知识可以累积涉及其他模态输入的当前用户界面状态，以进一步构建共享知识模型。共享知识可以用来使用当前和相关信息，构建一个或多个智能假

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 23 of 174

设，通过标识带有长期意义的信息，构建长期共享知识，并利用相关状态和措词信息，生成自适应响应。此外，由于协作性谈话模型化人类谈话，在心理上的适当时间量之后，短期会话数据会过期，从而人性化系统行为，降低基于过时的数据的上下文混淆的可能性，而同时还将来自过期的会话上下文的相关信息添加到长期知识模型中。长期共享知识一般可以是以用户为中心，而不是基于会话的，其中，可以随着时间的推移累积输入，以构建用户、环境、认识的、历史或其他长期知识模型。在用户参与协作性谈话过程中随时可以同时使用长期和短期共享知识。长期共享知识可以包括显式的和/或隐式的用户偏好，最近的上下文、请求、任务等等的历史，与词汇表相关的用户特定的行话和/或上下文的功能，最常使用的词语选择，或其他信息。长期共享知识可以用来使用当前和相关信息，构建一个或多个智能假设，当通过短期共享知识不可用时，利用适当的词语选择，生成自适应响应，精炼长期共享知识模型，标识特定任务的频率，标识用户经常感到困难的任务，或提供其他信息和/或分析，以生成更准确的会话响应。共享知识也可以用来修改未经提示的支持的级别（例如，对于新手与有经验的用户，经常被错认的用户等等）。如此，共享知识可以允许用户和语音用户界面共享假设和预期，如主题知识、谈话历史、词语使用、行话、语气或有助于人用户和系统之间的协作性谈话的其他假设和/或预期。

[013] 根据本发明的另一个方面，对于任何给定发声，可以标识谈话类型。通过分类和发展各种交换的概念模型，可以不断地对齐用户预期和域功能。通过考虑谈话目标、参与者角色，和/或信息在参与者之间的分配，可以生成关于谈话类型的一个或多个智能假设。基于谈话目标，参与者角色，以及信息的分配，智能假设可以考虑各种因素，将谈话（或发声）分类为可以彼此进行交互的一般的谈话类型，以形成很多变化和谈话类型的置换（例如，随着从一个参与者向另一个参与者重新分配信息，或者随着谈话目标基于信息的重新分配而改变，谈话类型可以动态地变化）。

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 24 of 174

[014] 根据本发明的另一个方面，智能假设可以包括发声中的用户的意图的一个或多个假设。此外，智能假设可以使用短期和/或长期共享知识，以主动地随着谈话的进行或随着时间的推移来构建和评估与用户的交互。这些假设可以模型化人与人的互动，以包括每一个假设的变化的确定程度。即，正如人依赖于参与者共享的知识来审查有多少信息以及有什么样的信息可用，智能假设可以利用标识的谈话类型和共享的知识来生成每一个假设的确定程度。

[015] 根据本发明的另一个方面，可以从智能假设生成句法上、语法上以及在上下文中敏感的“智能响应”，这些响应可以用来生成用户的谈话体验，而同时还引导用户以有利于识别方式应答。智能响应可以通过适应用户的说话方式、适当地使响应成帧，具有天然变化和/或个性（例如，通过改变语气、速度、定时、语气变化，措词，行话及口头或音频响应中的其他变化），产生谈话的感觉。

[016] 根据本发明的另一个方面，智能响应可以通过使用上下文记号和语法规则，适应用户的说话方式，以生成可以与用户协作的一个或多个句子。通过利用有关用户如何发出请求的声音的共享知识，可以使用用于识别请求的类似技术，模型化响应。智能响应可以根据统计对可能的响应进行分级和/或使响应随机化，产生构建与天然变化和谈话的感觉交换的机会。这相对于输入和输出不一致的的现有的语音用户界面而言，具有优点，在现有的语音用户界面中，输入是“谈话的”，而输出是“电脑语言”。

[017] 根据本发明的另一个方面，智能响应可以使响应成帧，以影响用户应答发声，以便于识别。例如，可以将响应模型化为更加可能导致完整的请求的发自用户的不当（illicit）的发声。如此，响应可以符合人的对话的协作性特征和天然的人的“模仿”刚刚听到的作为下一个发声的一部分的倾向。此外，当前上下文的知识可以增强响应，以生成更有意义的谈话响应。使响应成帧也可以根据人的模型，处理错误识别。例如，人经常记住许多最近的发声，特别是当一个或多个以前的发声被错误识别或无法识别时。谈话中的另一个参与者可能限

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 25 of 174

制对错误识别的或无法识别的发声的一部分的纠正，或在随后的发声和/或其他交互中，可以提供线索以指出初始的解释不正确。如此，通过存储和分析多个发声，随着谈话的进行，可以纠正前面的谈话的发声。

[018] 根据本发明的另一个方面，智能响应可以包括对用户的多模或跨模的响应。在一种实现方式中，响应可以识别并控制一个或多个设备和/或接口，用户可以通过使用最方便的任何一种输入方法或输入方法的组合来作出响应。

[019] 根据本发明的另一个方面，智能响应可以纠正谈话的过程，而不会中断会话流。即，尽管智能响应可以合理地“确定”，但是，智能响应有时也可能不正确。尽管现有的语音用户界面趋于不出现通常的会话失误（misstep），正常的人之间的交互则可能预期出现失误并相应地进行处理。如此，错误识别之后的响应可以在澄清而不是错误之后模型化，在随后的响应中可以选择词语，以使谈话继续进行，并确定与用户一起探查的适当的域。

[020] 基于下面的图形和具体实施方式，本发明的其他目标和优点对所属领域的技术人员是显而易见的。

附图说明

[021] 图 1 是根据本发明的一个方面的系统体系结构的示范性方框图。

[022] 图 2 是根据本发明的一个方面的会话语音引擎的示范性方框图。

[023] 图 3 是根据本发明的一个方面的协作性会话模型的示范性方框图。

具体实施方式

[024] 请参看图 1，该图显示了根据本发明的一个方面的用于实现协作会话语音用户界面的示范性系统体系结构。系统可以从用户那里接收输入，其中，在一种实现方式中，输入 105 可以包括由输入设备（例如，麦克风）接收到的人的发声，发声可以包括一个或多个

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 26 of 174

请求。输入 105 也可以是多模输入，其中，多模输入的至少一部分是发声。例如，输入设备可以包括麦克风和触摸屏设备的组合，而输入 105 可以包括发声，该发声包括涉及用户正在触摸的触摸屏设备上的显示器的一部分的请求。例如，触摸屏设备可以是导航设备，输入 105 可以包括发声“Give me directions to here”，其中，用户可以在导航设备的显示器上请求到所需目的地的方向。

[025] 输入 105 的发声分量可以由自动语音识别器 110 进行处理，以生成发声的一个或多个初步解释。自动语音识别器 110 可以使用当前技术已知的任何适用技术对发声进行处理。例如，在一种实现方式中，自动语音识别器 110 可以使用语音听写的技术，解释发声，以识别音素流，如标题为“Dynamic Speech Sharpening”的待审批的美国专利申请系列编号 11/513,269 所描述的，在此引用该申请的全部内容作为参考。 然后，可以将由自动语音识别器 110 生成的一个或多个初步解释提供到会话语音引擎 115，以便进一步进行处理。会话语音引擎 115 可以包括会话语言处理器 120 和/或语音搜索引擎 125，如下面的图 2 比较详细地描述的。会话语音引擎 115 可以与一个或多个数据库 130 进行通信，以生成自适应会话响应，响应可以作为输出 140 返回到用户。在一种实现方式中，输出 140 可以是多模输出和/或与一个或多个应用程序 145 的交互，以完成请求。例如，输出 140 可以包括音频响应和导航设备上的路线的显示的组合。例如，发声可以包括执行操作的请求，而输出 140 可以包括报告成功或失败的会话响应，以及操作的执行。此外，在各种实现方式中，自动语音识别器 110、会话语音引擎 115，和/或数据库 130 可以驻留在本地（例如，在用户设备上）、远程（例如，在服务器上），或者，也可以使用本地和远程处理的混合模式（例如，可以在本地对轻量级应用程序进行处理，而在远程对计算密集的应用程序进行处理）。

[026] 请参看图 2，该示范性方框图显示了根据本发明的一个方面的会话语音引擎 215。会话语音引擎 215 可以包括会话语言处理

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 27 of 174

器 220，该处理器使用自由形式的语音搜索模块 245、噪声耐量模块 250，和/或上下文判断进程 255，生成对一个请求或一系列请求的自适应会话响应。根据本发明的一个方面，模块 245-255 可以与语音搜索引擎 225 进行通信，语音搜索引擎 225 包括一个或多个上下文域代理 230 和/或一个或多个词汇表 235，以帮助解释发声和生成响应，如 Tom Freeman 和 Larry Baldwin 所著的“Enhancing the VUE™ (Voce-User-Experience) Through Conversational Speech”所描述的，在此引用该申请的全部内容作为参考。会话语音引擎 215 可以生成针对一个或多个请求的自适应会话响应，其中，请求可以取决于未说出口的假设、不完整的信息、由以前的发声建立的上下文、用户概况、历史概况、环境概况，或其他信息。此外，会话语音引擎 215 可以跟踪哪些请求已经完成，哪些请求正在被处理中，和/或哪些请求由于信息不完整或不准确而不能处理，并可以相应地生成响应。

[027] 根据本发明的一个方面，自由形式的语音搜索模块 245 可以理解使用典型的日常的语言（即，自由形式的）的发声，并可以解释人通常如何说话的变化，它们使用的词汇表，以及他们说话时所处的条件。由于诸如应力、分心，以及偶然发现之类的可变因素始终是不同的并且无限地变化的，自由形式的搜索模块 245 可以这样设计，目标是应该理解，没有人以同样的方式来到相同的人机界面的情况两次。如此，自由形式的搜索模块 245 可以实现模型化偶然的人的语音的一个或多个特点。在各种实现方式中，自由形式的搜索模块 245 可以包括，其中，自由形式的发声功能、一步访问功能、推理计划的操作功能、交替表达功能，和/或有缺点的语音功能。

[028] 自由形式的发声功能可以理解特殊的行话和/或俚语，容忍词语顺序的变化（例如，无论请求的主体先于或后于动词，可能不相关），并容忍唠叨的暂停（例如，“um”、“ah”、“eh”，及其他没有意义的发声）。例如，自由形式的发声功能可以以与自由形式的请求（名词可能在动词前面）的等效的方式对待形式化的英语“动词在名词之前”的请求。例如，可以等效地对待用户的发声“Change it to the

Squizz”和“You know, um, that Squizz channel, ah, switch it there”（其中，Squizz 是 XM Satellite Radio 上的一个频道）。不论是哪一种情况，自由形式的发声功能能够将“Squizz”识别为发声的主体，而将“Change it”或“switch it”识别为发声的动词或请求（例如，通过与上下文判断过程 255 或者其他功能协作，并标识相关上下文域代理 230 和/或词汇表 235，以解释发声）。

[029] 一步访问功能可以理解包括带有多个变量的复合请求的发声。例如，用户发声可以是“What is the forecast for Boston this weekend?” 一步访问功能可以将“weather”识别为上下文（例如，通过与上下文判断过程 255 或其他功能协作，并将“forecast”识别为“weather”的同义词），并搜索等于“Boston”的城市和等于“weekend”的时间。通过从单一发声中标识用于完成任务的所有相关信息，一步访问功能可以克服诸如命令与控制系统之类的现有的语音用户界面（它们使用口头菜单将人可以提供的信息限制于给定点，例如，电话目录服务的命令与控制系统可以说：“State please”，“City please”，“What listing”，等等）。此外，某些发声可以包括复合请求，一步访问功能可以将复合请求分解为多个子任务。例如，用户发声“I need to be at a meeting tomorrow in San Francisco at 8:00 am”可以分解成一组子任务，如：(1) 检查是否有会议之前的晚上的航班并预定一个航班，(2) 检查是否有宾馆，并预定宾馆，(3) 检查是否有小汽车并预定一辆汽车等等，其中，用户还可以进一步为各种任务指定偏好（例如，首先检查用户经常乘坐的航空公司是否有航班）。取决于有关用户的偏好和/或历史模式的共享知识的级别，一步访问功能可以根据请求推断另外的任务。例如，在上面的示例中，一步访问功能也可以检查天气预报，如果天气是“好”（如用户偏好所定义和/或根据历史模式推断），一步访问功能可以安排旧金山的首选的高尔夫球场的开球时间。

[030] 推理预定的操作功能可以从不完整的模糊的请求识别预定的请求。例如，当用户发出“Route <indecipherable> Chicago

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 29 of 174

<indecipherable> here”，其中，用户打算说**“Route calculation to Chicago from here”**，推理预定的操作功能可以模型化计算路线（起始点和目的地点）所需的东西。由于发声包括起始点和目的地点，可以推断计算从用户的当前位置到芝加哥的路线的请求。类似地，当推理预定的操作功能没有足够的信息来推断完整的请求时，可以生成自适应谈话响应以提示用户提供遗漏信息。例如，当发声包括股票行情的请求但没有公司名称（例如，**“Get me the stock price for <indecipherable>”**）时，响应可以是**“What company's stock quote do you want?”**然后，用户可以提供包括公司名称的发声，请求可以是完整的。在一种实现方式中，响应可以以最限制可能的解释的方式要求遗漏信息（例如，在需要城市和州的任务的请求中，可以首先要求州，因为州比城市的数量少）。此外，推理预定的操作功能可以通过在复合的任务和子任务级别维护上下文和识别相关和/或遗漏信息，模型化复合任务和/或请求。

[031] 交替表达功能可以识别名词和动词的常见的备选方案，以根据各种条件反映使用模式的变化。例如，用户可以基于年龄、社会经济背景、种族、用户的怪念头或其他因素，改变表达方式。如此，交替表达功能可以支持表达方式的变化，词序是不重要的或不曾预料到的。基于各种条件或人口资料的表达式的备选方案可以加载到上下文域代理 **230** 和/或词汇表 **235** 中，交替表达功能可以基于推断的或新发现的变化，更新上下文域代理 **230** 和/或词汇表 **235**。在一种实现方式中，谈话语音引擎 **215** 可以包括预订接口，以更新对上下文域代理 **230** 和/或词汇表 **235** 的更改（例如，储存库可以综合各种用户发声和部署全系统的更新）。在操作中，交替表达功能可以允许名词和/或动词以不同的方式表示，以给出简单化的，但典型的示例。例如，对 **Washington, D.C.** 的天气预报感兴趣的用户可能提供任何一个下面的发声，其中每一个都被等效地解释：**“What's the weather like in DC”**，**“Is it raining inside the Beltway”**，**“Gimme the forecast for the capital”**等等。类似地，**“Go to my home”**，**“Go home”**，**“Show route**

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 30 of 174

to home" 和"I would like to know my way home"这些发声都可以等效地解释，其中，用户概况可以包括用户的家庭地址，可以计算到家庭地址的导航路线。

[032] 有缺点的语音功能能够从矛盾的或由于别的原因不准确的信息推断请求，如当发声包括开始和停止、重新开始、口吃、run-on句子或其他有缺点的语音时。例如，用户可能有时改变主意，如此，在发声的过程中改变请求，尽管如此，有缺点的语音功能能够基于人语音的模型来推断请求。例如，对于发声"Well, I wanna .. . Mexi... no, steak restaurant please, I'm hungry"，现有的语音用户界面不会对关于人的语音的模型作出假设，不能推断用户是需要墨西哥餐厅还是牛排餐厅。有缺点的语音功能通过使用可能表示最后一个条件最有可能正确、或语调、着重点、强调，词语"not"的使用等人的理解的各种模型，或可能表示哪一个条件最有可能是正确的其他模型，来克服这些缺点。如此，在上面的示例中，有缺点的语音功能可以推断，用户需要牛排餐厅。

[033] 根据本发明的一个方面，噪声耐量模块 250 可以与有缺点的语音功能紧密关联，并可以操作以丢弃给定上下文中的没有含义的词或噪声，以便不会产生混淆。此外，噪声耐量模块 250 还可以过滤掉环境和非人类的噪声，以进一步降低混淆的可能性。在一种实现方式中，噪声耐量模块 250 可以与其他模块和功能协作，以过滤掉不适合上下文的词。例如，可以标识一个或多个上下文，可以过滤掉相对于系统功能没有含义的词，没有含义的人类的随机的发声及其他噪声。如此，噪声耐量模块 250 可以模型化现实世界的条件，以标识有意义的请求。例如，噪声耐量模块 250 还可以过滤掉一系列一个或多个麦克风内的其他人类对话和/或发声。例如，单一设备可以包括多个麦克风，或者多个设备可以各自包括一个或多个麦克风，噪声耐量模块可以校对输入，并通过比较来自各个麦克风的语音信号来协作地过滤声音。噪声耐量模块 250 也可以过滤掉多个麦克风内的非人类的环境噪音，词汇表之外的词（可能是说话者的歧义或词语误

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 31 of 174

用所造成的），或可能与目标请求不相关的其他噪声。噪声耐量模块 250 中的噪声模型可以基于人类的条件，定义性能基准。例如，当汽车以 65 英里每小时行驶时窗户毁损，汽车的驾驶员有 92% 的可能被乘客所理解，那么，噪声耐量模块 250 在那些条件下可以具有类似的性能。

[034] 根据本发明的一个方面，谈话语音引擎 215 可以包括上下文判断过程 255，该过程确定一个请求的一个或多个上下文，以确定谈话内的含义。可以通过让一个或多个上下文域代理竞争来确定给定发声的最适当的域，确定一个或多个上下文，如标题为“Systems and Methods for Responding to Natural Language Speech Utterance”的编号为 No. 11/197,504 的待审批的美国专利申请和标题为“Mobile Systems and Methods of Supporting Natural Language Human-Machine Interactions”的编号为 No. 11/212,693 的待审批的美国专利申请所描述的，这里引用了这两个申请作为参考。一旦一个给定的上下文域代理“赢得”了竞争，获胜的代理可以负责确定或推断进一步的上下文，并更新短期和长期共享知识。如果在上下文域代理之间存在僵持，则自适应会话响应可以提示用户消除僵持的代理之间的歧义。例如，用户的发声“What about traffic?”可能在各种上下文中具有不同的含义。即，当用户正在查询系统的媒体播放器时具有第一种含义（即，“traffic”将是歌手/作曲家 Steve Winwood 领导的摇滚乐队）当用户正在查询关于 Michael Douglas 电影的搜索界面时具有第二种含义（即，“traffic”将是 Steven Soderbergh 导演的电影），当用户正在查询导航设备以了解到机场的方向时具有第三种含义（即，“traffic”将与到机场的路上的道路的状况相关）。

[035] 此外，上下文判断过程 255 可以基于以前的发声和/或请求，推断预定的操作和/或上下文，而现有的系统独立地考虑每一个发声，潜在地一次又一次地产生相同的错误。例如，如果给定解释不正确，则可以从一个或多个自动语音识别器语法和/或从由上下文判断过程 255 确定的可能的随后的解释中删除不正确的解释作为潜在的解

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 32 of 174

释，从而确保，对于完全相同的发声，同一个错误不会重复。

[036] 上下文判断过程 255 可以通过连续地更新现有上下文的一个或多个模型，来克服现有的系统的缺点，其中，确定上下文的过程可以是谈话的副产品，这是不能先验地确定的。上下文判断过程 255 可以随着任务完成、部分地完成，被请求，等等，确定第一上下文域、变为第二上下文域，变回到第一上下文域，上下文堆栈可以跟踪谈话主题，并试图将当前发声匹配到最近的上下文中，次最近的主题等等，遍历上下文堆栈，直到可以确定最有可能的意图。例如，用户可能发出“What ' s the traffic report”，上下文判断过程 255 可以确定 Traffic 为上下文，并返回包括交通状况报告的输出，不正巧提及 Interstate-5 的交通状况。然后，用户可以发出“What about 1-5?”，上下文判断过程 255 可以知道，当前上下文是交通，可以搜索包括有关 Interstate-5 的信息的交通报告，可以作为输出返回表明 Interstate-5 拥挤的交通状况报告。然后，用户可以“Is there a faster way?”，上下文确定模块 255 可以知道，当前上下文仍是交通，并可以搜索交通不太拥挤的到指定的目的地的并避开 Interstate-5 的路线。此外，上下文判断过程 255 可以基于用户概况、环境概况、历史概况或其他信息，构建上下文，以进一步精炼上下文。例如，概况可以表明，Interstate-5 是星期一到星期五行驶的典型的路线。

[037] 当试图消除上下文之间的歧义时，概况特别有意义，词语在不同上下文中具有不同含义。例如，用户可以发出“What's the weather in Seattle?”，上下文判断过程 255 可以确定天气为上下文，并确定西雅图作为环境上下文。然后，用户可以发出“and Portland?”，上下文判断过程 255 可以基于俄勒冈州波特兰市和华盛顿州西雅图市之间的天气和环境相似性，返回俄勒冈州波特兰市的天气报告。然后，用户可以询问“What time does the game start?”，并可能搜索来自西雅图和/或波特兰的运动队的体育赛事，根据下面的图 3 中比较详细地描述的方法，以会话方式呈现结果。相关地，如果在第二次发声中用户最初发出了“What's the weather in Portsmouth, New

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 33 of 174

Hampshire”，则上下文判断过程 255 可以基于与 New Hampshire 的环境相似性，检索缅因州波特兰市的天气报告。此外，当环境概况、上下文共享知识，和/或其他短期和/或长期共享知识不能提供足够的信息消除各种可能性之间的歧义时，响应可以提示用户提供更进一步的信息（例如，“Did you mean Portland, Maine, or Portland, Oregon?”）。

[038] 上下文判断过程 255 可以与上下文域代理 230 协作，其中，每一个上下文域代理 230 都可以定义可能对用户有用的相关的功能的集合。此外，每一个上下文域代理 230 都可以包括相关的词汇表 235 和思想集合，以模型化词语分组，当一起评估时，可以消除一个上下文域与另一个上下文域的歧义 （例如，音乐上下文域代理 230 可以包括歌曲、音乐家、专辑等等的词汇表 235，而股票上下文域代理 230 可以包括公司名称、股票行情报价机符号，财务度量等等的词汇表 235）。如此，当搜索相关组合时删去不符上下文的词语和噪声字可以增强识别含义的准确性。相比之下，现有的系统试图向发声的每个组件分配含义（例如，包括不符上下文的词语和噪声字），导致几乎无限的可能的组合和混淆的可能性更大。此外，上下文域代理 230 可以包括进一步有助于解释发声、推断意图，完成不完整的请求等等的每一个条件的元数据（例如，Space Needle 词汇表单词可以包括 西雅图、界标、旅游、Sky City 餐厅等等的元数据）。给定消除歧义的条件，上下文判断过程 255 如此能够自动地确定其他完成请求所需的信息，忽略词序的重要性，并执行谈话语音的其他增强。

[039] 上下文域代理 230 也可以是自我意识的，向一个或多个生成的假设分配确定程度，其中，假设可以发展，以反映环境条件的变化，说话者的歧义，口音，或其他因素。在概念上，上下文域代理 230 可以被设计为模型化发声，因为耳背的人将在有噪声的一方。通过标识上下文，上下文内的功能、上下文内的词汇表，从历史来看在上下文中最常常执行什么任务，什么任务刚刚完成等等，上下文域代

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 34 of 174

理 230 可以根据相当弱的语音线索确定意图。此外，上下文堆栈可以是用于确定上下文的多个组件中的一个，如此，不对用户构成约束。所有上下文域都是可访问的，允许用户在任何时间切换上下文，不会产生混淆。如此，正如在人与人的谈话中，当发声明确时，可以快速地选择上下文域，没有菜单驱动的死路。例如，用户可以发出“Please call Rich Kennewick on his cell phone”，可以生成“Do you wish me to call Rich Kennewick on his cell?”的系统响应。用户可以此时决定稍后呼叫 Rich Kennewick，而相反，听一些音乐。如此，用户可以发出“No, play the Louis Armstrong version of Body and Soul from my iPod”，可以在通过媒体播放器播放 Body and Soul 时生成“Playing Body and Soul by Louis Armstrong”的系统响应。在此示例中，稍后的发声与第一次的发声没有上下文的联系，因为发声中的请求条件是明确的，可以轻松地切换上下文，而不依赖于上下文堆栈。

[040] 请参看图 3，该图显示了根据本发明的一个方面的示范性协作性会话模型 300。协作性会话模型 300 可以基于自由形式的语音搜索 245、噪声耐量 250，以及上下文判断 255，以实现会话人机界面，该会话人机界面反映了人如何彼此进行交互以及谈话中的他们的普通行为。简而言之，协作性会话模型 300 允许人和机器参与带有接受的目的或方向的谈话中，每一个参与者都对谈话有贡献。即，协作性谈话模型 300 包括了利用有关人依赖的发声的人的假定（无论是作为说话者和听者）的技术和进程流，从而创建了类似于日常的人与人的谈话的人机界面。在一种实现方式中，协作性谈话可以获取传入的数据（共享知识）305，以通知决定（智能假设构建）310，然后，可以精炼判断，并生成响应（适应性响应构建）315。

[041] 根据本发明的一个方面，共享的知识 305 包括有关传入的数据的短期和长期知识。短期知识可以在单一谈话过程中累积，而长期知识可以随着时间的推移累积，以构建用户概况、环境概况、历史概况或认识概况等等。

[042] 可以在会话输入累加器中保留在单一谈话过程中接收到

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 35 of 174

的输入。会话输入累加器可以包括跨模态意识，其中，除累积涉及用户发声、请求、位置等等的输入之外，会话输入累加器还可以累积涉及其他模态输入的当前用户界面状态，以进一步构建共享知识模型，和更准确的自适应响应（例如，如上文所描述的，当用户发出涉及触摸屏设备的一部分的请求时）。例如，会话输入累加器可以累积输入，包括每一个发声的识别文本，每一个发声的记录的语音文件，列表项目选择历史、图形用户界面处理历史，或其他输入数据。如此，会话输入累加器可以利用当前和相关信息填充智能假设生成器 310，通过标识带有长期意义的信息，构建长期共享知识，给适应性响应生成器 315 提供相关状态和词语使用信息，保留最近的上下文，以便与智能假设生成器 310 一起使用，和/或保留发声，以便在多遍评估过程中重新进行处理。此外，因为协作性谈话 300 模型化人类谈话，在心理上的适当时间量之后，短期会话数据会过期，从而人性化系统行为。例如，人大不可能回忆起两年之前的谈话的上下文，但是由于上下文可以通过机器识别，在预定的时间量之后会话上下文过期，降低基于过时的数据的上下文混淆的可能性。然而，可以向用户、历史、环境、认识的或其他长期知识模型中添加来自过期会话上下文的相关信息。

[043] 长期共享知识一般可以是以用户为中心的，而不是基于会话的。即，可以随着时间的推移累积输入，以构建用户、环境、认识的、历史或其他长期知识模型。在用户参与协作性谈话 300 过程中随时可以同时使用长期和短期共享知识（统称为共享知识 305）。长期共享知识可以包括显式的和/或隐式的用户偏好，最近使用的代理、上下文、请求、任务等等的历史，与词汇表相关的用户特定的行话和/或代理和/或上下文的功能，最常使用的词语选择，或其他信息。长期共享知识可以用来使用当前和相关信息，填充智能假设生成器 310，当适当的词语选择通过会话输入累加器不可用时，给适应性响应生成器 315 提供适当的词语选择，精炼长期共享知识模型，标识特定任务的频率，标识用户经常感到困难的任务，或提供其他信息和/或分析，以生成更准确的会话响应。

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 36 of 174

[044] 如上文所描述的，共享知识 305 可以用来填充智能假设生成器 310，以便用户和语音用户界面可以共享假设和预期，如主题知识、谈话历史、词语使用、行话、语气（例如，正式的、幽默的，简洁的等等），或有助于在人机界面中进行交互的其他假设和/或预期。

[045] 根据本发明的一个方面，成功的协作性谈话的一个组件可以是根据发声识别谈话的类型。通过分类和发展各种交换的概念模型，可以不断地对齐用户预期和域功能。通过考虑谈话目标、参与者角色，和/或信息在参与者之间的分配，智能假设生成器 310 可以生成关于谈话类型的假设。谈话目标可以广泛地包括：(1) 获取单独的信息片段或执行单独的任务，(2) 收集相关的信息片段，以作出决定，和/或 (3) 传播或收集大量的信息，以构建专业知识。参与者角色可以广泛地包括：(1) 控制谈话的主导，(2) 跟随主导并应请求提供输入的支持者，和/或 (3) 使用信息的消费者。在谈话开始时，信息可以由一个或多个参与者持有，其中，参与者可以持有大多数（或所有）信息，少量（或没有）信息，或者信息可以在多个参与者之间大致平均地分配。基于谈话目标，参与者角色，以及信息的分配，智能假设生成器 310 可以考虑各种因素，将谈话（或发声）分类为可以彼此进行交互的一般的谈话类型，以形成很多变化和谈话类型的置换（例如，随着从一个参与者向另一个参与者重新分配信息，或者随着谈话目标基于信息的重新分配而改变，谈话类型可以动态地变化）。

[046] 例如，在一种实现方式中，查询谈话可以包括获取单独的信息片段或执行特定的任务的谈话目标，其中，查询谈话的主导头脑中可能有特定目标，并可能引导谈话往下进行，以实现目标。其他参与者可能持有信息，并可能通过提供信息来支持主导。在启发式谈话中，谈话的主导可以控制谈话的支持者所需的信息。支持者的角色可以仅限于调节谈话的整个进展，并插入查询，以便澄清。在探讨性的谈话中，两个参与者都共享主导和支持者角色，谈话可能没有特定目标，或目标可以随着谈话的进行而即席实现。基于此模型，智能假设生成器 310 可以根据下面的图形广泛地对谈话（或发声）进行分类：

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 37 of 174

查询		
	参与者 A	*参与者 B*
	用户	语音用户界面
目标	获取信息/操作	提供信息/操作
角色	主导/消费者	支持者/施与者
信息分配	较少	较多

启发式		
	参与者 A	*参与者 B*
	用户	语音用户界面
目标	获得信息	提供信息
角色	跟随者/消费者	主导/施与者
信息分配	较少	较多

探讨性的		
	参与者 A	*参与者 B*
	用户	语音用户界面
目标	收集/共享信息	收集/共享信息
角色	跟随者/消费者和主导/施与者	跟随者/消费者和主导/施与者
信息分配	相等的或交替的	相等的或交替的

[047] 智能假设生成器 **310** 可以使用标识的谈话类型，以有助于生成一组针对发声中的用户的意图的假设。此外，智能假设生成器 **310** 可以使用来自会话输入累加器的短期共享知识，以主动地随着谈话的进行来构建和评估与用户的交互，并使用长期共享知识，以主动地随着时间的推移来构建和评估与用户的交互。如此，智能假设生成器 **310** 可以自适应地取得有关用户意图的一组 **n** 个最佳假设，可以

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 38 of 174

向适应性响应生成器 315 提供 n 个最佳假设。此外，智能假设生成器 310 可以通过计算每一个假设的确定程度来模型化人与人的互动。即，正如人依赖于参与者共享的知识来审查有多少信息以及有什么样的信息可用，智能假设生成器 310 可以利用标识的谈话类型和短期和长期共享知识来生成每一个假设的确定程度。

[048] 根据本发明的另一个方面，当发声包含完成请求或任务所需的所有信息（包括限定符）时，智能假设生成器 310 可以生成用户的意图的一个或多个显式的假设。每一个假设都可以具有对应的确定程度，可以用来确定响应中提供的未经提示的支持的级别。例如，响应可以包括一个确认，以确保发声不会被误解，或者，响应可以自适应地提示用户提供遗漏的信息。

[049] 根据本发明的另一个方面，当发声包含可能是丢失的必需的限定符或完成请求或任务所需的其他信息时，智能假设生成器 310 可以使用短期知识来生成用户的意图的一个或多个隐式假设。每一个假设都可以具有对应的确定程度。例如，当谈话开始时，存储在会话输入累加器中的短期知识可以是空的，随着谈话的进行，会话输入累加器可以构建谈话的历史。智能假设生成器 310 可以使用会话输入累加器中的数据来补充或推断有关当前发声的额外的信息。例如，智能假设生成器 310 可以基于与当前发声相关的许多以前的请求，来评估确定程度。在另一个示例中，当当前发声包含的信息不足以完成请求或任务时，会话输入累加器中的数据可以用来推断遗漏信息，以便可以生成假设。在再一个示例中，智能假设生成器 310 可以标识供适应性响应生成器 315 形成个性化的谈话的响应要使用的句法和/或语法。在再一个示例中，当当前发声包含完成请求或任务所需的阈值信息量时，可以依赖会话输入累加器中的数据来调整确定程度。

[050] 根据本发明的另一个方面，当发声缺少完成请求或任务所需的限定符或其他信息时，智能假设生成器 310 可以使用长期共享知识来生成用户的意图的一个或多个隐式假设。每一个假设都可以具有对应的确定程度。使用长期知识的过程可以基本上类似于使用短期

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 39 of 174

共享知识的过程，只是信息可以不受当前会话的约束，而输入机制可以包括来自不同于当前谈话内容的另外的来源的信息。例如，智能假设生成器 310 可以在任何时间使用来自长期共享知识的信息，甚至在启动新的谈话的情况下，而短期共享知识仅限于现有的谈话（当启动新的谈话时，没有短期共享知识可用）。长期共享知识可以来自多个来源，包括用户偏好或插入的数据源（例如，到远程数据库的预订接口），用户的专业知识（例如，基于错误的频率，请求的任务类型，等等，可以将用户标识为新手、中等，有经验，或其他类型的用户），也可以适用于其他代理的代理特定的信息和/或语言（例如，通过将信息与某一个代理分离，以将该信息归入其他代理），来自会话输入累加器的经常使用的主题，经常使用的动词、名词或其他词类，和/或来自会话输入累加器的其他句法信息，或者也可以使用其他长期共享知识的来源。

[051] 根据本发明的另一个方面，由智能假设生成器 310 生成的启用知识的发声，可以包括一个或多个显式的（由用户提供），和一个或多个隐式（由智能假设生成器 310 提供）上下文记号、限定符、条件，及其他信息，它们可以用来标识和评估相关任务。此时，智能假设生成器 310 可以向适应性响应生成器 315 提供输入。由适应性响应生成器 315 接收到的输入可以至少包括假设的排序的列表，包括显式的和/或隐式的假设，每一个假设都可以具有对应的确定程度。可以给假设分配四个确定程度中的一个： (1) “确信”，其中，上下文记号和限定符涉及一个任务，上下文和限定符涉及一个任务，而 ASR 置信水平超过预定的阈值；(2) “十分确信”，其中，上下文记号和限定符涉及一个以上的任务（选择等级最高的任务），条件涉及一个请求，和/或 ASR 置信水平低于预定的阈值； (3) “不确信”，其中，需要另外的上下文记号或限定符来表示任务或对任务进行分级；以及 (4) “无假设”，其中，很少或没有信息可以解密。每一个确定程度都还可以进一步被分类为显式的或隐式的，可以用来调整响应。由适应性响应生成器 310 接收到的输入也可以包括上下文、用

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 40 of 174

户句法和/或语法，上下文域代理特定的信息和/或偏好（例如，旅行上下文域代理可以知道，用户经常请求有关法国的信息，该信息可以与电影上下文域代理共享，以便响应有时候可以包括法语电影）。

[052] 根据本发明的另一个方面，适应性响应生成器 315 可以构建句法上、语法上以及在上下文中敏感的“智能响应”，这些响应可以与一个或多个代理一起使用，生成用户的谈话体验，而同时还引导用户以有利于识别方式应答。在一种实现方式中，智能响应可以包括通过输出设备（例如，说话者）播放的口头或音频应答，和/或通过设备、计算机或机器执行的操作（例如，下载 Web 页面，显示列表，执行应用程序等等）。在一种实现方式中，适当的响应可能不需要谈话适应，可以使用默认应答和/或针对给定任务随机地选择的响应集合。

[053] 根据本发明的另一个方面，适应性响应生成器 310 可以凭借由智能假设生成器 310 维护的信息，生成可以对上下文、当前发声的任务识别，用户知道有关主题的情况，应用程序已经知道有关主题的情况，关于用户偏好和/或相关标题的共享知识，适当的上下文措词（例如，行话），用户在最近的发声中发出的词语，谈话的发展和/或过程校正，谈话的语气，谈话的类型，响应的措词的天然变化，或其他信息敏感的响应。结果，适应性响应生成器 315 可以生成产生谈话的感觉的智能响应，适应在谈话期间累积的信息，维护跨模意识，保持谈话过程。

