Doc Code: A.NE.AFCP
Document Description: After Final Consideration Pilot Program Request

PTO/SB/434 (05-13)

CERTIFICATION AND REQUEST FOR CONSIDERATION UNDER THE AFTER FINAL CONSIDERATION PILOT PROGRAM 2.0

Practitioner Docket No.:	Application No.:	Filing Date:
N0484.71090US01	13/795,933-Conf. #2321	March 12, 2013

First Named Inventor:	Title:
William F. Ganong, III	METHODS AND APPARATUS FOR DETECTING A VOICE COMMAND

APPLICANT HEREBY CERTIFIES THE FOLLOWING AND REQUESTS CONSIDERATION UNDER THE AFTER FINAL CONSIDERATION PILOT PROGRAM 2.0 (AFCP 2.0) OF THE ACCOMPANYING RESPONSE UNDER 37 CFR 1.116.

1. The above-identified application is (i) an original utility, plant, or design nonprovisional application filed under 35 U.S.C. 111(a) [a continuing application (*e.g.*, a continuation or divisional application) is filed under 35 U.S.C. 111(a) and is eligible under (i)], or (ii) an international application that has entered the national stage in compliance with 35 U.S.C. 371(c).

2. The above-identified application contains an outstanding final rejection.

3. Submitted herewith is a response under 37 CFR 1.116 to the outstanding final rejection. The response includes an amendment to at least one independent claim, and the amendment does not broaden the scope of the independent claim in any aspect.

4. This certification and request for consideration under AFCP 2.0 is the only AFCP 2.0 certification and request filed in response to the outstanding final rejection.

5. Applicant is willing and available to participate in any interview requested by the examiner concerning the present response.

6. This certification and request is being filed electronically using the Office's electronic filing system (EFS-Web).

7. Any fees that would be necessary consistent with current practice concerning responses after final rejection under 37 CFR 1.116, *e.g.*, extension of time fees, are being concurrently filed herewith. [There is no additional fee required to request consideration under AFCP 2.0.]

8. By filing this certification and request, applicant acknowledges the following:

- Reissue applications and reexamination proceedings are not eligible to participate in AFCP 2.0.
- The examiner will verify that the AFCP 2.0 submission is compliant, *i.e.*, that the requirements of the program have been met (see items 1 to 7 above). For compliant submissions:
  - The examiner will review the response under 37 CFR 1.116 to determine if additional search and/or consideration (i) is necessitated by the amendment and (ii) could be completed within the time allotted under AFCP 2.0. If additional search and/or consideration is required but cannot be completed within the allotted time, the examiner will process the submission consistent with current practice concerning responses after final rejection under 37 CFR 1.116, *e.g.*, by mailing an advisory action.
  - If the examiner determines that the amendment does not necessitate additional search and/or consideration, or if the examiner determines that additional search and/or consideration is required and could be completed within the allotted time, then the examiner will consider whether the amendment places the application in condition for allowance (after completing the additional search and/or consideration, if required). If the examiner determines that the amendment does not place the application in condition for allowance, then the examiner will contact the applicant and request an interview.
    - The interview will be conducted by the examiner, and if the examiner does not have negotiation authority, a primary examiner and/or supervisory patent examiner will also participate.
    - If the applicant declines the interview, or if the interview cannot be scheduled within ten (10) calendar days from the date that the examiner first contacts the applicant, then the examiner will proceed consistent with current practice concerning responses after final rejection under 37 CFR 1.116.

Signature	/Thomas S. Chlebeck/	Date	2017-03-27
Name (Print/Typed)	Thomas S. Chlebeck	Practitioner Registration No.	70,616

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

☐ * Total of 1 forms are submitted.

Certificate of Electronic Filing under 37 CFR §1.8

I hereby certify that this paper (along with any paper referred to as being attached or enclosed) is being transmitted via the Office's electronic filing system in accordance with 37 CFR § 1.6(a)(4).

Dated: March 27, 2017 Electronic Signature for: /Gail Driscoll/

15337034.1

Electronic Patent Application Fee Transmittal

Application Number:	13795933
Filing Date:	12-Mar-2013
Title of Invention:	METHODS AND APPARATUS FOR DETECTING A VOICE COMMAND
First Named Inventor/Applicant Name:	William F. Ganong
Filer:	Thomas Stephen Chlebeck/Gail Driscoll
Attorney Docket Number:	N0484.71090US01

Filed as Large Entity

Filing Fees for Utility under 35 USC 111(a)

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

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Extension - 1 month with $0 paid	1251	1	200	200
Miscellaneous:				
			Total in USD ($)	**200**

Doc Code: TRAN.LET
Document Description: Transmittal Letter

USED IN LIEU OF PTO/SB/21 (07-09)

TRANSMITTAL FORM

(to be used for all correspondence after initial filing)

Application Number	13/795,933-Conf. #2321
Filing Date	March 12, 2013
First Named Inventor	William F. Ganong, III
Art Unit	2644
Examiner Name	H. C. Ho
Total Number of Pages in This Submission	
Attorney Docket Number	N0484.71090US01

ENCLOSURES (*Check all that apply*)

- [X] Fee Transmittal Form
  - [ ] Fee Attached
- [X] Amendment/Reply
  - [X] After Final
  - [ ] Affidavits/declaration(s)
- [X] Extension of Time Request
- [ ] Express Abandonment Request
- [ ] Information Disclosure Statement
- [ ] Certified Copy of Priority Document(s)
- [ ] Reply to Missing Parts/ Incomplete Application
  - [ ] Reply to Missing Parts under 37 CFR 1.52 or 1.53
- [ ] Drawing(s)
- [ ] Licensing-related Papers
- [ ] Petition
- [ ] Petition to Convert to a Provisional Application
- [ ] Power of Attorney, Revocation Change of Correspondence Address
- [ ] Terminal Disclaimer
- [ ] Request for Refund
- [ ] CD, Number of CD(s) _____
  - [ ] Landscape Table on CD
- [ ] After Allowance Communication to TC
- [ ] Appeal Communication to Board of Appeals and Interferences
- [ ] Appeal Communication to TC **(Appeal Notice, Brief, Reply Brief)**
- [ ] Proprietary Information
- [ ] Status Letter
- [X] Other Enclosure(s) (please Identify below):

Certification and Request for Consideration Under the After Final Consideration Pilot Program 2.0

Remarks

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT

Firm Name	WOLF, GREENFIELD & SACKS, P.C.		
Signature	/Thomas S. Chlebeck/		
Printed name	Thomas S. Chlebeck		
Date	2017-03-27	Reg. No.	70,616

Certificate of Electronic Filing under 37 CFR §1.8

I hereby certify that this paper (along with any paper referred to as being attached or enclosed) is being transmitted via the Office's electronic filing system in accordance with 37 CFR § 1.6(a)(4).

Dated: March 27, 2017 Electronic Signature for: /Gail Driscoll/

5337023.1

AMENDMENTS TO THE CLAIMS

Please amend the claims as shown below, with insertions indicated by underlining and deletions indicated by strikethrough and/or double brackets.

1. (Currently amended) A method of monitoring an acoustic environment of a mobile device for voice commands when the mobile device is operating in an idle mode, the method comprising:

receiving acoustic input from the environment of the mobile device while the mobile device is operating in the idle mode; and

detecting whether the acoustic input includes a voice command at least by: ~~based on~~

performing a plurality of processing stages on the acoustic input, wherein at least one of the plurality of processing stages is performed while the mobile device is operating in the idle mode; and

using at least one contextual cue to assist in detecting whether the acoustic input includes a voice command while the mobile device is operating in the idle mode, ~~using~~ the at least one contextual cue comprising information other than content of the acoustic input ~~to assist in detecting whether the acoustic input includes a voice command, wherein the at least one contextual cue comprises~~ and comprising a time the acoustic input was received, motion of the mobile device, location of the mobile device, occurrence of an event on the mobile device, recent activity on the mobile device, a stored sample of the acoustic environment of the mobile device and/or a power source of the mobile device.

2. (Previously presented) The method of claim 1, further comprising transitioning the mobile device from the idle mode to an active mode when it is detected that the acoustic input includes a voice command.

3. (Canceled)

4. (Previously presented) The method of claim 1, wherein the at least one contextual cue includes the time the acoustic input was received, the method further comprising influencing

Petitioner Exhibit 1004, Page 673 of 1518

evaluation of at least one of the plurality of processing stages based on the time the acoustic input was received.

5. (Previously presented) The method of claim 1, wherein the at least one contextual cue includes recent activity on the mobile device, the method further comprising influencing evaluation of at least one of the plurality of processing stages based on at least one of how recently the user activated the mobile device, how recently the user issued a voice command and how recently the user performed at least one action on the mobile device.