[054] 根据本发明的另一个方面，适应性响应生成器 315 可以通过适应用户的说话方式、适当地使响应成帧，具有天然变化和/或个性（例如，通过改变语气、速度、定时、语气变化，措词，行话及口头或音频响应中的其他变化），产生谈话的感觉。适应用户的说话方式的过程可以包括使用上下文记号和语法规则来生成一个或多个句子，用作可以与用户协作的响应集合。通过利用有关用户如何发出请求的短期（来自会话输入累加器）和长期（来自一个或多个配置文件）共享知识，可以使用用于识别请求的技术，模型化响应。适应性响应

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 41 of 174

生成器 315 可以根据统计对可能的响应进行分级和/或使响应随机化，产生构建与天然变化和谈话的感觉交换的机会。这相对于输入和输出不一致的的现有的语音用户界面而言，可能是显著的优点，在现有的语音用户界面中，输入是“谈话的”，而输出是“电脑语言”。 下面的示例可以演示响应如何适应用户的输入词语选择和说话方式：

用户	**Do you know *[mumbled words]* Seattle *[more mumbled words]*?**
语音用户界面	**Did you want Seattle sports scores, weather, traffic, or news?**

用户	**Find me *[mumbled words]* Seattle *[more mumbled words]*?**
语音用户界面	***I found* Seattle, did you want sports scores, weather, traffic, or news?**

用户	**Get me *[mumbled words]* Seattle *[more mumbled words]*?**
语音用户界面	**I've *got* Seattle, did you want me to *get* sports scores, weather, traffic, or news?**

[055] 根据本发明的另一个方面，适应性响应生成器 315 可以使响应成帧，以影响用户用可以被轻松地识别的发声作出应答。例如，用户可以发出“给我提供一些新闻”，而语音用户界面响应可以是“是哪些类别的新闻？Top news stories、国际新闻、政治新闻还是体育新闻？”响应可能是来自用户的非法的发声，如“Top news stories”或“国际新闻”，很可能导致完整的请求。如此，响应可以符合人的对话的协作性特征和天然的人的“模仿”刚刚听到的作为下一个发声的一部分的倾向。此外，当前上下文的知识可以增强响应，以生成更有意义的谈话响应，如在下面的交换中。

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 42 of 174

用户	**What's the weather like in Dallas?**

语音用户界面	**In Dallas, it's sunny and 90 degrees.**
用户	**What theaters are showing the movie "The Fantastic Four" there?**
语音用户界面	**10 theaters in Dallas are showing 'The Fantastic Four." Do you want show times for a particular theater?**

[056] 使响应成帧也可以根据人的模型，处理错误识别。例如，人经常记住许多最近的发声，特别是当一个或多个以前的发声被错误识别或无法识别时。谈话中的另一个参与者可能限制对错误识别的或无法识别的发声的一部分的纠正，或在随后的发声和/或其他交互中，可以提供线索以指出初始的解释不正确。如此，通过存储和分析多个发声，随着谈话的进行，可以纠正前面的谈话的发声。

[057] 根据本发明的另一个方面，适应性响应生成器 **315** 可以为用户生成多模或跨模的响应。在一种实现方式中，响应可以识别并控制一个或多个设备和/或接口，用户可以通过使用最方便的任何一种输入方法或输入方法的组合来作出响应。例如，在多模环境中，要求用户利用“**Yes**”或“**No**”引导发声的响应也可以可视地显示多种备选方案。

[058] 根据本发明的另一个方面，适应性响应生成器 **315** 可以纠正谈话的过程，而不会中断会话流。适应性响应生成器 **315** 可以基于假设的排序的列表和对应的确定程度，生成智能响应，可以用来纠正谈话的过程，而不会中断会话流。即，尽管智能响应可以合理地“确定”，但是，智能响应有时也可能不正确。尽管现有的语音用户界面趋于不出现常规失误，正常的人之间的交互可以预期出现失误，并相应地进行处理。如此，错误识别之后的响应可以在澄清而不是错误之后模型化，在随后的响应中可以选择词语，以使谈话继续进行，并

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 43 of 174

确定与用户一起探查的适当的域。例如，过程校正可以导致下面的交换：

用户	**Can you get *[mumbled words here]* Mariners *[more mumbled words]*[1]?**
语音用户界面	***I've got* the score for the Mariners game? 4-2 Mariners.**
用户	**No, the start time for tomorrow's game.**
语音用户界面	**Oh, 7:05.**

[059] 利用特定的示范性实现方式和实施例对本发明进行了描述。然而，所属领域的技术人员将认识到，在不偏离本发明的范围和精神的情况下，可以进行各种修改。因此，说明书和图形都只是示范性的，本发明的范围仅仅由所附的权利要求来确定。

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 44 of 174

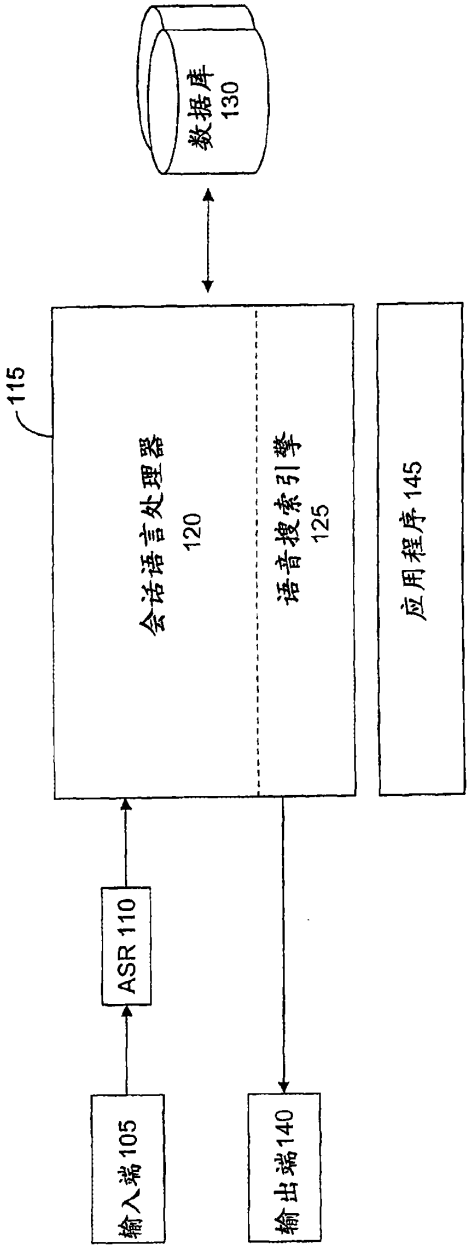


图1

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 45 of 174

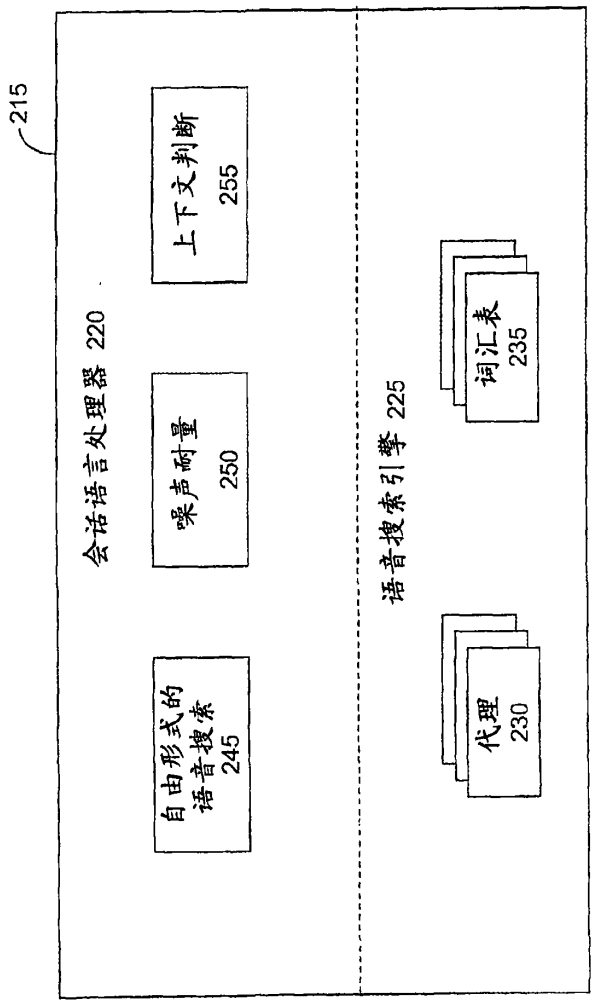


图2

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 46 of 174

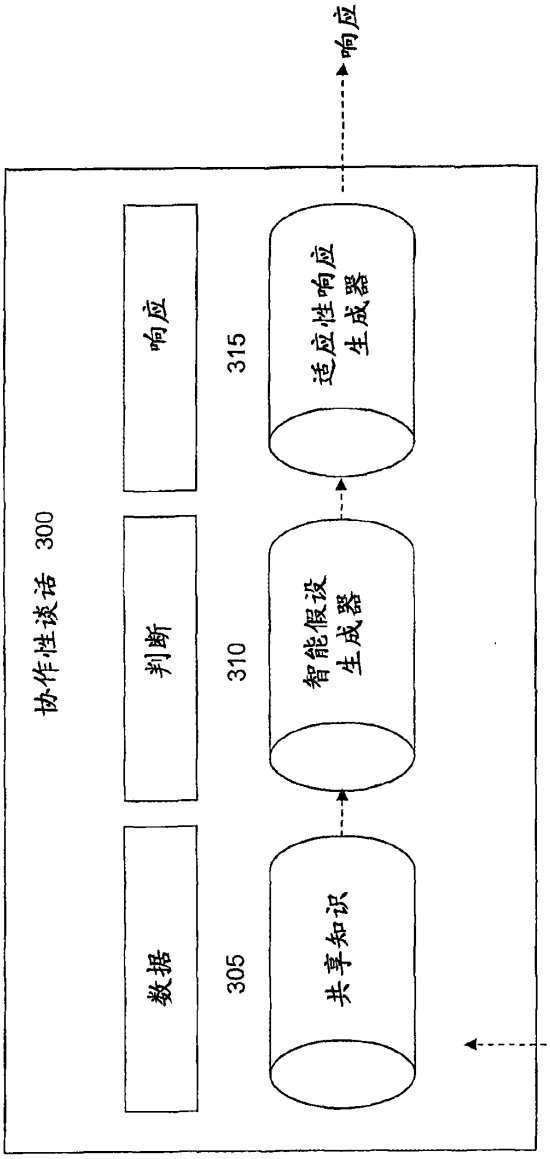


图3

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 47 of 174

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 033 701 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
06.09.2000 Bulletin 2000/36

(51) Int. Cl.[7]: G10L 15/18, G10L 15/26

(21) Application number: 00301441.2

(22) Date of filing: 24.02.2000

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: 01.03.1999 US 260156

(71) Applicant:
MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.
Kadoma-shi, Osaka-fu, 571 (JP)

(72) Inventors:
- Junqua, Jean-Claude
Santa barbara, CA 93110 (US)
- Contolini, Matteo
Santa Barbata, CA 93109 (US)

(74) Representative:
Franks, Robert Benjamin
Franks & Co.,
352 Omega Court,
Cemetery Road
Sheffield S11 8FT (GB)

(54) Apparatus and method using speech understanding for automatic channel selection in interactive television

(57) A speech understanding system for receiving a spoken request from a user and processing the request against a knowledge base of programming information for automatically selecting a television program is disclosed. The speech understanding system includes a knowledge extractor for receiving electronic programming guide (EPG) information and processing the EPG information for creating a program database. The system also includes a speech recognizer for receiving the spoken request and translating the spoken request into a text stream having a plurality of words. A natural language processor is provided for receiving the text stream and processing the words for resolving a semantic content of the spoken request. The natural language processor places the meaning of the words into a task frame having a plurality of key word slots. A dialogue manager analyzes the task frame for determining if a sufficient number of key word slots have been filled and prompts the user for additional information for filing empty slots. The dialog manager searches the program database using the key words placed within the task frame for selecting a program, and produces a signal for selecting a television channel associated with the program.

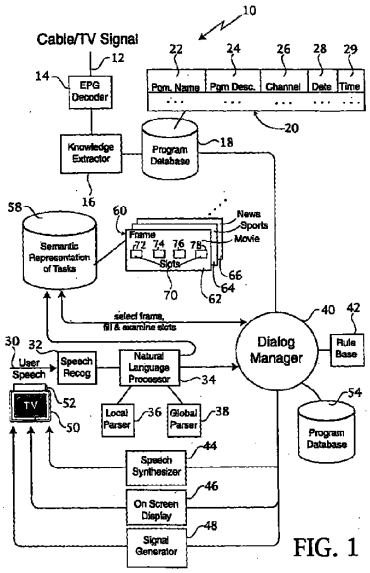


FIG. 1

EP 1 033 701 A2

Printed by Xerox (UK) Business Services
2.16.7 (HRS)/3.6

Description

BACKGROUND OF THE INVENTION

1. Technical Field

[0001] The present invention is directed to an apparatus and method using speech understanding for automating the channel selection process in interactive television. More particularly, the present invention is directed to a system which utilizes natural language processing for receiving a spoken channel selection request and processing the request against information extracted from an electronic programming guide for automatically selecting a channel.

2. Discussion

[0002] The design of interactive television allows two-way communication. Unlike conventional one-way television (TV) or radio broadcasts, most interactive TV enables viewers to respond by telephone. Interactive TV helps people understand complex programs and absorb large amounts of information quickly. Digital television is a new, more efficient method of television transmission which can broadcast several television channels into the space currently used to carry a single analog channel. This will create opportunities for many more new channels and program services.

[0003] In the future, it is expected that all television services will be transmitted digitally. By using digital technology, approximately 200 channels, perhaps more, will be made available. Digital transmission also has the potential to offer other advantages to the viewer, such as high definition or wide-screen pictures, CD-quality sound, and near "video-on-demand," where a film is shown with different start times on several different channels so that the viewer can choose a convenient time to start watching. Interactive services such as home banking, home shopping and connection to the internet could also be made available digitally through the television set

[0004] Viewers who choose to receive digital television by any of these methods will likely need to either buy or rent a special digital receiver or set-top box decoder which will enable digital television pictures to be reassembled on screen. Television sets will also be available with this decoder built in. The set-top box will include a processor that can be used to provide additional services such as speech recognition and speech understanding.

[0005] As the number of television channels increase, the viewer will have more and more difficulty making a channel selection. Instead of finding the desired channel by browsing through the entire listing of channels, channel selection can be made by understanding the semantic content of what the viewer wants to see through analyzing the content of the spoken request. By using natural language and dialogue, the viewer would be able to select the desired program and television channel using spoken requests.

[0006] In view of the above, it is desirable to create a knowledge representation of electronic program guide (EPG) information which is broadcasted as one of many television signals, and store this information in the TV or in the set-top box. It is also desirable to provide a system which can extract the semantics of the users spoken program selection request, and apply this request against the EPG knowledge database for searching for the desired program and switching the television to the corresponding channel. Finally, it is desirable to provide a system for interacting with the user by employing a dialogue phase for clarifying the user request or resolving ambiguities.

SUMMARY OF THE INVENTION

[0007] The method and apparatus of the present invention utilizes a speech understanding technique for automatically selecting channels on a television or video recording device after searching the content of the broadcasts and an electronic programming guide (EPG). This technique uses natural language and dialogue for selecting the desired TV program based upon spoken requests. Using the content of the EPG containing TV program information along with summaries of some events, such as movies, a dynamic recognition vocabulary is created. This vocabulary contains the key words and phrases that can be recognized and understood by the system. This vocabulary is completed by non-key words that are likely to be used by a viewer when selecting channels (e.g. "I would" or "please").

[0008] A database of grammar data structures containing a priori knowledge about the meaning of key words such as "sports" or "movies" is used by a natural language processor to understand the semantic content of the spoken request. Furthermore, a representation of the channel selection task semantics is used by a dialogue manager to help the user in performing the request. The natural language processor utilizes local and global parsing modules for identifying the semantically relevant portions of the spoken request and sends this information to the dialogue manager. The natural language processor also organizes the words and phrases, once understood, into a semantic representation of the spoken request.

[0009] The channel selection semantic representations contain information such as the program to select, the time, possibly the channel or network. etc. The viewer can select a channel but also ask information about what will be available on a particular day or time. The information can be provided to the user as audio responses or displayed on the television screen. Because of the natural language processor, the viewer can concentrate on his/her goal without worrying about the style of language he/she uses to communicate with

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 49 of 174

the speech understanding device. A split screen mode may be used when there is ambiguity to allow the user to further refine the request.

[0010] A history database of the user preferences (in terms of preferred sports or movie types) is built automatically by the dialogue manager. This history database may contain records of past spoken requests and dialogues which can be used by the dialogue manager to dynamically modify the language model of the speech understanding device to favor some of the words during recognition. This history database can also be used during the dialogue phase to rank the questions to ask according to the user preferences.

[0011] The speech understanding device can also be used for programming a video recording device by linking the channel selection to the record command of the recording device. For programs that are recorded frequently, macros can be established as an additional feature for the user. Similarly, reminders can be used to let the viewer know, for example, that a basketball game will be broadcasted tomorrow.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] Additional objects, advantages, and features of the present invention will become apparent from the following description and appended claims, taken in conjunction with the accompanying drawings in which:

Figure 1 is a schematic diagram showing the speech understanding and channel selection system according to a preferred embodiment of the present invention; and

Figure 2 is a schematic diagram disclosing the speech understanding technique performed by the natural language processor and the dialogue manager shown in Figure 1.

DETAILED DESCRIPTION OF THE INVENTION

[0013] In accordance with the teachings of the present invention, a system and method using speech understanding for automatically selecting a television channel is disclosed. Figure 1 illustrates the speech understanding and channel selection system 10 according to a preferred embodiment of the present invention. It is preferred that the channel selection system 10 is incorporated into a set-top decoder box 52. However, the system 10 can also be incorporated into a television 50, or alternatively into a satellite tuner or video playback/recording device.

[0014] As part of the present invention, a cable or television signal 12 provides electronic programming guide (EPG) information to the channel selection system 10. Although it is contemplated that the EPG information can also be downloaded via a telecommunications line from an internet based service provider or a dedicated dial-up EPG service provider. The television signal 12 is also made available for viewing and/or recording. An EPG decoder 14 receives the EPG information for display if desired. The EPG decoder 14 also converts and formats The EPG information into textual information which is communicated to a knowledge extractor 16. The knowledge extractor 16 is responsible for reorganizing the EPG information into a searchable format for storage within a program database 18.

[0015] The program database 18 includes a plurality of searchable program records 20 with a set of predefined fields, such as, but not limited to a program name field 22, a program description or subject matter field 24, a channel field 26, a date field 28, and a time field 29. The program database 18 is continually updated with new program records 20 as the information content of the EPG changes. Therefore, spoken requests can be processed at any time without waiting for updates to the program database 18. In addition, the expired program records 20 within the program database 18 are purged at periodic time intervals so that only a limited and manageable number of program records 20 are searched by the channel selection system 10 for satisfying the users spoken request.

[0016] On the speech processing side of the channel selection system 10, the spoken request and spoken information represented as user speech at 30 is received by a speech recognizer 32. The spoken words are processed by the speech recognizer 32 and converted into text. A suitable speech recognizer is that taught in Lee, K., Large Vocabulary Speaker Independent Continuous Speech Recognition; The Sphinx Systems, Ph.D. Thesis. Carnegie Mellon University, 1988. The text stream which is output from the speech recognizer 32 is provided to a natural language processor 34, which is primarily responsible for analyzing the text stream and resolving the semantic content and meaning of the spoken request. The speech understanding analysis executed by the natural language processor 34 is performed by a local parser module 36 and a global parser module 38. The details of the natural language processor 34 and its components are described in greater detail below.

[0017] A processor based dialogue manager 40 interacts with the various modules of the channel selection system 10, including the natural language processor 34. The dialogue manager 40 communicates with a task database 58 which contains semantic representation of the requested tasks. The task database 58 includes a plurality of predefined task frames 60 which contain a semantic representation of the tasks associated with the users spoken request. As shown, the task frames 60 include a movie task frame 62, a sports task frame 64 and a news task frame 66. While only three task frames 60 are shown, it should be understood that many other task frames can be designed for use with the present invention. Each task frame 60 includes a plurality of key word slots 70 for storing the understood

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 50 of 174

meaning of the key words which are parsed from the user's spoken request. As will be described in greater detail below, the frames and slots are filled with data by the global parser module 38 within the natural language processor 34.

[0018] The dialogue manager 40 retrieves programming records 20 from the program database 18 using the search criteria contained in the selected task frame 60. The search function performed by the dialogue manager 40 is assisted by a rule base 42, which will be described in greater detail below. A request history database 54 is maintained by the dialogue manager 40 for storing a history of the user preferences, such as preferred sports or movie types for viewing and/or recording.

[0019] The dialogue manager 40 has the ability to provide output to a speech synthesizer 44 which can produce an audible inquiry to the user. The dialogue manager 40 may also provide output to an on screen display (OSD) module 46 for presenting the inquiry to the user via a connected television screen 50. Finally, the dialogue manager 40 can provide output to a signal generator module 48 which can translate the output into the appropriate signal for changing the channel on the television 50 or set-top box 52. It is contemplated that as part of the present invention, the signal generator module 48 can produce a variety of commonly used infrared signals which are compatible with the remote command receiver found on most televisions, cable interface boxes, satellite receivers and video recording devices. In this fashion, the dialogue manager 40 can direct the signal generator module 48 to automatically change the television channel, or even program the video recording device to record a program from a desired channel at a particular time and day.

[0020] The operation of the natural language processor 34 is shown in Figure 2. As described above, the natural language processor 34 includes a local parser 36 and a global parser 38 for further analyzing and understanding the semantic content of the digitized words provided by the speech recognizer 32. The local parser 36 has the ability to analyze words, phrases, sentence fragments, and other types of spoken grammatical expressions. To simplify the explanation of the natural language processor 34, all of the grammatical expressions which can be recognized and understood will hereinafter be referenced to as words. Thus, the reference to words should be understood to include phrases, sentence fragments, and all other types of grammatical expressions.

[0021] The local parser 36 examines the words using a LR grammar module 86 to determine if the word is a key word or a non-key word. When a word is recognized as a key word, the word (or phrase, etc.) is "tagged" with a data structure which represents the understood meaning of the word. This examination is accomplished using a database of grammar data structures which comprise the vocabulary of the system. Thus, each recognizable word or phrase has an associated grammar data structure which represents the tag for the word. Once the correct grammar data structure is identified by the local parser 36, a tagging data structure for the word is generated, such as tagging data structure 82 or 84, defining the meaning of the word. The goal of the local parser 36 is to tag all of the spoken words, identified as key words, with the appropriate tagging data structure. The goal of the global parser 38 is to place all of the tagged words into the key word slots 70 of a chosen task frame 60.

[0022] In operation, the local parser 36 receives each word and using the LR grammar module 86 retrieves the grammar data structure associated with that word. The grammar structure for the word will tell the local parser 36 whether or not the word is a key word, and instruct the local parser 36 how to generate the appropriate tagging data structure 82, 84. If the word is not a key word, it is placed into a buffer in case further analysis by the global parser 38 is required. If the word is a key word, the grammar data structure will contain information on how to generate the tagging data structure. If the word is not a key word, the frame tag and slot tag fields will be empty, and the non-key word will be buffered. This frame and slot tag information allows the global parser 38 to place the key word into the appropriate slot 70 of the appropriate task frame 60. This process is assisted by the frame select and slot filler module 56. In the case of some key words, multiple frames may be applicable, and the tagging data structure 82, 84 will indicate that the same slot 70 of two different task frames 60 should be filled with the same key word. The correct task frame 60 can then be chosen during later iterations by the global parser 38.

[0023] An example of a spoken request might be "I would like to watch a movie tonight". This exemplary request contains several key words, namely, "watch", "movie" and "tonight". The remaining words are assumed to be non-key words. However, a dialogue phase may be necessary with this exemplary request in order to resolve the specifics of which movie the user would like to watch, and at what time the user would like to begin watching (or recording). As part of the present analysis technique, the local parser 36 would individually process the words "I" "would" "like" and "to", determine that these words are non-key words, and place these non-key words into a buffer (not shown). The local parser 36 then retrieves the grammar data structure for the word "watch," generates the tagging data structure, and tags the word with the tagging data structure. The tagged word is then passed to the global parser 38 which can determine that the users desired action is to watch a program, as opposed to record a program, or inquire as to what programs are on at a future date and/or time.

[0024] The tagging data structure for the word "movie", shown as data structure 82, will indicate that the movie task frame 62 should be selected. However, a

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 51 of 174

key word slot 70 will not be designated for the word "movie" because this key word is better associated with a task frame 60. The tagging data structure 84 for the word "tonight" will indicate that the semantic representation of this key word should be placed into the date slot 78 of any of the task frames 60. However, the global parser 38 will have already decided that the date slot 78 of the movie task frame 62 should be filled with the understood meaning of the word "tonight." Additionally, the tagging data structure can indicate that the time slot 80 of the movie task frame 62 should be filled with a time value of greater than or equal to 5:00 PM. This way, the dialogue manager 40 can recognize that the user wishes to search for programs with today's date which play in the evening.

[0025] At this point the local parser 36 has tagged all of the words within the spoken request, and the global parser 38, along with the frame select and slot filler module 56, has selected the appropriate task frame 60 for building the search request and filled the appropriate slots 70 with the understood meaning of the words. Next, the dialogue manager 40 can query the user for more specific information. The dialogue manager 40 knows which questions to ask the user based upon which key word slots 70 within the movie task frame 62 must be filled. For example, if the time slot 80 is empty, the dialogue manager 40 may ask the user "at what time would you like to watch a movie?". If the user responds with a spoken time, or time range, the local parser 36 will tag the key words relating to time using the technique described above, and the global parser 38 will place these key words into the time slot 80 of the movie task frame 62.

[0026] The global parser 38 is primarily responsible for analyzing the tagging data structure generated by the local parser 36, for identifying the meaning of the word within the context of the spoken request and then placing the meaning of the word in the appropriate slot 70. The global parser 38 is comprised of many decision tree structures 88. A particular decision tree 88 is utilized once the context of the spoken command is determined. Each decision tree 88 has a starting point, and terminates at a particular action. The action at the terminus of the decision tree 88 instructs the global parser 38 where to place the word, or how to resolve the particular ambiguity. In the case of the present invention, the action will typically instruct the global parser 38 as to which task frame 60 should be selected, or into which key word slot 70 a particular tagged word should be placed.

[0027] A rule base 42 assists the dialogue manager 40 in determining which combinations of filled key word slots 70 provide enough information to perform a search within the program database 15. For example, if the time key word slot 80 of the movie task frame 62 is filled, the dialogue manager 40 can search the program database 18 for all movies that begin at the requested time or during a time range. However, if the search produces more than a predetermined number of movies at the requested time, the dialogue manager 40 may ask the user "what type of movie would you like to watch?". At this point the dialogue manager 40 is attempting to fill the subject key word slot 74 or genre key word slot 76 within the movie task frame 62. If the user responds with a spoken subject or genre, the local parser 36 will tag the key words relating to the subject or genre using the technique described above. These newly tagged words will then be passed to the global parser 38 and placed into the appropriate slots 70 of the movie task frame 62.

[0028] For example, if the user responds with "action movie", the global parser 38 will place the word "action" into the genre slot 76, and again the dialogue manager 40 will narrow its search. If only one action movie is available at the desired time, the dialogue manager 40 will instruct the signal generator 48 to automatically switch the television 50 or set-top box 52 to the channel appearing in the program record 20. If several action movies are available at the desired time, the dialogue manager 40 may list all of the available action movies via the OSD module 46. At this point, the user may select the desired movie by number or channel. As an alternative feature of the present invention, the dialogue manager 40 may provide a confirmation of the users request as feedback to the user prior to switching the channel.

[0029] As the user learns to provide the dialogue manager 40 with a complete set of information within one spoken request, such as "I would like to watch the Detroit Red Wings hockey game tonight', or "I would like to record the program Nova on PBS this Tuesday", the natural language processor 34 can fill enough key word slots 70 to permit a search to be performed, and the spoken request fully satisfied by the dialogue manager 40. In the case of the request to watch the hockey game, the dialogue manager 40 will complete the search and produce the appropriate signal for changing the channel of the set-top box 52 or television 50 based upon the information contained within thq channel field 26 of the programming record 20. In the case of the request to record the desired program from PBS, the dialogue manager 40 will complete the search and retrieve the date, time and channel information from the programming record 20 and produce the appropriate signal via signal generator module 48 for programming the video recording device.

[0030] As part of the present invention, it is further contemplated that the dialogue manager 40 can receive feedback signals from the video recording device in cases where the device is already programmed to record a different program at the same time, or that a blank tape must be inserted into the recording device. In this manner, various conflicts can be resolved while the user is present.

[0031] The foregoing discussion discloses and describes exemplary embodiments of the present invention. One skilled in the art will readily recognize

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 52 of 174

from such discussion, and from the accompanying drawings and claims, that various changes. modifications, and variations can be made therein without departing from the spirit and scope of the invention as defined in the following claims.

Claims

1. A speech understanding system for receiving a spoken request from a user and processing the request against a knowledge base of programming information for automatically selecting a television program comprising:

   a knowledge extractor for receiving electronic programming guide (EPG) information and processing the EPG information for creating a program database;
   a speech recognizer for receiving the spoken request and translating the spoken request into a text stream having a plurality of words;
   a natural language processor for receiving the text stream and processing the words for resolving a semantic content of the spoken request, the natural language processor placing the meaning of the words into a task frame having a plurality of key word slots; and
   a dialogue manager for analyzing the task frame for determining if a sufficient number of key word slots have been filled and prompting the user for additional information for filing empty slots, the dialog manager searching the program database using the key words placed within the task frame for selecting a program, and producing a signal for selecting a television channel associated with the television program.

2. The system of Claim 1 wherein the natural language processor includes a local parser for analyzing the words within the text stream and identifying key words, the local parser utilizing a LR grammar database for resolving the meaning of the words.

3. The system of Claim 2 wherein the local parser generates a tagging data structure for each key word, the tagging data structure representing the meaning of the keyword.

4. The system of Claim 3 wherein the natural language processor includes a global parser for receiving the tagging data structure for each key word and for selecting the task frame associated with the meaning of the spoken request

5. The system of Claim 4 wherein the global parser interacts with a plurality of decision trees for determining which task frame is associated with the meaning of the spoken request.

6. The system of Claim 4 wherein the global parser receives the tagging data structure for each key word and places a meaning of the key word into a key word slot.

7. The system of Claim 1 wherein the dialogue manager interacts with a rule base for determining when the task frame contains enough information for searching the program database.

8. The system of Claim 1 wherein the dialogue manager prompts the user for additional information through a speech synthesizer.

9. The system of Claim 1 wherein the dialogue manager prompts the user for additional information through a display system.

10. The system of Claim 1 wherein a signal generator is connected to the dialogue manager for receiving a command and generating a signal for operating a remote receiver.

11. The system of Claim 1 wherein the program database includes a plurality of program records relating to each program and channel combination available for selection by the user.

12. A speech understanding system for receiving a spoken request from a user and processing the request against a knowledge base of programming information for automatically selecting a television program comprising:

    a knowledge extractor for receiving electronic programming guide (EPG) information and processing the EPG information for creating a program database;
    a speech recognizer for receiving the spoken request and translating the spoken request into a text stream having a plurality of words;
    a natural language processor for receiving the text stream and processing the words for resolving a semantic content of the spoken request the semantic content being represented by a task frame having a plurality of key word slots;
    the natural language processor including a local parser for analyzing the words using a database of grammar structures and generating a tagging data structure representing the meaning of the word;
    the natural language processor including a global parser for selecting a task frame and placing the meaning of the words contained in the tagging data structure into the key word slots

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 53 of 174

within the task frame; and

a dialogue manager for analyzing the task frame for determining if a sufficient number of key word slots have been filled and prompting the user for additional information until a predetermined combination of slots have been filled, the dialog manager searching the program database using the key words placed within the task frame for selecting a program, and producing a signal for selecting a television channel associated with the television program.

13. The system of Claim 12 wherein the global parser receives the tagging data structure for each key word and selects the task frame associated with the meaning of the spoken request.

14. The system of Claim 12 wherein the global parser interacts with a plurality of decision trees for determining which task frame is associated with the meaning of the spoken request.

15. The system of Claim 12 wherein the tagging data structure for each tagged word includes a frame tag indicating which task frame the word is associated with and a slot tag indicating which key word slot the meaning of the word is associated with.

16. The system of Claim 12 wherein the dialogue manager interacts with a rule base for determining when the task frame contains the predetermined combination of filled slots.

17. The system of Claim 12 wherein the EPG information is received from one of a television signal, a cable television signal, a satellite television signal and a telecommunication signal.

18. A method for understanding a spoken request and selecting a television channel based upon the spoken request comprising:

receiving electronic programming guide (EPG) information;
processing the EPG information for producing a program database;
receiving the spoken request from a user and processing the spoken request with a natural language processor for extracting a semantic representation of the task requested by the user;
building a search request of key words identified by the natural language processor.
performing a search within the program database using the key words; and
generating a signal for selecting a television channel associated with the desired program.

19. The method of Claim 18 further including the step of initiating a dialog phase using a dialog manager for prompting a user for additional information.

20. The method of Claim 19 further including the step of analyzing the user request using a database of grammar structures and generating one or more tagging data structure representing the meaning of the spoken request.