6. (Previously presented) The method of claim 1, wherein the at least one contextual cue includes an occurrence of an event on the mobile device, the method further comprising influencing evaluation of at least one of the plurality of processing stages based on whether the acoustic input is received proximate in time with the occurrence of the event.

7. (Previously presented) The method of claim 1, wherein the at least one contextual cue includes motion of the mobile device received as an input from an accelerometer of the mobile device, the method further comprising influencing evaluation of at least one of the plurality of processing stages based on an assessment of whether the user moved the mobile device proximate in time with receiving the acoustic input based, at least in part, on the input from the accelerometer.

8. (Previously presented) The method of claim 1, wherein the at least one contextual cue includes the location of the mobile device, the method further comprising influencing evaluation of at least one of the plurality of processing stages based on determining whether the mobile device is located in at least one of a vehicle, a home or a place of work of the user of the mobile device.

9. (Previously presented) The method of claim 1, wherein the at least one contextual cue includes the power source of the mobile device, the method further comprising influencing evaluation of at least one of the plurality of processing stages based on at least one of determining

Petitioner Exhibit 1004, Page 674 of 1518

whether the mobile device is being powered using the battery or connected to wall power, and a state of the battery if the mobile device is being powered using the battery.

10. (Previously presented) The method of claim 1, wherein the at least one contextual cue includes the sample of the acoustic environment obtained by the mobile device by storing acoustic input previously received by the mobile device and determined not to include a voice command, the method further comprising subtracting the sample of the acoustic environment from the acoustic input.

11. (Original) The method of claim 1, wherein a sequence of performing the plurality of processing stages is determined, at least in part, by the at least one contextual cue.

12. (Original) The method of claim 1, wherein the mobile device includes a primary processor and a secondary processor, and wherein at least one first processing stage of the plurality of processing stages is performed by the secondary processor prior to engaging the primary processor to process the acoustic input.

13. (Original) The method of claim 12, wherein at least one second processing stage of the plurality of processing stages is performed by the primary processor only after the at least one first processing stage is performed and neither concludes that the acoustic input includes a voice command nor concludes that the acoustic input is the result of spurious acoustic activity.

14. (Previously presented) The method of claim 12, wherein performing the at least one first processing stage includes performing at least one voice activity detection processing stage including performing at least one of spectral analysis on the acoustic input to evaluate whether a spectrum of the acoustic input is indicative of voice activity, periodicity analysis to evaluate whether signal periodicity of the acoustic input is indicative of voice activity, and phone loop analysis to evaluate whether the acoustic input includes speech.

Petitioner Exhibit 1004, Page 675 of 1518

15. (Original) The method of claim 12, wherein performing the at least one first processing stage comprises performing limited vocabulary automatic speech recognition on the acoustic input to detect whether the acoustic input includes an explicit voice trigger.

16. (Original) The method of claim 12, wherein performing the at least one second processing stage comprises performing at least one speech processing stage including at least one of automatic speech recognition on the acoustic input and at least one natural language processing stage to understand the meaning of the voice command.

17. (Original) The method of claim 1, wherein the mobile device is capable of connecting to at least one server via a network, the method further comprising transmitting at least a portion of the acoustic information to the at least one server via the network for processing by the at least one server.

18. (Original) The method of claim 1, further comprising performing one or more default actions when a specific task could not be ascertained from the acoustic input.

19. (Currently amended) At least one non-transitory computer readable medium encoded with instructions that, when executed on at least one processor, perform a method of monitoring an acoustic environment of a mobile device to receive acoustic input from the environment of the mobile device while the mobile device is operating in an idle mode, the method comprising:

detecting whether the acoustic input includes a voice command at least by: ~~based on~~

performing a plurality of processing stages on the acoustic input, wherein at least one of the plurality of processing stages is performed while the mobile device is operating in the idle mode; and

using at least one contextual cue to assist in detecting whether the acoustic input includes a voice command while the mobile device is operating in the idle mode, ~~using~~ the at least one contextual cue comprising information other than content of the acoustic input ~~to assist in detecting whether the acoustic input includes a voice command~~, ~~wherein the at least~~

Petitioner Exhibit 1004, Page 676 of 1518

~~one contextual cue comprises~~ and comprising a time the acoustic input was received, motion of the mobile device, location of the mobile device, occurrence of an event on the mobile device, recent activity on the mobile device, a stored sample of the acoustic environment of the mobile device and/or a power source of the mobile device.

20. (Currently amended) A mobile device comprising:

at least one input to receive acoustic input from an environment of the mobile device while the mobile device is operating in an idle mode; and

at least one processor configured to:

detect whether the acoustic input includes a voice command at least by: ~~based on~~

performing a plurality of processing stages on the acoustic input, wherein at least one of the plurality of processing stages is performed while the mobile device is operating in the idle mode; and

using at least one contextual cue to assist in detecting whether the acoustic input includes a voice command while the mobile device is operating in the idle mode, ~~use~~ the at least one contextual cue comprising information other than content of the acoustic input ~~to assist in detecting whether the acoustic input includes a voice command, wherein the at least one contextual cue comprises~~ and comprising a time the acoustic input was received, motion of the mobile device, location of the mobile device, occurrence of an event on the mobile device, recent activity on the mobile device, a stored sample of the acoustic environment of the mobile device and/or a power source of the mobile device.

Petitioner Exhibit 1004, Page 677 of 1518

REMARKS

In response to the Office Action dated November 25, 2016, Assignee respectfully requests reconsideration. Claims 1, 2, and 4-20 were previously pending in this application. By this amendment, claims 1, 19, and 20 have been amended. No claims have been added or canceled. As a result, claims 1, 2, and 4-20 are pending for examination with claims 1, 19, and 20 being independent. No new matter has been added.

Telephone Interview

The Assignee's representative, Thomas S. Chlebeck, thanks Examiner Ho for his courtesy extended in granting and conducting a telephone interview on January 20, 2017. The substance of that interview is summarized herein.

Assignee's representative discussed how the claims distinguish over the references. In particular, Assignee's representative discussed techniques described in the specification to provide examples of how *both processing stages* ***on acoustic input*** (received by a mobile device in a low power mode) ***and*** one or more *contextual cues* ***other than the acoustic input*** (as recited in the claims) may be used to detect whether the acoustic input includes a voice command.

Assignee's representative also described some examples provided in the specification of how contextual cues (e.g., time, location, motion of the mobile device, recent activity, etc.) can be used. For example, the time at which acoustic input is received can be used as indication of whether acoustic input is a voice command (e.g., acoustic input received when people normally are sleeping, such as at 3 a.m.) is less likely to include a voice command than acoustic input received during waking hours. Location may also be used as a contextual cue since voice commands may be more likely to be spoken in some locations than others (e.g., at a loud airport or sports arena). Contextual cues may be used alone or together, for example, if the mobile device is in a car it may make it more likely that acoustic input includes a voice command even if the time suggests otherwise (e.g., during a late drive). As another example, because a user often moves a mobile device prior to speaking to it (e.g., such as when a user raises the device towards the mouth, picks up the device, etc.), motion of the device may also be used as a contextual cue to assess whether acoustic input includes a voice command. Other examples are also provided in the specification.

Petitioner Exhibit 1004, Page 678 of 1518

Assignee's representative then explained that Oh's check for touch/button input at S109 in FIG. 5 and paragraph 0129 (which the Examiner had believed might teach a contextual cue as claimed—specifically an occurrence of an event on the mobile device), is only relevant *if no voice command is inputted*, as FIG. 5 clearly shows. The purpose of S109 is to allow the controller 180 of Oh to perform other functions—even when no voice command is inputted—by performing the function selected by the touch/button input at S109. The Examiner appeared to appreciate that a touch/button input that occurs only if a voice command is *not* inputted cannot teach a contextual cue used to assist in detecting whether acoustic input includes a voice command.

The Examiner stated that he found the order of the limitations in the previously pending claims confusing. Assignee's representative offered to clarify the limitations so that the above-emphasized features (how ***both*** *processing stages* ***on acoustic input and*** one or more *contextual cues* ***other than the acoustic input*** may be used to detect whether the acoustic input includes a voice command) were more clearly together. The Examiner agreed that doing so should clearly distinguish from the cited art, especially because the Examiner's understanding of the application had improved. The Examiner suggested the Assignee's representatives submit corresponding amendments under the After Final Consideration Pilot 2.0 program so that he could perform additional searching and call if needed. The amendments herein are thus submitted.

Rejections Under 35 U.S.C. §103

The Office Action rejects independent claims 11, 19, and 20 under 35 U.S.C. § 103(a) as purportedly being unpatentable over Hung (U.S. Patent Publication No. 2014/0012573) in view of Oh (U.S. Patent Publication No. 2010/0009719) and Freeman (U.S. Patent No. 8,977,255). The Assignee respectfully requests reconsideration.

As discussed and agreed during the interview, neither Hung, nor Oh, nor Freeman, alone or in any combination, disclose or suggest at least "*detecting whether the acoustic input includes a voice command* at least *by*: performing a plurality of ***processing stages on the acoustic input***, wherein at least one of the plurality of processing stages is performed while the mobile device is operating *in the idle mode*; and ***using at least one contextual cue to assist in detecting whether the acoustic input includes a voice command*** while the mobile device is operating in the idle mode, the

Petitioner Exhibit 1004, Page 679 of 1518

at least one contextual cue comprising information ***other than content of the acoustic input***, and comprising a time the acoustic input was received, motion of the mobile device, location of the mobile device, occurrence of an event on the mobile device, recent activity on the mobile device, a stored sample of the acoustic environment of the mobile device and/or a power source of the mobile device," as recited by amended claim 1. Independent claims 19 and 20 recite similar limitations.

Therefore, even if one of ordinary skill in the art would have been motivated to combine Hung, Oh, and Freeman (a contention with which the Assignee disagrees), the alleged combination fails to disclose or suggest the above emphasized limitations in claims 1, 19, and 20. Therefore, claims 1, 19, and 20 patentably distinguish over the alleged combination and are in allowable condition. Claims 2 and 4-18 depend from claim 1 and are allowable for at least the same reasons.

General Comments on Dependent Claims

Because each of the dependent claims depends from a base claim that is believed to be in condition for allowance, the Assignee believes that it is unnecessary at this time to argue the further distinguishing features of all of the dependent claims. However, the Assignee does not necessarily concur with the interpretation of the dependent claims as set forth in the Office Action, nor does the Assignee concur that the basis for the rejection of any of the dependent claims is proper. Therefore, the Assignee reserves the right to specifically address in the future the further patentability of the dependent claims not specifically addressed herein.

Petitioner Exhibit 1004, Page 680 of 1518

CONCLUSION

In view of the foregoing remarks, the application is believed to be in condition for allowance. A notice to this effect is respectfully requested. If the Examiner believes for any reason that the application is not in condition for allowance, he is requested to contact the undersigned to discuss any outstanding issues relating to the allowability of the application.

If this response is not considered timely filed and if a request for an extension of time is otherwise absent, the Assignee hereby requests any necessary extension of time. If there is a fee occasioned by this response, including an extension fee, the Director is hereby authorized to charge any deficiency or credit any overpayment in the fees filed, asserted to be filed or which should have been filed herewith to our Deposit Account No. 23/2825 under Docket No. N0484.71090US01 from which the undersigned is authorized to draw.

Dated: 2017-03-27

Respectfully submitted,

By /Thomas S. Chlebeck/
Thomas S. Chlebeck
Registration No.: 70,616
WOLF, GREENFIELD & SACKS, P.C.
600 Atlantic Avenue
Boston, Massachusetts 02210-2206
617.646.8000

Petitioner Exhibit 1004, Page 681 of 1518

Electronic Acknowledgement Receipt

EFS ID:	28745011
Application Number:	13795933
International Application Number:	
Confirmation Number:	2321
Title of Invention:	METHODS AND APPARATUS FOR DETECTING A VOICE COMMAND
First Named Inventor/Applicant Name:	William F. Ganong
Customer Number:	91337
Filer:	Thomas Stephen Chlebeck/Gail Driscoll
Filer Authorized By:	Thomas Stephen Chlebeck
Attorney Docket Number:	N0484.71090US01
Receipt Date:	27-MAR-2017
Filing Date:	12-MAR-2013
Time Stamp:	14:36:06
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	CARD
Payment was successfully received in RAM	$200
RAM confirmation Number	032817INTEFSW14375800
Deposit Account	
Authorized User	

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Transmittal Letter	N048471090US01-TRN-TSC.pdf	25338 5a0d01a5a27c5119a792e7c051124041cef0 2c90	no	1
Warnings:					
Information:					
2	Extension of Time	N048471090US01-EXT-TSC.pdf	23258 d05d2c63aa0a7b300f1d2a7db7f3697de36 bc8e5	no	1
Warnings:					
Information:					
3	Fee Worksheet (SB06)	N048471090US01-FEE-TSC.pdf	31710 7f8ad8df4b984f72b0f5a10662f9c23b802b 1078	no	1
Warnings:					
Information:					
4	After Final Consideration Program Request	N048471090US01-AFCP-TSC. pdf	37704 c0991808c2b674f62644e54f479369bd1f20 175b	no	1
Warnings:					
Information:					
5		N048471090US01-116AMN-TSC.pdf	58995 0cf843a61faa2a82e84e6abf3da15cfba2aef 1df	yes	10

Multipart Description/PDF files in .zip description

Document Description	Start	End
Response After Final Action	1	1
Claims	2	6
Applicant Arguments/Remarks Made in an Amendment	7	10

Warnings:					
Information:					
6	Fee Worksheet (SB06)	fee-info.pdf	30886 429fbfcd226ae9826914d4ba8ae379a32b1cfec1	no	2
Warnings:					
Information:					
			Total Files Size (in bytes):	207891	

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 13/795,933	Filing Date 03/12/2013	☐ To be Mailed

ENTITY: ☒ LARGE ☐ SMALL ☐ MICRO

APPLICATION AS FILED – PART I

	(Column 1)	(Column 2)		
FOR	NUMBER FILED	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

		(Column 1)		(Column 2)	(Column 3)		
AMENDMENT	**03/27/2017**	CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)
	Total (37 CFR 1.16(i))	* 19	Minus	** 20	= 0	x $80 =	0
	Independent (37 CFR 1.16(h))	* 3	Minus	***3	= 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**

		(Column 1)		(Column 2)	(Column 3)		
AMENDMENT		CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	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
DELEACHES YOUNG

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.

PTO/SB/17 (03-13)
Approved for use through 04/30/2017. 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

FEE TRANSMITTAL

Complete if known	
Application Number	13/795,933-Conf. #2321
Filing Date	March 12, 2013
First Named Inventor	William F. Ganong, III
Examiner Name	H. C. Ho
Art Unit	2644
Practitioner Docket No.	N0484.71090US01

☐ Applicant asserts small entity status. See 37 CFR 1.27.

☐ Applicant certifies micro entity status. See 37 CFR 1.29. Form PTO/SB/15A or B or equivalent must either be enclosed or have been submitted previously.

TOTAL AMOUNT OF PAYMENT ($) 2,100.00

METHOD OF PAYMENT (check all that apply)

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

☐ Deposit Account Deposit Account Number: 23/2825 Deposit Account Name: Wolf, Greenfield & Sacks, P.C.

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 filing fee**

☒ Charge any additional fee(s) or underpayment of fee(s) under 37 CFR 1.16 and 1.17 ☒ Credit any overpayment of fee(s)

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

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	
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	2,160	1,080	540	
Provisional	260	130	65	0	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 files 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 ($)**	**Fee Paid ($)**
_____ - or HP =	_____ x	_____ =	_____

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

Indep. Claims	**Extra Claims**	**Fee ($)**	**Fee Paid ($)**
_____ - or HP =	_____ x	_____ =	_____

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

Multiple Dependent Claims	
Fee ($)	**Fee Paid ($)**

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) ($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 ($)**	**Fee Paid ($)**
_____ - 100 =	_____ /50 =	_____ (round **up** to a whole number) x	_____ =	_____

4. OTHER FEE(S)

	Fees Paid ($)
Non-English specification, $130 fee (no small or micro entity discount)	
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): 1252 Extension for response within second month	400.00 (one-month previously paid 3/28/17)
1820 Request for continued examination (RCE) - 2nd and ...	1,700.00

SUBMITTED BY

Signature	/Thomas S. Chlebeck/	Registration No. (Attorney/Agent)	70,616	Telephone	617.646.8000
Name (Print/Type)	Thomas S. Chlebeck			Date	2017-04-25

Certificate of Electronic Filing under 37 CFR §1.8

I hereby certify that this paper (along with any paper referred to as being attached or enclosed) is being transmitted via the Office's electronic filing system in accordance with 37 CFR § 1.6(a)(4).