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 54 of 174

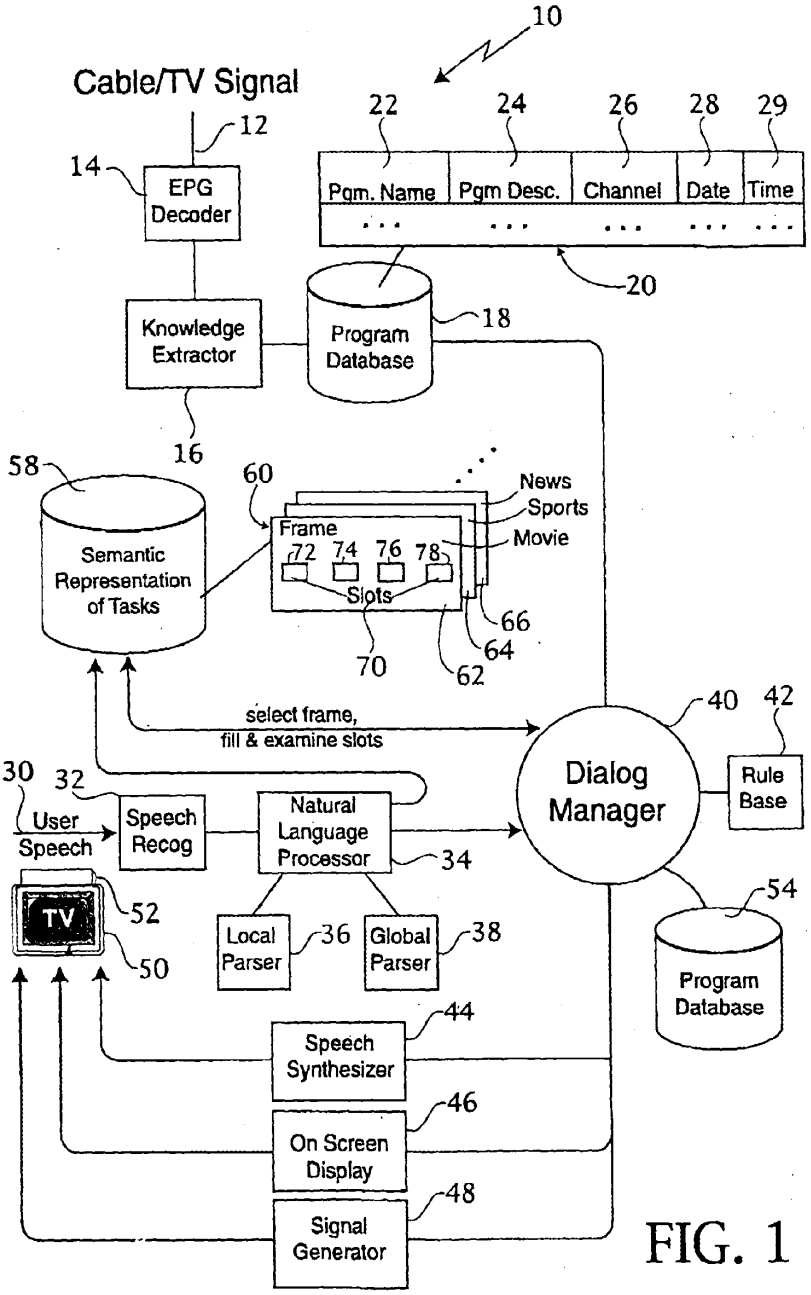


FIG. 1

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 55 of 174

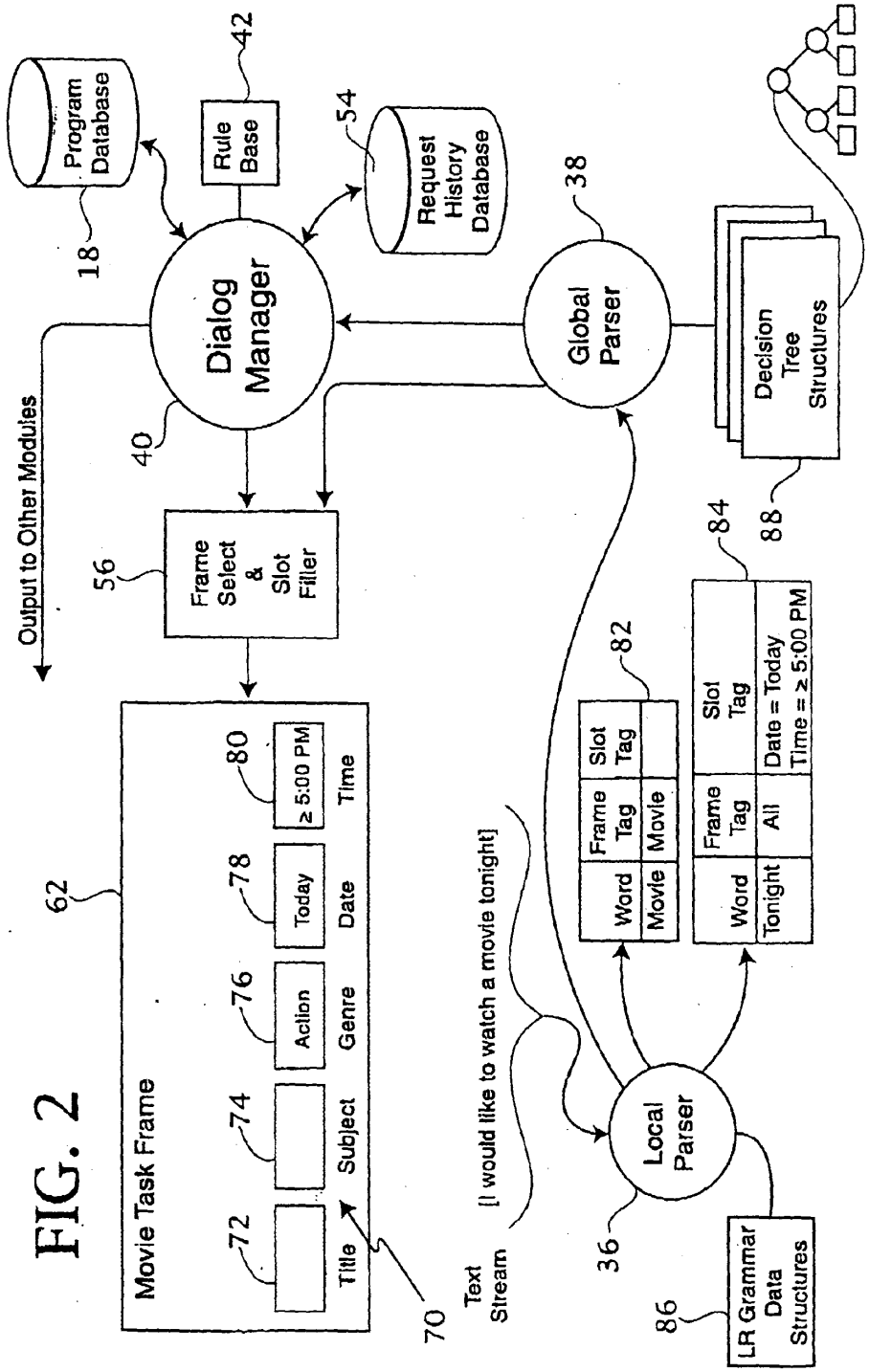


FIG. 2

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 56 of 174

Electronic Acknowledgement Receipt

EFS ID:	31660865
Application Number:	14149296
International Application Number:	
Confirmation Number:	4038
Title of Invention:	INTERACTIVE SERVER, CONTROL METHOD THEREOF, AND INTERACTIVE SYSTEM
First Named Inventor/Applicant Name:	Ji-hye CHUNG
Customer Number:	23373
Filer:	Peter A. McKenna/Todd Yarborough
Filer Authorized By:	Peter A. McKenna
Attorney Docket Number:	Q203469
Receipt Date:	31-JAN-2018
Filing Date:	07-JAN-2014
Time Stamp:	15:13:01
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		Q203469IDSasfiled.pdf	56274 3c0077d8f89b12595a5d22e348e4eb45b63c80b9	yes	4

	Multipart Description/PDF files in .zip description				
	Document Description		**Start**	**End**	
	Transmittal Letter		1	2	
	Information Disclosure Statement (IDS) Form (SB08)		3	4	
Warnings:					
Information:					
2	Foreign Reference	Q203469RefCN101535983AwithABS.pdf	7469842 3542061338c6ca4a96814bb9b51c68a2c010b0cd	no	37
Warnings:					
Information:					
3	Foreign Reference	Q203469RefEP1033701A2.pdf	1714232 9478aed1ab12a4a15e3f6865d76f2cfbf6209527	no	9
Warnings:					
Information:					
4	Other Reference-Patent/App/Search documents	Q203469CNOA11617CNAppNo2014100067482withtranslation.pdf	7894185 95c81c4029b9ad059a9d062e8307d1bfbff6f4a7	no	33
Warnings:					
Information:					
		Total Files Size (in bytes):	17134533		

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.

PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of

Ji-hye CHUNG, et al.

Appln. No.: 14/149,296

Confirmation No.: 4038

Filed: January 7, 2014

Docket No: Q203469

Group Art Unit: 2658

Examiner: OPSASNICK, MICHAEL N

For: INTERACTIVE SERVER, CONTROL METHOD THEREOF, AND INTERACTIVE SYSTEM

INFORMATION DISCLOSURE STATEMENT
UNDER 37 C.F.R. §§ 1.97 and 1.98

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

Sir:

In accordance with the duty of disclosure under 37 C.F.R. § 1.56, Applicant hereby notifies the U.S. Patent and Trademark Office of the documents which are listed on the attached PTO/SB/08 (modified) form and/or listed herein and which the Examiner may deem material to patentability of the claims of the above-identified application.

One copy of each of the listed documents is submitted herewith, except for the following: U.S. patents and/or U.S. patent publications; and co-pending non-provisional U.S. applications filed after June 30, 2003.

The present Information Disclosure Statement is being filed after the later of three months from the application's filing date and the mailing date of the first Office Action on the merits, but before a Final Office Action, Notice of Allowance, or an action that otherwise closes prosecution in the application (whichever is earlier), and therefore Applicant is filing

concurrently herewith a Statement Under 37 C.F.R. § 1.97(e). No fee under 37 C.F.R. § 1.17(p) is required.

For foreign language documents, please see column "T" in the "Foreign Patent Documents" section of the attached PTO/SB/08 form.

The submission of the listed documents is not intended as an admission that any such document constitutes prior art against the claims of the present application. Applicant does not waive any right to take any action that would be appropriate to antedate or otherwise remove any listed document as a competent reference against the claims of the present application.

The USPTO is directed and authorized to charge all required fees, except for the Issue Fee and the Publication Fee, to Deposit Account No. 19-4880. Please also credit any overpayments to said Deposit Account.

Respectfully submitted,

/Peter A. McKenna/

Peter A. McKenna
Registration No. 38,551

SUGHRUE MION, PLLC
Telephone: 202.293.7060
Facsimile: 202.293.7860

WASHINGTON OFFICE
23373
CUSTOMER NUMBER

Date: January 31, 2018

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 60 of 174

PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of

Ji-hye CHUNG, et al.

Appln. No.: 14/149,296

Confirmation No.: 4038

Filed: January 7, 2014

Docket No: Q203469

Group Art Unit: 2658

Examiner: OPSASNICK, MICHAEL N

For: INTERACTIVE SERVER, CONTROL METHOD THEREOF, AND INTERACTIVE SYSTEM

AMENDMENT UNDER 37 C.F.R. § 1.111

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

Sir:

In response to the Office Action dated September 28, 2017, please amend the above-identified application as follows on the accompanying pages.

TABLE OF CONTENTS

AMENDMENTS TO THE CLAIMS 2

REMARKS 15

AMENDMENTS TO THE CLAIMS

This listing of claims will replace all prior versions and listings of claims in the application:

LISTING OF CLAIMS:

1-22. (canceled).

23. (currently amended): A method of controlling a display apparatus, the method comprising:

displaying a content received through a server selected by a user command on a display of the display apparatus;

in response to a first uttered voice being received, displaying a first response message transmitted from a server apparatus with the content on the display based on first response information corresponding to the first uttered voice; and

in response to a second uttered voice being received while the content and the first response message are displayed on the display, displaying a second response message on the display based on second response information corresponding to the second uttered voice with the content,

wherein the second response information is generated based on an execution language and a command corresponding to the second uttered voice and utterance history information, ~~and~~ the utterance history information ~~includes~~ including the first response information, and

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 62 of 174

wherein upon a condition that the execution language and the command are not stored in the display apparatus, the execution language and the command are estimated based on utterance history information that is stored within a preset time period from a time when the second uttered voice is received, and the second response message is generated based on the estimated execution language and the estimated command.

24. (previously presented): The method of claim 23, wherein the utterance history information further comprises information on the first uttered voice.

25. (previously presented): The method of claim 23, further comprising:

determining whether the first uttered voice includes control information for controlling a function of the display apparatus; and

performing a function corresponding to the first uttered voice, in response to the first uttered voice including the control information.

26. (previously presented): The method of claim 23, further comprising:

determining whether the second uttered voice includes control information for controlling a function of the display apparatus; and

performing a function corresponding to the second uttered voice, in response to the second uttered voice including the control information.

27. (previously presented): The method of claim 25, wherein the control information includes information for controlling a volume of the display apparatus, and the function is for adjusting a volume.

28. (previously presented): The method of claim 23, wherein the second response information is generated based on the second uttered voice and utterance history information which is most recently obtained prior to receiving the second uttered voice.

29. (previously presented): The method of claim 23, wherein the outputting the first response message comprises, in response to an uttered voice requesting a contents search being received, outputting a contents search result message based on contents search information, and the outputting the second response message comprises, in response to an uttered voice related to the contents being received, outputting a response message corresponding to an uttered voice related to the contents, based on response information related to a content execution, and

wherein the response information related to a content execution is generated based on the uttered voice related to a content and the contents search information.

30. (previously presented): The method of claim 23, wherein the first response information comprises an execution language and a command corresponding to the first uttered voice, or comprises information obtained based on the execution language and the command.

31. (previously presented): The method of claim 30, wherein the first uttered voice comprises a first uttered element indicating an execution subject and a second uttered element indicating an execution command,

wherein the execution language corresponds to the first uttered element and the command corresponds to the second uttered element.

32. (previously presented): The method of claim 23, wherein the second response information comprises information obtained based on the execution language and the command.

33. (previously presented): The method of claim 32, wherein the second uttered voice comprises a first uttered element indicating an execution subject and a second uttered element indicating an execution command,

wherein the execution language corresponds to the first uttered element and the command corresponds to the second uttered element.

34. (previously presented): The method of claim 23, wherein the utterance history information comprises at least one uttered voice received prior to receiving the first uttered voice and response information corresponding to the at least one uttered voice.

35. (previously presented): The method of claim 31, wherein the execution language and the command are combined and generated as an execution command script.

36. (currently amended): A method for controlling a display apparatus, comprising:

displaying a content received through a server selected by a user command on a display of the display apparatus;

in response to a first uttered voice being received, displaying a first response message transmitted from a server apparatus with the content on the display based on first response information corresponding to the first uttered voice;

in response to a second uttered voice being received while the content and the first response message are displayed on the display, determining whether the second uttered voice includes control information for controlling a function of the display apparatus;

in response to the second uttered voice including the control information, performing a function corresponding to the second uttered voice; and

in response to the second uttered voice not including the control information, displaying a second response message on the display based on second response information corresponding to the second uttered voice with the content,

wherein the second response information is generated based on an execution language and a command corresponding to the second uttered voice and utterance history information, ~~and~~ the utterance history information ~~includes~~ including information regarding the first uttered voice, and

wherein upon a condition that the execution language and the command are not stored in the display apparatus, the execution language and the command corresponding to the second uttered voice are estimated based on utterance history information that is stored within a preset time period from a time when the second uttered voice is received, and the second response message is generated based on the estimated execution language and the estimated command.

37. (previously presented): The method of claim 36, wherein the utterance history information further comprises the first response information.

38. (previously presented): The method of claim 36, further comprising:

determining whether the first uttered voice includes control information for controlling a function of the display apparatus; and

in response to the first uttered voice including the control information, performing a function corresponding to the first uttered voice.

39. (canceled).

40. (previously presented): The method of claim 38, wherein the control information comprises information for controlling a volume of the display apparatus and the function is a function for adjusting a volume.

41. (currently amended): A display apparatus comprising:

a display configured to display a content <u>received through a server</u> selected by a user command;

a communicator configured to communicate with a server apparatus; and

a controller configured to control the communicator to transmit an uttered voice signal regarding received first uttered voice to the server apparatus to receive first response information corresponding to the first uttered voice, and display a first response message <u>transmitted from the server apparatus</u> with the content on the display based on the received first response information,

determine whether received second uttered voice includes control information for controlling a function of the display apparatus,

in response to the second uttered voice including the control information, perform a function corresponding to the second uttered voice,

in response to the second uttered voice not including the control information, control the communicator to transmit an uttered voice signal regarding the second uttered voice to the server apparatus to receive second response information corresponding to the second uttered voice, and display a second response message on the display based on the received second response information with the content,

wherein the second response information is generated based on an execution language and a command corresponding to the second uttered voice and utterance history information, ~~and wherein~~ the utterance history information ~~includes~~ including the first response information, and

wherein upon a condition that the execution language and the command are not stored in the display apparatus, the execution language and the command corresponding to the second uttered voice are estimated based on utterance history information that is stored within a preset time period from a time when the second uttered voice is received, and the second response message is generated based on the estimated execution language and the estimated command.

42. (previously presented): The apparatus of claim 41, wherein the utterance history information further comprises information on the first uttered voice.

43. (previously presented): The apparatus of claim 41, wherein the controller determines whether the first uttered voice includes control information for controlling a function of the display apparatus, and in response to the first uttered voice including the control information, performs a function corresponding to the first uttered voice.

44. (canceled).

45. (previously presented): The apparatus of claim 43, wherein the control information includes information for controlling a volume of the display apparatus and the function is a function for adjusting a volume.

46. (previously presented): The apparatus of claim 41, further comprising:

a storage;

wherein the controller stores the utterance history information to the storage.

47. (previously presented): The apparatus of claim 41, wherein the second response information is generated based on the second uttered voice and utterance history information which is most recently obtained prior to receiving the second uttered voice.

48. (previously presented): The apparatus of claim 41, wherein the controller, in response to an uttered voice requesting a contents search being received, outputs a contents search result message based on contents search information, and in response to an uttered voice related to the contents execution being received, outputs a response message corresponding to an

uttered voice related to the contents execution, based on response information related to a content execution, and the response information related to a content execution is generated based on at least one of the uttered voice related to the contents and the contents search information.

49. (previously presented): The apparatus of claim 41, wherein the first response information comprises an execution language and a command corresponding to the first uttered voice, or comprises information obtained based on the execution language and the command.

50. (previously presented): The apparatus of claim 49, wherein the first uttered voice comprises a first uttered element indicating an execution subject and a second uttered element indicating an execution command,

wherein the execution language corresponds to the first uttered element and the command corresponds to the second uttered element.

51. (previously presented): The apparatus of claim 41, wherein the second response information comprises information obtained based on the execution language and the command.

52. (previously presented): The apparatus of claim 51, wherein the second uttered voice comprises a first uttered element indicating an execution subject and a second uttered element indicating an execution command,

wherein the execution language corresponds to the first uttered element and the command corresponds to the second uttered element.

53. (previously presented): The apparatus of claim 41, wherein the utterance history information comprises at least one uttered voice received prior to receiving the first uttered voice and response information corresponding to the at least one uttered voice.

54. (previously presented): The apparatus of claim 50, wherein the execution language and the command are combined and generated as an execution command script.

55. (currently amended): A display apparatus comprising:

a display configured to display a content received through a server selected by a user command;

a communicator configured to communicate with a server apparatus; and

a controller configured to control the communicator to transmit an uttered voice signal regarding received first uttered voice to the server apparatus to receive first response information corresponding to the first uttered voice, and display a first response message transmitted from the server apparatus with the content on the display based on the received first response information,

determine whether received second uttered voice includes control information for controlling a function of the display apparatus,

in response to the second uttered voice including the control information, perform a function corresponding to the second uttered voice, and

in response to the second uttered voice not including the control information, control the communicator to transmit an uttered voice signal regarding the second uttered voice to the server apparatus to receive second response information corresponding to the second uttered voice, and

display a second response message on the display based on the received second response information with the content,

wherein the second response information is generated based on an execution language and a command corresponding to the second uttered voice and utterance history information, ~~and wherein~~ the utterance history information including ~~includes~~ information regarding the first uttered voice, and

wherein upon a condition that the execution language and the command are not stored in the display apparatus, the execution language and the command corresponding to the second uttered voice are estimated based on utterance history information that is stored within a preset time period from a time when the second uttered voice is received, and the second response message is generated based on the estimated execution language and the estimated command.

56. (previously presented): The apparatus of claim 55, wherein the utterance history information further comprises the first response information.

57. (previously presented): The apparatus of claim 55, wherein the controller determines whether the first uttered voice includes control information for controlling a function of the display apparatus, and in response to the first uttered voice including the control information, performs a function corresponding to the first uttered voice.

58. (canceled).

59. (previously presented): The apparatus of claim 57, wherein the control information includes information for controlling a volume of the display apparatus and the function is a function for adjusting a volume.

60. (previously presented): The method of claim 26, wherein the control information includes information for controlling a volume of the display apparatus, and the function is for adjusting a volume.

61. (previously presented): The method of claim 36, wherein the control information comprises information for controlling a volume of the display apparatus and the function is a function for adjusting a volume.

62. (previously presented): The apparatus of claim 41, wherein the control information includes information for controlling a volume of the display apparatus and the function is a function for adjusting a volume.

63. (previously presented): The apparatus of claim 55, wherein the control information includes information for controlling a volume of the display apparatus and the function is a function for adjusting a volume.

64. (previously presented): The method of claim 23, wherein the first response message is displayed in a portion of the displayed content.

65. (previously presented): The method of claim 64, wherein, in response to receiving the first uttered voice requesting a search of a content, displaying a list as the first response message, the list comprising at least one content searched according to the first uttered voice.

66. (previously presented): The method of claim 65, wherein, in response to receiving the second uttered voice selecting a content from the list, displaying a notice message indicating that the content is selected as the second response message and displaying the selected content on the display.

67. (previously presented): The method of claim 23, wherein, in response to a second uttered voice being received while the content and the first response message are displayed on the display, displaying the second response message with the content on the display.

REMARKS

Reconsideration and allowance of the claims are respectfully requested. Upon entry of this Amendment, claims 23, 36, 41, and 55 are amended. Claims 23-38, 40-43, 45-57, and 59-67 are all the claims pending in the Application.

Claim rejections -- 35 U.S.C. §103

Claims 23-26, 28-35, 36-38, and 63-67 are rejected under pre-AIA 35 U.S.C. 103(a) as allegedly being unpatentable over US 2012/0265528 to Gruber, in view of US 2008/0091406 to Baldwin.

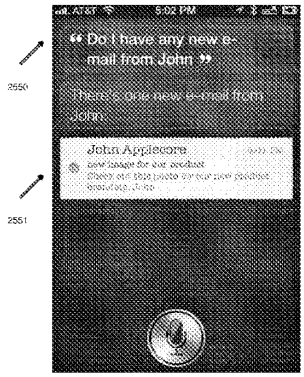
Independent claim 23

Claim 23 has been amended to recite, among others, the features of:

> displaying a content received through a server selected by a user command on a display of the display apparatus;
> in response to a first uttered voice being received, displaying a first response message transmitted from a server apparatus with the content on the display based on first response information corresponding to the first uttered voice; and
> in response to a second uttered voice being received while the content and the first response message are displayed on the display, displaying a second response message on the display based on second response information corresponding to the second uttered voice with the content,
> wherein the second response information is generated based on an execution language and a command corresponding to the second uttered voice and utterance history information, the utterance history information including the first response information

Accordingly, Applicant has amended claim 23 to clarify that the displayed content is received through a server, and that the content is displayed along with a first response message that is transmitted from (i.e. received through) a server apparatus. In other words, the claimed server and the claimed server apparatus are different from each other.

Gruber relates to a smartphone, and according to Gruber, only the input utterance voice 2550 and the response message 2551 corresponding thereto are provided on a screen, but a broadcast screen is not provided together. See FIG. 25A of Gruber, reproduced below:



<FIG. 25A of Gruber>

That is, the content is not provided together with the response message in response to a first uttered voice, as claimed.

Gruber fails to disclose, teach or reasonably suggest that the content is received from a different server than a server from which the first response message is received as claimed.

Baldwin has been cited solely for its teachings related to time periods, and not for the features discussed above. Indeed, Baldwin fails to cure the deficiencies of Gruber discussed above, and also fails to provide any information that would suggest a modification of Gruber to arrive at the features discussed above.

Accordingly, amended claim 23 is patentable over the Gruber and Baldwin combination.

Independent claim 36

Claim 36 has been amended to recite features similar to, but not necessarily coextensive with, those of claim 23 and thus claim 36 is patentable over the Gruber and Baldwin combination for similar reasons as claim 23 discussed above.

Remaining claims

The remaining rejected claims are patentable based on their respective dependencies.

Claims 27, 40, 60, and 61 are rejected under pre-AIA 35 U.S.C. 103(a) as allegedly being unpatentable over Gruber and Baldwin, in further view of US 2002/0198714 to Zhou. This rejection is respectfully traversed.

Claims 27, 40, 60, and 61 each depends from one of claims 23 and 36, which have been shown above to be patentable over the Gruber and Baldwin combination. Zhou has not been cited for the features discussed above, and indeed fails to provide any information that would suggest a modification of Gruber and Baldwin to arrive at the features discussed above. Thus, claims 27, 40, 60, and 61 are patentable based on their respective dependencies.

Claims 41-43, 46, 47-54, and 55-57 are rejected under pre-AIA 35 U.S.C. 103(a) as allegedly being unpatentable over US 2011/0313775 to Laligand, in view of Gruber, and in further view of Baldwin.

Independent claim 41

Claim 41 has been amended to recite, among others, the features of:

> a display configured to display a content received through a server selected by a user command;
> a communicator configured to communicate with a server apparatus; and
> a controller configured to control the communicator to transmit an uttered voice signal regarding received first uttered voice to the server apparatus to receive first response information corresponding to the first uttered voice, and display a first response message transmitted from the server apparatus with the content on the display based on the received first response information,
> determine whether received second uttered voice includes control information for controlling a function of the display apparatus,
> in response to the second uttered voice including the control information, perform a function corresponding to the second uttered voice,
> in response to the second uttered voice not including the control information, control the communicator to transmit an uttered voice signal regarding the second uttered voice to the server apparatus to receive second response information corresponding to the second uttered voice, and display a second response message on the display based on the received second response information with the content

Similar to claim 23 discussed above, Applicant has amended claim 41 to clarify that the displayed content is received through a server, and that the content is displayed along with a first response message that is transmitted from (i.e. received through) a server apparatus. In other words, the claimed server and the claimed server apparatus are different from each other.

Similar to Gruber, Laligand is directed to a smartphone, and also fails to disclose, teach or reasonably suggest receiving the content from one server and a first response message transmitted from another, different server as claimed.

As discussed above, Gruber also fails to disclose, teach or reasonably suggest that the content is received from a different server than a server from which the first response message is received as claimed.

Baldwin has been cited solely for its teachings related to time periods, and not for the features discussed above. Indeed, Baldwin fails to cure the deficiencies of Laligand and Gruber discussed above, and also fails to provide any information that would suggest a modification of Laligand and Gruber to arrive at the features discussed above.

Accordingly, amended claim 41 is patentable over the Laligand, Gruber and Baldwin combination.

Independent claim 55

Claim 55 has been amended to recite features similar to, but not necessarily coextensive with, those of claim 41 and thus claim 55 is patentable over the Gruber and Baldwin combination for similar reasons as claim 41 discussed above.

Remaining claims

The remaining rejected claims are patentable based on their respective dependencies.

Claims 45, 59, 62, and 63 are rejected under pre-AIA 35 U.S.C. 103(a) as allegedly being unpatentable over Laligand in view of Gruber, Baldwin, and further in view of Zhou. This rejection is respectfully traversed.

Claims 45, 59, 62, and 63 each depends from one of claims 41 and 55, which have been shown above to be patentable over the Laligand, Gruber and Baldwin combination. Zhou has not been cited for the features discussed above, and indeed fails to provide any information that would suggest a modification of Laligand, Gruber and Baldwin to arrive at the features discussed above. Thus, claims 45, 59, 62, and 63 are patentable based on their respective dependencies.

Conclusion

In view of the above, reconsideration and allowance of this application are now believed to be in order, and such actions are hereby solicited. If any points remain in issue which the Examiner feels may be best resolved through a personal or telephone interview, the Examiner is kindly requested to contact the undersigned at the telephone number listed below.

The USPTO is directed and authorized to charge all required fees, except for the Issue Fee and the Publication Fee, to Deposit Account No. 19-4880. Please also credit any overpayments to said Deposit Account.

Respectfully submitted,

/ Kevin C. Kunzendorf /

Kevin C. Kunzendorf
Registration No. 58,308

SUGHRUE MION, PLLC
Telephone: 202.293.7060
Facsimile: 202.293.7860

WASHINGTON OFFICE
23373
CUSTOMER NUMBER

Date: December 27, 2017

Electronic Acknowledgement Receipt

EFS ID:	31348642
Application Number:	14149296
International Application Number:	
Confirmation Number:	4038
Title of Invention:	INTERACTIVE SERVER, CONTROL METHOD THEREOF, AND INTERACTIVE SYSTEM
First Named Inventor/Applicant Name:	Ji-hye CHUNG
Customer Number:	23373
Filer:	Kevin Christopher Kunzendorf/Yuko Tanaka
Filer Authorized By:	Kevin Christopher Kunzendorf
Attorney Docket Number:	Q203469
Receipt Date:	27-DEC-2017
Filing Date:	07-JAN-2014
Time Stamp:	20:10:53
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		Q203469Amendment111.pdf	942189 8f1fa15b74c273c34be9326a4f70aa967ec9569a	yes	20

	Multipart Description/PDF files in .zip description		
	Document Description	**Start**	**End**
	Amendment/Req. Reconsideration-After Non-Final Reject	1	1
	Claims	2	14
	Applicant Arguments/Remarks Made in an Amendment	15	20
Warnings:			
Information:			
	Total Files Size (in bytes):	942189	

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/SB/06 (09-11)
Approved for use through 1/31/2014. OMB 0651-0032
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.

PATENT APPLICATION FEE DETERMINATION RECORD Substitute for Form PTO-875	Application or Docket Number 14/149,296	Filing Date 01/07/2014	☐ To be Mailed

ENTITY: ☒ LARGE ☐ SMALL ☐ MICRO

APPLICATION AS FILED – PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	RATE ($)	FEE ($)
☐ BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A	
☐ SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A	
☐ EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	minus 20 =	*	X $ =	
INDEPENDENT CLAIMS (37 CFR 1.16(h))	minus 3 =	*	X $ =	
☐ APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $310 ($155 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).			
☐ MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))				
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL	

APPLICATION AS AMENDED – PART II

AMENDMENT

12/27/2017	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	* 42	Minus	** 42	= 0	X $80 =	0
Independent (37 CFR 1.16(h))	* 4	Minus	***4	= 0	X $420 =	0
☐ Application Size Fee (37 CFR 1.16(s))						
☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))						
					TOTAL ADD'L FEE	0

AMENDMENT

	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	*	Minus	**	=	X $ =	
Independent (37 CFR 1.16(h))	*	Minus	***	=	X $ =	
☐ Application Size Fee (37 CFR 1.16(s))						
☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))						
					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 number found in the appropriate box in column 1.

LIE
VICTOR BARLOW

This collection of information is required by 37 CFR 1.16. 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, 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.

UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 CFR 1.53(b))

Attorney Docket No.	Q203469
First Named Inventor	Ji-hye CHUNG
Title	INTERACTIVE SERVER, CONTROL METHOD THEREOF, AND INTERACTIVE SYSTEM
Express Mail Label No.	

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

ADDRESS TO: **Commissioner for Patents**
P.O. Box 1450
Alexandria VA 22313-1450

1. ☑ **Fee Transmittal Form** (PTO/SB/17 or equivalent)
2. ☐ **Applicant asserts small entity status.** See 37 CFR 1.27.
3. ☐ **Applicant certifies micro entity status.** See 37 CFR 1.29. Applicant must attach form PTO/SB/15A or B or equivalent.
4. ☑ **Specification** [Total Pages 46]
   Both the claims and abstract must start on a new page.
   (See MPEP § 608.01(a) for information on the preferred arrangement)
5. ☑ Drawing(s) (35 U.S.C. 113) [Total Sheets 7]
6. **Inventor's Oath or Declaration** [Total Sheets 3]
   (including substitute statements under 37 CFR 1.64 and assignments serving as an oath or declaration under 37 CFR 1.63(e))
   a. ☑ Newly executed (original or copy)
   b. ☐ A copy from a prior application (37 CFR 1.63(d))
7. ☑ **Application Data Sheet.** **See note below.* See 37 CFR 1.76 (PTO/AIA/14 or equivalent)
8. ☐ **CD-ROM or CD-R** in duplicate, large table or Computer Program (*Appendix*)
   ☐ Landscape Table on CD
9. ☐ **Nucleotide and/or Amino Acid Sequence Submission**
   (if applications, items a. – c. are required)
   a. ☐ Computer Readable Form (CRF)
   b. Specification Sequence Listing on:
      i. ☐ CD-ROM or CD-R (2 copies); or
      ii. ☐ Paper
   c. ☐ Statements verifying identity of above copies

ACCOMPANYING APPLICATION PAPERS

10. ☐ **Assignment Papers** (cover sheet & document(s))
    Name of Assignee SAMSUNG ELECTRONICS CO., LTD.
11. ☐ **37 CFR 3.73(c) Statement** *(when there is an assignee)* ☑ **Power of Attorney**
12. ☐ **English Translation Document** *(if applicable)*
13. ☑ **Information Disclosure Statement** *(PTO/SB/08 or PTO-1449)*
    ☑ Copies of citations attached
14. ☐ **Preliminary Amendment**
15. ☐ **Return Receipt PostCard** *(MPEP § 503)(Should be specifically itemized)*
16. ☐ **Certified Copy of Priority Document(s)** *(if foreign priority is claimed)*
17. ☐ **Nonpublication Request** Under 35 U.S.C. 122(b)(2)(B)(i). Applicant must attach form PTO/SB/35 or equivalent.
18. ☑ **Priority is claimed from:**
    KR 10-2013-0001829 2013-01-07
19. ☑ **Other**: PTO/SB/38 Form

Note: (1) Benefit claims under 37 CFR 1.78 and foreign priority claims under 1.55 must be included in an Application Data Sheet (ADS).
(2) For applications filed under 35 U.S.C. 111, the application must contain an ADS specifying the applicant if the applicant is an assignee, person to whom the inventor is under an obligation to assign, or person who otherwise shows sufficient proprietary interest in the matter. See 37 CFR 1.46(b).

☑ The USPTO is directed and authorized to charge all required fees, except for the Issue Fee and the Publication Fee, to Deposit Account No. 19-4880. Please also credit any overpayments to said Deposit Account.

20. **If a CONTINUING APPLICATION**, *check appropriate box, and supply the requisite information below and in the first sentence of the specification following the title, or in an Application Data Sheet under 37 CFR 1.76*:

☐ Continuation ☐ Divisional ☐ Continuation-in-part (CIP) of prior application No:

Prior application information: *Examiner:* *Art Unit:*

21. CORRESPONDENCE ADDRESS

☑ The address associated with Customer Number: WASHINGTON OFFICE 23373 CUSTOMER NUMBER OR ☐ Correspondence address below

Signature	/Peter A. McKenna/	Date	January 7, 2014
Name	Peter A. McKenna	Registration No.	38,551

FEE TRANSMITTAL

Complete if Known	
Application Number	Not Yet Assigned
Filing Date	January 7, 2014
First Named Inventor	Ji-hye CHUNG
Examiner Name	Not Yet Assigned
Art Unit	Not Yet Assigned
Docket No.	Q203469

☐ 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 attached or have been submitted previously.

TOTAL AMOUNT OF PAYMENT ($)1,760

METHOD OF PAYMENT (check all that apply)

☐ Check ☑ Credit Card ☐ Money Order ☐ None ☐ Other (please identify):

☐ Deposit Account Deposit Account Number: 19-4880 Deposit Account Name: Sughrue Mion, PLLC

For the above-identified deposit account, the Director is hereby authorized to: (check all that apply)

☐ Charge fee(s) indicated below ☐ Charge fee(s) indicated below, **except for the Issue Fee and the Publication Fee**

☐ Charge fee(s) indicated below, **except for the filing fee, the Issue Fee and the Publication Fee**

☑ Charge any additional fee(s), except for the Issue Fee and the Publication Fee, or underpayments of fee(s) under 37 CFR 1.16 and 1.17 ☑ Credit any overpayment of fees(s)

FEE CALCULATION

1. BASIC FILING, SEARCH, AND EXAMINATION FEES (U = undiscounted fee; S = small entity fee; M = micro entity fee)

	FILING FEES			SEARCH FEES			EXAMINATION FEES			
Application Type	**U($)**	**S($)**	**M($)**	**U($)**	**S($)**	**M($)**	**U($)**	**S($)**	**M(&)**	**Fees Paid ($)**
Utility	280	140*	70	600	300	150	720	360	180	1,600
Design	180	90	45	120	60	30	460	230	115	
Plant	180	90	45	380	190	95	580	290	145	
Reissue	280	140	70	600	300	150	2160	1080	540	
Provisional	260	130	65	0	0		0	0	0	

* The $140 small entity status filing fee for a utility application is further reduced to $70 for a small entity status applicant who filed the application via EFS-Web.

2. EXCESS CLAIM FEES

Fee Description	Undiscounted Fee ($)	Small Entity Fee ($)	Micro Entity Fee ($)
Each claim over 20 (including Reissues)	80	40	20
Each independent claim over 3 (including Reissues)	420	210	105
Multiple dependent claims	780	390	195

Total Claims		Extra Claims		Fee ($)		Fee Paid ($)
22	- 20 or HP =	2	X	80	=	160

HP = highest number of total claims paid for, if greater than 20.

Multiple Dependent Claims	
Fee($)	Fee Paid ($)
780	0

Indep. Claims		Extra Claims		Fee ($)		Fee Paid ($)
3	- 3 or HP =	0	X	420	=	0

HP = highest number of independent claims paid for, if greater than 3.

3. APPLICATION SIZE FEE

If the specification and drawings exceed 100 sheets of paper (excluding electronically filed sequence or computer listings under 37 CFR 1.52(e)), the application size fee due is $400 ($200 for small entity or $ 100 for micro entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).