Dated: April 25, 2017 Electronic Signature: /Trish McDonald/

Doc Code: TRAN.LET
Document Description: Transmittal Letter

USED IN LIEU OF PTO/SB/21 (07-09)

TRANSMITTAL FORM

(to be used for all correspondence after initial filing)

Application Number	13/795,933-Conf. #2321
Filing Date	March 12, 2013
First Named Inventor	William F. Ganong, III
Art Unit	2644
Examiner Name	H. C. Ho
Attorney Docket Number	N0484.71090US01

Total Number of Pages in This Submission

ENCLOSURES (*Check all that apply*)

- [x] Fee Transmittal Form
- [ ] Fee Attached
- [ ] Amendment/Reply
- [ ] After Final
- [ ] Affidavits/declaration(s)
- [x] Extension of Time Request
- [ ] Express Abandonment Request
- [ ] Information Disclosure Statement
- [ ] Certified Copy of Priority Document(s)
- [ ] Reply to Missing Parts/ Incomplete Application
- [ ] Reply to Missing Parts under 37 CFR 1.52 or 1.53
- [ ] Drawing(s)
- [ ] Licensing-related Papers
- [ ] Petition
- [ ] Petition to Convert to a Provisional Application
- [ ] Power of Attorney, Revocation Change of Correspondence Address
- [ ] Terminal Disclaimer
- [ ] Request for Refund
- [ ] CD, Number of CD(s) _____
- [ ] Landscape Table on CD
- [ ] After Allowance Communication to TC
- [ ] Appeal Communication to Board of Appeals and Interferences
- [ ] Appeal Communication to TC **(Appeal Notice, Brief, Reply Brief)**
- [ ] Proprietary Information
- [ ] Status Letter
- [x] Other Enclosure(s) (please Identify below):

Request for Continued Examination (RCE)

Remarks

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT

Firm Name	WOLF, GREENFIELD & SACKS, P.C.		
Signature	/Thomas S. Chlebeck/		
Printed name	Thomas S. Chlebeck		
Date	2017-04-25	Reg. No.	70,616

Certificate of Electronic Filing under 37 CFR §1.8

I hereby certify that this paper (along with any paper referred to as being attached or enclosed) is being transmitted via the Office's electronic filing system in accordance with 37 CFR § 1.6(a)(4).

Dated: April 25, 2017 Electronic Signature: /Trish McDonald/

5401332.1

Electronic Patent Application Fee Transmittal

Application Number:	13795933
Filing Date:	12-Mar-2013
Title of Invention:	METHODS AND APPARATUS FOR DETECTING A VOICE COMMAND
First Named Inventor/Applicant Name:	William F. Ganong
Filer:	Thomas Stephen Chlebeck/Trish McDonald
Attorney Docket Number:	N0484.71090US01

Filed as Large Entity

Filing Fees for Utility under 35 USC 111(a)

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

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Extension - 2 months with $200 paid	1252	1	400	400
Miscellaneous:				
RCE- 2ND AND SUBSEQUENT REQUEST	1820	1	1700	1700
		Total in USD ($)		**2100**

Electronic Acknowledgement Receipt

EFS ID:	29026676
Application Number:	13795933
International Application Number:	
Confirmation Number:	2321
Title of Invention:	METHODS AND APPARATUS FOR DETECTING A VOICE COMMAND
First Named Inventor/Applicant Name:	William F. Ganong
Customer Number:	91337
Filer:	Thomas Stephen Chlebeck/Trish McDonald
Filer Authorized By:	Thomas Stephen Chlebeck
Attorney Docket Number:	N0484.71090US01
Receipt Date:	25-APR-2017
Filing Date:	12-MAR-2013
Time Stamp:	18:25:45
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	CARD
Payment was successfully received in RAM	$2100
RAM confirmation Number	042617INTEFSW18261500
Deposit Account	232825
Authorized User	Wolf Greenfield

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

37 CFR 1.17 (Patent application and reexamination processing fees)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Miscellaneous Incoming Letter	N048471090US01-TRN-TSC.pdf	25051 9e1a86933e1955b3f7037791b2bdc0dd957755fe	no	1
Warnings:					
Information:					
2	Extension of Time	N048471090US01-EXT-TSC.pdf	23332 8f08b1d6e708970c6b98946d671cad94d2e926a1	no	1
Warnings:					
Information:					
3	Fee Worksheet (SB06)	N048471090US01-FEE-TSC.pdf	32169 b093254f82226c8be98db7d073b152e8d6a18679	no	1
Warnings:					
Information:					
4	Request for Continued Examination (RCE)	N048471090US01-RCE-TSC.pdf	27037 920e600ab06897dc4314bb29c2b44a5ca17eb589	no	1
Warnings:					
This is not a USPTO supplied RCE SB30 form.					
Information:					
5	Fee Worksheet (SB06)	fee-info.pdf	32678 b732a58c7c537b07c01f2c8d09311972c729f5b8	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	140267		

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.

Request for Continued Examination (RCE) Transmittal

Address to:
Mail Stop RCE
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Application Number	13/795,933-Conf. #2321
Filing Date	March 12, 2013
First Named Inventor	William F. Ganong, III
Art Unit	2644
Examiner Name	H. C. Ho
Attorney Docket Number	N0484.71090US01

This is a Request for Continued Examination (RCE) under 37 CFR 1.114 of the above-identified application.

Request for Continued Examination (RCE) practice under 37 CFR 1.114 does not apply to any utility or plant application filed prior to June 8, 1995, to any international application that does not comply with the requirements of 35 U.S.C 371, or to any design application. See Instruction Sheet for RCEs (not to be submitted to the USPTO on page 2.)

1. Submission required under 37 CFR 1.114 Note: If the RCE is proper, any previously filed unentered amendments and amendments enclosed with the RCE will be entered in the order in which they were filed unless applicant instructs otherwise. If applicant does not wish to have any previously filed unentered amendment(s) entered, applicant must request non-entry of such amendment(s).
   - a. [x] Previously submitted. If a final Office action is outstanding, any amendments filed after the final Office action may be considered as a submission even if this box is not checked.
     - i. [ ] Consider the arguments in the Appeal Brief or Reply Brief previously filed on _____
     - ii. [x] Other Amendment After Final filed March 27, 2017
   - b. [ ] Enclosed
     - i. [ ] Amendment/Reply
     - ii. [ ] Affidavit(s)/ Declaration(s)
     - iii. [ ] Information Disclosure Statement (IDS)
     - iv. [x] Other Fee Transmittal
2. Miscellaneous
   - a. [ ] Suspension of action on the above-identified application is requested under 37 CFR 1.103(c) for a period of _____ months. (Period of suspension shall not exceed 3 months; Fee under 37 CFR 1.17(i) required)
   - b. [ ] Other _____
3. Fees The RCE fee under 37 CFR 1.17(e) is required by 37 CFR 1.114 when the RCE is filed.
   - a. [ ] The Director is hereby authorized to charge the following fees, any underpayment of fees, or credit any overpayments, to Deposit Account No. 23/2825 .
     - i. [x] RCE fee required under 37 CFR 1.17(e)
     - ii. [x] Extension of time fee (37 CFR 1.136 and 1.17)
     - iii. [ ] Other _____
   - b. [ ] Check in the amount of $ _____ enclosed
   - c. [x] Payment by credit card ~~(Form PTO-2038 enclosed)~~

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

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT REQUIRED

Signature	/Thomas S. Chlebeck/	Date	2017-04-25
Name (Print/Type)	Thomas S. Chlebeck	Registration No.	70,616

~~CERTIFICATE OF MAILING OR TRANSMISSION~~

~~I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to: Mail Stop RCE, Commissioner for Patents, P. O. Box 1450, Alexandria, VA 22313-1450 or facsimile transmitted to the U.S. Patent and Trademark Office on the date shown below.~~

~~Signature~~			
~~Name (Print/Type)~~		~~Date~~	

Certificate of Electronic Filing under 37 CFR §1.8

I hereby certify that this paper (along with any paper referred to as being attached or enclosed) is being transmitted via the Office's electronic filing system in accordance with 37 CFR § 1.6(a)(4).

Dated: April 25, 2017 Electronic Signature: /Trish McDonald/

5401333.1

PTO/AIA/22 (03-13)
Approved for use through 7/31/2016. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PETITION FOR EXTENSION OF TIME UNDER 37 CFR 1.136(a)	Docket Number (Optional) N0484.71090US01
Application Number 13/795,933-Conf. #2321	Filed March 12, 2013
For METHODS AND APPARATUS FOR DETECTING A VOICE COMMAND	
Art Unit 2644	Examiner H. C. Ho

This is a request under the provisions of 37 CFR 1.136(a) to extend the period for filing a reply in the above-identified application.

The requested extension and fee are as follows (check time period desired and enter the appropriate fee below):

		Fee	Small Entity Fee	Micro Entity Fee	
☐	One month (37 CFR 1.17(a)(1))	$200	$100	$50	$
X	Two months (37 CFR 1.17(a)(2))	$600	$300	$150	$ 400.00 ($200 previously paid)
☐	Three months (37 CFR 1.17(a)(3))	$1,400	$700	$350	$
☐	Four months (37 CFR 1.17(a)(4))	$2,200	$1,100	$550	$
☐	Five months (37 CFR 1.17(a)(5))	$3,000	$1,500	$750	$

☐ Applicant asserts small entity status. See 37 CFR 1.27.

☐ Applicant certifies micro entity status. See 37 CFR 1.29.
Form PTO/SB/15A or B or equivalent must either be enclosed or have been submitted previously.

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

X Payment by credit card. ~~Form PTO-2038 is attached.~~

☐ The Director has already been authorized to charge fees in this application to a Deposit Account.

X The Director is hereby authorized to charge any fees which may be required, or credit any overpayment, to Deposit Account Number 23/2825 .

X Payment made via EFS-Web.

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

I am the

☐ applicant.

☐ attorney or agent of record. Registration number _____.

X attorney or agent acting under 37 CFR 1.34. Registration number 70,616 .

/Thomas S. Chlebeck/	2017-04-25
Signature	Date
Thomas S. Chlebeck	617.646.8000
Typed or printed name	Telephone Number

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

☐ * Total of 1 forms are submitted.

Certificate of Electronic Filing under 37 CFR §1.8

I hereby certify that this paper (along with any paper referred to as being attached or enclosed) is being transmitted via the Office's electronic filing system in accordance with 37 CFR § 1.6(a)(4).

Dated: April 25, 2017 Electronic Signature: /Trish McDonald/

5402135.1

UNITED STATES PATENT AND TRADEMARK OFFICE

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

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
13/795,933	03/12/2013	William F. Ganong III	N0484.71090US01	2321

91337 7590 05/09/2017
Nuance Communications, Inc.
c/o Wolf, Greenfield & Sacks, P.C.
600 Atlantic Avenue
Boston, MA 02210-2206

EXAMINER	
HO, HUY C	

ART UNIT	PAPER NUMBER
2644	

NOTIFICATION DATE	DELIVERY MODE
05/09/2017	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):

Patents_eOfficeAction@wolfgreenfield.com
N0484_eOfficeAction@WolfGreenfield.com
IP.Inbox@nuance.com

Examiner-Initiated Interview Summary	Application No.	Applicant(s)	
	13/795,933	GANONG ET AL.	
	Examiner	Art Unit	
	HUY C. HO	2644	

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

(1) *HUY C. HO.* (3)____.

(2) *THOMAS CHLEBECK (Reg. No. 70,616).* (4)____.

Date of Interview: *10 April 2017.*

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

Exhibit shown or demonstration conducted: ☐ Yes ☒ No.
If Yes, brief description: *N/A.*

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

Claim(s) discussed: *1*.

Identification of prior art discussed: *Hung, Oh and Freeman.*

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

New proposed amendments after Final Rejection to claim 1 have shown to overcome the prior art rejection. However since the amendments were made after the Final Rejection, the amended claims are subject to be further searched and considered.

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

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

☐ Attachment

/HUY C HO/ Primary Examiner, Art Unit 2644	

Advisory Action Before the Filing of an Appeal Brief	**Application No.** 13/795,933	**Applicant(s)** GANONG ET AL.	
	Examiner HUY C. HO	**Art Unit** 2644	**AIA (First Inventor to File) Status** No

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

THE REPLY FILED 27 March 2017 FAILS TO PLACE THIS APPLICATION IN CONDITION FOR ALLOWANCE.

NO NOTICE OF APPEAL FILED

1. ☒ The reply was filed after a final rejection. No Notice of Appeal has been filed. To avoid abandonment of this application, applicant must timely file one of the following replies: (1) an amendment, affidavit, or other evidence, which places the application in condition for allowance; (2) a Notice of Appeal (with appeal fee) in compliance with 37 CFR 41.31; or (3) a Request for Continued Examination (RCE) in compliance with 37 CFR 1.114 if this is a utility or plant application. Note that RCEs are not permitted in design applications. The reply must be filed within one of the following time periods:
   a) ☒ The period for reply expires 3 months from the mailing date of the final rejection.
   b) ☐ The period for reply expires on: (1) the mailing date of this Advisory Action; or (2) the date set forth in the final rejection, whichever is later. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of the final rejection.
   c) ☐ A prior Advisory Action was mailed more than 3 months after the mailing date of the final rejection in response to a first after-final reply filed within 2 months of the mailing date of the final rejection. The current period for reply expires months from the mailing date of *the prior Advisory Action* or SIX MONTHS from the mailing date of the final rejection, whichever is earlier.
   Examiner Note: If box 1 is checked, check either box (a), (b) or (c). ONLY CHECK BOX (b) WHEN THIS ADVISORY ACTION IS THE FIRST RESPONSE TO APPLICANT'S FIRST AFTER-FINAL REPLY WHICH WAS FILED WITHIN TWO MONTHS OF THE FINAL REJECTION. ONLY CHECK BOX (c) IN THE LIMITED SITUATION SET FORTH UNDER BOX (c). See MPEP 706.07(f).

Extensions of time may be obtained under 37 CFR 1.136(a). The date on which the petition under 37 CFR 1.136(a) and the appropriate extension fee have been filed is the date for purposes of determining the period of extension and the corresponding amount of the fee. The appropriate extension fee under 37 CFR 1.17(a) is calculated from: (1) the expiration date of the shortened statutory period for reply originally set in the final Office action; or (2) as set forth in (b) or (c) above, if checked. Any reply received by the Office later than three months after the mailing date of the final rejection, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

NOTICE OF APPEAL

2. ☐ The Notice of Appeal was filed on ___. A brief in compliance with 37 CFR 41.37 must be filed within two months of the date of filing the Notice of Appeal (37 CFR 41.37(a)), or any extension thereof (37 CFR 41.37(e)), to avoid dismissal of the appeal. Since a Notice of Appeal has been filed, any reply must be filed within the time period set forth in 37 CFR 41.37(a).

AMENDMENTS

3. ☒ The proposed amendments filed after a final rejection, but prior to the date of filing a brief, will not be entered because
   a) ☒ They raise new issues that would require further consideration and/or search (see NOTE below);
   b) ☐ They raise the issue of new matter (see NOTE below);
   c) ☐ They are not deemed to place the application in better form for appeal by materially reducing or simplifying the issues for appeal; and/or
   d) ☐ They present additional claims without canceling a corresponding number of finally rejected claims.
   NOTE: *See Continuation Sheet.* (See 37 CFR 1.116 and 41.33(a)).
4. ☐ The amendments are not in compliance with 37 CFR 1.121. See attached Notice of Non-Compliant Amendment (PTOL-324).
5. ☐ Applicant's reply has overcome the following rejection(s): _____.
6. ☐ Newly proposed or amended claim(s) _____ would be allowable if submitted in a separate, timely filed amendment canceling the non-allowable claim(s).
7. ☒ For purposes of appeal, the proposed amendment(s): (a) ☒ will not be entered, or (b) ☐ will be entered, and an explanation of how the new or amended claims would be rejected is provided below or appended.

AFFIDAVIT OR OTHER EVIDENCE

8. ☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.
9. ☐ The affidavit or other evidence filed after final action, but before or on the date of filing a Notice of Appeal will not be entered because applicant failed to provide a showing of good and sufficient reasons why the affidavit or other evidence is necessary and was not earlier presented. See 37 CFR 1.116(e).
10. ☐ The affidavit or other evidence filed after the date of filing the Notice of Appeal, but prior to the date of filing a brief, will not be entered because the affidavit or other evidence failed to overcome all rejections under appeal and/or appellant fails to provide a showing of good and sufficient reasons why it is necessary and was not earlier presented. See 37 CFR 41.33(d)(1).
11. ☐ The affidavit or other evidence is entered. An explanation of the status of the claims after entry is below or attached.

REQUEST FOR RECONSIDERATION/OTHER

12. ☐ The request for reconsideration has been considered but does NOT place the application in condition for allowance because: _____.
13. ☐ Note the attached Information *Disclosure Statement*(s). (PTO/SB/08) Paper No(s). _____
14. ☒ Other: PTO-2323, PTO-413B and PTO-892 are included herein.