Total Sheets		Extra Sheets		Number of each additional 50 or fraction thereof		Fee($)	Fee Paid ($)
53	- 100 =	0	/ 50	0	(round up to a whole number) x	400	0
53 *.75	40 - 100 =	0	/ 50	0	EFS Fee Discount	400	0

4. OTHER FEE(S) Fees Paid ($)

Non-English Specification, $140 fee ($ 70 for small entity or $35 for micro entity)

Non-electronic filing fee under 37 CFR 1.16(t) for a utility application, $400 fee ($200 small or micro entity)

Other (e.g., late filing surcharge):

SUBMITTED BY

Signature	/Peter A. McKenna/	Registration No. (Attorney/Agent) 38,551	Telephone: 202.293.7060
Name (Print/Type)	Peter A. McKenna		Date: January 7, 2014

FIG. 1

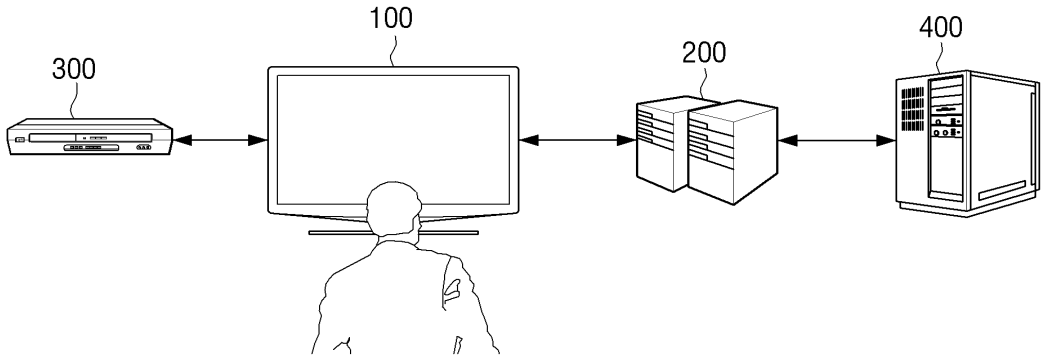


FIG. 2

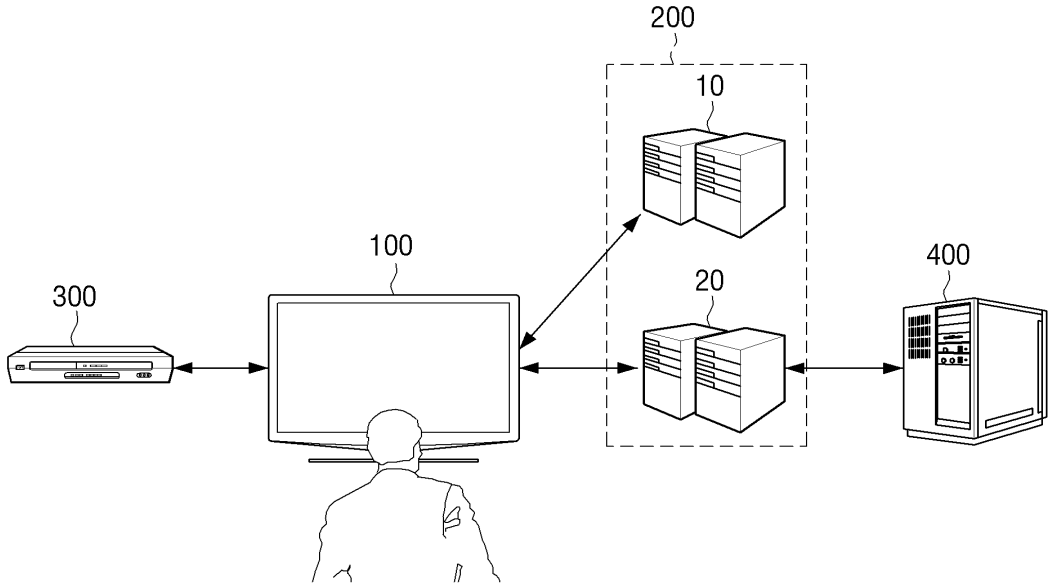


FIG. 3

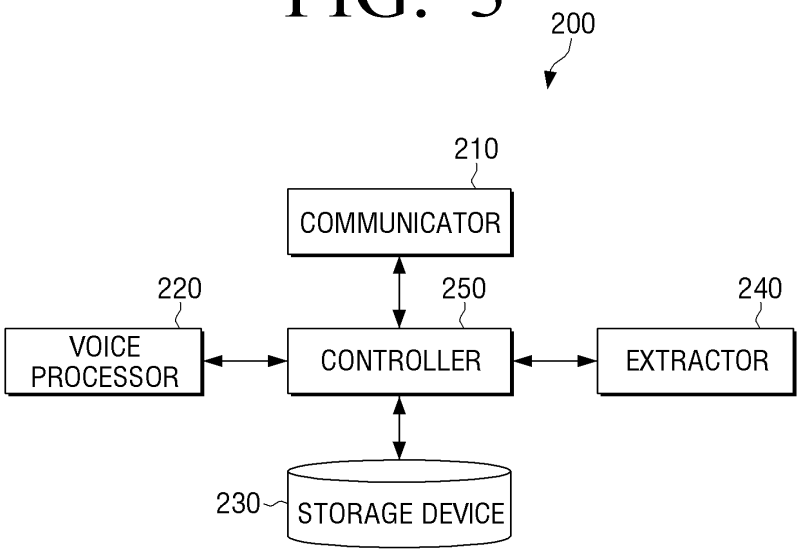

200
210
COMMUNICATOR
220
250
240
VOICE PROCESSOR
CONTROLLER
EXTRACTOR
230
STORAGE DEVICE

FIG. 4

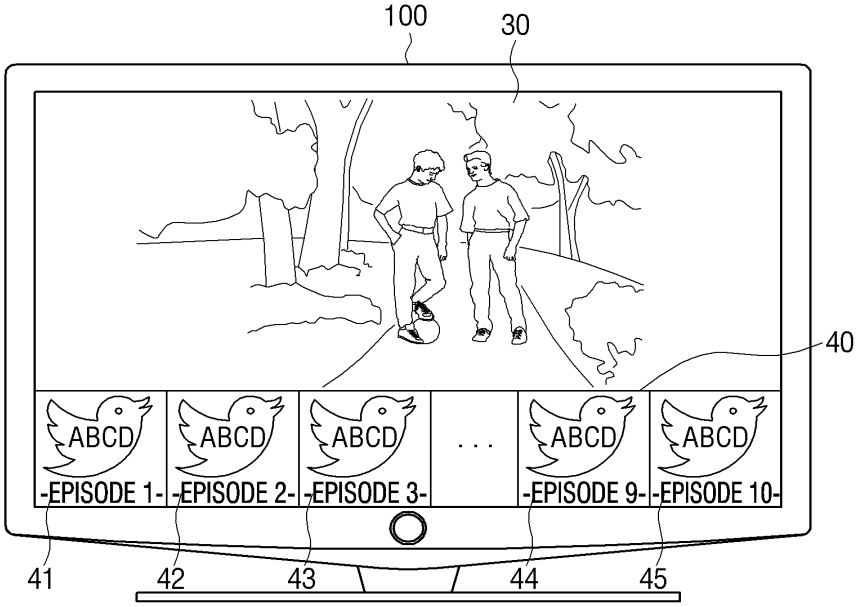

100
30
40
ABCD
-EPISODE 1-
ABCD
-EPISODE 2-
ABCD
-EPISODE 3-
. . .
ABCD
-EPISODE 9-
ABCD
-EPISODE 10-
41
42
43
44
45

FIG. 5

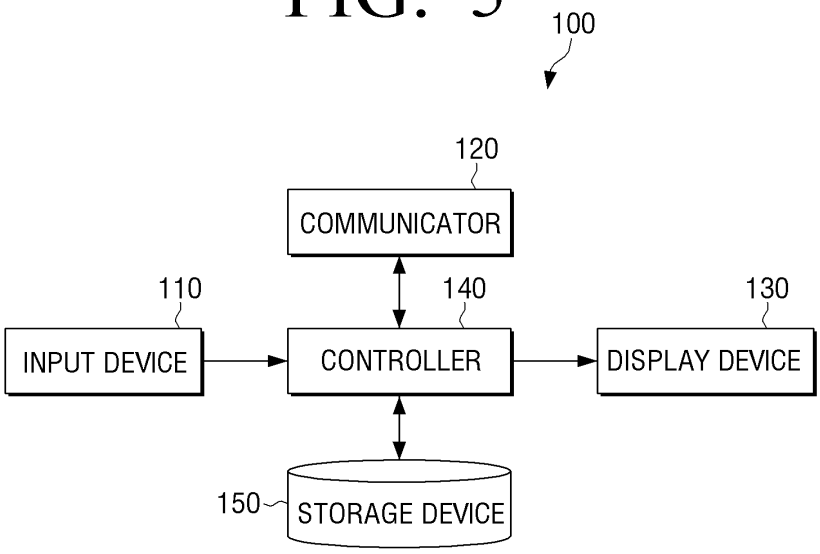

100
120
COMMUNICATOR
110
140
130
INPUT DEVICE
CONTROLLER
DISPLAY DEVICE
150
STORAGE DEVICE

FIG. 6

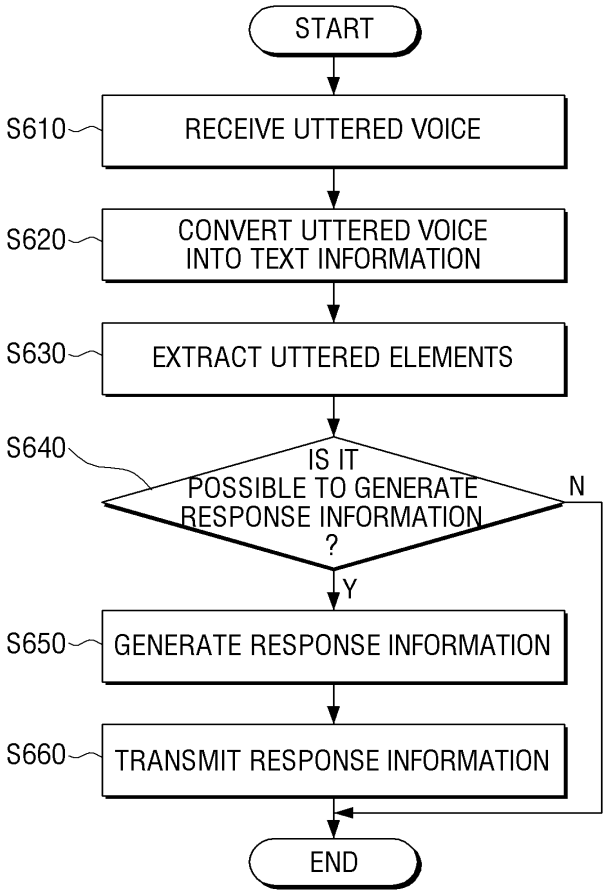
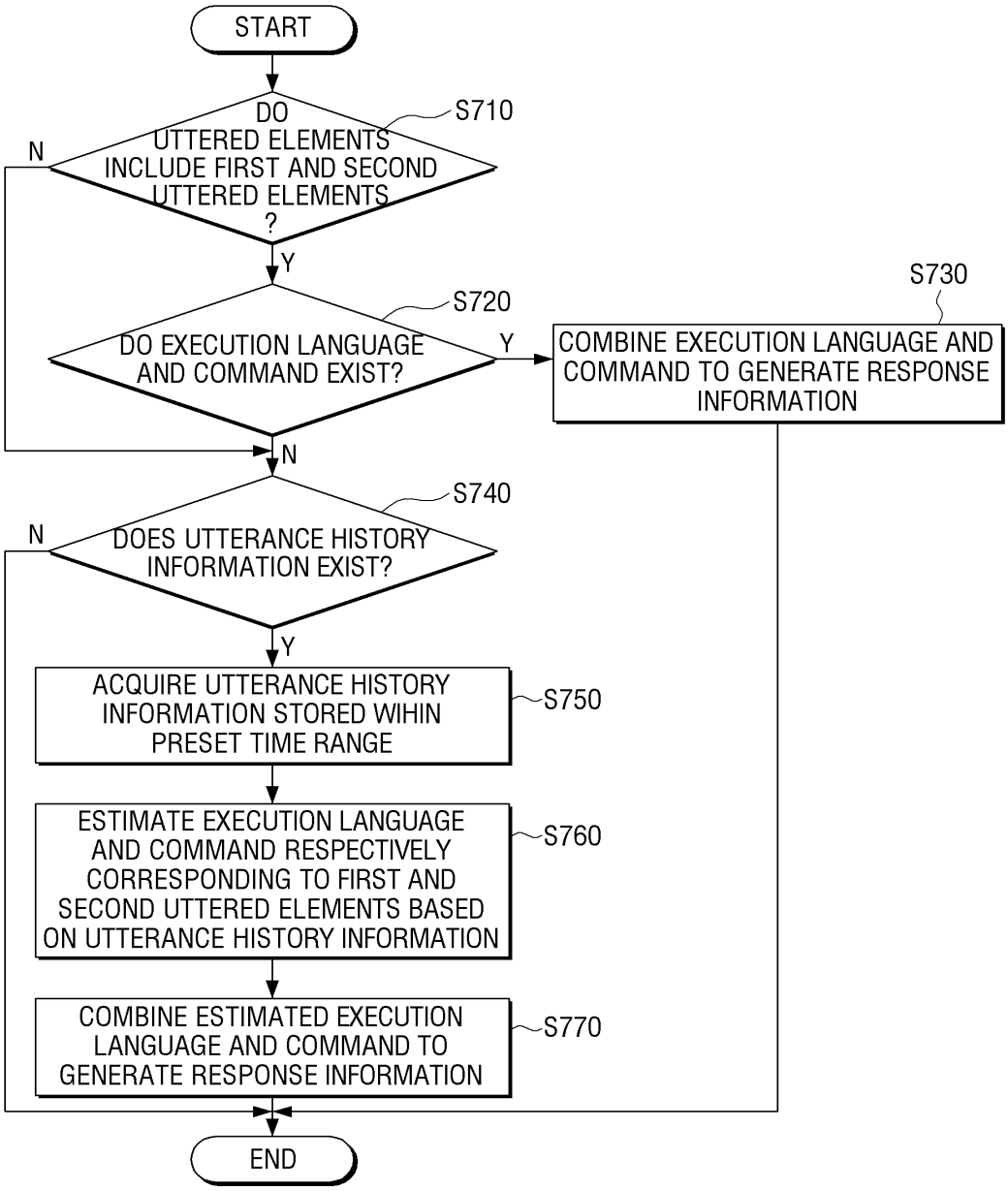

START
S610
RECEIVE UTTERED VOICE
S620
CONVERT UTTERED VOICE INTO TEXT INFORMATION
S630
EXTRACT UTTERED ELEMENTS
S640
IS IT POSSIBLE TO GENERATE RESPONSE INFORMATION ?
N
Y
S650
GENERATE RESPONSE INFORMATION
S660
TRANSMIT RESPONSE INFORMATION
END

FIG. 7


START
DO UTTERED ELEMENTS INCLUDE FIRST AND SECOND UTTERED ELEMENTS ?
S710
N
Y
DO EXECUTION LANGUAGE AND COMMAND EXIST?
S720
Y
COMBINE EXECUTION LANGUAGE AND COMMAND TO GENERATE RESPONSE INFORMATION
S730
N
DOES UTTERANCE HISTORY INFORMATION EXIST?
S740
N
Y
ACQUIRE UTTERANCE HISTORY INFORMATION STORED WIHIN PRESET TIME RANGE
S750
ESTIMATE EXECUTION LANGUAGE AND COMMAND RESPECTIVELY CORRESPONDING TO FIRST AND SECOND UTTERED ELEMENTS BASED ON UTTERANCE HISTORY INFORMATION
S760
COMBINE ESTIMATED EXECUTION LANGUAGE AND COMMAND TO GENERATE RESPONSE INFORMATION
S770
END

PTO/AIA/14 (03-13)
Approved for use through 01/31/2014. OMB 0651-0032
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 contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	Q203469
	Application Number	
Title of Invention	INTERACTIVE SERVER, CONTROL METHOD THEREOF, AND INTERACTIVE SYSTEM	

The application data sheet is part of the provisional or nonprovisional application for which it is being submitted. The following form contains the bibliographic data arranged in a format specified by the United States Patent and Trademark Office as outlined in 37 CFR 1.76.
This document may be completed electronically and submitted to the Office in electronic format using the Electronic Filing System (EFS) or the document may be printed and included in a paper filed application.

Secrecy Order 37 CFR 5.2

☐ Portions or all of the application associated with this Application Data Sheet may fall under a Secrecy Order pursuant to 37 CFR 5.2 (Paper filers only. Applications that fall under Secrecy Order may not be filed electronically.)

Inventor Information:

Inventor 1 Remove

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Ji-hye		CHUNG	

Residence Information (Select One) ○ US Residency ◉ Non US Residency ○ Active US Military Service

City	Seoul	Country of Residence i	KR

Mailing Address of Inventor:

Address 1	113-804, Jamsil els Apt., 19, Jamsil 2-dong		
Address 2	Songpa-gu		
City	Seoul	**State/Province**	
Postal Code		**Country** i	KR

Inventor 2 Remove

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Cheong-jae		LEE	

Residence Information (Select One) ○ US Residency ◉ Non US Residency ○ Active US Military Service

City	Suwon-si	Country of Residence i	KR

Mailing Address of Inventor:

Address 1	905-1501, Byeokjeokgol 9-danji Jugong Apt.		
Address 2	Yeongtong 2-dong, Yeongtong-gu, Gyeonggi-do		
City	Suwon-si	**State/Province**	
Postal Code		**Country** i	KR

Inventor 3 Remove

Legal Name

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	Q203469
	Application Number	
Title of Invention	INTERACTIVE SERVER, CONTROL METHOD THEREOF, AND INTERACTIVE SYSTEM	

Prefix	Given Name	Middle Name	Family Name	Suffix
	Hye-jeong		LEE	

Residence Information (Select One) ○ US Residency ◉ Non US Residency ○ Active US Military Service

City	Seoul	Country of Residence i	KR

Mailing Address of Inventor:

Address 1	101-404, Ssangyong Apt., Pungnap 2-dong		
Address 2	Songpa-gu		
City	Seoul	**State/Province**	
Postal Code		**Country** i	KR

Inventor 4 Remove

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Yong-wook		SHIN	

Residence Information (Select One) ○ US Residency ◉ Non US Residency ○ Active US Military Service

City	Osan-si	Country of Residence i	KR

Mailing Address of Inventor:

Address 1	1190-5, Woolim Apt., Galgot-dong		
Address 2	Gyeonggi-do		
City	Osan-si	**State/Province**	
Postal Code		**Country** i	KR

All Inventors Must Be Listed - Additional Inventor Information blocks may be generated within this form by selecting the **Add** button. Add

Correspondence Information:

Enter either Customer Number or complete the Correspondence Information section below. For further information see 37 CFR 1.33(a).

☐ **An Address is being provided for the correspondence Information of this application.**

Customer Number	23373		
Email Address	sughrue@sughrue.com	Add Email	Remove Email

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	Q203469
	Application Number	
Title of Invention	INTERACTIVE SERVER, CONTROL METHOD THEREOF, AND INTERACTIVE SYSTEM	

Application Information:

Title of the Invention	INTERACTIVE SERVER, CONTROL METHOD THEREOF, AND INTERACTIVE SYSTEM		
Attorney Docket Number	Q203469	**Small Entity Status Claimed** ☐	
Application Type	Nonprovisional		
Subject Matter	Utility		
Total Number of Drawing Sheets (if any)	7	**Suggested Figure for Publication (if any)**	

Publication Information:

☐ Request Early Publication (Fee required at time of Request 37 CFR 1.219)

☐ **Request Not to Publish.** I hereby request that the attached application not be published under 35 U.S.C. 122(b) and certify that the invention disclosed in the attached application **has not and will not** be the subject of an application filed in another country, or under a multilateral international agreement, that requires publication at eighteen months after filing.

Representative Information:

Representative information should be provided for all practitioners having a power of attorney in the application. Providing this information in the Application Data Sheet does not constitute a power of attorney in the application (see 37 CFR 1.32).
Either enter Customer Number or complete the Representative Name section below. If both sections are completed the customer Number will be used for the Representative Information during processing.

Please Select One:	◉ Customer Number	◯ US Patent Practitioner	◯ Limited Recognition (37 CFR 11.9)
Customer Number	23373		

Domestic Benefit/National Stage Information:

This section allows for the applicant to either claim benefit under 35 U.S.C. 119(e), 120, 121, or 365(c) or indicate National Stage entry from a PCT application. Providing this information in the application data sheet constitutes the specific reference required by 35 U.S.C. 119(e) or 120, and 37 CFR 1.78.

Prior Application Status			Remove
Application Number	Continuity Type	Prior Application Number	Filing Date (YYYY-MM-DD)

Additional Domestic Benefit/National Stage Data may be generated within this form by selecting the **Add** button. Add

Foreign Priority Information:

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	Q203469
	Application Number	
Title of Invention	INTERACTIVE SERVER, CONTROL METHOD THEREOF, AND INTERACTIVE SYSTEM	

This section allows for the applicant to claim priority to a foreign application. Providing this information in the application data sheet constitutes the claim for priority as required by 35 U.S.C. 119(b) and 37 CFR 1.55(d). When priority is claimed to a foreign application that is eligible for retrieval under the priority document exchange program (PDX) [i] the information will be used by the Office to automatically attempt retrieval pursuant to 37 CFR 1.55(h)(1) and (2). Under the PDX program, applicant bears the ultimate responsibility for ensuring that a copy of the foreign application is received by the Office from the participating foreign intellectual property office, or a certified copy of the foreign priority application is filed, within the time period specified in 37 CFR 1.55(g)(1).

Remove

Application Number	Country [i]	Filing Date (YYYY-MM-DD)	Access Code [i] (if applicable)
10-2013-0001829	KR	2013-01-07	

Additional Foreign Priority Data may be generated within this form by selecting the **Add** button. Add

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications

☐ This application (1) claims priority to or the benefit of an application filed before March 16, 2013 and (2) also contains, or contained at any time, a claim to a claimed invention that has an effective filing date on or after March 16, 2013.

Authorization to Permit Access:

☒ Authorization to Permit Access to the Instant Application by the Participating Offices

If checked, the undersigned hereby grants the USPTO authority to provide the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), the World Intellectual Property Office (WIPO), and any other intellectual property offices in which a foreign application claiming priority to the instant patent application is filed access to the instant patent application. See 37 CFR 1.14(c) and (h). This box should not be checked if the applicant does not wish the EPO, JPO, KIPO, WIPO, or other intellectual property office in which a foreign application claiming priority to the instant patent application is filed to have access to the instant patent application.

In accordance with 37 CFR 1.14(h)(3), access will be provided to a copy of the instant patent application with respect to: 1) the instant patent application-as-filed; 2) any foreign application to which the instant patent application claims priority under 35 U.S.C. 119(a)-(d) if a copy of the foreign application that satisfies the certified copy requirement of 37 CFR 1.55 has been filed in the instant patent application; and 3) any U.S. application-as-filed from which benefit is sought in the instant patent application.

In accordance with 37 CFR 1.14(c), access may be provided to information concerning the date of filing this Authorization.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	Q203469
	Application Number	
Title of Invention	INTERACTIVE SERVER, CONTROL METHOD THEREOF, AND INTERACTIVE SYSTEM	

Applicant Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

Applicant 1 Remove

If the applicant is the inventor (or the remaining joint inventor or inventors under 37 CFR 1.45), this section should not be completed. The information to be provided in this section is the name and address of the legal representative who is the applicant under 37 CFR 1.43; or the name and address of the assignee, person to whom the inventor is under an obligation to assign the invention, or person who otherwise shows sufficient proprietary interest in the matter who is the applicant under 37 CFR 1.46. If the applicant is an applicant under 37 CFR 1.46 (assignee, person to whom the inventor is obligated to assign, or person who otherwise shows sufficient proprietary interest) together with one or more joint inventors, then the joint inventor or inventors who are also the applicant should be identified in this section. Clear

(●) Assignee ( ) Legal Representative under 35 U.S.C. 117 ( ) Joint Inventor

( ) Person to whom the inventor is obligated to assign. ( ) Person who shows sufficient proprietary interest

If applicant is the legal representative, indicate the authority to file the patent application, the inventor is:

Name of the Deceased or Legally Incapacitated Inventor :

If the Applicant is an Organization check here. [X]

Organization Name	SAMSUNG ELECTRONICS CO., LTD.		
Mailing Address Information:			
Address 1	129, Samsung-ro, Yeongtong-gu, Gyeonggi-do		
Address 2			
City	Suwon-si	**State/Province**	
Country i	KR	Postal Code	443-742
Phone Number		Fax Number	
Email Address			

Additional Applicant Data may be generated within this form by selecting the Add button. Add

Non-Applicant Assignee Information:

Providing assignment information in this section does not subsitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	Q203469
	Application Number	
Title of Invention	INTERACTIVE SERVER, CONTROL METHOD THEREOF, AND INTERACTIVE SYSTEM	

Assignee 1

Complete this section only if non-applicant assignee information is desired to be included on the patent application publication in accordance with 37 CFR 1.215(b). Do not include in this section an applicant under 37 CFR 1.46 (assignee, person to whom the inventor is obligated to assign, or person who otherwise shows sufficient proprietary interest), as the patent application publication will include the name of the applicant(s).

Remove

If the Assignee is an Organization check here. ☐

Prefix	**Given Name**	Middle Name	**Family Name**	Suffix

Mailing Address Information:

Address 1	
Address 2	

City		**State/Province**	
Country i		Postal Code	
Phone Number		Fax Number	
Email Address			

Additional Assignee Data may be generated within this form by selecting the Add button. Add

Signature:

Remove

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

Signature	/Peter A. McKenna/			Date (YYYY-MM-DD)	2014-01-07
First Name	Peter A.	Last Name	McKenna	Registration Number	38551

Additional Signature may be generated within this form by selecting the Add button. Add

This collection of information is required by 37 CFR 1.76. 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 23 minutes to complete, including gathering, preparing, and submitting the completed application data sheet 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.**

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 the Freedom of Information Act requires disclosure of these records.

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 inspections 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.

No documents available for this priority number.

Espacenet

Bibliographic data: KR20090000442 (A) — 2009-01-07

GENERAL DIALOGUE SERVICE APPARATUS AND METHOD

Inventor(s): CHANG DU SEONG [KR] ± (CHANG, DU SEONG)

Applicant(s): KT CORP [KR] ± (KT CORPORATION)

Classification: **- international:** ***G06F17/00; G06F17/20***
- cooperative:

Application number: KR20070064510 20070628

Priority number(s): KR20070064510 20070628

Abstract of KR20090000442 (A)

A device and a method for providing a universal conversation service by using an information state and an example-based inference model according to a category of dialogue are provided to increase quality of the dialogue by inferring the dialogue and making a system response to all kinds of user questions intelligently based on information obtained from entire dialogue, which is included in a purpose and general category. A semantic interpreter(22) analyzes a user intention by interpreting the meaning of an input query in a classified category. A response inference engine(24) infers a present system intention for the user intention classified into each category. A system state controller(23) generates present system state information from the previous user/system intention and the accumulated system state information. A response template generator(28) generates a system response template for a target category from the preset user/system intention. A system response generator() generates the system response from the system response template for the target and general category.

Last updated: 13.03.2013 Worldwide Database 5.8.6.6; 93p

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 100 of 174

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2009-0000442
(43) 공개일자 2009년01월07일

(51) Int. Cl.
G06F 17/00 (2006.01) *G06F 17/20* (2006.01)
(21) 출원번호 10-2007-0064510
(22) 출원일자 2007년06월28일
심사청구일자 없음

(71) 출원인
주식회사 케이티
경기 성남시 분당구 정자동 206
(72) 발명자
장두성
경기 안양시 동안구 달안동 샛별한양아파트 303동 1303호
(74) 대리인
특허법인 신성

전체 청구항 수 : 총 22 항

(54) 범용 대화서비스 장치 및 그 방법

(57) 요 약

1. 청구범위에 기재된 발명이 속한 기술분야

본 발명은 범용 대화서비스 장치 및 그 방법과, 상기 방법을 실현시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체에 관한 것임.

2. 발명이 해결하려고 하는 기술적 과제

본 발명은 목적 영역의 대화와 일반 영역의 대화가 혼재되어 있는 영역(즉, 사용자의 다양한 자연어 질의)에서 대화 전반에 걸쳐 습득된 정보를 이용하여 대화를 추론하고, 모든 형태의 사용자 질의에 대해 지능적으로 시스템 답변을 만들어 대화 품질을 높일 수 있는, 범용 대화서비스 장치 및 그 방법과, 상기 방법을 실현시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공하는데 그 목적이 있음.

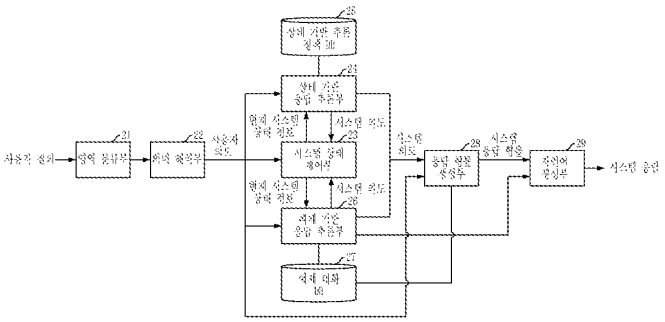
3. 발명의 해결방법의 요지

본 발명은, 대화서비스 장치에 있어서, 입력 질의어를 영역별로 분류하기 위한 영역분류수단; 상기 분류된 영역에서 입력 질의어의 의미를 해석하여 현재 사용자 의도를 분석하기 위한 의미해석수단; 현재 시스템 상태 정보를 참조하여, 영역별로 분류된 현재 사용자 의도에 대한 현재 시스템 의도를 추론하기 위한 응답추론수단; 이전 사용자/시스템 의도와 그 이전까지 누적된 시스템 상태 정보로부터 상기 현재 시스템 상태 정보를 생성하기 위한 시스템 상태 제어수단; 상기 현재 사용자 의도와 상기 현재 시스템 의도로부터 '목적 영역에 대한 시스템 응답 형틀'을 생성하기 위한 응답 형틀 생성수단; 및 상기 '목적 영역에 대한 시스템 응답 형틀' 및/또는 '일반 영역에 대한 시스템 응답 형틀'로부터 시스템 응답을 생성하기 위한 시스템 응답 생성수단을 포함한다.

4. 발명의 중요한 용도

본 발명은 대화서비스 등에 이용됨.

대 표 도 - 도2



VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 101 of 174

특허청구의 범위

청구항 1

대화서비스 장치에 있어서,

입력 질의어를 영역별로 분류하기 위한 영역분류수단;

상기 분류된 영역에서 입력 질의어의 의미를 해석하여 현재 사용자 의도를 분석하기 위한 의미해석수단;

현재 시스템 상태 정보를 참조하여, 영역별로 분류된 현재 사용자 의도에 대한 현재 시스템 의도를 추론하기 위한 응답추론수단;

이전 사용자/시스템 의도와 그 이전까지 누적된 시스템 상태 정보로부터 상기 현재 시스템 상태 정보를 생성하기 위한 시스템 상태 제어수단;

상기 현재 사용자 의도와 상기 현재 시스템 의도로부터 '목적 영역에 대한 시스템 응답 형틀'을 생성하기 위한 응답 형틀 생성수단; 및

상기 '목적 영역에 대한 시스템 응답 형틀' 및/또는 '일반 영역에 대한 시스템 응답 형틀'로부터 시스템 응답을 생성하기 위한 시스템 응답 생성수단

을 포함하는 범용 대화서비스 장치.

청구항 2

제 1 항에 있어서,

상기 현재 시스템 상태 정보는,

직전의 발화에서 시스템이 취한 의도(이전 시스템 의도), 그동안 대화에서 누적된 사용자의 의도(이전 사용자 의도)로 구성되는 것을 특징으로 하는 범용 대화서비스 장치.

청구항 3

제 2 항에 있어서,

상기 현재 시스템 상태 정보는,

초기 시스템 상태 정보와, 이를 바탕으로 이전 사용자 의도A에 대한 이전 시스템 의도B의 대화 1단계의 상태값(코딩값)과, 제1 시스템 상태 정보 '(사용자 의도A, 시스템 의도B)'를 바탕으로 이전 사용자 의도C에 대한 이전 시스템 의도D의 대화 2단계의 상태값(코딩값)과, 제2 시스템 상태 정보 '(사용자 의도A, 시스템 의도B), (사용자 의도C, 시스템 의도D)'를 바탕으로 현재 사용자 의도E에 대한 현재 시스템 의도F의 대화 3단계 상태값(코딩값)으로 이루어지는 것을 특징으로 하는 범용 대화서비스 장치.

청구항 4

제 1 항에 있어서,

상기 목적/일반 영역에 대한 시스템 응답 형틀은,

일부 어휘만을 교체하여 실제 자연어 문장을 만들어 낼 수 있는 생성 틀인 것을 특징으로 하는 범용 대화서비스 장치.

청구항 5

제 1 항에 있어서,

상기 질의어는,

언어적인 정보 및 비언어적인 정보를 포함하는 멀티모달 입력 중 하나인 것을 특징으로 하는 범용 대화서비스 장치.

청구항 6

제 1 항 내지 제 5 항 중 어느 한 항에 있어서,

상기 시스템 상태 제어수단은,

상기 이전/현재 시스템 상태 정보를 영역이 혼용된 대화의 기간 내에 통합 사용 및 유지·관리할 수 있는 것을 특징으로 하는 범용 대화서비스 장치.

청구항 7

제 6 항에 있어서,

상기 응답추론수단은,

목적 영역에 대한 상기 현재 사용자 의도를 입력받아, 상기 현재 시스템 상태 정보와 상태 기반 추론 정책을 기반으로 목적 영역에 대한 상기 현재 시스템 의도를 추론하기 위한 상태 기반 응답 추론수단; 및

일반 영역에 대한 상기 현재 사용자 의도를 입력받아, 상기 현재 시스템 상태 정보와 예제 대화 DB를 기반으로 일반 영역에 대한 상기 현재 시스템 의도 및 상기 '일반 영역에 대한 시스템 응답 형틀'을 추출하기 위한 예제 기반 응답 추론수단

을 포함하는 범용 대화서비스 장치.

청구항 8

제 7 항에 있어서,

상기 응답 형틀 생성수단은,

상기 예제 대화 DB를 참조하여, 목적 영역에 대한 상기 현재 사용자 의도와 상기 현재 시스템 의도로부터 상기 '목적 영역에 대한 시스템 응답 형틀'을 생성하는 것을 특징으로 하는 범용 대화서비스 장치.

청구항 9

제 8 항에 있어서,

상기 '목적/일반 영역에 대한 시스템 응답 형틀'이 상기 예제 대화 DB에 없는 경우, 예외처리용 응답 형틀을 참조하는 것을 특징으로 하는 범용 대화서비스 장치.

청구항 10

제 6 항에 있어서,

상기 시스템 응답 생성수단은,

상기 '목적/일반 영역에 대한 시스템 응답 형틀'로부터 자연어로 된 시스템 응답을 생성하는 것을 특징으로 하는 범용 대화서비스 장치.

청구항 11

제 7 항에 있어서,

상기 상태 기반 추론 정책은,

시스템의 각 상태에서 취할 수 있는 시스템의 의도(시스템이 어떠한 발화나 행동을 할 것인가를 나타내는 정보)를 수작업 혹은 통계적으로 나열한 정책인 것을 특징으로 하는 범용 대화서비스 장치.

청구항 12

대화서비스 장치에 있어서,

입력 질의어를 영역별로 분류하기 위한 영역분류수단;

상기 분류된 영역에서 입력 질의어의 의미를 해석하여 현재 사용자 의도를 분석하기 위한 의미해석수단;

목적 영역에 대한 현재 사용자 의도를 입력받아, 현재 시스템이 가지고 있는 상태 정보(현재 시스템 상태 정보)와 상태 기반 추론 정책을 기반으로 목적 영역에 대한 현재 시스템 의도를 추론하기 위한 상태 기반 응답 추론수단;

예제 대화 DB를 참조하여, 목적 영역에 대한 상기 현재 사용자 의도와 상기 현재 시스템 의도로부터 목적 영역에 대한 시스템 응답 형틀을 생성하기 위한 응답 형틀 생성수단;

일반 영역에 대한 상기 현재 사용자 의도를 입력받아, 상기 현재 시스템 상태 정보와 상기 예제 대화 DB를 기반으로 일반 영역에 대한 현재 시스템 의도 및 시스템 응답 형틀을 추출하기 위한 예제 기반 응답 추론수단;

이전 사용자/시스템 의도와 그 이전까지 누적된 시스템 상태 정보로부터 상기 현재 시스템 상태 정보를 생성하기 위한 시스템 상태 제어수단; 및

목적 영역 및/또는 일반 영역에 대한 상기 시스템 응답 형틀로부터 자연어로 된 시스템 응답을 생성하기 위한 자연어 생성수단

을 포함하는 범용 대화서비스 장치.