STATUS OF CLAIMS

15. The status of the claim(s) is (or will be) as follows:
    Claim(s) allowed: .
    Claim(s) objected to: .
    Claim(s) rejected: 1, 2 and 4-20.
    Claim(s) withdrawn from consideration: .

/HUY C HO/
Primary Examiner, Art Unit 2644

Continuation of 3. NOTE: Independent claims 1, 19 and 20 were amended after the Final Rejection, thus it raises new issue and needs further consideration and search. New found references are included herein.

Notice of References Cited	Application/Control No. 13/795,933	Applicant(s)/Patent Under Reexamination GANONG ET AL.	
	Examiner HUY C. HO	Art Unit 2644	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-2012/0034904 A1	02-2012	LeBeau; Michael J.	G10L15/265	455/414.1
*	B	US-2012/0035924 A1	02-2012	JITKOFF; John Nicholas	G01C21/3608	704/235
*	C	US-2010/0121636 A1	05-2010	Burke; Dave	G06F3/0346	704/233
*	D	US-2012/0127072 A1	05-2012	KIM; Hyeran	G10L15/24	345/156
	E	US-				
	F	US-				
	G	US-				
	H	US-				
	I	US-				
	J	US-				
	K	US-				
	L	US-				
	M	US-				

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	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.

Examiner-Initiated Interview Summary	Application No.	Applicant(s)	
	13/795,933	GANONG ET AL.	
	Examiner	Art Unit	
	HUY C. HO	2644	

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

(1) *HUY C. HO.* (3)____.

(2) *THOMAS CHLEBECK (Reg. No. 70,616).* (4)____.

Date of Interview: *10 April 2017.*

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

Exhibit shown or demonstration conducted: ☐ Yes ☒ No.
If Yes, brief description: *N/A.*

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

Claim(s) discussed: *1*.

Identification of prior art discussed: *Hung, Oh and Freeman.*

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

New proposed amendments after Final Rejection to claim 1 have shown to overcome the prior art rejection. However since the amendments were made after the Final Rejection, the amended claims are subject to be further searched and considered.

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

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

☐ Attachment

/HUY C HO/ Primary Examiner, Art Unit 2644	

AFCP 2.0 Decision	Application No.	Applicant(s)
	13/795,933	GANONG ET AL.
	Examiner	Art Unit
	HUY C. HO	2644

This is in response to the After Final Consideration Pilot request filed 03/27/2017.

1. **Improper Request** – The AFCP 2.0 request is improper for the following reason(s) and the after final amendment submitted with the request will be treated under pre-pilot procedure.

   ☐ An AFCP 2.0 request form PTO/SB/434 (or equivalent document) was not submitted.

   ☐ A non-broadening amendment to at least one independent claim was not submitted.

   ☐ A proper AFCP 2.0 request was submitted in response to the most recent final rejection.

   ☐ Other:

2. **Proper Request**

   A. After final amendment submitted with the request will not be treated under AFCP 2.0.
   The after final amendment cannot be reviewed and a search conducted within the guidelines of the pilot program.

   ☐ The after final amendment will be treated under pre-pilot procedure.

   B. Updated search and/or completed additional consideration.
   The examiner performed an updated search and/or completed additional consideration of the after final amendment within the time authorized for the pilot program. The result(s) of the updated search and/or completed additional consideration are:

   ☐ 1. All of the rejections in the most recent final Office action are overcome and a Notice of Allowance is issued herewith.

   ☐ 2. The after final amendment would not overcome all of the rejections in the most recent final Office action. See attached interview summary for further details.

   ☐ 3. The after final amendment was reviewed, and it raises a new issue(s). See attached interview summary for further details.

   ☒ 4. The after final amendment raises new issues, but would overcome all of the rejections in the most recent final Office action. A decision on determining allowability could not be made within the guidelines of the pilot. See attached interview summary for further details, including any newly discovered prior art.

   ☒ 5. Other: Please see the Interview Summary and details from the Advisory Action.

   Examiner Note: Please attach an interview summary when necessary as described above.

Petitioner Exhibit 1004, Page 701 of 1518

04/19/2017

DO NOT ENTER: /H.C.H/

Docket No.: N0484.71090US01
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

First Named Inventor:	William F. Ganong, III
Application No.:	13/795,933
Confirmation No.:	2321
Filed:	March 12, 2013
For:	METHODS AND APPARATUS FOR DETECTING A VOICE COMMAND
Examiner:	H. C. Ho
Art Unit:	2644

Certificate of Electronic Filing under 37 CFR §1.8
I hereby certify that this paper (along with any paper referred to as being attached or enclosed) is being transmitted via the Office's electronic filing system in accordance with 37 CFR § 1.6(a)(4).

Dated: March 27, 2017 Electronic Signature for: /Gail Driscoll/

AMENDMENT AFTER FINAL ACTION UNDER 37 C.F.R. § 1.116

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

Dear Sir:

INTRODUCTORY COMMENTS

In response to the Office Action dated November 25, 2016, finally rejecting claims 1, 2, and 4-20, please amend the above-identified U.S. patent application as follows:

Amendments to the Claims are reflected in the listing of claims which begins on page 2 of this paper.

Remarks/Arguments begin on page 7 of this paper.

DOCKET NO.: N0484.71090US01

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

First Named Inventor:	William F. Ganong
Application No.:	13/795,933
Confirmation No.:	2321
Filed:	March 12, 2013
For:	METHODS AND APPARATUS FOR DETECTING A VOICE COMMAND
Examiner:	Huy C. Ho
Art Unit:	2644

CERTIFICATE OF ELECTRONIC FILING UNDER 37 C.F.R. § 1.8

The undersigned hereby certifies that this paper, along with any paper referred to as being attached or enclosed, is being transmitted via the Office electronic filing system in accordance with § 1.6(a)(4), on the 30th day of May, 2017.

___/Kyle D. Spencer/___
Kyle D. Spencer

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

STATEMENT FILED PURSUANT TO THE DUTY OF DISCLOSURE UNDER 37 C.F.R. §§ 1.56, 1.97 AND 1.98

Sir:

Pursuant to the duty of disclosure under 37 C.F.R. §§ 1.56, 1.97 and 1.98, the undersigned requests consideration of this Information Disclosure Statement.

PART I: Compliance with 37 C.F.R. § 1.97

This Information Disclosure Statement has been filed before the mailing of a first Office Action after the filing of a Request for Continued Examination under 37 C.F.R. § 1.114.

No fee or certification is required.

PART II: Information Cited

The undersigned hereby makes of record in the above-identified application the information listed on the attached form PTO-1449 (modified PTO/SB/08). The order of

Petitioner Exhibit 1004, Page 703 of 1518

presentation of the references should not be construed as an indication of the importance of the references.

PART III: Remarks

Documents cited anywhere in the Information Disclosure Statement are enclosed unless otherwise indicated. It is respectfully requested that:

1. The Examiner consider completely the cited information, along with any other information, in reaching a determination concerning the patentability of the present claims;

2. The enclosed form PTO-1449 (modified PTO/SB/08) be signed by the Examiner to evidence that the cited information has been fully considered by the United States Patent and Trademark Office during the examination of this application;

3. The citations for the information be printed on any patent which issues from this application.

By submitting this Information Disclosure Statement, the undersigned makes no representation that a search has been performed, of the extent of any search performed, or that more relevant information does not exist.

By submitting this Information Disclosure Statement, the undersigned makes no representation that the information cited in the Statement is, or is considered to be, material to patentability as defined in 37 C.F.R. § 1.56(b).

By submitting this Information Disclosure Statement, the undersigned makes no representation that the information cited in the Statement is, or is considered to be, in fact, prior art as defined by 35 U.S.C. § 102.

Notwithstanding any statements by the undersigned, the Examiner is urged to form his or her own conclusion regarding the relevance of the cited information.

Petitioner Exhibit 1004, Page 704 of 1518

An early and favorable action is hereby requested.

The Director is hereby authorized to charge any deficiency or credit any overpayment in the fees occasioned by the filing of this Information Disclosure Statement to our Deposit Account No. 23/2825 under Docket No. N0484.71090US01 from which the undersigned is authorized to draw.

Respectfully submitted,

By: /Thomas S. Chlebeck/
Thomas S. Chlebeck, Reg. No. 70,616
Wolf, Greenfield & Sacks, P.C.
600 Atlantic Avenue
Boston, Massachusetts 02210-2206
Telephone: (617) 646-8000

Docket No.: N0484.71090US01
Date: 2017-05-30
xNDDx

Petitioner Exhibit 1004, Page 705 of 1518

Doc Code: TRAN.LET
Document Description: Transmittal Letter

USED IN LIEU OF PTO/SB/21 (07-09)

TRANSMITTAL FORM

(to be used for all correspondence after initial filing)

Application Number	13/795,933-Conf. #2321
Filing Date	March 12, 2013
First Named Inventor	William F. Ganong
Art Unit	2644
Examiner Name	Huy C. Ho
Attorney Docket Number	N0484.71090US01

Total Number of Pages in This Submission

ENCLOSURES (*Check all that apply*)

- [ ] Fee Transmittal Form
  - [ ] Fee Attached
- [ ] Amendment/Reply
  - [ ] After Final
  - [ ] Affidavits/declaration(s)
- [ ] Extension of Time Request
- [ ] Express Abandonment Request
- [x] Information Disclosure Statement
- [ ] Certified Copy of Priority Document(s)
- [ ] Reply to Missing Parts/ Incomplete Application
  - [ ] Reply to Missing Parts under 37 CFR 1.52 or 1.53
- [ ] Drawing(s)
- [ ] Licensing-related Papers
- [ ] Petition
- [ ] Petition to Convert to a Provisional Application
- [ ] Power of Attorney, Revocation Change of Correspondence Address
- [ ] Terminal Disclaimer
- [ ] Request for Refund
- [ ] CD, Number of CD(s) _____
  - [ ] Landscape Table on CD
- [ ] After Allowance Communication to TC
- [ ] Appeal Communication to Board of Appeals and Interferences
- [ ] Appeal Communication to TC **(Appeal Notice, Brief, Reply Brief)**
- [ ] Proprietary Information
- [ ] Status Letter
- [x] Other Enclosure(s) (please Identify below):

  Form PTO – 1449

Remarks

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT

Firm Name	WOLF, GREENFIELD & SACKS, P.C.		
Signature	/Thomas S. Chlebeck/		
Printed name	Thomas S. Chlebeck		
Date	2017-05-30	Reg. No.	70,616

Certificate of Electronic Filing under 37 CFR §1.8

I hereby certify that this paper (along with any paper referred to as being attached or enclosed) is being transmitted via the Office's electronic filing system in accordance with 37 CFR § 1.6(a)(4).

Dated: May 30, 2017 Signature: ___/Kyle D. Spencer/___ (Kyle D. Spencer)

FORM PTO-1449/A and B (modified PTO/SB/08)	APPLICATION NO.: 13/795,933	ATTY. DOCKET NO.: N0484.71090US01
INFORMATION DISCLOSURE STATEMENT BY APPLICANT	FILING DATE: March 12, 2013	CONFIRMATION NO.: 2321
	FIRST NAMED INVENTOR: William F. Ganong	
Sheet 1 of 1	GROUP ART UNIT: 2644	EXAMINER: Huy C. Ho

U.S. PATENT DOCUMENTS

Examiner's Initials #	Cite No.	U.S. Patent Document Number	U.S. Patent Document Kind Code	Name of Patentee or Applicant of Cited Document	Date of Publication or Issue of Cited Document MM-DD-YYYY
		9,646,610	B2	Macho	05-09-2017
		2015-0340042	A1	Sejnoha et al.	11-26-2015
		2016-0358605	A1	Ganong et al.	12-08-2016

FOREIGN PATENT DOCUMENTS

Examiner's Initials #	Cite No.	Foreign Patent Document Office/ Country	Foreign Patent Document Number	Foreign Patent Document Kind Code	Name of Patentee or Applicant of Cited Document	Date of Publication of Cited Document MM-DD-YYYY	Translation (Y/N)

OTHER ART – NON PATENT LITERATURE DOCUMENTS

Examiner's 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.	Translation (Y/N)

[NOTE – No copies of U.S. patents, published U.S. patent applications, or pending, unpublished patent applications stored in the USPTO's Image File Wrapper (IFW) system, are included. See 37 CFR § 1.98 and 1287OG163. Copies of all other patent(s), publication(s), unpublished, pending U.S. patent applications, or other information listed are provided as required by 37 CFR § 1.98 unless 1) such copies were provided in an IDS in an earlier application that complies with 37 CFR § 1.98, and 2) the earlier application is relied upon for an earlier filing date under 35 U.S.C. § 120.]

EXAMINER:	DATE CONSIDERED:

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

5463454.1

Electronic Acknowledgement Receipt

EFS ID:	29338505
Application Number:	13795933
International Application Number:	
Confirmation Number:	2321
Title of Invention:	METHODS AND APPARATUS FOR DETECTING A VOICE COMMAND
First Named Inventor/Applicant Name:	William F. Ganong
Customer Number:	91337
Filer:	Thomas Stephen Chlebeck/Kyle Spencer
Filer Authorized By:	Thomas Stephen Chlebeck
Attorney Docket Number:	N0484.71090US01
Receipt Date:	30-MAY-2017
Filing Date:	12-MAR-2013
Time Stamp:	13:30:22
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	Miscellaneous Incoming Letter	N048471090US01-TRN-TSC.pdf	24629 75b2a0a5c8dd4a98dfc3359cabca4237e960ca97	no	1

Warnings:

Information:					
2		N048471090US01-IDS-TSC.pdf	28384 e86fef4162a9d4a4e33769a553d566a74abf367e	yes	4

	Multipart Description/PDF files in .zip description		
	Document Description	**Start**	**End**
	Transmittal Letter	1	3
	Information Disclosure Statement (IDS) Form (SB08)	4	4

Warnings:	
Information:	
Total Files Size (in bytes):	53013

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.

Doc Code: TRAN.LET
Document Description: Transmittal Letter

USED IN LIEU OF PTO/SB/21 (07-09)

TRANSMITTAL FORM

(to be used for all correspondence after initial filing)

Application Number	13/795,933-Conf. #2321
Filing Date	March 12, 2013
First Named Inventor	William F. Ganong, III
Art Unit	2644
Examiner Name	Huy C. Ho
Attorney Docket Number	N0484.71090US01

Total Number of Pages in This Submission

ENCLOSURES (*Check all that apply*)

- [ ] Fee Transmittal Form
  - [ ] Fee Attached
- [ ] Amendment/Reply
  - [ ] After Final
  - [ ] Affidavits/declaration(s)
- [ ] Extension of Time Request
- [ ] Express Abandonment Request
- [x] Information Disclosure Statement
- [ ] Certified Copy of Priority Document(s)
- [ ] Reply to Missing Parts/ Incomplete Application
  - [ ] Reply to Missing Parts under 37 CFR 1.52 or 1.53
- [ ] Drawing(s)
- [ ] Licensing-related Papers
- [ ] Petition
- [ ] Petition to Convert to a Provisional Application
- [ ] Power of Attorney, Revocation Change of Correspondence Address
- [ ] Terminal Disclaimer
- [ ] Request for Refund
- [ ] CD, Number of CD(s) _____
  - [ ] Landscape Table on CD
- [ ] After Allowance Communication to TC
- [ ] Appeal Communication to Board of Appeals and Interferences
- [ ] Appeal Communication to TC **(Appeal Notice, Brief, Reply Brief)**
- [ ] Proprietary Information
- [ ] Status Letter
- [x] Other Enclosure(s) (please Identify below):

  Form PTO – 1449

  Copies of Cited References

Remarks

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT

Firm Name	WOLF, GREENFIELD & SACKS, P.C.		
Signature	/Thomas S. Chlebeck/		
Printed name	Thomas S. Chlebeck		
Date	2017-10-05	Reg. No.	70,616

Certificate of Electronic Filing under 37 CFR §1.8

I hereby certify that this paper (along with any paper referred to as being attached or enclosed) is being transmitted via the Office's electronic filing system in accordance with 37 CFR § 1.6(a)(4).