청구항 13

대화서비스 방법에 있어서,

입력 질의어를 영역별로 분류하는 단계;

상기 분류된 영역에서 입력 질의어의 의미를 해석하여 현재 사용자 의도를 분석하는 단계;

이전 사용자/시스템 의도와 그 이전까지 누적된 시스템 상태 정보로부터 현재 시스템 상태 정보를 생성하는 단계;

상기 현재 시스템 상태 정보와 상태 기반 추론 정책을 기반으로 목적 영역의 현재 사용자 의도에 대한 현재 시스템 의도를 추론하는 단계;

목적 영역에 대한 상기 현재 사용자 의도와 상기 현재 시스템 의도로부터 목적 영역에 대한 시스템 응답 형틀을 생성하는 단계;

상기 현재 시스템 상태 정보와 예제 대화 DB를 기반으로 일반 영역의 현재 사용자 의도에 대한 일반 영역의 현재 시스템 의도 및 시스템 응답 형틀을 추출하는 단계; 및

목적 영역 및/또는 일반 영역에 대한 상기 시스템 응답 형틀로부터 시스템 응답을 생성하는 단계

를 포함하는 범용 대화서비스 방법.

청구항 14

제 13 항에 있어서,

목적/일반 영역에 대한 상기 시스템 응답 형틀이 상기 예제 대화 DB에 없는 경우, 예외처리용 응답 형틀을 참조하여 시스템 응답 형틀을 생성하는 단계

를 더 포함하는 범용 대화서비스 방법.

청구항 15

제 13 항 또는 제 14 항에 있어서,

상기 현재 시스템 상태 정보는,

직전의 발화에서 시스템이 취한 의도(이전 시스템 의도), 그동안 대화에서 누적된 사용자의 의도(이전 사용자 의도)로 구성되는 것을 특징으로 하는 범용 대화서비스 방법.

청구항 16

제 15 항에 있어서,

상기 현재 시스템 상태 정보는,

초기 시스템 상태 정보와, 이를 바탕으로 이전 사용자 의도A에 대한 이전 시스템 의도B의 대화 1단계의 상태값(코딩값)과, 제1 시스템 상태 정보 '(사용자 의도A, 시스템 의도B)'를 바탕으로 이전 사용자 의도C에 대한 이전 시스템 의도D의 대화 2단계의 상태값(코딩값)과, 제2 시스템 상태 정보 '(사용자 의도A, 시스템 의도B), (사용자 의도C, 시스템 의도D)'를 바탕으로 현재 사용자 의도E에 대한 현재 시스템 의도F의 대화 3단계 상태값(코딩값)으로 이루어지는 것을 특징으로 하는 범용 대화서비스 방법.

청구항 17

제 15 항에 있어서,

상기 목적/일반 영역에 대한 시스템 응답 형틀은,

일부 어휘만을 교체하여 실제 자연어 문장을 만들어 낼 수 있는 생성 틀인 것을 특징으로 하는 범용 대화서비스 방법.

청구항 18

제 17 항에 있어서,

상기 질의어는,

언어적인 정보 및 비언어적인 정보를 포함하는 멀티모달 입력 중 하나인 것을 특징으로 하는 범용 대화서비스 방법.

청구항 19

제 17 항에 있어서,

상기 목적 영역에 대한 시스템 응답 형틀을 생성하는 과정에서는,

상기 예제 대화 DB를 참조하여, 목적 영역에 대한 상기 현재 사용자 의도와 상기 현재 시스템 의도로부터 상기 '목적 영역에 대한 시스템 응답 형틀'을 생성하는 것을 특징으로 하는 범용 대화서비스 방법.

청구항 20

제 17 항에 있어서,

상기 시스템 응답을 생성하는 과정에서는,

상기 '목적/일반 영역에 대한 시스템 응답 형틀'로부터 자연어로 된 시스템 응답을 생성하는 것을 특징으로 하는 범용 대화서비스 장치.

청구항 21

제 17 항에 있어서,

상기 상태 기반 추론 정책은,

시스템의 각 상태에서 취할 수 있는 시스템의 의도(시스템이 어떠한 발화나 행동을 할 것인가를 나타내는 정보)를 수작업 혹은 통계적으로 나열한 정책인 것을 특징으로 하는 범용 대화서비스 방법.

청구항 22

프로세서를 구비한 범용 대화서비스 장치에,

입력 질의어를 영역별로 분류하는 기능;

상기 분류된 영역에서 입력 질의어의 의미를 해석하여 현재 사용자 의도를 분석하는 기능;

이전 사용자/시스템 의도와 그 이전까지 누적된 시스템 상태 정보로부터 현재 시스템 상태 정보를 생성하는 기능;

상기 현재 시스템 상태 정보와 상태 기반 추론 정책을 기반으로 목적 영역의 현재 사용자 의도에 대한 현재 시

스템 의도를 추론하는 기능;

목적 영역에 대한 상기 현재 사용자 의도와 상기 현재 시스템 의도로부터 목적 영역에 대한 시스템 응답 형틀을 생성하는 기능;

상기 현재 시스템 상태 정보와 예제 대화 DB를 기반으로 일반 영역의 현재 사용자 의도에 대한 일반 영역의 현재 시스템 의도 및 시스템 응답 형틀을 추출하는 기능; 및

목적 영역 및/또는 일반 영역에 대한 상기 시스템 응답 형틀로부터 시스템 응답을 생성하는 기능

을 실현시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

명 세 서

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

<12> 본 발명은 대화의 영역에 따라 정보상태 및 예제 기반 추론 모델을 같이 이용한 대화서비스 장치 및 그 방법과, 상기 방법을 실현시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체에 관한 것으로, 보다 상세하게는 목적 영역의 대화와 일반 영역의 대화가 혼재된 사용자의 다양한 자연어 질의에 대해 지능적으로 시스템 답변이나 재질문을 함으로써 대화 품질을 높일 수 있는 범용 대화서비스 장치 및 그 방법과, 상기 방법을 실현시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체에 관한 것이다.

<13> 특히, 본 발명은 대화서비스 시스템[사용자 질의에 대해 응답하는 '네트워크(예로서, 유무선 인터넷(WEB, WAP/ME), 인트라넷, 랜(LAN), 무선랜(Wireless LAN), 전화망(PSTN), 동기/비동기 이동통신망, 휴대 인터넷(WiBro), 지상파 방송망, DMB(Digital Multimedia Broadcasting) 방송망 등)를 통한 클라이언트-서버 인터페이스 구조', '사용자-로봇 인터페이스 구조' 등]에서 입력되는 사용자 질의에 대해 호텔/극장/열차/항공권 예매나 홈네트워크 기기제어 등과 같이 시스템이 설정한 특정 목적 영역과, 채팅 등과 같이 일반 대화가 혼용되어 사용될 때, 보다 유용하게 활용될 수 있다.

<14> 음성 인터페이스는 컴퓨터로 하여금 사람의 말을 알아듣게 하고, 원하는 정보를 말로 주고 받도록 가능하게 해주는 사용자 중심의 휴먼-컴퓨터 인터페이스 기술이다. 현재, 음성인식 기술이 활발하게 적용되고 분야는 전화 기반 고객 서비스 부문으로, 컴퓨터가 고객과 직접 대화하며 필요한 정보를 주는 방식으로 활용되고 있다.

<15> 특히, 지능형 로봇, 텔레매틱스, 디지털 홈, 차세대 PC 등 유비쿼터스 환경에서 사용자 인터페이스로 음성 인터페이스가 자리매김하고 있다.

<16> 현재, 음성 인터페이스는 사용자-시스템 간의 '대화'하는 방식을 끊임없이 변화시키고 있다.

<17> 하지만, 현재의 대화서비스(예로서, 전화 기반 고객응대 서비스, 로봇의 음성인식 서비스 등)는 미리 정해진 범위의 대화 이외에는 불가능하기 때문에 응용폭이 제한된다는 단점이 있다.

<18> 사용자가 대화서비스 장치(일 예, 네트워크 기반의 서버, 지능형 로봇)와의 자연스러운 의사 소통을 위해서는, 사용자의 음성 질의를 대화서비스 장치가 듣고 이해하여 사용자가 원하는 대답을 음성 또는 디스플레이 화면을 통해 제공할 수 있어야 한다. 이를 위해서는 음성 질의 문장을 인식할 수 있는 음성인식 기능, 대화 처리를 통해 사용자의 질의에 대한 적절한 응답을 생성하는 기능, 음성 합성을 통해 음성을 결과로 출력하는 기능, 검색된 정보에 대한 디스플레이 기능 등이 필요하다.

<19> 따라서, 현재까지 대화서비스 기술에 대한 연구는 대화서비스 장치의 음성인식 성능을 향상시키고자 하는 연구에 초점이 맞춰져 있다. 그러므로, 음성인식 성능을 향상시키기 위해서는, 대화서비스 장치가 사용자의 질의에 대해 키워드를 인식하고 키워드에 대한 적절한 응답을 제공하면, 음성인식 성능이 높지 않더라도 대화를 진행하는 사용자가 불편하지 않게 되므로, 음성인식의 성능 향상에 대한 노력뿐만 아니라, 대화의 자연스러움을 위한 대화 처리에 대한 연구가 병행될 필요가 있다.

<20> 하지만, 현재까지는 사용자의 음성 질의에 대한 대화서비스 장치의 인식 성능이 높지 않기 때문에 사용자-대화서비스 장치 간의 자연스러운 대화(지능형 대화)의 진행이 어렵고, 이로 인해 대화서비스 장치의 품질이 떨어져

VB Assets Exhibit No. 2005
Samsung Elec. and Samsung Elec. Am. v. VB Assets, LLC
IPR2025-00868, U.S. Patent No. 8,515,765
Page 106 of 174

사용자에게 질 높은 지능형 대화서비스를 제공할 수 없는 문제점이 있었다.

<21> 통상, 대화서비스 장치에서 이루어지는 사용자와 시스템 간의 대화는, 홈네트워크 기기 제어 목적이나 열차 예약 목적과 같이 대화서비스 제공자가 제공하는 특정 목적 영역에 관한 대화와, 채팅과 같은 일반적인 영역의 대화로 구분할 수 있다.

<22> 대화서비스 시 목적 영역에서 수행되는 대화는, 사용자와 시스템(대화서비스 장치)이 여러 번에 걸쳐 질의와 응답을 주고 받으며, 특정한 목적을 표현하는 정보의 집합을 구성해 나간다. 예를 들어 설명하면, 열차 예약을 목적으로 하는 서비스 영역에서 이루어지는 대화는 '열차 예약'이라는 공동의 목적을 위해 '출발역', '도착역', '일자', '시간', '열차종류', '예약편수' 등의 정보 집합을 대화를 통해 구성해 나간다. 이때, 대화는 이들 정보 집합을 만들기 위해서 사용자의 발화에서 부족한 정보를 시스템(대화서비스 장치)이 사용자에게 요청하기도 하고, 사용자가 시스템(대화서비스 장치)에게 자신이 필요한 정보를 문의하기도 하는 내용으로 이루어진다.

<23> 이러한 목적 영역에서 시스템(대화서비스 장치)이 응답을 생성하기 위해서는, 직전에 사용자가 발화한 질의의 의도(의미) 이외에 현재의 시스템의 상태, 즉 시스템이 가지고 있는 정보의 현황에 기반한 응답 추론이 이루어져야 한다. 이러한 정보의 현황을 응답 추론에 반영하기 위하여 수작업으로 정의한 상태들의 전이를 유한 상태 오토마타(FSA : Finite State Automata)로 표현하여, 이를 이용한 방법이 사용되어 왔다.

<24> 다른 방법으로는, 각 상태에서 최종 목적 상태까지 도달하기 위한 계획(Plan)을 미리 준비하고 현재 상태에 맞는 적절한 계획을 선택하는 방법도 사용되었으며, 최근에는 시스템이 가지고 있는 정보의 상태 분류와 이들 간의 전이 모델을 통계적 훈련에 의해 모델링하는 마르코프 의사결정과정(MDP : Markov Decision Process), 부분관찰 마르코프 의사결정과정(POMDP : Partially Observable Markov Decision Process) 등이 도입되고 있다.

<25> 그러나, 앞서 언급된 응답 추론 방법들은 한정된 목적 영역의 사용자 질의에서 최적의 효율을 낼 수 있도록 작성 혹은 학습할 수 있다는 장점이 있는 반면, 예정된 영역이 아닌 일반 영역의 사용자 발화에서는 지능적인 시스템 발화를 만들어 낼 수 없다는 한계가 있다.

<26> 한편, 채팅 등과 같은 일반 영역에서의 사용자 발화는 그 발화의 목적이 매우 다양하여 모든 가능한 상태를 미리 정의하는 것이 사실상 불가능하다. 이러한 채팅영역에 대한 시스템 응답을 생성하기 위해서는, 미리 구축된 예제 DB에서 사용자의 질의에 가장 유사한 대화를 찾아 해당하는 시스템의 응답 예제를 시스템의 응답 발화로 삼는 방법을 주로 사용한다. 이 방법은 일정 수준 이상의 예제 DB가 구축되면, 여러 형태의 사용자 질의에 적절한 응답을 생성할 수 있다는 장점이 있다. 반면에, 사용자에게 보충 질문을 하는 것과 같은 목적 의도를 가진 시스템 응답을 생성하지 못하므로 목적 영역의 대화 관리 방법으로는 부족하다는 한계가 있다.

<27> 특히, 대화서비스가 처리할 수 있는 영역을 특정한 영역(예로서, 목적 영역 또는 일반 영역)으로만 한정하는 것은 대화서비스를 사용하는 사용자의 발화 자유도와 서비스 완성도에 부정적인 영향을 준다. 즉, 홈네트워크나 텔레매틱스와 같은 특정 목적 영역을 주로 사용하는 대화서비스라고 하더라도, 사용자는 이 서비스를 사용하면서 서비스의 목적 영역에서 벗어나는 예를 들면 "오늘 어때?" 등과 같은 발화를 할 수도 있으며, 그때마다 시스템(대화서비스 장치)이 "죄송합니다. 다시 말해 주십시오." 혹은 "적합하지 않은 발화입니다."라고 한다면 지능적인 대화시스템이라고 할 수 없을 것이다. 이러한 경우에, "한번 즐겨봅시다."와 같이 적절한 시스템 응답을 준다면, 보다 지능적인 대화라 할 수 있다.

<28> 이처럼 예약이나 기기제어 등과 같은 목적 영역의 대화와 채팅 등과 같은 일반 영역의 대화가 혼재되어 있는 영역에서는, 기존의 계획(Plan) 기반 방법이나 FSN, MDP 등을 이용한 응답 추론 방법은 추론의 대상이 되는 영역이 너무나 방대하여 추론 규칙이나 대화의 상태 및 전이도를 작성하거나 학습하는 것이 거의 불가능하다. 또한, 예제 기반 대화 추론 방법으로는 대화 전반에 걸쳐 습득된 정보를 이용하여 목적 영역의 대화를 진행하기에 부족함이 있다.

<29> 따라서, 모든 형태의 사용자 질의(목적 영역의 대화와 일반 영역의 대화가 혼용되어 사용되는 질의어)에 대해 지능적으로 시스템 답변을 만들어 높은 대화 품질을 가지는 범용의 대화서비스 방안이 절실히 요구된다.

발명이 이루고자 하는 기술적 과제

<30> 본 발명은 상기 문제점을 해결하기 위하여 제안된 것으로, 목적 영역의 대화와 일반 영역의 대화가 혼재되어 있는 영역(즉, 사용자의 다양한 자연어 질의)에서 대화 전반에 걸쳐 습득된 정보를 이용하여 대화를 추론하고, 모든 형태의 사용자 질의에 대해 지능적으로 시스템 답변을 만들어 대화 품질을 높일 수 있는, 범용 대화서비스 장치 및 그 방법과, 상기 방법을 실현시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공하

는데 그 목적이 있다.

<31> 본 발명의 다른 목적 및 장점들은 하기의 설명에 의해서 이해될 수 있으며, 본 발명의 실시예에 의해 보다 분명하게 알게 될 것이다. 또한, 본 발명의 목적 및 장점들은 특허 청구 범위에 나타낸 수단 및 그 조합에 의해 실현될 수 있음을 쉽게 알 수 있을 것이다.

발명의 구성 및 작용

<32> 상기 목적을 달성하기 위한 본 발명은, 대화서비스 장치에 있어서, 입력 질의어를 영역별로 분류하기 위한 영역분류수단; 상기 분류된 영역에서 입력 질의어의 의미를 해석하여 현재 사용자 의도를 분석하기 위한 의미해석수단; 현재 시스템 상태 정보를 참조하여, 영역별로 분류된 현재 사용자 의도에 대한 현재 시스템 의도를 추론하기 위한 응답추론수단; 이전 사용자/시스템 의도와 그 이전까지 누적된 시스템 상태 정보로부터 상기 현재 시스템 상태 정보를 생성하기 위한 시스템 상태 제어수단; 상기 현재 사용자 의도와 상기 현재 시스템 의도로부터 '목적 영역에 대한 시스템 응답 형틀'을 생성하기 위한 응답 형틀 생성수단; 및 상기 '목적 영역에 대한 시스템 응답 형틀' 및/또는 '일반 영역에 대한 시스템 응답 형틀'로부터 시스템 응답을 생성하기 위한 시스템 응답 생성수단을 포함하여 이루어진 것을 특징으로 한다.

<33> 한편, 본 발명은, 대화서비스 장치에 있어서, 입력 질의어를 영역별로 분류하기 위한 영역분류수단; 상기 분류된 영역에서 입력 질의어의 의미를 해석하여 현재 사용자 의도를 분석하기 위한 의미해석수단; 목적 영역에 대한 현재 사용자 의도를 입력받아, 현재 시스템이 가지고 있는 상태 정보(현재 시스템 상태 정보)와 상태 기반 추론 정책을 기반으로 목적 영역에 대한 현재 시스템 의도를 추론하기 위한 상태 기반 응답 추론수단; 예제 대화 DB를 참조하여, 목적 영역에 대한 상기 현재 사용자 의도와 상기 현재 시스템 의도로부터 목적 영역에 대한 시스템 응답 형틀을 생성하기 위한 응답 형틀 생성수단; 일반 영역에 대한 상기 현재 사용자 의도를 입력받아, 상기 현재 시스템 상태 정보와 상기 예제 대화 DB를 기반으로 일반 영역에 대한 현재 시스템 의도 및 시스템 응답 형틀을 추출하기 위한 예제 기반 응답 추론수단; 이전 사용자/시스템 의도와 그 이전까지 누적된 시스템 상태 정보로부터 상기 현재 시스템 상태 정보를 생성하기 위한 시스템 상태 제어수단; 및 목적 영역 및/또는 일반 영역에 대한 상기 시스템 응답 형틀로부터 자연어로 된 시스템 응답을 생성하기 위한 자연어 생성수단을 포함하여 이루어진 것을 특징으로 한다.

<34> 다른 한편, 본 발명은, 대화서비스 방법에 있어서, 입력 질의어를 영역별로 분류하는 단계; 상기 분류된 영역에서 입력 질의어의 의미를 해석하여 현재 사용자 의도를 분석하는 단계; 이전 사용자/시스템 의도와 그 이전까지 누적된 시스템 상태 정보로부터 현재 시스템 상태 정보를 생성하는 단계; 상기 현재 시스템 상태 정보와 상태 기반 추론 정책을 기반으로 목적 영역의 현재 사용자 의도에 대한 현재 시스템 의도를 추론하는 단계; 목적 영역에 대한 상기 현재 사용자 의도와 상기 현재 시스템 의도로부터 목적 영역에 대한 시스템 응답 형틀을 생성하는 단계; 상기 현재 시스템 상태 정보와 예제 대화 DB를 기반으로 일반 영역의 현재 사용자 의도에 대한 일반 영역의 현재 시스템 의도 및 시스템 응답 형틀을 추출하는 단계; 및 목적 영역 및/또는 일반 영역에 대한 상기 시스템 응답 형틀로부터 시스템 응답을 생성하는 단계를 포함하여 이루어진 것을 특징으로 한다.

<35> 또한, 본 발명은, 목적/일반 영역에 대한 상기 시스템 응답 형틀이 상기 예제 대화 DB에 없는 경우, 예외처리용 응답 형틀을 참조하여 시스템 응답 형틀을 생성하는 단계를 더 포함하여 이루어진 것을 특징으로 한다.

<36> 또 다른 한편, 본 발명은, 프로세서를 구비한 범용 대화서비스 장치에, 입력 질의어를 영역별로 분류하는 기능; 상기 분류된 영역에서 입력 질의어의 의미를 해석하여 현재 사용자 의도를 분석하는 기능; 이전 사용자/시스템 의도와 그 이전까지 누적된 시스템 상태 정보로부터 현재 시스템 상태 정보를 생성하는 기능; 상기 현재 시스템 상태 정보와 상태 기반 추론 정책을 기반으로 목적 영역의 현재 사용자 의도에 대한 현재 시스템 의도를 추론하는 기능; 목적 영역에 대한 상기 현재 사용자 의도와 상기 현재 시스템 의도로부터 목적 영역에 대한 시스템 응답 형틀을 생성하는 기능; 상기 현재 시스템 상태 정보와 예제 대화 DB를 기반으로 일반 영역의 현재 사용자 의도에 대한 일반 영역의 현재 시스템 의도 및 시스템 응답 형틀을 추출하는 기능; 및 목적 영역 및/또는 일반 영역에 대한 상기 시스템 응답 형틀로부터 시스템 응답을 생성하는 기능을 실현시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공한다.

<37> 상술한 목적, 특징 및 장점은 첨부된 도면과 관련한 다음의 상세한 설명을 통하여 보다 분명해 질 것이며, 그에 따라 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 본 발명의 기술적 사상을 용이하게 실시할 수 있을 것이다. 또한, 본 발명을 설명함에 있어서 본 발명과 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에 그 상세한 설명을 생략하기로 한다. 이하, 첨부된 도면을

참조하여 본 발명에 따른 바람직한 일실시예를 상세히 설명하기로 한다.

<38> 도 1 은 본 발명에 따른 대화서비스 장치가 적용되는 음성인식 시스템의 일실시예 구성도이다.

<39> 대화시스템의 입력으로는 음성인식, 키보드 입력 등을 통해 입력된 언어적인 정보 뿐만 아니라, 터치인식, 화상인식, 동작인식 등의 결과로 기호화된 비언어적인 정보를 포함한 언어적/비언어적 멀티모달 입력을 받는 것이 가능하다.

<40> 사용자로부터 질의(문자 또는 음성 질의어 등)가 입력되면(사용자가 네트워크(예로서, 유무선 인터넷(WEB, WAP/ME), 인트라넷, 랜(LAN), 무선랜(Wireless LAN), 전화망(PSTN), 동기/비동기 이동통신망, 휴대 인터넷(WiBro), 지상파 방송망, DMB(Digital Multimedia Broadcasting) 방송망 등)를 통해 질의어가 입력되거나, 직접 질의어가 입력됨 등), 입력이 음성일 경우(이때, 직접 입력되는 내용이 문자이면 음성인식 기능을 수행하지는 않음), 음성인식 시스템의 입력 처리부(12)는 음성인식 기능을 수행하여 음성 인식 결과 문장을 대화서비스 장치(20)로 출력한다. 그러면, 대화서비스 장치(20)는 음성 인식부(12)로부터 사용자 질의(사용자의 다양한 자연어 질의 문장, 즉 목적 영역의 대화와 일반 영역의 대화가 혼재되어 있음)를 입력받아 대화 전반에 걸쳐 습득된 정보를 이용하여 대화를 추론하고, 모든 형태의 사용자 질의에 대해 지능적으로 시스템 답변(응답 문장)을 생성한 후 음성 합성부(13)로 출력한다(음성 형식으로 출력). 그리고, 디스플레이부(14)는 응답 결과(시스템 답변)를 디스플레이할 필요가 있는 경우, 대화서비스 장치(20)로부터 사용자 질의에 대한 응답 문장을 입력받아 사용자에게 디스플레이(문자 형식으로 출력)한다.

<41> 여기서, 입력 처리부(12)는 음성인식을 위해 음성 데이터베이스를 구비하고, 대화서비스 장치(20)는 입력에 적절한 응답 메시지(대화 예제)를 저장하고 있는 예제 대화 데이터베이스를 구비하고 있다. 이때, 음성 데이터베이스에는 사용자의 음성발화 데이터에 대한 문자(열)의 특징값과 해당 특징값에 대응하는 문자(열)이 저장 및 관리된다. 또한, 예제 대화 데이터베이스에는 대화의 예제가 저장 및 관리된다.

<42> 상기에서, 음성인식 시스템과 사용자 간에는 유무선 인터넷(WEB, WAP/ME), 인트라넷, 랜(LAN), 무선랜(Wireless LAN), 전화망(PSTN), 동기/비동기 이동통신망, 휴대 인터넷(WiBro), 지상파 방송망, DMB(Digital Multimedia Broadcasting) 방송망 등을 통해 통신회선을 연결하여 상호 간에 대화서비스와 관련된 통신이 이루어지도록 연결 설정된다.

<43> 따라서, 네트워크를 통해 연결된 사용자 단말기(11)로는 통신중재장치, 모니터, 키보드(혹은 키패드), 마우스, 음성을 입력할 수 있는 장치 등이 포함되어진 개인용컴퓨터(PC), 노트북 뿐만 아니라, 일반 유무선 전화기, 셀룰러폰 혹은 PCS폰 등의 이동통신 단말기, IMT-2000, UMTS 등과 같은 차세대 이동통신 단말기, 개인휴대단말기(PDA), DMB폰 등 통신망 기반의 대화서비스(질의/응답)가 가능한 단말기면 족하다.

<44> 호텔/극장/열차/항공권 예매나 홈네트워크 기기제어, 채팅 등과 같이 시스템이 설정한 특정 목적 영역과 일반 대화 영역이 혼재되어 있는 환경에서 대화서비스를 원하는 사용자는 단말기(11)로 네트워크[예로서, 유무선 인터넷(WEB, WAP/ME), 인트라넷, 랜(LAN), 무선랜(Wireless LAN), 전화망(PSTN), 동기/비동기 이동통신망, 휴대 인터넷(WiBro), 지상파 방송망, DMB(Digital Multimedia Broadcasting) 방송망 등]를 통해 질의어(문자 또는 음성발화 등)를 입력하거나, 직접 질의어(문자 또는 음성발화 등)를 입력하여 본 발명에 따른 범용 대화서비스를 수행할 수 있다. 이때, 사용자 단말기(11)에는 음성 입력 및 출력을 위한 사운드카드, 마이크 그리고 스피커 등이 구비될 수 있고, 문자 입력 및 출력을 위한 키보드, 디스플레이부 등이 구비될 수 있다. 즉, 사용자가 단말기(11)의 입력장치(예를 들면, 키보드, 마이크 등)를 통해 문자 혹은 음성발화를 입력하게 되면, 음성인식 서비스 시스템은 음성인식 기능을 포함한 입력 처리부(12)와, 의미 분석 및 적절한 응답 메시지를 출력할 수 있는 대화서비스 장치(20), 그리고 해당 응답 메시지를 음성 및/또는 문자 형식으로 출력해주는 출력 처리부[음성 합성부(13), 디스플레이부(14)]를 통해, 목적 영역의 대화와 일반 영역의 대화가 혼재되어 있는 영역에서 대화 전반에 걸쳐 습득된 정보를 이용하여 대화를 추론하고 모든 형태의 사용자 질의에 대해 지능적으로 시스템 답변을 만들어 높은 대화 품질을 갖도록 한다.

<45> 여기서, 입력 처리부(12)는 네트워크를 통해 또는 직접 입력 전달된 사용자의 입력(문자 또는 음성발화 등) 데이터를 1차로 처리한다. 이때, 입력 데이터가 문자 데이터인 경우에는 따옴표 등의 불필요한 요소를 제거하는 과정을 거치고, 입력 데이터가 음성발화 데이터인 경우에는 음성인식 기능을 통해 해당 문자열로 변환하는 과정을 거친다. 이때, 음성인식 기능시, 음성발화 데이터에서 특징을 추출하고 분석하여, 음성의 특징값과 해당 특징값에 대응되는 문자열을 저장·관리하고 있는 음성 데이터베이스에서 가장 근접한 결과를 추출하게 된다. 이렇게 입력 처리부(12)에서 변환된 사용자 입력 데이터(인식에 불필요한 요소를 제거한 문자 데이터, 문자화된

음성발화 데이터)는 문자열로 대화서비스 장치(20)로 전달된다.

<46> 또한, 대화서비스 장치(20)는 사용자가 문자 혹은 음성발화를 입력했을 때 해당 문장에 대해 의미 분석을 하고 그 정보를 검색키로 하여 사용자의 질의에 맞는 시스템 응답을 문자열로 생성하여 출력 처리부(13,14)로 전달한다.

<47> 또한, 출력 처리부(13,14)에서는 대화서비스 장치(20)에서 생성된 문자열의 시스템 응답(시스템 응답 메시지)을 음성 합성부(13)를 통해 음성 데이터로 변환하여 출력하고, 문자 데이터를 디스플레이부(14)를 통해 디스플레이한다.

<48> 만약, 사용자가 사용자 단말기(11)를 통해 질의어를 입력한 경우, 사용자 단말기(11)에서는 전달된 음성 데이터를 스피커를 통해 출력하고, 문자 데이터를 화면상에 출력하게 된다. 이 경우, 출력 처리부(13,14)는 문자열의 시스템 발화(시스템 응답)를 그대로(문자 데이터) 사용자 단말기(11)로 전송하거나, 문자열의 시스템 발화를 음성으로 변환하여 음성 데이터를 사용자 단말기(11)로 전송하거나, 문자열의 시스템 발화를 음성으로 변환하여 음성 데이터를 문자 데이터와 함께 사용자 단말기(11)로 전송한 것이다.

<49> 이후, 사용자는 시스템 응답(시스템 발화) 메시지를 보고 다음 입력을 진행하게 되고, 음성인식 시스템은 상기의 과정을 반복적으로 수행하여 대화서비스를 지능형으로 제공하게 된다.

<50> 그럼, 본 발명에 따른 대화서비스 장치(20)에 대해 보다 상세하게 살펴보기로 한다.

<51> 대화서비스 장치(20)는 사용자의 자연어 질의(문자 또는 음성발화 등)에 대해 답변이나 재질문을 함으로써 대화를 진행하는데, 이때 대화서비스 장치(20)는 입력되는 사용자의 질의가 호텔예약이나 홈네트워크 기기제어 등과 같이 시스템이 설정한 특정한 목적 영역에 한정되지 않고, 채팅과 같은 일반 대화와 혼용되어 사용될 때, 입력되는 사용자 질의(사용자의 다양한 자연어 질의 문장, 즉 목적 영역의 대화와 일반 영역의 대화가 혼재되어 있음)에 대해 목적 영역 대화와 일반 영역 대화로 분류하고, 사용자 질의의 영역에 따라 적합한 응답 추론부(상태 기반 응답 추론부(24) or 예제 기반 응답 추론부(26))를 사용하고, 시스템 상태 제어부(23)를 이용하여 시스템 상태 정보를 영역이 혼용된 대화의 기간 내에 통합 사용/유지할 수 있도록 하여 생성되는 시스템 응답의 품질을 높일 수 있다[범용 대화관리 기능].

<52> 먼저, 도 1을 참조하여, 본 발명에 따른 범용 대화서비스 장치(20)의 구성을 살펴보면, 입력 질의어(문자 데이터 또는 음성발화 데이터 중 어느 하나임)를 영역(목적 영역, 일반 영역)별로 분류하기 위한 영역 분류부(21)와, 분류된 영역(목적 영역 또는 일반 영역)의 특화된 의미(목적 영역 의미 해석, 일반 영역 의미 해석)를 해석하여 현재 사용자 의도를 분석하기 위한 의미 해석부(22)와, 현재 시스템 상태 정보(직전의 발화에서 시스템이 취한 의도(이전 시스템 의도), 그동안 대화에서 누적된 사용자의 의도(이전 사용자 의도)로 구성됨)를 참조하여, 영역별로 분류된 현재 사용자 의도에 대한 현재 시스템 의도를 추론하기 위한 응답 추론부(상태 기반 응답 추론부(24), 예제 기반 응답 추론부(26))와, 이전 사용자/시스템 의도와 그 이전까지 누적된 시스템 상태 정보로부터 현재 시스템 상태 정보를 생성하기 위한 시스템 상태 제어부(23)와, 현재 사용자 의도와 현재 시스템 의도로부터 목적 영역에 대한 시스템 응답 형틀(일부 어휘만을 교체하여 실제 자연어 문장을 만들어 낼 수 있는 생성 틀임)을 생성하기 위한 응답 형틀 생성부(28)와, 목적 영역에 대한 시스템 응답 형틀 및/또는 일반 영역에 대한 시스템 응답 형틀로부터 시스템 응답을 생성하기 위한 자연어 생성부(29)를 포함한다.

<53> 이하에서는, 대화 1단계에서 사용자의 의도를 '사용자 의도A'로 정의하고, '사용자 의도A'에 대한 시스템 의도를 '시스템 의도B'로 정의한다. 또한, 대화 2단계에서 사용자의 의도를 '사용자 의도C'로 정의하고, '사용자 의도C'에 대한 시스템 의도를 '시스템 의도D'로 정의한다. 또한, 대화 3단계에서 사용자의 의도를 '사용자 의도E'로 정의하고, '사용자 의도E'에 대한 시스템 의도를 '시스템 의도F'로 정의한다.

<54> 따라서, 상기 '현재 시스템 상태 정보'는, 초기 시스템 상태 정보와, 이를 바탕으로 사용자 의도A에 대한 시스템 의도B의 대화 1단계의 상태값(코딩값)과, 제1 시스템 상태 정보 "(사용자 의도A, 시스템 의도B)"를 바탕으로 사용자 의도C에 대한 시스템 의도D의 대화 2단계의 상태값(코딩값), 제2 시스템 상태 정보 "(사용자 의도A, 시스템 의도B), (사용자 의도C, 시스템 의도D)"를 바탕으로 사용자 의도E에 대한 시스템 의도F의 대화 3단계 상태값(코딩값)으로 이루어진다.

<55> 즉, '시스템 의도B'는 사용자 의도A와 초기 시스템 상태 정보를 바탕으로 생성되고, '시스템 의도D'는 사용자 의도C와 제1 이전 시스템 상태 정보(초기 시스템 상태 정보를 바탕으로 한 사용자 의도A에 대한 시스템 의도B의 상태값(코딩값))를 바탕으로 생성되며, '시스템 의도F'는 사용자 의도E와 제2 이전 시스템 상태 정보(제1 이전

시스템 상태 정보를 바탕으로 한 사용자 의도C에 대한 시스템 의도D의 상태값)를 바탕으로 생성된다.

<56> 상기 이전/현재 시스템 상태 정보는 시스템 상태 제어부(23)에 의해 관리되며, 시스템 상태 제어부(23)에서는 현재 시스템 상태 정보를 영역이 혼용된 대화의 기간 내에 통합 사용 및 유지·관리할 수 있다.

<57> 상기에서, 응답 추론부(24)는, 목적 영역에 대한 현재 사용자 의도를 입력받아, 현재 시스템 상태 정보와 상태 기반 추론 정책(상태 기반 추론 정책 DB(25))을 기반으로 목적 영역에 대한 현재 시스템 의도를 추론하기 위한 상태 기반 응답 추론부(24)와, 일반 영역에 대한 현재 사용자 의도를 입력받아, 현재 시스템 상태 정보와 예제 대화 DB(27)를 기반으로 일반 영역에 대한 현재 시스템 의도 및 시스템 응답 형틀을 추출하기 위한 예제 기반 응답 추론부(26)를 포함한다. 이때, 예제 기반 응답 추론부(26)에서는 일반 영역에 대한 시스템 응답 형틀이 예제 대화 DB(27)에 없는 경우, 예외처리용 응답 형틀을 참조하여 시스템 응답 형틀을 생성하게 된다.

<58> 또한, 응답 형틀 생성부(28)는, 예제 대화 DB(27)를 참조하여, 목적 영역에 대한 현재 사용자 의도와 현재 시스템 의도로부터 목적 영역에 대한 시스템 응답 형틀을 생성한다. 이때, 응답 형틀 생성부(28)에서는 목적 영역에 대한 시스템 응답 형틀이 예제 대화 DB(27)에 없는 경우, 예외처리용 응답 형틀을 참조하여 시스템 응답 형틀을 생성하게 된다.

<59> 따라서, 자연어 생성부(29)에서는 목적 영역 및 일반 영역에 대한 시스템 응답 형틀로부터 자연어로 된 시스템 응답을 생성할 수 있다.

<60> 이상의 내용을 정리해보면, 본 발명에 따른 범용 대화서비스 장치(20)는, 입력 질의어(문자 데이터 또는 음성발화 데이터 중 어느 하나임)를 영역(목적 영역, 일반 영역)별로 분류하기 위한 영역 분류부(21)와, 분류된 영역(목적 영역 또는 일반 영역)의 특화된 의미(목적 영역 의미 해석, 일반 영역 의미 해석)를 해석하여 현재 사용자 의도를 분석하기 위한 의미 해석부(22)와, 목적 영역에 대한 현재 사용자 의도를 입력받아, 현재 시스템이 가지고 있는 상태 정보(직전의 발화에서 시스템이 취한 의도(이전 시스템 의도), 그동안 대화에서 누적된 사용자의 의도(이전 사용자 의도)로 구성되는 현재 시스템 상태 정보)와 상태 기반 추론 정책(상태 기반 추론 정책 DB(25))을 기반으로 목적 영역에 대한 현재 시스템 의도를 추론하기 위한 상태 기반 응답 추론부(24)와, 예제 대화 DB(27)를 참조하여, 목적 영역에 대한 현재 사용자 의도와 현재 시스템 의도로부터 목적 영역에 대한 시스템 응답 형틀(일부 어휘만을 교체하여 실제 자연어 문장을 만들어 낼 수 있는 생성 틀임)을 생성하기 위한 응답 형틀 생성부(28)와, 일반 영역에 대한 현재 사용자 의도를 입력받아, 현재 시스템 상태 정보와 예제 대화 DB(27)를 기반으로 일반 영역에 대한 현재 시스템 의도 및 시스템 응답 형틀(일부 어휘만을 교체하여 실제 자연어 문장을 만들어 낼 수 있는 생성 틀임)을 추출하기 위한 예제 기반 응답 추론부(26)와, 이전 사용자/시스템 의도와 그 이전까지 누적된 시스템 상태 정보로부터 현재 시스템 상태 정보를 생성하기 위한 시스템 상태 제어부(23)와, 목적 영역 및/또는 일반 영역에 대한 시스템 응답 형틀로부터 자연어로 된 시스템 응답을 생성하기 위한 자연어 생성부(29)를 포함한다.

<61> 본 발명에 따른 범용 대화서비스 장치(20)는 대화의 영역(목적 영역, 일반 영역)에 따라 시스템 상태 정보 및 예제 기반 응답 추론 모델을 같이 이용한다. 이때, 영역 분류부(21)를 통해 목적 영역과 일반 영역으로 사용자의 질의(문자 또는 음성발화 데이터 등)를 분류하여 처리한다. 그리고, 목적 영역과 일반 영역으로 분류된 사용자 질의의 영역에 따라 응답 추론부(상태 기반 응답 추론부(25), 예제 기반 응답 추론부(26))를 적용한다.

<62> 이때, 시스템 상태 제어부(23)에서는 시스템 상태 정보를 영역이 혼용된 대화의 기간 내에 통합 사용/유지할 수 있다. 이의 구체적인 구현 방법으로, 현재의 시스템 상태 정보는 시스템 상태 제어부(23)에 의해 관리되며, 직전의 발화에서 시스템이 취한 의도들(이전 시스템 의도), 그동안 대화에서 누적된 사용자의 의도들(이전 사용자 의도)로 상태 정보가 이루어지며, 상태 기반 응답 추론부(24)와 예제 기반 응답 추론부(26)에서는 응답 추론의 입력으로 현재 시스템이 가지고 있는 상태 정보(현재 시스템 상태 정보)를 사용한다. 이때, 목적 영역으로 분류된 사용자 질의의 영역에 대해서는 상태 기반 응답 추론부(23)를 적용하고, 일반 영역으로 분류된 사용자 질의의 영역에 대해서는 예제 기반 응답 추론부(26)를 적용한다.

<63> 여기서, 상태 기반 응답 추론부(23)에서는 분석된 사용자 발화의 의미(현재 사용자 의도)와 현재 시스템이 가지고 있는 상태 정보(현재 시스템 상태 정보), 상태 기반 추론 정책을 입력으로 현재의 시스템 의도, 즉 시스템이 어떠한 발화나 행동을 할 것인가를 결정한다. 이때, 현재 시스템 상태 정보는 시스템 상태 제어부(23)에 의해 관리되며(다음 시스템 상태 정보로서 활용될 수 있음), 생성된 현재 시스템 의도는 응답 형틀 생성부(28)에 의해 예제 대화 DB(27)와 예외처리용 응답 형틀(도면에 도시되지 않음)을 참조하여 자연어 생성에 사용될 응답 형틀로 변형되어, 자연어 생성부(29)에서 일부 어휘의 교체와 어미/조사의 재배열 등을 통하여 사용자에게 전달된

시스템 발화(시스템 응답)으로 변환된다.

<64> 한편, 예제 기반 응답 추론부(26)에서는 분석된 사용자 발화의 의미(현재 사용자 의도)와 현재 시스템이 가지고 있는 상태 정보(현재 시스템 상태 정보)를 참조하여 예제 대화 DB(27)에서 탐색하고, 해당 예문의 시스템 답변(시스템 응답)을 시스템 응답 형틀로 삼고, 이는 다시 자연어 생성부(29)에서 일부 어휘의 교체와 어미/조사의 재배열 등을 통하여 사용자에게 전달될 시스템 발화(시스템 응답)으로 변환된다.

<65> 상기와 같은 구성을 갖는 본 발명에 따른 범용 대화서비스 장치(20)의 동작을 도 3을 참조하여 살펴보면 다음과 같다.

<66> 사용자 질의는 열차예약과 같이 서비스가 제공하고자 하는 목적 영역 질의와, 채팅과 같은 그 밖의 일반 영역 질의로 분류할 수 있다.

<67> 따라서, 대화서비스 장치(20)에 이러한 영역이 혼용되어 발화되는 사용자 질의(문자 또는 음성발화 데이터 등)가 들어오면(301), 가장 먼저 영역 분류부(21)에서 목적 영역 질의와 일반 영역 질의로 질의의 영역을 판별한다(302). 이때, 영역을 판별하는 방법론으로는, 대화의 영역이 부착되어 수집된 대량의 사용자 질의 집합에서 훈련된 분류(Classification) 기법을 사용한다.

<68> 일 예로, 영역 분류부(21)에서는 음성언어 주제 판별 방식을 사용할 수 있다.

<69> 참고적으로, 음성언어 주제 판별 방식을 살펴보면, 사용자 음성을 인식하여 문장을 출력하고, 인식된 문장에서 속성을 추출한다. 이때, 속성 추출은 인식된 발화 문장에서 어휘 정보(주요 단어와 그 품사)를 추출하고, 각 주요 단어의 상/하위 관계 등을 고려하여 개념 정보('상/하위 단어')를 추출하거나, 주요 단어들 간에 어휘/구문 관계를 고려하여 문맥 정보를 추출하거나, 발화 문장이 적어도 2개의 문장으로 이루어진 경우 발화 문장 간에 수사 관계를 고려하여 '문장간 수사 관계'를 추출한다. 다만, 단어의 어휘 정보와 개념 정보, 단어들 간의 문맥 정보, 문장들 간의 수사 관계는 각각 독립적인 분류 속성으로 이의 일부 혹은 전체를 분류 속성으로 사용 가능하다. 이후, 사용자 질의에서 추출된 속성과 미리 훈련된 주제별 속성 집합을 비교하여 인식된 문장의 속성에 가장 가까운 속성 집합의 주제 코드를 찾고, 그 유사도를 계산한다. 만약, 계산된 유사도가 미리 설정된 임계치보다 높거나 같은 경우 찾아진 주제 코드를 입력 음성의 주제 코드로 최종 확정하여 출력하고, 임계치보다 낮은 경우 주제 판별이 실패한 것으로 판단하여 주제 판별 실패 코드를 출력한다.

<70> 예를 들면, 사용자로부터의 음성으로부터 얻을 수 있는 문장이 "객실 예약을 하고 싶은데요."일 때, 이 문장으로부터 얻을 수 있는 주요 단어들은 '객실', '예약', '하다'이다. 이들 단어로부터 추출 가능한 상위어는 '객실'의 상위어인 '호텔', '예약'의 상위어인 '고객대응'이다. 이때, 추출된 구문 관계는 '예약'과 '하다'의 목적격 관계인 '예약-을-하다'이다. 이러한 과정을 거쳐, 이 문장의 속성은 {'객실', "예약', '하다', '호텔', '고객대응', '예약-을-하다'}로 표현된다.

<71> 또 다른 예로, 사용자의 음성이 "어제 예약을 했는데요. 그런데, 오늘 일이 생겨서요."일 때, '(예약-을-하다) - 반전 - (일-이-생기다)'와 같은 문장간 수사 관계를 추출할 수 있다. 이 음성의 속성은 이 문장간 수사 관계를 포함하여 {'어제', '예약', '하다', '오늘', '일', '생기다', '날짜', '고객대응', '예약-을-하다', '일-이-생기다', '(예약-을-하다) - 반전 - (일-이-생기다)'}로 표현된다.

<72> 이와 같은 음성언어 주제 판별 방식을 통해 영역 분류부(21)는 사용자의 발화 음성을 인식하여, 인식된 발화 문장에서 주요 어휘를 추출한 후 각 주요 단어의 상/하위 관계 등을 고려하여 개념 정보를 추출하거나, 인식된 발화 문장에서 주요 단어를 추출하고 주요 단어들 간에 어휘/구문 관계를 고려하여 문맥 정보를 추출하거나, 발화 문장이 적어도 2개의 문장으로 이루어진 경우 인식된 발화 문장들에서 주요 단어를 추출하고 발화 문장 간에 수사 관계를 고려하여 '문장 간 수사 관계'를 추출함으로써, 추출된 속성들로 표현된 문장을 주제별로 분류할 수 있다.

<73> 이후, 영역 분류부(21)에서 영역(목적 영역 또는 일반 영역)이 분류된 사용자 질의는 영역에 특화된 의미 해석부(22)를 통하여 해당 영역에서 필요한 의미를 분석한다(303,304). 즉, 목적 영역으로 분류된 사용자 질의의 영역에 대해서는 목적 영역 의미 해석을 통하여 사용자의 의도(현재 사용자 의도)를 파악하고(303), 일반 영역으로 분류된 사용자 질의의 영역에 대해서는 일반 영역 의미 해석을 통하여 사용자의 의도(현재 사용자 의도)를 파악한다(304).

<74> 예를 들어, "대전에 가는 열차를 예약해 주세요."라는 사용자 발화에 대해서, [열차 예약]이라는 목적 영역으로 분류되었으면(302), 이 목적 영역을 해석하는 의미 해석부(21)에 의해 {사용자의도="열차예약", 도착역="대전",

문형="요청"} 등과 같이 분석된다(303).

<75> 이후, 분석된 사용자 발화의 의미(현재 사용자 의도)를 입력으로 상태 기반 응답 추론부(24)는 시스템의 의도, 즉 "확인(출발역=서울)"과 같이 시스템이 어떠한 발화나 행동을 할 것인가를 결정한다(306).

<76> 이때, 상태 기반 응답 추론부(24)는 현재 시스템이 가지고 있는 상태 정보(현재 시스템 상태 정보)(309)와 상태 기반 추론 정책(상태 기반 추론 정책 DB(25))을 기반으로 시스템 의도를 추론한다(308). 여기서, 상태 기반 추론 정책이란, 시스템의 각 상태에서 취할 수 있는 시스템의 의도를 수작업 혹은 통계적으로 나열한 정책 규칙이다. 그리고, 현재의 시스템 상태 정보(309)는 시스템 상태 제어부(23)에 의해 관리되며, 직전의 발화에서 시스템이 취한 의도, 그동안 대화에서 누적된 사용자의 의도들로 정보 상태가 이루어진다.

<77> 상기에서, 현재 시스템 상태 정보(309)는 다음과 같이 생성된다.

<78> 최초 대화 1단계에서 사용자 의도A[예로서, "열차를 예약해 주세요"라는 사용자 발화에 대해 {사용자의도="열차예약", 문형="요청"}](305)에 대해 시스템 의도B[예로서, "확인(출발역=서울)"](308)가 생성되었다고 가정하고, 대화 2단계에서 사용자 의도C[예로서, "서울에서 출발합니다"라는 사용자 발화에 대해 {사용자의도="출발역(서울)", 문형="응답"}](305)에 대해 시스템 의도D[예로서, "확인(도착역=대전)"](308)가 생성되었다고 가정하며, 대화 3단계에서 사용자 의도E[예로서, "대전에 도착합니다"라는 사용자 발화에 대해 {사용자의도="도착역(대전)", 문형="응답"}](305)에 대해 시스템 의도F[예로서, "예약완료되었습니다"](308)가 생성되었다고 가정하자.

<79> 이 경우, 사용자 의도A(305)에 대해서는 초기 시스템 상태 정보(309)를 바탕으로 이전 시스템 의도B(308)가 생성된다. 그러나, 사용자 의도C(305)에 대해서는 이전까지 누적된 제1 시스템 상태 정보[(사용자 의도A, 시스템 의도B)](309)를 바탕으로 시스템 의도D(308)가 생성될 것이다. 그리고, 사용자 의도E(305)에 대해서는 이전까지 누적된 제2 시스템 상태 정보[(사용자 의도A, 시스템 의도B), (사용자 의도C, 시스템 의도D)](309)를 바탕으로 시스템 의도F(308)가 생성될 것이다.

<80> 이처럼, '시스템 의도B'(308)는 초기 시스템 상태 정보(309)를 바탕으로 한 사용자 의도A(305)에 대한 시스템 의도이고, '시스템 의도D'(308)는 제1 이전 시스템 상태 정보(초기 시스템 상태 정보를 바탕으로 한 사용자 의도A에 대한 시스템 의도B의 상태값(코딩값))(309)를 바탕으로 한 사용자 의도C(305)에 대한 시스템 의도이며, '시스템 의도F'(308)는 제2 이전 시스템 상태 정보(제1 이전 시스템 상태 정보를 바탕으로 한 사용자 의도C에 대한 시스템 의도D의 상태값)(309)를 바탕으로 한 사용자 의도E(305)에 대한 시스템 의도가 될 것이다.

<81> 따라서, 시스템 상태 제어부(23)는 이전 사용자 의도(예로서, 사용자 의도C) 및 이전 시스템 의도(예로서, 사용자 의도C에 대한 시스템 의도D)와 그 이전까지 누적된 시스템 상태 정보(예로서, 사용자 의도A에 대한 시스템 의도B)로부터 현재 시스템 상태 정보[(사용자 의도A, 시스템 의도B), (사용자 의도C, 시스템 의도D)]를 생성한다(309). 그러면, 상태 기반 응답 추론부(24)에서는 현재 시스템 상태 정보[(사용자 의도A, 시스템 의도B), (사용자 의도C, 시스템 의도D)](309)와 목적 영역에 대한 현재 사용자 의도(예로서, 사용자 의도E)(305), 상태 기반 추론 정책(상태 기반 추론 정책 DB(25))을 바탕으로 목적 영역에 대한 현재 시스템 의도(예로서, 시스템 의도F)(308)를 추론할 수 있게 된다. 이때, 현재 시스템 의도F(308)는 대화 단계가 지속되는 경우(예로서, 대화 4단계) 다음 시스템 상태 정보[(사용자 의도A, 시스템 의도B), (사용자 의도C, 시스템 의도D), (사용자 의도E, 시스템 의도F)]에 반영된다.

<82> 이와 같이 생성된 시스템의 의도(308)는 응답 형틀 생성부(28)에서 자연어 생성에 사용될 시스템 응답 형틀로 변형된다(310). 이때, 예제 대화 DB(27)를 참조하여 시스템 응답 형틀을 생성하는데, 만약 목적 영역에 대한 시스템 응답 형틀이 예제 대화 DB(27)에 없는 경우, 예외처리용 응답 형틀을 참조한다.

<83> 여기서, 응답 형틀이란 일부 어휘만을 교체하여 실제 자연어 문장을 만들어 낼 수 있는 생성 틀을 의미하며, 시스템 의도에 대해 "$(출발역=서울)에서 출발하십니까?"와 같은 시스템 응답 형틀을 생성한다(310).

<84> 이후, 자연어 생성부(29)에서는 시스템 응답 형틀(311)로부터 일부 어휘의 교체와 어미/조사의 재배열 등을 통하여 실제 사용자에게 전달될 시스템 발화(시스템 응답)를 생성한다(312,313).

<85> 한편, "서울에서 출발하십니까?"와 같은 시스템 발화(시스템 응답)에 대하여 사용자는 "예 맞습니다." 혹은 "서울에서부터 2장 예매해 주세요." 등과 같이 시스템이 원하는 [열차 예약] 목적 영역의 발화를 할 수도 있지만, "오늘 어때?"와 같이 다소 서비스 목적과 동떨어진 발화(일반 영역 대화)를 할 수 있다(301). 이러한 경우, 영역 분류부(21)에 의해 일반 영역의 발화로 분류되어(302) 일반 영역을 해석하는 의미 해석부(22)에 의해 그 의미가 분석된다(304). 이 경우, 해석된 사용자 의도는 {사용자의도="인사", 문형="HOW"}와 같다.

<86> 이후, 예제 기반 응답 추론부(26)는 사용자 발화(사용자 의도)와 가장 가까운 의미를 가진 예문을 해석된 사용자 의도(현재 사용자 의도)와 현재의 시스템 상태 정보(현재 시스템 상태 정보)(309)를 참고하여 예제 대화 DB(27)에서 탐색하여 해당 예문의 시스템 답변(시스템 응답)을 시스템 응답 형틀로 삼는다(311).

<87> 여기서, 예제 대화 DB(27)를 검색할 때, 사용자 발화의 의미(사용자 의도)뿐만이 아니라, "사용자의 요청에 대한 확인을 요청"하고 있는 시스템의 현재 상태 정보(현재 시스템 상태 정보)(309)를 참고하여 이러한 상황과 가장 가까운 발화의 답변 "모르니까 물어보지"를 검색한다. 이때, 예제 대화 DB(27)에는 시스템의 응답 형틀의 형태로 시스템 발화(시스템 응답)가 저장되어 있으므로 예제 기반 응답 추론부(26)의 추론 결과는 시스템 응답 형틀이다(311). 이는(311) 다시 자연어 생성부(29)에서 일부 어휘의 교체와 어미/조사의 재배열 등을 통하여 실제 사용자에게 전달될 시스템의 발화(시스템 응답)로 생성된다(312,313). 여기서, 예제 기반 응답 추론시(307), 만약 일반 영역에 대한 시스템 응답 형틀이 예제 대화 DB(27)에 없는 경우, 예외처리용 응답 형틀을 참조한다.

<88> 이때, 현재 시스템 상태 정보(309)는 상기 목적 영역에 대한 응답 추론 방식과 같이 일반 영역의 사용자 의도A(305)에 대해서는 초기 시스템 상태 정보(309)를 바탕으로 이전 시스템 의도B(308)가 생성된다. 그리고, 일반 영역의 사용자 의도C(305)에 대해서는 이전까지 누적된 제1 시스템 상태 정보[(사용자 의도A, 시스템 의도B)](309)를 바탕으로 시스템 의도D(308)가 생성될 것이다. 또한, 일반 영역의 사용자 의도E(305)에 대해서는 이전까지 누적된 제2 시스템 상태 정보[(사용자 의도A, 시스템 의도B), (사용자 의도C, 시스템 의도D)](309)를 바탕으로 시스템 의도F(308)가 생성될 것이다.

<89> 따라서, 시스템 상태 제어부(23)는 이전 사용자 의도(예로서, 사용자 의도C) 및 이전 시스템 의도(예로서, 사용자 의도C에 대한 시스템 의도D)와 그 이전까지 누적된 시스템 상태 정보(예로서, 사용자 의도A에 대한 시스템 의도B)로부터 현재 시스템 상태 정보[(사용자 의도A, 시스템 의도B), (사용자 의도C, 시스템 의도D)]를 생성한다(309). 그러면, 예제 기반 응답 추론부(26)에서는 현재 시스템 상태 정보[(사용자 의도A, 시스템 의도B), (사용자 의도C, 시스템 의도D)](309)와 일반 영역에 대한 현재 사용자 의도(예로서, 사용자 의도E)(305), 예제 대화 DB(27)를 바탕으로 일반 영역에 대한 현재 시스템 의도(예로서, 시스템 의도F)(308)를 추론할 수 있게 된다. 이때, 현재 시스템 의도F(308)는 대화 단계가 지속되는 경우(예로서, 대화 4단계) 다음 시스템 상태 정보[(사용자 의도A, 시스템 의도B), (사용자 의도C, 시스템 의도D), (사용자 의도E, 시스템 의도F)]에 반영된다.

<90> 상술한 바와 같은 본 발명의 방법은 프로그램으로 구현되어 컴퓨터로 읽을 수 있는 형태로 기록매체(씨디롬, 램, 롬, 플로피 디스크, 하드 디스크, 광자기 디스크 등)에 저장될 수 있다. 이러한 과정은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있으므로 더 이상 상세히 설명하지 않기로 한다.

<91> 이상에서 설명한 본 발명은, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 있어 본 발명의 기술적 사상을 벗어나지 않는 범위 내에서 여러 가지 치환, 변형 및 변경이 가능하므로 전술한 실시예 및 첨부된 도면에 의해 한정되는 것이 아니다.

발명의 효과

<92> 상기와 같은 본 발명은, 목적 영역의 대화와 일반 영역의 대화가 혼재되어 있는 영역에서 대화 전반에 걸쳐 습득된 정보를 이용하여 대화를 추론할 수 있고, 모든 형태의 사용자 질의에 대해 지능적으로 시스템 답변을 만들어 높은 대화 품질을 가지는 범용의 대화서비스를 제공할 수 있는 효과가 있다.

도면의 간단한 설명

<1> 도 1 은 본 발명에 따른 대화서비스 장치가 적용되는 음성인식 시스템의 일실시예 구성도,

<2> 도 2 는 본 발명에 따른 범용 대화서비스 장치의 일실시예 구성도,

<3> 도 3 은 본 발명에 따른 범용 대화서비스 방법을 보여주는 일실시예 설명도이다.

<4> * 도면의 주요 부분에 대한 부호 설명

<5> 11 : 사용자 단말기 12 : 음성 인식부

<6> 13 : 음성 합성부 14 : 디스플레이부

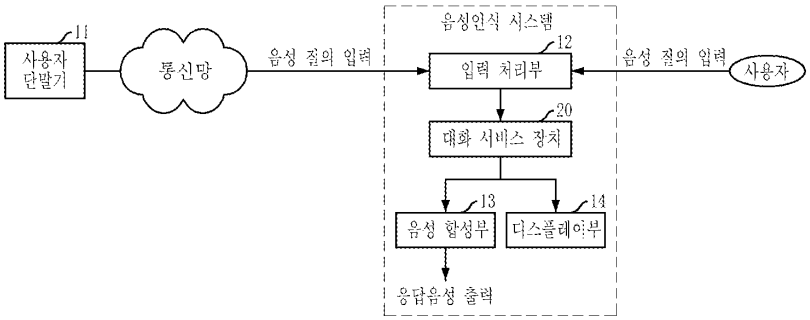
<7> 20 : 대화서비스 장치 21 : 영역 분류부

<8> 22 : 의미 해석부 23 : 시스템 상태 제어부

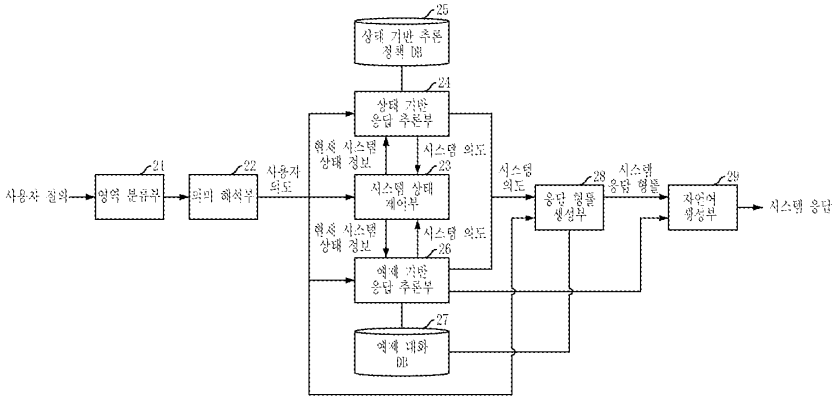
<9>	24 : 상태 기반 응답 추론부	25 : 상태 기반 추론 정책 DB
<10>	26 : 예제 기반 응답 추론부	27 : 예제 대화 DB
<11>	28 : 응답 형틀 생성부	29 : 자연어 생성부

도면

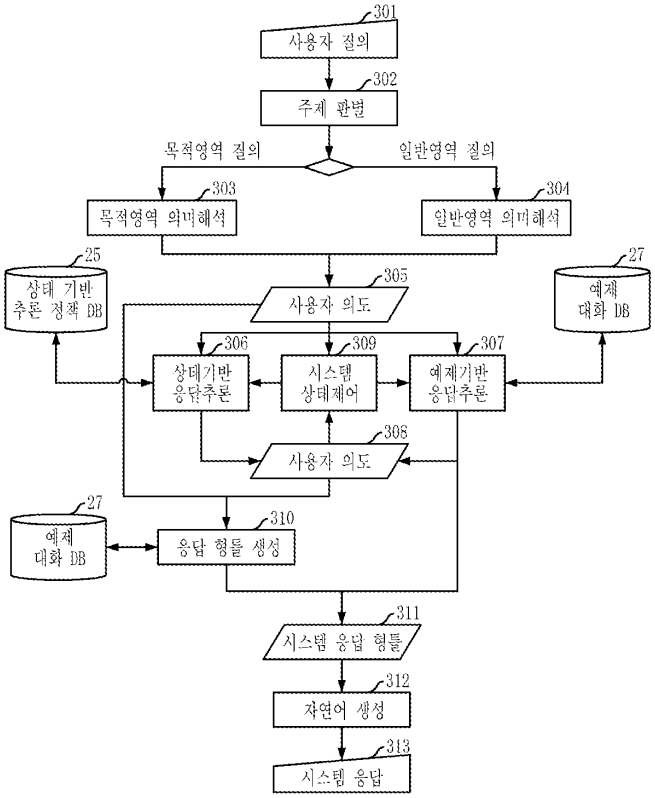
도면1



도면2



도면3


301
사용자 질의
302
주제 판별
목적영역 질의
일반영역 질의
303
목적영역 의미해석
304
일반영역 의미해석
25
상태 기반
추론 정책 DB
305
사용자 의도
27
예제
대화 DB
306
상태기반
응답추론
309
시스템
상태제어
307
예제기반
응답추론
308
사용자 의도
27
예제
대화 DB
310
응답 형틀 생성
311
시스템 응답 형틀
312
자연어 생성
313
시스템 응답

INTERACTIVE SERVER, CONTROL METHOD THEREOF, AND INTERACTIVE SYSTEM

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application claims priority from Korean Patent Application No. 10-2013-0001829, filed on January 7, 2013, in the Korean Intellectual Property Office, the disclosure of which is incorporated herein by reference in its entirety.

BACKGROUND

1. Field

[0002] Apparatuses, methods, and systems consistent with exemplary embodiments generally relate to providing an interactive server, a control method thereof, and an interactive system, and more particularly, to providing an interactive server which provides response information corresponding to various utterances of a user, a control method thereof, and an interactive system.

2. Description of the Related Art

[0003] In general, a display apparatus which is capable of recognizing voices in an interactive system collects uttered voices of a user and transmits the collected uttered voices to an external server connected through a network. When the external server receives the uttered voices, the external server parses the uttered voices to comprehend intensions of the user, generates response information corresponding to the uttered voices, and transmits the response information to the display apparatus. Therefore, the display apparatus executes a function corresponding to the uttered voices of the user or provides information corresponding to the uttered voices of the user based on the response information received from the external server.

[0004] However, the conventional interactive system has a limit to parsing of the uttered voices of the user and comprehending of the intention of the user based on the parsed result. For example, in the case of a clear uttered voice of the user like "Show first content.", the external server parses the uttered voice to correctly comprehend an intention of the user, generates response information corresponding to the uttered voice, and transmits the response information to the display apparatus. Therefore, the display apparatus displays the first content requested by the user based on the response information.

[0005] However, in the case of an ambiguous uttered voice of the user like "Show this.", the external server does not clearly comprehend an intention of the user from the ambiguous uttered voice. In other words, the interactive system comprehends an intention of the user only with respect to a predefined uttered voice and performs an operation corresponding to the predefined uttered voice or provides information corresponding to the predefined uttered voice, thereby restricting utterances of the user.

SUMMARY

[0006] Exemplary embodiments address at least the above problems and/or disadvantages and other disadvantages not described above. Also, the exemplary embodiments are not required to overcome the disadvantages described above, and an exemplary embodiment may not overcome any of the problems described above.

[0007] The exemplary embodiments provide an interactive server which performs operations corresponding to various utterances of a user in an interactive system, a control method thereof, and the interactive system.

[0008] According to an aspect of the exemplary embodiments, there is provided an interactive server including: a communicator configured to communicate with a apparatus to receive a first uttered voice signal; a storage device configured to store utterance history

information of a second uttered voice signal received from the apparatus before the first uttered voice signal is received; an extractor configured to extract uttered elements from the received first uttered voice signal; and a controller configured to generate response information based on at least one from among the utterance history information stored in the storage device and the extracted uttered elements, and transmit the response information to the apparatus.

[0009] If an execution language and a command respectively corresponding to the extracted uttered elements are stored in the storage device, the controller may combine the execution language and the command respectively corresponding to the uttered elements to generate the response information. If an execution language and a command respectively corresponding to the extracted uttered elements are not stored in the storage device, the controller may estimate an execution language and a command respectively corresponding to the extracted uttered elements based on the utterance history information stored in the storage device and combine the estimated execution language and command to generate the response information.

[0010] The first uttered voice signal may comprise a first uttered element and a second uttered element. If the execution language and the command respectively corresponding to the extracted uttered elements are not stored in the storage device, the controller may acquire utterance history information stored within a preset time range based on a time when the second uttered voice signal is received and estimate an execution language corresponding to the first uttered element and a command corresponding to the second uttered element as the execution language and the command corresponding to the first uttered voice signal within the acquired utterance history information.

[0011] If the utterance history information stored in the storage device does not include the utterance history information stored within the preset time based on the time

when the second uttered voice signal is received, the controller may not generate the response information.

[0012] If the first uttered voice signal does not include at least one of the first uttered element and the second uttered element, the controller may generate response information corresponding to the second uttered voice signal based on the utterance history information stored within the preset time range based on the time when the second uttered voice signal is received.

[0013] The interactive server may further include a voice processor configured to convert the received first uttered voice signal into text information.

[0014] According to another aspect of the exemplary embodiments, there is provided a method of controlling an interactive server. The method may comprise: receiving a first uttered voice signal from an apparatus; extracting uttered elements from the received first uttered voice signal; generating response information based on stored utterance history information and the extracted uttered elements; and transmitting the generated response information to the apparatus. The stored utterance history information may be information which is generated based on a second uttered voice signal received from the apparatus before the first uttered voice signal is received.

[0015] The method may further include: determining whether an execution language and a command respectively corresponding to the extracted uttered elements exist. If the execution language and the command respectively corresponding to the uttered elements exist, the execution language and the command respectively corresponding to the uttered elements may be combined to generate response information, and if the execution language and the command respectively corresponding to the uttered elements do not exist, an execution language and a command corresponding to the extracted uttered elements may be

estimated based on the stored utterance history information and may be combined to generate the response information.

[0016] The first uttered voice signal may comprise first and second uttered elements. If the execution language and the command respectively corresponding to the uttered elements do not exist, utterance history information stored within a preset time range may be acquired based on a time when the second uttered voice signal is received, and the execution language corresponding to a first uttered element and the command corresponding to the second uttered element may be estimated as the execution language and a command corresponding to the uttered voice signal within the acquired utterance history information.

[0017] If the utterance history information does not comprise the utterance history information stored within the preset time range based on the time when the second uttered voice signal is received, the response information may not be generated.

[0018] If the first uttered voice signal may not comprise at least one of the first and second uttered elements, response information corresponding to the first uttered voice signal may be generated based on the utterance history information stored within the preset time range based on the time when the second uttered voice signal is received.

[0019] The method may further comprise: converting the received first uttered voice signal into text information.

[0020] According to another aspect of the exemplary embodiments, there is provided an interactive system comprising: an apparatus which, if a first uttered voice of a user is input, performs a function corresponding to the input uttered voice; and an interactive server which, if a first voice signal is received from the apparatus, extracts uttered elements from the first uttered voice signal, generates response information based on at least one from among stored utterance history information and the extracted uttered elements, and transmits the response

information to the apparatus. The apparatus may perform a function corresponding to an intention of the user based on the response information received from the interactive server.

[0021] The interactive server may generate the response information based on an execution language and a command respectively corresponding to the uttered elements if the execution language and the command respectively corresponding to the extracted uttered elements exist, and, if the execution language and the command respectively corresponding to the extracted uttered elements do not exist, estimate an execution language and a command respectively corresponding to the extracted uttered elements based on the stored utterance history information and combine the estimated execution language and command to generate the response information.

[0022] According to an aspect of an exemplary embodiment, the time referred to above may be related to a time difference between utterances of the first uttered voice signal and the second uttered voice signal. According to an aspect of another exemplary embodiment, the apparatus referred to above may be a display apparatus.

[0023] According to yet another aspect of an exemplary embodiment, the first uttered voice signal may be uttered by a first user and the utterance history information may relate to an uttered voice input of a second user.

[0024] According to various exemplary embodiments of the present general inventive concept, in an interactive system, an interactive server may comprehend intentions of a user with respect to various uttered voices of the user, generate response information of the various uttered voices, and transmit the response information to a display apparatus.

BRIEF DESCRIPTION OF THE DRAWINGS

[0025] The above and/or other aspects will be more apparent by describing certain exemplary embodiments with reference to the accompanying drawings, in which:

[0026] FIG. 1 is a view illustrating an interactive system which provides response information appropriate for a uttered voice of a user according to an exemplary embodiment of the present general inventive concept;

[0027] FIG. 2 is a view illustrating an interactive system which provides response information appropriate for a uttered voice of a user according to another exemplary embodiment of the present general inventive concept;

[0028] FIG. 3 is a block diagram illustrating an interactive server according to an exemplary embodiment of the present general inventive concept;

[0029] FIG. 4 is a view illustrating a method of generating uttered information based on a display state of an object displayed on a screen of a display apparatus according to an exemplary embodiment of the present general inventive concept;

[0030] FIG. 5 is a block diagram illustrating a display apparatus according to an exemplary embodiment of the present general inventive concept;

[0031] FIG. 6 is a flowchart illustrating a method of controlling an interactive server according to an exemplary embodiment of the present general inventive concept; and

[0032] FIG. 7 is a flowchart illustrating a method of generating response information through an interactive server according to an exemplary embodiment of the present general inventive concept.

DETAILED DESCRIPTION OF THE EXEMPLARY EMBODIMENTS

[0033] Exemplary embodiments are described in greater detail with reference to the accompanying drawings.

[0034] In the following description, the same drawing reference numerals are used for the same elements even in different drawings. The matters defined in the description, such as detailed construction and elements, are provided to assist in a comprehensive understanding

of the exemplary embodiments. Thus, it is apparent that the exemplary embodiments can be carried out without those specifically defined matters. Also, well-known functions or constructions are not described in detail since they would obscure the exemplary embodiments with unnecessary detail.

[0035] FIG. 1 is a view illustrating an interactive system which provides response information appropriate for an uttered voice of a user according to an exemplary embodiment of the present general inventive concept.

[0036] Referring to FIG. 1, the interactive system according to the present exemplary embodiment includes a display apparatus 100 and an interactive server 200.

[0037] The display apparatus 100 may be realized as various types of electronic apparatuses which surf the Internet, such as a smart TV, a portable phone such as a smart phone, a desktop PC, a notebook PC, a navigation system, etc. If an uttered voice of a user is input, the display apparatus 100 performs an operation corresponding to the uttered voice of the user. In detail, the display apparatus 100 outputs a response message corresponding to the uttered voice of the user or performs a function corresponding to the uttered voice of the user. In other words, the display apparatus 100 parses the input uttered voice to determine whether the input uttered voice is internally performable and performs the function corresponding to the input uttered voice according to the determination result or performs an operation based on response information received from the interactive server 200.

[0038] For example, if an uttered voice "Turn up volume." is input, the display apparatus 100 adjusts a volume based on one of pre-stored control information pieces corresponding to the input uttered voice.

[0039] As another example, if an uttered voice "How's the weather today?" is input, the display apparatus 100 transmits an uttered voice signal of the input uttered voice to the interactive server 200. Here, the uttered voice may be an analog signal. Therefore, the

display apparatus 100 converts the uttered voice that is the analog signal into a digital signal (hereinafter referred to as an uttered voice) and transmits the digital signal to the interactive server 200. The display apparatus 100 outputs a result of the today's weather as a voice or a text format image based on the response information received from the interactive server 200.

[0040] As another example, if an uttered voice "Show ○○○ (content name)." is input, the display apparatus 100 transmits the input uttered voice to the interactive server 200. If response information is received from the interactive server 200, the display apparatus 100 displays a search list of "○○○(content name)" based on the received response information. Here, the search list may include content information of each of episodes. Therefore, the user checks the content information of the episodes displayed on the search list to make an utterance for selecting ○○○ (content name) of an episode that the user wants to view. For example, the user makes an utterance "Show first one." to view "○○○ (content name)" corresponding to the first one of the content information displayed on the search list. Therefore, the display apparatus 100 receives the uttered voice "Show first one." and transmits the uttered voice "Show first one." to the interactive server 200. If response information corresponding to the uttered voice is received from the interactive server 200, the display apparatus 100 receives "○○○ (content name)" corresponding to the first content information from a content providing server 300 and displays "○○○ (content name)" based on the received response information.

[0041] The interactive server 200 provides response information appropriate for an intention of the user based on an uttered voice of the user received from the display apparatus 100. In detail, if an uttered voice of the user is received from the display apparatus 100, the interactive server 200 extracts uttered elements from the received uttered voice, and generates and transmits response information related to the uttered voice of the user based on the extracted uttered elements. As described above, the uttered voice of the user received from

the display apparatus 100 may be a digital signal. Therefore, if the uttered voice converted into the digital signal is received, the interactive server 200 generates the uttered voice as text information, parses the text information to extract uttered elements, and generates response information corresponding to the uttered voice of the user based on the extracted uttered elements. The generation of the text information from the uttered voice converted into the digital signal is a well-known technology, and thus a detailed description thereof will be omitted herein.

[0042] The uttered elements are core keywords of the uttered voice of the user for performing an operation requested by the user and are classified into uttered elements extracted according to a domain, a user action, and a feature. As described above, if text information is generated with respect to the uttered voice of the user "How's the weather today?", the interactive server 200 extracts uttered elements "today," "weather," and "How." The interactive server 200 classifies the uttered elements "today" and "weather" as uttered elements (hereinafter referred to as first uttered elements) of a feature (hereinafter referred to as an object to be executed) and classifies the uttered element "How?" as an uttered element (hereinafter referred to as a second uttered element) of a user action (hereinafter referred to as an execution command). The interactive server 200 classifies an uttered element (hereinafter referred to as a third uttered element) of a domain as belonging to a web search domain based on the extracted uttered elements. When the first through third uttered elements are classified from the text information of the uttered voice of the user, the interactive server 200 is provided with weather information from an external server 400, which provides various types of contents, and generates response information including the weather information, and transmits the response information to the display apparatus 100. Therefore, the display apparatus 100 displays information about a today's weather through at least one of a voice and a text based on the response information received from the interactive server 200.

[0043] If text information of an uttered voice "Select first one." is generated, the interactive server 200 extracts a first uttered element "first one" and a second uttered element "Select" from the generated text information. If the first and second uttered elements are extracted, the interactive server 200 determines whether to generate response information appropriate for an intention of the user based on the extracted first and second uttered elements. According to an exemplary embodiment, the interactive server 200 may determine whether to generate response information appropriate for an intention of the user according to whether information corresponding to the extracted first and second uttered elements exists in a database (DB). Here, the information corresponding to the first and second uttered elements may be respectively an execution language corresponding to the first uttered element and a command corresponding to the second uttered element. Therefore, the interactive server 200 determines whether the execution language and the command respectively corresponding to the first and second uttered elements exist in the DB. If it is determined that the information corresponding to the first and second uttered elements does not exist in the DB, the interactive server 200 generates response information based on one of stored uttered history information just before the uttered voice "Select first one." is received and the extracted first and second uttered elements and transmits the response information to the display apparatus 100. For example, an uttered voice "Search for ○○○ (content name)." may be received just before the uttered voice "Select first one." is received, and utterance history information about the uttered voice "○○○ (content name)" may be stored. Therefore, the interactive server 200 generates response information related to a content execution of "○○○ (content name)" based on previous utterance history information and transmits the response information to the display apparatus 100. Therefore, the display apparatus 100 displays a first content of "○○○ (content name)" based on the response information received from the interactive server 200.

[0044] The interactive server 200 may include a first server 10 which generates an uttered voice of the user converted into a digital signal as text information and a second server 20 which generates response information corresponding to the uttered voice generated as the text information. An interactive system which provides response information appropriate for an uttered voice of a user through the display apparatus 100 and the first and second servers 10 and 20 will now be described in detail.

[0045] FIG. 2 is a view illustrating an interactive system which provides response information appropriate for an uttered voice of a user, according to another exemplary embodiment of the present general inventive concept.

[0046] As shown in FIG. 2, if an uttered voice is input from a user, a display apparatus 100 converts the input uttered voice into a digital signal and transmits the digital signal to the first server 10. If the uttered voice converted into the digital signal is received, the first server 10 generates text information of the uttered voice of the user according to a particular pattern of stored various uttered voices and transmits the text information to the display apparatus 100.

[0047] The display apparatus 100 receives the text information of the uttered voice of the user from the first server 100 and transmits the text information of the uttered voice of the user to the second server 20. The second server 20 receives the text information of the uttered voice of the user, parses the received text information to extract uttered elements, generates response information for performing an operation corresponding to the uttered voice of the user based on the extracted uttered elements, and transmits the response information to the display apparatus 100.

[0048] An operation of providing response information corresponding to an uttered voice of a user in an interactive system including the display apparatus 100 and the

interactive server 200 has been schematically described in brief. Elements of the display apparatus 100 and elements of the interactive server 200 will now be described in detail.

[0049] FIG. 3 is a block diagram illustrating an interactive server 200 according to an exemplary embodiment of the present general inventive concept.

[0050] Referring to FIG. 3, the interactive server 200 includes a communicator 210, a voice processor 220, a storage device 230, an extractor 240, and a controller 250.

[0051] The communicator 210 communicates with the display apparatus 100 which provides an uttered voice of a user. In particular, the communicator 210 receives a digital signal (hereinafter referred to as an uttered voice) of an uttered voice including a first uttered element indicating an object to be executed and a second uttered element indicating an execution command. Here, the first uttered element indicating the object to be executed is an uttered element classified as a feature in the uttered voice of the user. Also, the second uttered element indicating the execution command is an uttered element classified as a user action in the uttered voice of the user. For example, in an uttered voice "Search for action movie.", "action" may be a first uttered element classified as a feature, and "Search" may be a second uttered element classified as a user action.

[0052] If the uttered voice including the first and second uttered elements is received as described above, the voice processor 220 converts the received uttered voice into text information. According to an exemplary embodiment, the voice processor 220 may convert the received uttered voice of the user into a text by using a Speech-to-Text (STT) algorithm. However, the present general inventive concept is not limited thereto, and thus the communicator 210 may receive the text information of the uttered voice of the user from the display apparatus 100. In this case, the display apparatus 100 receives the text information of the uttered voice of the user from a voice recognizing apparatus (not shown) such as the first server 10 and transmits the text information of the uttered voice of the user to the interactive

server 200. Therefore, the communicator 210 receives the text information of the uttered voice of the user from the display apparatus 100. If the text information of the uttered voice of the user is received from the display apparatus 100, a configuration of the voice processor 220 may be omitted.

[0053] If the uttered voice of the user is converted into the text information or the text information of the uttered voice of the user is received from the display apparatus 100, the extractor 240 extracts uttered elements from the uttered voice converted into the text information. As described above, the extractor 240 may extract a first uttered element indicating a feature and a second uttered element indicating an execution command from the uttered voice converted into the text information.

[0054] As in the above-described example, the extractor 240 may extract first uttered elements "action" and "movie" and a second uttered element "Search" from an uttered voice "Search for action movie." As another example, the extractor 240 may extract a first uttered element "this" and a second uttered element "Execute" from an uttered voice "Execute this."

[0055] If the uttered elements as described above are extracted, the controller 250 generates response information based on utterance history information stored in the storage device 230 and the extracted uttered elements and transmits the generated response information to the display apparatus 100. Here, the storage device 230 is a storage medium which stores various types of programs necessary for operating the interactive server 200 and may be realized as a memory, a hard disk drive (HDD), or the like. For example, the storage device 230 may include a read only memory (ROM) which stores a program for performing an operation of the controller 250, a random access memory (RAM) which temporarily stores data generated according to the performance of the operation of the controller 250, etc. The storage device 230 may further include an electrically erasable and programmable ROM (EEROM) which stores various types of reference data. In particular, the storage device 230

stores the utterance history information of the uttered voice received from the display apparatus 100. Here, the utterance history information refers to information which is generated based on an uttered voice received before a current uttered voice is received from the display apparatus 100. In detail, the storage device 230 stores the utterance history information which includes at least one of the text information generated from the uttered voice received from the display apparatus 100, time information about when the corresponding uttered voice is received, response information generated based on the corresponding uttered voice, and time information about when the response information is generated.

[0056] Therefore, the controller 250 generates response information based on at least one of previous utterance history information stored in the storage device 230 before a current uttered voice is received and the uttered elements extracted from the received uttered voice and transmits the response information to the display apparatus 100. In detail, the controller 250 determines whether information corresponding to the uttered elements extracted by the extractor 240 exists in the storage device 230. If it is determined that the information corresponding to the extracted uttered elements exists in the storage device 230, the controller 250 generates response information based on the information corresponding to the extracted uttered elements. Here, the information corresponding to the extracted uttered elements may include at least one of an execution language corresponding to a first uttered element indicating an object to be executed and a command corresponding to a second uttered element indicating an execution command. The execution language and the command may be stored in the storage device 230. In other words, the storage device 230 may further include execution languages defined according to first uttered elements and commands defined according to second uttered elements. For example, a first uttered element indirectly indicating an object "this" to be executed may be defined as an execution

language "$this$". A first uttered element indirectly indicating an object "next" to be executed may be defined as an execution language "$this$". A first uttered element indirectly indicating an object "third" to be executed may be defined as an execution language "$3rd$". A first uttered element directly indicating an object to be executed like "○○○ (content name)" may be defined as an execution language "$○○○(content name)$". A second uttered element indicating an execution command may be classified into a content search, a content recommendation, a device function execution, a device user interface (UI) control, and a web search based on a first uttered element directly indicating an object to be executed and an execution area which is to provide response information corresponding to an uttered voice. A command may be defined according to the classified types. Therefore, if uttered elements are extracted from an uttered voice of the user, the controller 250 determines whether an execution language and a command corresponding to the extracted uttered elements have been stored in the storage device 230. If it is determined that the execution language and the command corresponding to the extracted uttered elements have been stored in the storage device 230, the controller 250 acquires the execution language and the command corresponding to the extracted uttered elements and combines the acquired execution language and command to generate response information.

[0057] For example, the extractor 240 may extract a first uttered element "○○○ (content name)" and a second uttered element "Search" from an uttered voice "Search for ○○○(content name)." If the first and second uttered elements are extracted from the uttered voice "Search for ○○○(content name)", the controller 250 determines that an execution area is a content based on the extracted first uttered element. The controller 250 also acquires an execution language "$○○○(content name)$" corresponding to the first uttered element "○○○(content name)" with reference to execution languages defined according to first uttered elements stored in the storage device 230. The controller 250 determines that a type of the

second uttered element "Search" is related to a content search based on the execution area determined as the content. Therefore, the controller 250 acquires a command "content_search" corresponding to the content search from commands defined according to second uttered elements stored in the storage device 230. The controller 250 combines the execution language "$○○○(content name)$" and the command “content_search" respectively corresponding to the first and second uttered elements to generate an execution command script "content_search($○○○(content name)$)" If the execution command script is generated, the controller 250 transmits response information including the generated execution command script to the display apparatus 100. Therefore, the display apparatus 100 receives content information of episodes of "○○○(content name)" from the content providing server 300 based on the execution command script included in the response information received from the interactive server 200 and displays a content list of the received content information of the episodes on a screen.

[0058] However, the present general inventive concept is not limited thereto, and thus the controller 250 may acquire content information of episodes of "○○○(content name)" from electronic program guide (EPG) information stored in the storage device 230 based on the execution command script ""content_search($○○○(content name)$)" or may receive the content information from the external server 400 and transmit the content information to the display apparatus 100.

[0059] As another example, the extractor 240 may extract a first uttered element "bedtime reservation" and a second uttered element "Make" from an uttered voice "Make bedtime reservation." If the first and second uttered elements are extracted from the uttered voice "Make bedtime reservation.", the controller 250 determines that an execution area is a device control based on the extracted first uttered element. The controller 250 also acquires an execution language "$bedtime reservation$" corresponding to the first uttered element

"bedtime reservation" with reference to execution languages defined according to first uttered elements stored in the storage device 230. The controller 250 determines that a type of the second uttered element "Make" corresponds to a device function execution based on the execution area determined as the device control. Therefore, the controller 250 acquires a command "device_execute" corresponding to the device function execution from commands defined according to second uttered elements stored in the storage device 230. The controller 250 combines the execution language "$bedtime reservation$" and the command "device_execute" respectively corresponding to the first and second uttered elements to generate an execution command script ""device_execute($bedtime reservation$)". If the execution command script is generated, the controller 250 transmits response information including the generated execution command script to the display apparatus 100. Therefore, the display apparatus 100 sets a bedtime reservation based on the execution command script of the response information received from the interactive server 200 and outputs a notification message showing that the bedtime reservation has been set through at least one of a voice and a text.

[0060] As another example, if uttered information including an uttered voice "Select first one." and execution information is received, the extractor 240 extracts a first uttered element "first one" and a second uttered element "Select" from the uttered voice "Select first one." If the first and second uttered elements are extracted from the uttered voice "Select first one.", the controller 250 acquires an execution language "$1st$" corresponding to the first uttered element "first one" with reference to execution languages defined according to first uttered elements stored in the storage device 230. The controller 250 determines that an execution area is a device UI control based on the execution information of the received uttered information and determines that a type of the second uttered element "Select" corresponds to the device UI control based on the determined execution area. Therefore, the

controller 250 acquires a command "device_select" corresponding to the device UI control from commands defined according to second uttered elements stored in the storage device 230. If the execution language "$1st$" corresponding to the first uttered element "first one" is acquired, the controller 250 converts the execution language into an execution language corresponding to an acquired execution language "$1st$" based on the execution information of the uttered information received from the display apparatus 100. The controller 250 combines the execution language converted from the acquired execution language "$1st$" and the command to generate an execution command script.

[0061] An operation of transmitting an uttered voice and uttered information including execution information from the display apparatus 100 to the interactive server 200 and an operation of generating response information based the uttered information received from the display apparatus 100 in the interactive server 200 as in the above-described examples will now be described.

[0062] FIG. 4 is a view illustrating a method of generating uttered information based on a display state of an object displayed on a screen of a display apparatus according to an exemplary embodiment of the present general inventive concept.

[0063] As shown in FIG. 4, the display apparatus 100 receives and displays a first content 30 through a channel requested by a user. The display apparatus 100 also displays a content list 40 of a second content requested by the user based on a user command on a screen. Content information 41 through 45 of episodes of the second content which have been broadcasted until now are displayed on the content list 40. The display apparatus 100 allocates identification information to the content information 41 through 45 displayed on the content list 40 based on preset priorities. When the first content 30 and the content list 40 of the second content are displayed on the screen as described above, the user makes an utterance to view the second content corresponding to a particular episode with reference to

the content information 41 through 45 displayed on the content list 40. For example, the user makes an utterance "Select third one." to view the second content corresponding to the content information 43 of a third episode. If such an uttered voice is input, the display apparatus 100 determines whether a currently input uttered voice has been input within a preset time, based on an input time of a previous uttered voice. If it is determined the currently input uttered voice has been input within the preset time, the display apparatus 100 transmits an uttered voice "Select third one." to the interactive server 200. If it is determined that the currently input uttered voice has not been input within the preset time, the display apparatus 100 transmits the uttered voice "Select third one." and uttered information including the identification information (the above-described execution information) respectively allocated to the content information 41 through 45 of the episodes displayed on the content list 40 to the interactive server 200.

[0064] If the uttered information as described above is received, the extractor 240 extracts a first uttered element "third one" and a second uttered element "Select" from the uttered voice of the received uttered information. If the first and second uttered elements are extracted, the controller 250 determines whether an execution language and a command respectively corresponding to the extracted first and second uttered elements have been stored in the storage device 230. As described above, the storage device 230 may store execution languages and commands defined according to first and second uttered elements. Therefore, the controller 250 acquires an execution language "$3rd$" corresponding to the first uttered element "third one" with reference to the execution languages defined according to the first uttered elements stored in the storage device 230. The controller 250 determines that the execution area is a device UI control based on the identification information of the received uttered information and determines that a type of the second uttered element "Select" is a device UI control based on the determined execution area. Therefore, the controller 250

acquires a command "device_select" corresponding to the device UI control from commands defined according to second uttered elements stored in the storage device 230. If an execution language "$3rd$" corresponding to the first uttered element "third one" is acquired, the controller 250 converts the execution language into an execution language corresponding to an acquired execution language "$3rd$" based on the identification information included in the uttered information received from the display apparatus 100. For example, if identification information of the content information 43 corresponding to the execution language "$3rd$" is "#3", the controller 250 may convert the execution language "$3rd$" into an execution language "#3" converted from an acquired execution language "$3rd$". The controller 250 combines the execution language "#3" and the command "device_select" to generate an execution command script "device_select(#3)" and transmits response information including the execution command script "device_select(#3)" to the display apparatus 100. Therefore, the display apparatus 100 receives and displays the second content of an episode corresponding to the content information 43 from the content providing server 300 based on the execution command script included in the response information received from the interactive server 200.

[0065] If an execution language and a command corresponding to uttered elements extracted from an uttered voice of a user exist as described above, the controller 250 generates response information corresponding to an intention of the user based on the execution language and the command corresponding to the uttered elements.

[0066] If the execution language and the command corresponding to the uttered elements extracted from the uttered voice of the user do not exist, the controller 250 estimates an execution language and a command corresponding to the extracted uttered elements based on utterance history information stored in the storage device 230. The controller 250

combines the estimated execution language and command to generate response information appropriate for an intention of the user.

[0067] According to an exemplary embodiment, if the execution language and the command corresponding to the extracted uttered elements, i.e., first and second uttered elements, exist, the controller 250 acquires utterance history information stored within a preset time range based on a time when a current uttered voice is received, from utterance history information stored in the storage device 230. The controller 250 estimates an execution language corresponding to the first uttered element and a command corresponding to the second uttered element as an execution language and a command corresponding to the received uttered voice within the acquired utterance history information. If the utterance history information stored within the preset time range is not stored based on the time when the current uttered voice is received, among the utterance history information stored in the storage device 230, the controller 250 does not generate response information.

[0068] As described above, the storage device 230 may store utterance history information including at least one of text information generated from an uttered voice received from the display apparatus 100, time information about when the corresponding uttered voice is received, response information generated based on the corresponding uttered voice, and time information about when the response information is generated. Therefore, if at least one of an execution language and a command corresponding to first and second uttered elements extracted from a currently received uttered voice is not stored in the storage device 230, the controller 250 estimates an execution language and a command corresponding to the first and second uttered elements based on the utterance history information stored in the storage device 230.

[0069] For example, the extractor 240 may extract a first uttered element "first one" and a second uttered element "Select" from an uttered voice "Select first one." If the first and

second uttered elements are extracted as described above, the controller 250 determines whether an execution language and a command respectively corresponding to the first and second uttered elements have been stored in the storage device 230. In detail, the controller 250 acquires an execution language "$1st$" corresponding to the uttered element "first one" with reference to execution languages defined according to first uttered elements stored in the storage device 230. The controller 250 determines whether it is possible to determine an execution area based on the first uttered element. If it is determined that it is impossible to determine the execution area from the first uttered element "first one", the controller 250 does not acquire a command corresponding to the second uttered element "Select" from the storage device 230 which stores commands defined according to second uttered elements. Therefore, the controller 250 determines that an execution area and an execution command for providing a service of the uttered voice is unclear from the uttered voice "Select first one." According to the determination result, the controller 250 determines whether utterance history information has been stored within a preset time range among utterance history information stored in the storage device 230 based on a time when the uttered voice "Select first one." is received. In other words, the controller 250 determines whether a time difference between the received time of the uttered voice "Select first one." and the latest previous utterance history information belongs to a preset time range among utterance history information stored in the storage device 230. If it is determined that the latest previous utterance history information does not belong to the preset time range, the controller 250 does not generate response information with respect to the uttered voice "Select first one." If it is determined that the latest previous utterance history information belongs to the preset time range, the extractor 240 extracts first and second uttered elements from text information of the uttered voice included in the latest previous utterance history information.

[0070] For example, if the uttered voice included in the previous utterance history information is "Search for "○○○ (content name)"", the extractor 240 extracts a first uttered element "○○○ (content name)" and a second uttered element "Search". The controller 250 estimates the first uttered element "first one" as the first uttered element "○○○ (content name)" based on the first uttered element "○○○ (content name)" extracted from the uttered voice of the previous utterance history information. The controller 250 determines that an execution area is a device UI control based on the estimated first uttered element and acquires a command "device_select" corresponding to the device UI control from commands defined according to second uttered elements stored in the storage device 230. The controller 250 combines the execution language "$1st$" corresponding to the acquired first uttered element and the command "device_select" corresponding to the second uttered element to generate an execution command script "device_select($1st$)". The controller 250 transmits response information including the generated execution command script to the display apparatus 100. Therefore, the display apparatus 100 outputs a response message "You have selected first content." through at least one of a voice and a text based on the response information received from the interactive server 200. The display apparatus 100 also receives a first content from the content providing server 300 based on the execution command script of the response information and displays the first content.

[0071] According to another exemplary embodiment, if a received uttered voice does not include at least one of first and second uttered elements, the controller 250 may generate response information based on utterance history information stored within a preset tine range among utterance history information stored in the storage device 230 based on a time when the uttered voice is received.

[0072] For example, the extractor 240 does not extract first and second uttered elements from an uttered voice "Yes." In this case, the controller 250 determines that the

uttered voice "Yes." is unclear and acquires previous utterance history information stored within a preset time range from utterance history information stored in the storage device 230 based on a time when the uttered voice "Yes." is received. For example, if an uttered voice included in the acquired previous utterance history information is "Turn off TV.", the controller 250 extracts a first uttered element "TV" and a second uttered element "Turn off". The controller 250 estimates the currently received uttered voice "Yes." as "confirm" based on the extracted first and second uttered elements. Therefore, the controller 250 determines that an execution area is a device control based on the first uttered element "TV" extracted from the previous utterance history information. The controller 250 acquires execution language "TV" corresponding to the first uttered element "TV" with reference to execution languages defined according to first uttered elements stored in the storage device 230. The controller 250 determines that a type of the second uttered element "Turn off" is a device function execution based on the execution area determined as the device control. Therefore, the controller 150 acquires a command corresponding to the device function execution from commands defined according to second uttered elements stored in the storage device 230. The controller 250 acquires a command "device_execute_turn off" corresponding to a power control from commands acquired to correspond to the device function execution. The controller 250 combines the execution language "TV" corresponding to the first uttered element and the command "device_execute_turn off" corresponding to the second uttered element to generate an execution command script "device_exectue_turn off(TV)". If the execution command script is generated, the controller 250 transmits response information including the generated execution command script to the display apparatus 100. Therefore, the display apparatus 100 turns off a TV based on the execution command script included in the response information received from the interactive server 200.

[0073] The elements of the interactive server 200 according to the present general inventive concept have been described in detail. Elements of the display apparatus 100 which receive an uttered voice of a user will now be described in detail.

[0074] FIG. 5 is a block diagram illustrating a display apparatus 100 according to an exemplary embodiment of the present general inventive concept.

[0075] Referring to FIG. 5, the display apparatus 100 includes an input device 110, a communicator 120, a display device 130, a controller 140, and a storage device 150.

[0076] The input device 110 receives an uttered voice from a user. In detail, if an analog uttered voice of the user is input, the input device 110 samples the input analog uttered voice of the user to convert the analog uttered voice into a digital signal. Here, if the input uttered voice of the user includes noise (e.g., an air conditioner sound, a cleaner sound, or the like), the input device 110 removes the noise from the uttered voice of the user and then converts the uttered voice of the user from which the noise has been removed, into a digital signal. The input device 110 also receives various types of user controls and transmits the various types of user controls to the controller 130. In this case, the input device 110 receives a user control command through a touch pad, a keypad including various types of functional keys, numeral keys, special keys, letter keys, etc., or a touch screen.

[0077] The communicator 120 transmits the uttered voice of the user input through the input device 110 to a server apparatus (hereinafter referred to as an interactive server) and receives response information corresponding to the transmitted uttered voice. Here, the communicator 120 may include various types of communication modules such as a near field communication (NFC) module (not shown), a wireless communication module (not shown), etc. Here, the NFC module (not shown) communicates with the interactive server 200 and an external server (not shown) providing a content and may be a Bluetooth communication module, a Zigbee communication module, or the like. Here, the interactive server 200 and

the external server are positioned in a near field. The wireless communication module is connected to an external network according to a wireless communication protocol such as WiFi, Institute of Electrical and Electronics Engineers (IEEE), or the like to perform a communication. The wireless communication module may further include a mobile communication module which accesses a mobile communication network according to various types of mobile communication standards, such as 3rd Generation (3G), 3rd Generation Partnership Project (3GPP), Long term Evolution (LTE), etc., to perform a communication.

[0078] The display device 130 may be realized as a liquid crystal display (LCD), an organic light-emitting display (OLED), a plasma display panel (PDP), or the like to provide various types of display screens through the display apparatus 100. In particular, the display device 130 displays a content or content information corresponding to the uttered voice of the user based on the response information received from the interactive server 200.

[0079] The controller 140 controls elements of the display apparatus 100. In particular, if the response information including an execution command script generated through a combination of an execution language and a command is received from the interactive server 200, the controller 140 performs an operation based on the execution command script included in the received response information.

[0080] If response information of an uttered voice of the user "Search for ○○○ (content name)." is received from the interactive server 20 as in the above-described example, the controller 140 receives content information of episodes of "○○○(content name)" from the content providing server 300 through the communicator 120 based on an execution command script of the received response information and displays a content list of the content information of the episodes on a screen through the display device 130. In detail, the response information received from the interactive server 200 may include an execution

command script "content_search($○○○(content name)$)". Therefore, the controller 140 parses the execution command script "content_search($○○○(content name)$)" to classify an execution language as "$○○○(content name)$" and a command as "content_search". If the execution language and the command are classified from the execution command script as described above, the controller 140 controls the display device 130 to receive content information of episodes of "○○○ (content name)" from the content providing server 300 based on the classified execution language and display a content list of the content information of the episodes on a screen. However, the present general inventive concept is not limited thereto, and thus the communicator 120 may receive response information including content information of episodes of "○○○ (content name)" from the interactive server 200. In this case, the controller 140 may display the content information of the episodes of "○○○ (content name)" included in the response information received from the interactive server 200 on the screen through the display device 130.

[0081] If response information with respect to an uttered voice of a user "Make bedtime reservation." is received from the interactive server 200 as in the another above-described example, the controller 140 automatically sets a bedtime reservation to a preset time based on an execution command script included in the received response information. In detail, the response information received from the interactive server 200 may include an execution command script "device_execute($bedtime reservation$)". Therefore, the controller 140 parses the execution command script "device_execute($bedtime reservation$)" to classify an execution language as "$$bedtime reservation$$" and a command as "device_execute". If the execution language and the command are classified from the execution command script, the controller 140 automatically sets the bedtime reservation to a preset time based on the classified execution language. If the bedtime reservation is automatically set as described above, the controller 140 outputs a text format notification

message "Bedtime reservation has been completely set." on the screen through the display device 130 based on the received response information.

[0082] As shown in FIG. 4, the display device 130 may receive and display the first content 130 through a channel requested by the user according to a control command of the controller 140. The display device 130 may display the content list 40 of the second content requested by the user on the screen based on a user command. If the content list 40 including the content information 41 through 45 of the episodes of the second content is displayed on the screen as described above, the controller 140 may allocate identification information to each of the content information 41 through 45 of the episodes based on preset priorities. If the first content 30 and the content list 40 of the second content are displayed on the screen, an uttered voice for viewing the second content corresponding to a particular episode may be input from the user. For example, the user may make an utterance "Select third one." to view the second content corresponding to the content information 43 of the third episode. If this uttered voice is input, the controller 140 determines whether utterance history information of the uttered voice "Select third one." has been stored in the storage device 150. Here, the utterance history information may include an uttered voice input from a previous user and time information about when the corresponding uttered voice is input. If it is determined that the utterance history information has not been stored in the storage device 150, the controller 140 transmits uttered information including the uttered voice "Select third one." and the identification information allocated to the content information 41 through 45 of the episodes to the interactive server 200.

[0083] If it is determined that the utterance history information has been stored in the storage device 150, the controller 140 acquires the latest utterance history information based on the time information included in the stored utterance history information. The controller 140 determines whether the uttered voice "Select third one." has been input within a preset

time based on the time information included in the acquired utterance history information. If it is determined that the uttered voice "Select third one." has been input within the preset time, the controller 140 transmits the uttered voice "Select third one." to the interactive server 200 through the communicator 120. If it is determined that the uttered voice "Select third one." has not been input within the preset time, the controller 140 transmits the uttered information including the uttered voice "Select third one." and the identification information (the above-described execution information) allocated to the content information 41 through 45 of the episodes displayed on the content list 40 to the interactive server 200.

[0084] If response information with respect to the uttered voice "Select third one." is received from the interactive server 200, the controller 140 receives the second content corresponding to the content information 43 of the third episode from the content providing server 300 based on the execution command script included in the received response information and displays the second content through the display device 130.

[0085] The elements of the display apparatus 100 which recognize an uttered voice of a user and perform operations based on response information corresponding to the recognized uttered voice have been described in detail. A method of controlling the interactive server 200 which provides response information corresponding to an uttered voice of a user and a method of controlling the display apparatus 100 which perform an operation based on response information will now be described in detail.

[0086] FIG. 6 is a flowchart illustrating a method of controlling an interactive server according to an exemplary embodiment of the present general inventive concept.

[0087] Referring to FIG. 6, in operation S610, the interactive server 200 receives an uttered voice signal (hereinafter referred to as an uttered voice) from the display apparatus 100. Here, an uttered voice of a user is a voice signal which is converted from an analog voice signal into a digital signal. In operation 620, the interactive server 200 converts the

received uttered voice into text information. According to an exemplary embodiment, the interactive server 200 may convert the uttered voice of the user into a text by using an STT algorithm. However, the present general inventive concept is not limited thereto, and thus the interactive server 200 may receive the text information of the uttered voice of the user from the display apparatus 100. In this case, the display apparatus 100 receives the text information of the uttered voice of the user from an automatic speech recognition (ASR) server such as the first server 10 and transmits the text information to the interactive server 200. Therefore, the interactive server 200 may receive the text information of the uttered voice of the user from the display apparatus 100.

[0088] If the uttered voice of the user is converted into the text information or the text information of the uttered voice of the user is received from the display apparatus 100, the interactive server 200 extracts uttered elements from the uttered voice converted into the text information in operation S630. In other words, the interactive server 200 extracts a first uttered element indicating a feature and a second uttered element indicating an execution command from the uttered voice converted into the text information. If at least one of the first and second uttered elements is extracted, the interactive server 200 determines whether it is possible to generate response information appropriate for an intention of the user based on the extracted uttered element in operation S640. If it is determined in operation S640 that it is impossible to generate the response information based on the extracted uttered element, the interactive server 200 does not generate the response information from the uttered voice of the user. If it is determined in operation S640 that it is possible to generate the response information based on the extracted uttered element, the interactive server 200 generates the response information from the uttered voice of the user in operation S650. In operation S660, the interactive server 200 transmits the response information to the display apparatus 100.

[0089] In detail, if the first and second uttered elements are extracted from the uttered voice of the user, the interactive server 200 determines whether an execution language and a command respectively corresponding to the extracted first and second uttered elements have been stored. Here, the first uttered element is an uttered element indicating a feature, and the second uttered element is an uttered element indicating an execution command. Therefore, the interactive server 200 determines whether an execution language corresponding to the first uttered element indicating the feature and a command corresponding to the second uttered element indicating the execution command have been stored. If it is determined that the execution language and the command respectively corresponding to the extracted first and second uttered elements have been stored, the interactive server 200 combines the execution language and the command respectively corresponding to the extracted first and second uttered elements to generate an execution command script. If it is determined that the execution language and the command respectively corresponding to the extracted first and second uttered elements have not been stored, the interactive server 200 determines whether it is possible to estimate an execution language and a command respectively corresponding to the extracted first and second uttered elements based on stored utterance history information. If it is determined that it is possible to estimate the execution language and the command respectively corresponding to the extracted first and second uttered elements based on the stored utterance history information, the interactive server 200 combines the estimated execution language and command to generate an execution command script. If it is determined that it is impossible to estimate the execution language and the command respectively corresponding to the extracted first and second uttered elements based on the stored utterance history information, the interactive server 200 does not generate the response information of the uttered voice of the user. As described above, the interactive server 200 determines whether it is possible to generate response information of an uttered voice of a

user based on uttered elements extracted from the uttered voice of the user and stored utterance history information and transmits the generated response information to the display apparatus 100 according to the determination result.

[0090] A method of generating response information based on uttered elements extracted from an uttered voice of a user and stored utterance history information in the interactive server 200 will now be described in detail with reference to FIG. 7.

[0091] FIG. 7 is a flowchart illustrating a method of generating response information in an interactive server according to an exemplary embodiment of the present general inventive concept.

[0092] Referring to FIG. 7, an uttered voice of a user is converted into text information or an uttered voice converted into text information is received from the display apparatus 100 through operation 620 described above. If the uttered voice of the user is converted into the text information or is received as described above, the interactive server 200 extracts uttered elements from the uttered voice converted into the text information. If the uttered elements are extracted, the interactive server 200 parses the extracted uttered elements to determine whether the uttered elements include a first uttered element indicating a feature and a second uttered element indicating an execution command in operation S710. If it is determined in operation S710 that the extracted uttered elements do not include at least one of the first and second uttered elements, the interactive server 200 determines whether previous utterance history information has been stored through operation S740 that will be described later.

[0093] If the first and second uttered elements are extracted from the uttered voice of the user, the interactive server 200 determines whether an execution language and a command respectively corresponding to the extracted first and second uttered elements exist in operation S720. If it is determined in operation S720 that the execution language and the

command respectively corresponding to the extracted first and second uttered elements exist, the interactive server 200 combines the execution language and the command respectively corresponding to the extracted first and second uttered elements to generate response information in operation S730. If it is determined in operation S720 that the execution language and the command respectively corresponding to the extracted first and second uttered elements do not exist, the interactive server 200 determines whether utterance history information of the uttered voice exists before the uttered voice of the user is received in operation S740. If it is determined in operation S740 that the utterance history information does not exist, the interactive server 200 does not generate the response information with respect to the uttered voice of the user. If it is determined in operation S740 that the utterance history information exists, the interactive server 200 acquires utterance history information stored within a preset time range based on a time when the uttered voice of the user is received, in operation S750. In operation S760, the interactive server 200 estimates an execution language corresponding to the first uttered element extracted from the uttered voice of the user and a command corresponding to the second uttered element extracted from the uttered voice of the user as an execution language and a command corresponding to the uttered voice of the user within the acquired utterance history information. If the execution language and the command are extracted from the uttered voice of the user as described above, the interactive server 200 combines the estimated execution language and command to generate response information in operation S770.

[0094] A method of generating response information of an uttered voice of a user in the interactive server 200 if uttered elements extracted from the uttered voice of the user do not include at least one of first and second uttered elements will now be described.

[0095] For example, the interactive server 200 does not extract first and second uttered elements from an uttered voice "Yes." In this case, the interactive server 200

determines that the uttered voice "Yes." is unclear and acquires previous utterance history information stored within a preset time range from stored utterance history information based on a time when the uttered voice "Yes." is received. For example, if an uttered voice included in the acquired previous utterance history information is "Turn off TV.", the interactive server 200 extracts a first uttered element "TV" and a second uttered element "Turn off". The interactive server 200 estimates the currently received uttered voice "Yes." as "confirm" based on the extracted first and second uttered elements. Therefore, the interactive server 200 determines that an execution area is a device control based on the first uttered element "TV" extracted from the previous utterance history information. The interactive server 200 acquires an execution language "TV" corresponding to the first uttered element "TV" with reference to execution languages defined according to stored first uttered elements. The interactive server 200 determines that a type of the second uttered element "Turn off" is related to a device function execution based on the execution area determined as the device control. Therefore, the interactive server 200 acquires a command corresponding to the device function execution from commands defined according to stored second uttered elements. The interactive server 200 acquires a command "device_execute_turn off" corresponding to a power control from the commands acquired to correspond to the device function execution. The interactive server 200 combines the execution language "TV" and the command "device_execute_turn off" corresponding to the second uttered element to generate an execution command script "device_execute_turn off(TV)". If the execution command script is generated as described above, the interactive server 200 generates response information including the generated execution command script.

[0096] A method of generating response information of an uttered voice of a user in the interactive server 200 if an execution language and a command respectively

corresponding to first and second uttered elements extracted from the uttered voice of the user exist will now be described.

[0097] If an uttered voice "Search for ○○○ (content name)" is input, the interactive server 200 extracts a first uttered element "○○○ (content name)" and a second uttered element "Search" from the uttered voice "Search for ○○○ (content name)". If the first and second uttered elements are extracted as described above, the interactive server 200 determines that an execution area is a content based on the extracted first uttered element. The interactive server 200 acquires an execution language "$○○○(content name)$" corresponding to the first uttered element "○○○ (content name)" from commands defined according to stored first uttered elements. The interactive server 200 determines that a type of the second uttered element "Search" is related to a content search based on the execution area determined as the content. Therefore, the interactive server 200 acquires a command "content_search" corresponding to the content search from commands defined according to stored second uttered elements.

[0098] If the execution language and the command respectively corresponding to the first and second uttered elements are acquired as described above, the interactive server 200 combines the execution language "$○○○(content name)$" and the command "content_search" to generate an execution command script "content_search($○○○(content name)$)". If the execution command script is generated, the interactive server 200 generates response information including the generated execution command script.

[0099] A method of generating response information of an uttered voice of a user in the interactive server 200 if an execution language and a command respectively corresponding to first and second uttered elements extracted from the uttered voice of the user exist, will now be described.

[0100] For example, interactive server 200 may extract a first uttered element "first one" and a second uttered element "Select" from an uttered voice "Select first one." If the first and second uttered elements are extracted as described above, the interactive server 200 determines whether an execution language and a command respectively corresponding to the first and second uttered elements have been stored. In detail, the interactive server 200 acquires an execution language "$1st$" corresponding to the first uttered element "first one" with reference to execution languages defined according to stored first uttered elements. The interactive server 200 determines whether it is possible to determine an execution area based on the extracted first uttered element. If it is determined that it is impossible to determine the execution area based on the first uttered element "first one", the interactive server 200 does not acquire a command corresponding to the second uttered element "Select" from commands defined according to stored second uttered elements. Therefore, the interactive server 200 determines that an execution area and an execution command for providing a service of the uttered voice is unclear from the uttered voice "Select first one." According to the determination result, the interactive server 200 determines whether stored utterance history information includes utterance history information stored within a preset time range based on a time when the voice "Select first one." was uttered. In other words, the interactive server 200 determines whether a time difference between the time when the uttered voice "Select first one." and the latest stored previous utterance history information belongs to a preset time range, among the stored utterance history information. If it is determined that the latest stored previous utterance history information does not belong to the preset time range, the interactive server 200 does not generate response information of the uttered voice "Select first one." If it is determined that the latest stored previous utterance history information belongs to the preset time range, the interactive server 200 extracts first and second uttered elements

from text information of the uttered voice included in the previous utterance history information.

[0101] For example, if an uttered voice included in previous utterance history information is "○○○ (content name)", the interactive server 200 extracts a first uttered element "○○○(content name)" and a second uttered element "Search". The interactive server 200 estimates the first uttered element "first one" as the first uttered element "○○○(content name)" based on the first uttered element "○○○(content name)" of the first and second uttered elements extracted from the uttered voice included in the previous utterance history information. The interactive server 200 determines that an execution area is a device UI control based on the estimated first uttered element and acquires a command "device_select" corresponding to the device UI control from commands defined according to stored second uttered elements. The interactive server 200 combines the execution language "$1st$" corresponding to the acquired first uttered element and the command "device_select" corresponding to the second uttered element to generate an execution command script "device_select($1st$)". If the execution command script is generated, the interactive server 200 generates response information including the generated execution command script.

[0102] As described above, the interactive server 200 according to the present general inventive concept parses an uttered voice of a user to comprehend an intention of the user and determines whether it is possible to generate response information based on the comprehended uttered intention of the user. If it is determined that it is impossible to generate the response information based on the uttered intention of the user, the interactive server 200 generates response information of a currently uttered voice of the user based on a previously uttered voice of the user. Therefore, the user makes an utterance in a manner the user wants to express without being limited to a particular utterance, to be provided with a service of the utterance.

[0103] The foregoing exemplary embodiments and advantages are merely exemplary and are not to be construed as limiting. The present teaching can be readily applied to other types of apparatuses. Also, the description of the exemplary embodiments is intended to be illustrative, and not to limit the scope of the claims, and many alternatives, modifications, and variations will be apparent to those skilled in the art.

What is Claimed is:

1. An interactive server comprising:

a communicator configured to communicate with a apparatus to receive a first uttered voice signal;

a storage device configured to store utterance history information of a second uttered voice signal received from the apparatus before the first uttered voice signal is received;

an extractor configured to extract uttered elements from the received first uttered voice signal; and

a controller configured to generate response information based on at least one from among the utterance history information stored in the storage device and the extracted uttered elements, and transmit the response information to the apparatus.

2. The interactive server of claim 1, wherein:

if an execution language and a command respectively corresponding to the extracted uttered elements are stored in the storage device, the controller combines the execution language and the command respectively corresponding to the uttered elements to generate the response information; and

if the execution language and the command respectively corresponding to the extracted uttered elements are not stored in the storage device, the controller estimates the execution language and the command respectively corresponding to the extracted uttered elements based on the utterance history information stored in the storage device and combines the estimated execution language and the command to generate the response information.

3. The interactive server of claim 2, wherein:

the first uttered voice signal comprises first and second uttered elements; and

if the execution language and the command respectively corresponding to the extracted uttered elements are not stored in the storage device, the controller acquires the utterance history information stored within a preset time range based on a time when the second uttered voice signal is received and estimates an execution language corresponding to the first uttered element and a command corresponding to the second uttered element as the execution language and the command corresponding to the first uttered voice signal within the acquired utterance history information.

4. The interactive server of claim 3, wherein if the utterance history information stored in the storage device does not comprise the utterance history information stored within the preset time based on the time when the second uttered voice signal is received, the controller does not generate the response information.

5. The interactive server of claim 3, wherein if the first uttered voice signal does not comprise at least one of the first uttered element and the second uttered element, the controller generates response information corresponding to the second uttered voice signal based on the utterance history information stored within the preset time range based on the time when the second uttered voice signal is received.

6. The interactive server of claim 1, further comprising:

a voice processor configured to convert the received first uttered voice signal into text information.

7. A method of controlling an interactive server, the method comprising:

receiving a first uttered voice signal from an apparatus;

extracting uttered elements from the received first uttered voice signal;

generating response information based on at least one from among stored utterance history information and the extracted uttered elements; and

transmitting the generated response information to the apparatus,

wherein the stored utterance history information is information which is generated based on a second uttered voice signal received from the apparatus before the first uttered voice signal is received.

8. The method of claim 7, further comprising:

determining whether an execution language and a command respectively corresponding to the extracted uttered elements exist,

wherein if the execution language and the command respectively corresponding to the uttered elements exist, the execution language and the command respectively corresponding to the uttered elements are combined to generate response information, and if the execution language and the command respectively corresponding to the uttered elements do not exist, the execution language and the command corresponding to the extracted uttered elements are estimated based on the stored utterance history information and are combined to generate the response information.

9. The method of claim 8, wherein:

the first uttered voice signal comprises first and second uttered elements; and

if the execution language and the command respectively corresponding to the uttered elements do not exist, utterance history information stored within a preset time range is

acquired based on a time when the second uttered voice signal is received, and the execution language corresponding to the first uttered element and the command corresponding to the second uttered element are estimated as the execution language and a command corresponding to the uttered voice signal within the acquired utterance history information.

10. The method of claim 9, wherein if the utterance history information does not comprise the utterance history information stored within the preset time range based on the time when the second uttered voice signal is received, the response information is not generated.

11. The method of claim 9, wherein if the first uttered voice signal does not comprise at least one of the first uttered element and the second uttered element, response information corresponding to the first uttered voice signal is generated based on the utterance history information stored within the preset time range based on the time when the second uttered voice signal is received.

12. The method of claim 7, further comprising:

converting the received first uttered voice signal into text information.

13. An interactive system comprising:

an apparatus which, if an uttered voice of a user is input, performs a function corresponding to the input uttered voice; and

an interactive server which, if a first uttered voice signal is received from the apparatus, extracts uttered elements from the first uttered voice signal, generates response

information based on at least one from among stored utterance history information and the extracted uttered elements, and transmits the response information to the apparatus,

wherein the apparatus performs a function corresponding to an intention of the user based on the response information received from the interactive server, and the stored utterance history information is information which is generated based on a second uttered voice signal received from the apparatus before the first uttered voice signal is received.

14. The interactive system of claim 13, wherein the interactive server generates the response information based on an execution language and a command respectively corresponding to the uttered elements if the execution language and the command respectively corresponding to the extracted uttered elements exist, and, if the execution language and the command respectively corresponding to the extracted uttered elements do not exist, estimates the execution language and the command respectively corresponding to the extracted uttered elements based on the stored utterance history information and combines the estimated execution language and command to generate the response information.

15. The interactive server of claim 3, wherein the time is related to a time difference between utterance of the first uttered voice signal and the second uttered voice signal.

16. The method of claim 9, wherein the time is related to a time difference between utterance of the first uttered voice signal and the second uttered voice signal.

17. The interactive server of claim 1, wherein the first uttered voice signal is uttered by a first user and the utterance history information of the second uttered voice signal relates to an uttered voice input of a second user.

18. The method of claim 7, wherein the first uttered voice signal is uttered by a first user and the utterance history information relates to an uttered voice input of a second user.

19. The interactive system of claim 13, wherein the first uttered voice signal is uttered by a first user and the utterance history information relates to an uttered voice input of a second user.

20. The interactive server of claim 1, wherein the apparatus comprises a display apparatus.

21. The method of claim 7, wherein the apparatus comprises a display apparatus.

22. The interactive system of claim 13, wherein the apparatus comprises a display apparatus.

ABSTRACT OF THE DISCLOSURE

An interactive server, a control method thereof, and an interactive system are provided. The interactive server includes: a communicator which communicates with a display apparatus to receive an uttered voice signal; a storage device which stores utterance history information of a second uttered voice signal received from the display apparatus before the first uttered voice signal is received; an extractor which extracts uttered elements from the received first uttered voice signal; and a controller which generates response information based on the utterance history information stored in the storage device and the extracted uttered elements and transmits the response information to the display apparatus. Therefore, the interactive server comprehends intentions of the user with respect to various uttered voices of the user to generate response information according to the intentions and transmits the response information to the display apparatus.

Docket No.: Q203469,SEC's
Ref.: CA-201209-047-1-US0
Only After Sept 16, 2012
US Rights Only

ASSIGNMENT WITH DECLARATION FOR PATENT APPLICATION (37 CFR 1.63)

특허 출원 관련 양도 및 선언 (연방규칙집 제 37 조 1.63 항)

Korean Language Assignment with Declaration

이하 양도인으로 지칭되는 하기 서명 발명자인 본인(들)은 아래에 명시된출원에서 서술한특정 개선을 발명하였으므로 그리고

_____에 소재한 _____(양수인)는(은) 출원 및 발명의 그리고 이에 따라 획득될 일체의 미국 특허에 대한 모든 권리, 소유권 및 이권을 획득하기를 원하므로,

이제 대가를 위한 그 수령을 인정합니다.

상기 양도인인 본인(들)은 상기 명칭의 양수인, 그 승계자 및 그 양수인에게 모든 분할과 그 연속을 포함하여 미국 용으로 공개된 출원과 발명 그리고 그에 대해 승인될 수 있는 미국의 모든 특허증서와 미합중국 코드 제 35 장 제§119 항에 의거한 우선권 청구권을 포함한 모든 그 재발행에 대한 전체 권리, 소유권 및 이권을 판매, 양도 및 이전하며 그리고 본인(들)은 출원서에 제시된 발명에 대해 승인된 일체의 특허증서를 양수인, 그 승계자 또는 그 양수인에게 발급하도록 미국 특허청장에게 요청하는 바이며 또한 본인(들)은 양수인의 요청할 경우 미국 출원과 관련하여 양수인이 필요하다고 생각하는 모든 서류를 더 이상의 보수 없이 서명날인할 것입니다.

(공인은 기록을 위해 요구되지는 않지만 미합중국 코드 제 35 장 제§261 항에 의거 서명날인에 대한 일응의 증거입니다)

본인은 아래 성명의 발명자로서 다음과 같이 선언합니다.

본 양도 및 선언서는 다음사항을 위한 것입니다:

☒ 첨부 출원 또는

☐ 미합중국 출원 또는 PCT
국제출원 번호:_____
출원일:_____

본 출원의 제목은 다음과 같습니다:

_____

_____

_____

Whereas, I/We, the undersigned inventor(s) hereinafter called assignor(s), have invented certain improvements described in the application identified below; and

Whereas, SAMSUNG ELECTRONICS CO., LTD. of 129, Samsung-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do, 443-742, Republic of Korea, (assignee), desires to acquire the entire right, title, and interest in the application and invention, and to any United States patents to be obtained therefor;

Now therefore, for valuable consideration, receipt whereof is hereby acknowledged,

I/We, the above named assignor(s), hereby sell, assign and transfer to the above named assignee, its successors and assigns, the entire right, title and interest in the application and the invention disclosed therein for the United States of America, including all divisions, and continuations thereof, and all Letters Patent of the United States that may be granted thereon, and all reissues thereof, including the right to claim priority under 35 USC 119, and I/we request the Director of the U.S. Patent and Trademark Office to issue any Letters Patent granted upon the invention set forth in the application to the assignee, its successors and assigns; and I/we will execute without further consideration all papers deemed necessary by the assignee in connection with the United States application when called upon to do so by the assignee.

(Legalization not required for recording but is prima facie evidence of execution under 35 USC 261)

As the below named inventor, I hereby declare that:

This assignment with declaration is directed to:

☒ The attached application, or

☐ United States Application or PCT International Application Number _____ filed on _____.

The application is entitled:

INTERACTIVE SERVER, CONTROL METHOD THEREOF, AND INTERACTIVE SYSTEM

Korean Language Assignment with Declaration

위에 표시된 출원은 본인이 신청하였거나 또는 신청하도록 허가를 받았습니다.

본인은, 본인이 본 출원에서 청구된 발명의 원 발명자 또는 원 공동 발명자라고 믿습니다.

본인은 본 양도 및 선언서 제출과 관련된 출원의 내용을 검토하였으며 아울러 이해합니다.

본인은 연방 규정 코드 제 37 장 제 1.56 항에 규정된 바에 따라, 특허성의 판단에 중요한 본인이 아는 모든 정보를 특허청에 공개할 의무가 있음을 알고 있습니다.

본 양도 및 선언서 내에 의도적 허위 진술이 있을 경우 미합중국 코드 제 18 장의 1001 항에 의거 벌금이나 5 년 이하의 징역 또는 두 가지 처벌을 모두 받을 수 있음을 인정합니다.

해당관청의 특허출원 공개 허용 권한

☒ 이 칸에 표시한 경우, 하기의 서명자는 상기 언급된 출원에 대해 우선권을 주장하는 외국 특허청구 사항이 출원되어 있는 유럽특허청(EPO), 일본특허청(JPO), 한국특허청(KIPO), 세계지적재산권기구(WIPO) 및 기타 특허청들에게 상기 언급된 특허 출원 정보를 공개할지 여부를 결정할 권한을 미국특허청(USPTO)에 허여합니다. 연방규정 코드 제 37 장 제 1.14(c)항과 (h)항을 참조하십시오. 만약 상기 출원자가 상기 언급된 출원에 대해 우선권을 주장하는 외국 특허청구 사항이 출원되어 있는 EPO, JPO, KIPO, 또는 기타 특허청들이 본 특허출원 내용을 열람하는 것을 원하지 않는다면 이 칸에 표시하지 않습니다.

연방규정 코드 제 37 장 제 1.14(h)(3)항에 따라, 다음의 경우에 한해 출원된 내용의 사본을 열람할 수 있습니다: 1) 상기 언급된 특허 출원, 2) 미합중국 코드 제 35 장 119(a)-(d)항에 따라 상기 언급된 특허출원이 우선권을 주장하는 대상인 기타 외국 특허출원으로, 그 외국 특허출원 사본이 연방규정 코드 제 37 장 제 1.55 항의 사본 증명 요건에 부합되며, 상기 언급된 특허출원에 이미 출원되어 있는 경우, 3) 상기 언급된 특허출원이 부수적인 이익을 주장하고자 하는 대상인 기타 미국 특허출원.

연방규정 코드 제 37 장 제 1.14(c) 항에 따라, "해당관청의 특허출원 공개 허용 권한" 서류 제출 일자에 관한 정보를 입수할 수 있습니다.

The above identified application was made or was authorized to be made by me.

I believe that I am the original inventor or an original joint inventor of a claimed invention in the application.

I have reviewed and understand the contents of the application for which this assignment with declaration is being submitted.

I am aware of the duty to disclose to the Office all information known to me to be material to patentability as defined in 37 CFR 1.56.

I hereby acknowledge that any willful false statement made in this assignment with declaration is punishable under 18 USC 1001 by fine or imprisonment of not more than five (5) years, or both.

Authorization To Permit Access To Application by Participating Office

☒ If checked, the undersigned hereby grants the USPTO authority to provide the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), the World Intellectual Property Office (WIPO), and any other intellectual property offices in which a foreign application claiming priority to the above-identified application is filed access to the above-identified patent application. See 37 CFR 1.14(c) and (h). This box should not be checked if the applicant does not wish the EPO, JPO, KIPO, or other intellectual property office in which a foreign application claiming priority to the above-identified application is filed to have access to the application.

In accordance with 37 CFR 1.14(h)(3), access will be provided to a copy of the application-as-filed with respect to: 1) the above-identified patent application-as-filed, 2) any foreign application to which the above-identified application claims priority under 35 USC 119(a)-(d) if a copy of the foreign application that satisfies the certified copy requirement of 37 CFR 1.55 has been filed in the above-identified patent application, and 3) any U.S. application-as-filed from which benefit is sought in the above-identified patent application.

In accordance with 37 CFR 1.14(c), access may be provided to information concerning the date of filing the Authorization to Permit Access to Application by Participating Office.

STATEMENT OF ACCURATE TRANSLATION IN ACCORDANCE WITH 37 CFR §1.69(b):

The assignment with declaration is an accurate translation of the corresponding English language assignment with declaration.

Signature

Date 8/30/2012

NAME OF SOLE OR FIRST INVENTOR: 단독 혹은 최초 발명자의 성명			
Given Name (first and middle [if any]) 이름(성씨는 제외) Ji-hye		Family Name or Surname 성(姓) CHUNG	
Inventor's signature 발명자의 서명		Date 일자	12.30.2013
Residence: 거주지: Seoul, Republic of Korea			
Mailing Address: 우편 주소: 113-804, Jamsil els Apt., 19, Jamsil 2-dong, Songpa-gu, Seoul, Republic of Korea			
NAME OF SECOND INVENTOR: 두번째 발명자의 성명			
Given Name (first and middle [if any]) 이름(성씨는 제외) Cheong-jae		Family Name or Surname 성(姓) LEE	
Inventor's signature 발명자의 서명		Date 일자	12.30.2013
Residence: 거주지: Suwon-si, Republic of Korea			
Mailing Address: 우편 주소: 905-1501, Byeokjeokgol 9-danji Jugong Apt., Yeongtong 2-dong, Yeongtong-gu, Suwon-si, Gyeonggi-do, Republic of Korea			
NAME OF THIRD INVENTOR: 세번째 발명자의 성명			
Given Name (first and middle [if any]) 이름(성씨는 제외) Hye-jeong		Family Name or Surname 성(姓) LEE	
Inventor's signature 발명자의 서명		Date 일자	12.30.2013
Residence: 거주지: Seoul, Republic of Korea			
Mailing Address: 우편 주소: 101-404, Ssangyong Apt., Pungnap 2-dong, Songpa-gu, Seoul, Republic of Korea			
NAME OF FOURTH INVENTOR: 네번째 발명자의 성명			
Given Name (first and middle [if any]) 이름(성씨는 제외) Yong-wook		Family Name or Surname 성(姓) SHIN	
Inventor's signature 발명자의 서명		Date 일자	12.30.2013
Residence: 거주지: Osan-si, Republic of Korea			
Mailing Address: 우편 주소: 1190-5, Woolim Apt., Galgot-dong, Osan-si, Gyeonggi-do, Republic of Korea			

TRANSMITTAL FOR POWER OF ATTORNEY TO ONE OR MORE REGISTERED PRACTITIONERS	
NOTE: This form is to be submitted with the Power of Attorney by Applicant form (PTO/AIA/82B or equivalent) to identify the application to which the Power of Attorney is directed, in accordance with 37 CFR 1.5. If the Power of Attorney by Applicant form is not accompanied by this transmittal form or an equivalent, the Power of Attorney will not be recognized in the application.	
Application Number	14/149,296
Filing Date	January 7, 2014
First Named Inventor	Ji-hye CHUNG
Title	INTERACTIVE SERVER, CONTROL METHOD THEREOF, AND INTERACTIVE SYSTEM
Art Unit	Not Yet Assigned
Examiner Name	Not Yet Assigned
Attorney Docket Number	Q203469

SIGNATURE of Applicant or Patent Practitioner			
Signature	/Peter A. McKenna/	Date	January 7, 2014
Name	Peter A. McKenna	Telephone	202.293.7060
Registration Number	38,551		

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

☑ *Total of 1 forms are submitted.

PTO/AIA/82B(07-12)
Approved for use through 11/30/2014. 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 BY APPLICANT

I hereby revoke all previous powers of attorney given in the application identified in the attached transmittal letter.

[■] I hereby appoint Practitioner(s) associated with the following Customer Number as my/our attorney(s) or agent(s), and to transact all business in the United States Patent and Trademark Office connected therewith for the application referenced in the attached transmittal letter (form PTO/AIA/82A or equivalent): 23373

OR

[ ] I hereby appoint Practitioner(s) named below as my/our attorney(s) or agent(s), and to transact all business in the United States Patent and Trademark Office connected therewith for the application referenced in the attached transmittal letter (form PTO/AIA/82A or equivalent):

Name	Registration Number	Name	Registration Number

Please recognize or change the correspondence address for the application identified in the attached transmittal letter to:

[■] The address associated with the above-mentioned Customer Number.
OR
[ ] The address associated with Customer Number:
OR

[ ] Firm or Individual Name					
Address					
City		State		Zip	
Country					
Telephone		Email			

I am the Applicant:

[ ] Inventor or Joint Inventor
[ ] Legal Representative of a Deceased or Legally Incapacitated Inventor
[■] Assignee or Person to Whom the Inventor is Under an Obligation to Assign
[ ] Person Who Otherwise Shows Sufficient Proprietary Interest (*e.g.*, a petition under 37 CFR 1.46(b)(2) was granted in the application or is concurrently being filed with this document)

SIGNATURE of Applicant for Patent

Signature	Heungmo Lee	Date	2012.10.11
Name	Heungmo Lee	Telephone	+82-31-279-4761
Title and Company	Vice President, Samsung Electronics Co., Ltd.		

NOTE: Signature - This form must be signed by the applicant in accordance with 37 CFR 1.33. See 37 CFR 1.4 for signature requirements and certifications. Submit multiple forms for more than one signature, see below *.

[ ] *Total of _____ forms are submitted.

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:	17842948
Application Number:	14149296
International Application Number:	
Confirmation Number:	4038
Title of Invention:	INTERACTIVE SERVER, CONTROL METHOD THEREOF, AND INTERACTIVE SYSTEM
First Named Inventor/Applicant Name:	Ji-hye CHUNG
Customer Number:	23373
Filer:	Peter A. McKenna/Catherine Larson
Filer Authorized By:	Peter A. McKenna
Attorney Docket Number:	Q203469
Receipt Date:	07-JAN-2014
Filing Date:	
Time Stamp:	16:19:09
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	Power of Attorney	Q203469PowerofAttorneyasfiled.pdf	145996 edcf6566264abb6b5eac167db868ffb85a2e5b00	no	2

Warnings:

Information:

Total Files Size (in bytes):	145996

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.

PATENT ASSIGNMENT COVER SHEET

Electronic Version v1.1
Stylesheet Version v1.2

EPAS ID: PAT2670631

SUBMISSION TYPE:	NEW ASSIGNMENT
NATURE OF CONVEYANCE:	ASSIGNMENT

CONVEYING PARTY DATA

Name	Execution Date
JI-HYE CHUNG	12/30/2013
CHEONG-JAE LEE	12/30/2013
HYE-JEONG LEE	12/30/2013
YONG-WOOK SHIN	12/30/2013

RECEIVING PARTY DATA

Name:	SAMSUNG ELECTRONICS CO., LTD.
Street Address:	129, SAMSUNG-RO, YEONGTONG-GU, GYEONGGI-DO
City:	SUWON-SI
State/Country:	KOREA, REPUBLIC OF
Postal Code:	443-742

PROPERTY NUMBERS Total: 1

Property Type	Number
Application Number:	14149296

CORRESPONDENCE DATA

Fax Number: (202)293-7860
Phone: 202-293-7060
Email: sughrue@sughrue.com, clarson@sughrue.com
Correspondence will be sent via US Mail when the email attempt is unsuccessful.
Correspondent Name: SUGHRUE MION, PLLC
Address Line 1: 2100 PENNSYLVANIA AVENUE, N.W.
Address Line 2: SUITE 800
Address Line 4: WASHINGTON, DISTRICT OF COLUMBIA 20037

ATTORNEY DOCKET NUMBER:	Q203469
NAME OF SUBMITTER:	CATHERINE LARSON - TEAMS DEPARTMENT

Signature:	/Catherine Larson/
Date:	01/07/2014
	This document serves as an Oath/Declaration (37 CFR 1.63).

Total Attachments: 3
source=Q203469CombinedDecandAssignmentExecuted#page1.tif
source=Q203469CombinedDecandAssignmentExecuted#page2.tif
source=Q203469CombinedDecandAssignmentExecuted#page3.tif

Docket No.: Q203469,SEC's
Ref.: CA-201209-047-1-US0
Only After Sept 16, 2012
US Rights Only

ASSIGNMENT WITH DECLARATION FOR PATENT APPLICATION (37 CFR 1.63)

특허 출원 관련 양도 및 선언 (연방규칙집 제 37 조 1.63 항)

Korean Language Assignment with Declaration

이하 양도인으로 지칭되는 하기 서명 발명자인 본인(들)은 아래에 명시된 출원에서 서술한 특정 개선을 발명하였으므로 그리고

_____에 소재한 _____(양수인)는(은) 출원 및 발명의 그리고 이에 따라 획득될 일체의 미국 특허에 대한 모든 권리, 소유권 및 이권을 획득하기를 원하므로,

이제 대가를 위한 그 수령을 인정합니다.

상기 양도인인 본인(들)은 상기 명칭의 양수인, 그 승계자 및 그 양수인에게 모든 분할과 그 연속을 포함하여 미국 용으로 공개된 출원과 발명 그리고 그에 대해 승인될 수 있는 미국의 모든 특허증서와 미합중국 코드 제 35 장 제§119 항에 의거한 우선권 청구권을 포함한 모든 그 재발행에 대한 전체 권리, 소유권 및 이권을 판매, 양도 및 이전하며 그리고 본인(들)은 출원서에 제시된 발명에 대해 승인된 일체의 특허증서를 양수인, 그 승계자 또는 그 양수인에게 발급하도록 미국 특허청장에게 요청하는 바이며 또한 본인(들)은 양수인의 요청할 경우 미국 출원과 관련하여 양수인이 필요하다고 생각하는 모든 서류를 더 이상의 보수 없이 서명날인할 것입니다.

(공인은 기록을 위해 요구되지는 않지만 미합중국 코드 제 35 장 제§261 항에 의거 서명날인에 대한 일응의 증거입니다)

본인은 아래 성명의 발명자로서 다음과 같이 선언합니다.

본 양도 및 선언서는 다음사항을 위한 것입니다:

☒ 첨부 출원 또는

☐ 미합중국 출원 또는 PCT
국제출원 번호:_____
출원일:_____

본 출원의 제목은 다음과 같습니다:

_____

_____

_____

Whereas, I/We, the undersigned inventor(s) hereinafter called assignor(s), have invented certain improvements described in the application identified below; and

Whereas, SAMSUNG ELECTRONICS CO., LTD. of 129, Samsung-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do, 443-742, Republic of Korea, (assignee), desires to acquire the entire right, title, and interest in the application and invention, and to any United States patents to be obtained therefor;

Now therefore, for valuable consideration, receipt whereof is hereby acknowledged,

I/We, the above named assignor(s), hereby sell, assign and transfer to the above named assignee, its successors and assigns, the entire right, title and interest in the application and the invention disclosed therein for the United States of America, including all divisions, and continuations thereof, and all Letters Patent of the United States that may be granted thereon, and all reissues thereof, including the right to claim priority under 35 USC 119, and I/we request the Director of the U.S. Patent and Trademark Office to issue any Letters Patent granted upon the invention set forth in the application to the assignee, its successors and assigns; and I/we will execute without further consideration all papers deemed necessary by the assignee in connection with the United States application when called upon to do so by the assignee.

(Legalization not required for recording but is prima facie evidence of execution under 35 USC 261)

As the below named inventor, I hereby declare that:

This assignment with declaration is directed to:

☒ The attached application, or

☐ United States Application or PCT International Application Number _____ filed on _____.

The application is entitled:

INTERACTIVE SERVER, CONTROL METHOD THEREOF, AND INTERACTIVE SYSTEM

Korean Language Assignment with Declaration

위에 표시된 출원은 본인이 신청하였거나 또는 신청하도록 허가를 받았습니다.

본인은, 본인이 본 출원에서 청구된 발명의 원 발명자 또는 원 공동 발명자라고 믿습니다.

본인은 본 양도 및 선언서 제출과 관련된 출원의 내용을 검토하였으며 아울러 이해합니다.

본인은 연방 규정 코드 제 37 장 제 1.56 항에 규정된 바에 따라, 특허성의 판단에 중요한 본인이 아는 모든 정보를 특허청에 공개할 의무가 있음을 알고 있습니다.

본 양도 및 선언서 내에 의도적 허위 진술이 있을 경우 미합중국 코드 제 18 장의 1001 항에 의거 벌금이나 5 년 이하의 징역 또는 두 가지 처벌을 모두 받을 수 있음을 인정합니다.

해당 관청의 특허출원 공개 허용 권한

☒ 이 칸에 표시한 경우, 하기의 서명자는 상기 언급된 출원에 대해 우선권을 주장하는 외국 특허청구 사항이 출원되어 있는 유럽특허청(EPO), 일본특허청(JPO), 한국특허청(KIPO), 세계지적재산권기구(WIPO) 및 기타 특허청들에게 상기 언급된 특허 출원 정보를 공개할지 여부를 결정할 권한을 미국특허청(USPTO)에 허여합니다. 연방규정 코드 제 37 장 제 1.14(c)항과 (h)항을 참조하십시오. 만약 상기 출원자가 상기 언급된 출원에 대해 우선권을 주장하는 외국 특허청구 사항이 출원되어 있는 EPO, JPO, KIPO, 또는 기타 특허청들이 본 특허출원 내용을 열람하는 것을 원하지 않는다면 이 칸에 표시하지 않습니다.

연방규정 코드 제 37 장 제 1.14(h)(3)항에 따라, 다음의 경우에 한해 출원된 내용의 사본을 열람할 수 있습니다: 1) 상기 언급된 특허 출원, 2) 미합중국 코드 제 35 장 119(a)-(d)항에 따라 상기 언급된 특허출원이 우선권을 주장하는 대상인 기타 외국 특허출원으로, 그 외국 특허출원 사본이 연방규정 코드 제 37 장 제 1.55 항의 사본 증명 요건에 부합되며, 상기 언급된 특허출원에 이미 출원되어 있는 경우, 3) 상기 언급된 특허출원이 부수적인 이익을 주장하고자 하는 대상인 기타 미국 특허출원.

연방규정 코드 제 37 장 제 1.14(c) 항에 따라, "해당관청의 특허출원 공개 허용 권한" 서류 제출 일자에 관한 정보를 입수할 수 있습니다.

The above identified application was made or was authorized to be made by me.

I believe that I am the original inventor or an original joint inventor of a claimed invention in the application.

I have reviewed and understand the contents of the application for which this assignment with declaration is being submitted.

I am aware of the duty to disclose to the Office all information known to me to be material to patentability as defined in 37 CFR 1.56.

I hereby acknowledge that any willful false statement made in this assignment with declaration is punishable under 18 USC 1001 by fine or imprisonment of not more than five (5) years, or both.

Authorization To Permit Access To Application by Participating Office

☒ If checked, the undersigned hereby grants the USPTO authority to provide the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), the World Intellectual Property Office (WIPO), and any other intellectual property offices in which a foreign application claiming priority to the above-identified application is filed access to the above-identified patent application. See 37 CFR 1.14(c) and (h). This box should not be checked if the applicant does not wish the EPO, JPO, KIPO, or other intellectual property office in which a foreign application claiming priority to the above-identified application is filed to have access to the application.

In accordance with 37 CFR 1.14(h)(3), access will be provided to a copy of the application-as-filed with respect to: 1) the above-identified patent application-as-filed, 2) any foreign application to which the above-identified application claims priority under 35 USC 119(a)-(d) if a copy of the foreign application that satisfies the certified copy requirement of 37 CFR 1.55 has been filed in the above-identified patent application, and 3) any U.S. application-as-filed from which benefit is sought in the above-identified patent application.

In accordance with 37 CFR 1.14(c), access may be provided to information concerning the date of filing the Authorization to Permit Access to Application by Participating Office.

STATEMENT OF ACCURATE TRANSLATION IN ACCORDANCE WITH 37 CFR §1.69(b):

The assignment with declaration is an accurate translation of the corresponding English language assignment with declaration.

Signature

Date 8/30/2012

NAME OF SOLE OR FIRST INVENTOR: 단독 혹은 최초 발명자의 성명			
Given Name (first and middle [if any]) 이름(성씨는 제외) Ji-hye		Family Name or Surname 성(姓) CHUNG	
Inventor's signature 발명자의 서명		Date 일자	12.30.2013
Residence: 거주지: Seoul, Republic of Korea			
Mailing Address: 우편 주소: 113-804, Jamsil els Apt., 19, Jamsil 2-dong, Songpa-gu, Seoul, Republic of Korea			
NAME OF SECOND INVENTOR: 두번째 발명자의 성명			
Given Name (first and middle [if any]) 이름(성씨는 제외) Cheong-jae		Family Name or Surname 성(姓) LEE	
Inventor's signature 발명자의 서명		Date 일자	12.30.2013
Residence: 거주지: Suwon-si, Republic of Korea			
Mailing Address: 우편 주소: 905-1501, Byeokjeokgol 9-danji Jugong Apt., Yeongtong 2-dong, Yeongtong-gu, Suwon-si, Gyeonggi-do, Republic of Korea			
NAME OF THIRD INVENTOR: 세번째 발명자의 성명			
Given Name (first and middle [if any]) 이름(성씨는 제외) Hye-jeong		Family Name or Surname 성(姓) LEE	
Inventor's signature 발명자의 서명		Date 일자	12.30.2013
Residence: 거주지: Seoul, Republic of Korea			
Mailing Address: 우편 주소: 101-404, Ssangyong Apt., Pungnap 2-dong, Songpa-gu, Seoul, Republic of Korea			
NAME OF FOURTH INVENTOR: 네번째 발명자의 성명			
Given Name (first and middle [if any]) 이름(성씨는 제외) Yong-wook		Family Name or Surname 성(姓) SHIN	
Inventor's signature 발명자의 서명		Date 일자	12.30.2013
Residence: 거주지: Osan-si, Republic of Korea			
Mailing Address: 우편 주소: 1190-5, Woolim Apt., Galgot-dong, Osan-si, Gyeonggi-do, Republic of Korea			