Dated: October 5, 2017 Electronic Signature: /Trish McDonald/

5675627.1

FORM PTO-1449/A and B (modified PTO/SB/08)	APPLICATION NO.: 13/795,933	ATTY. DOCKET NO.: N0484.71090US01
INFORMATION DISCLOSURE STATEMENT BY APPLICANT	FILING DATE: March 12, 2013	CONFIRMATION NO.: 2321
	FIRST NAMED INVENTOR: William F. Ganong	
Sheet 1 of 1	GROUP ART UNIT: 2644	EXAMINER: Huy C. Ho

U.S. PATENT DOCUMENTS

Examiner's Initials #	Cite No.	U.S. Patent Document Number	U.S. Patent Document Kind Code	Name of Patentee or Applicant of Cited Document	Date of Publication or Issue of Cited Document MM-DD-YYYY
		2003-0216909	A1	Davis et al.	11-20-2003

FOREIGN PATENT DOCUMENTS

Examiner's Initials #	Cite No.	Foreign Patent Document Office/ Country	Foreign Patent Document Number	Foreign Patent Document Kind Code	Name of Patentee or Applicant of Cited Document	Date of Publication of Cited Document MM-DD-YYYY	Translation (Y/N)
		CN	105009204	A	Amazon Tech Inc.	10-28-2015	Y-abstract only
		EP	2932500	B1	Amazon Tech Inc.	02-15-2017	

OTHER ART – NON PATENT LITERATURE DOCUMENTS

Examiner's 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.	Translation (Y/N)
		Chinese Office Action and Translation thereof for Chinese Application No. 201480013903.1 dated July 28, 2017 (N0484.71090CN00).	Y

[NOTE – No copies of U.S. patents, published U.S. patent applications, or pending, unpublished patent applications stored in the USPTO's Image File Wrapper (IFW) system, are included. See 37 CFR § 1.98 and 1287OG163. Copies of all other patent(s), publication(s), unpublished, pending U.S. patent applications, or other information listed are provided as required by 37 CFR § 1.98 unless 1) such copies were provided in an IDS in an earlier application that complies with 37 CFR § 1.98, and 2) the earlier application is relied upon for an earlier filing date under 35 U.S.C. § 120.]

EXAMINER:	DATE CONSIDERED:

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

5675612.1

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

(11) EP 2 932 500 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
15.02.2017 Bulletin 2017/07

(21) Application number: **13811335.2**

(22) Date of filing: **09.12.2013**

(51) Int Cl.:
G10L 15/26 *(2006.01)* ***G10L 15/28*** *(2013.01)*
G10L 25/78 *(2013.01)*

(86) International application number:
PCT/US2013/073913

(87) International publication number:
WO 2014/093238 (19.06.2014 Gazette 2014/25)

(54) **SPEECH RECOGNITION POWER MANAGEMENT**

SPRACHERKENNUNGSLEISTUNGSVERWALTUNG

GESTION D'ALIMENTATION PAR RECONNAISSANCE DE LA PAROLE

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(30) Priority: **11.12.2012 US 201213711510**

(43) Date of publication of application:
21.10.2015 Bulletin 2015/43

(73) Proprietor: **Amazon Technologies, Inc.**
Reno, NV 89507 (US)

(72) Inventors:
- **BASYE, Kenneth, John**
  Sutton, MA 01590 (US)
- **SECKER-WALKER, Hugh, Evan**
  Seattle, WA 98109-5210 (US)
- **DAVID, Tony**
  Seattle, WA 98109-5210 (US)
- **KNESER, Reinhard**
  Seattle, WA 98109-5210 (US)
- **ADAMS, Jeffrey, Penrod**
  Seattle, WA 98109-5210 (US)
- **SALVADOR, Stan, Weidner**
  Seattle, WA 98109-5210 (US)
- **KRISHNAMOORTHY, Mahesh**
  Seattle, WA 98109-5210 (US)

(74) Representative: **D Young & Co LLP**
120 Holborn
London EC1N 2DY (GB)

(56) References cited:
US-A- 5 983 186
US-A1- 2003 216 909
US-A1- 2010 277 579
US-A1- 2010 292 987

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Printed by Jouve, 75001 PARIS (FR)

EP 2 932 500 B1

Description

BACKGROUND

[0001] Computing devices may include speech recognition capabilities. For example, a computing device can capture audio input and recognize speech using an acoustic model and a language model. The acoustic model is used to generate hypotheses regarding which sound subword units (e.g., phonemes, etc.) correspond to speech based on the acoustic features of the speech. The language model is used to determine which of the hypotheses generated using the acoustic model is the most likely transcription of the speech based on lexical features of the language in which the speech is spoken. A computing device may also be capable of processing the recognized speech for specific speech recognition applications. For example, finite grammars or natural language processing techniques may be used to process speech. US 2003/216909 discloses (see the Abstract) that a subset of values is used to discriminate voice activity in a signal. The subset of values belongs to a larger set of values representing a segment of a signal, the larger set of values being suitable for speech recognition.

SUMMARY

[0002] The subject matter claimed herein is defined in the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0003] The various aspects and many of the attendant advantages of the present disclosure will become more readily appreciated as the same become better understood by reference to the following detailed description when taken in conjunction with the accompanying drawings, wherein:

FIG. 1 is a schematic diagram depicting an illustrative power management subsystem.

FIG. 2 is a schematic diagram depicting an illustrative user computing device including a power management subsystem.

FIG. 3 is flow diagram depicting an illustrative routine for speech recognition power management which may be implemented by the power management subsystem of FIG. 1.

FIG. 4A, FIG. 4B, and FIG. 4C are state diagrams depicting an illustrative operation of a distributed speech recognition system.

FIG. 5 is a pictorial diagram depicting an illustrative user interface that may be provided by a user computing device that includes a power management subsystem.

DETAILED DESCRIPTION

[0004] In some current approaches to speech recognition, speech recognition capabilities are allocated among one or more computing devices in a distributed computing environment. In a particular example of these approaches, a first computing device may be configured to capture audio input, and may transmit the audio input over a network to a second computing device. The second computing device may perform speech recognition on the audio input and generate a transcription of speech included in the audio input. The transcription of the speech then may be transmitted over the network from the second computing device back to the first computing device. In other current approaches, the first computing device may be configured to capture audio input and transcribe the audio input on its own.

[0005] In these and other current approaches, the first computing device may be configured to remain in a persistently active state. In such a persistently active state, the first computing device may continuously maintain a network connection to the second computing device. The first computing device may also continue to power any hardware used to implement its own speech recognition capabilities. One drawback of these approaches, among others, is that the first computing device may consume unacceptable amounts of energy to maintain the persistently active state. Such energy demands may prove especially problematic for mobile computing devices that rely on battery power. Still other problems are present in current approaches.

[0006] Accordingly, aspects of the present disclosure are directed to power management for speech recognition. A computing device may be provided with a power management subsystem that selectively activates or deactivates one or more modules of the computing device. This activation may be responsive to an audio input that includes one or more pre-designated spoken words, sometimes referred to herein as "keywords." A keyword that prompts the activation of one or more components may be activated is sometimes referred to herein as a "wakeword," while a keyword that prompts the deactivation of one or more components is sometimes referred to herein as a "sleepword." In a particular example, the computing device may include a selectively activated network interface module that, when activated, consumes energy to provide the computing device with connectivity to a second computing device, such as a speech recognition server or other computing device. The power management subsystem may process an audio input to determine that the audio input includes a wakeword, and activate the network interface module in response to determining that the audio input comprises the wakeword. With the network interface module activated, the power management subsystem may cause transmission of the audio input to a speech recognition server for processing.

[0007] The power management subsystem may itself include one or more selectively activated modules. In

Petitioner Exhibit 1004, Page 713 of 1518

some embodiments, one or more of the selectively activated modules are implemented as dedicated hardware (such as an integrated circuit, a digital signal processor or other type of processor) that may be switched from a low-power, deactivated state with relatively lesser functionality, to a high-power, activated state with relatively greater functionality, and vice versa. In other embodiments, one or more modules are implemented as software that includes computer-executable code carried out by one or more general-purpose processors. A software module may be activated (or deactivated) by activating (or deactivating) a general-purpose processor configured to or capable of carrying out the computer-executable code included in the software. In still further embodiments, the power management system includes both one or more hardware modules and one or more software modules.

[0008] The power management subsystem may further include a control module in communication with the one or more selectively activated modules. Such a control module is sometimes referred to herein as a "power management module," and may include any of the hardware or software described above. The power management module may cause the activation or deactivation of a module of the power management subsystem. In some embodiments, the power management module activates or deactivates one or more modules based at least in part on a characteristic of audio input obtained by an audio input module included in the computing device. For example, a module of the power management subsystem may determine one or more values, which values may include, for example, an energy level or volume of the audio input; a score corresponding to a likelihood that speech is present in the audio input; a score corresponding to a likelihood that a keyword is present in the speech; and other values. The module may communicate the one or more values to the power management module, which may either communicate with another module to cause activation thereof or communicate with the module from which the one or more values were received to cause deactivation of that module and/or other modules. In other embodiments, however, a first selectively activated module may communicate directly with a second selectively activated module to cause activation thereof. In such embodiments, no power management module need be present. In still further embodiments, a power management subsystem may be provided with one or more modules, wherein at least some of the one or more modules are in communication with each other but not with the power management module.

[0009] In an example implementation, the power management subsystem may include an audio detection module, which may be configured to determine an energy level or volume of an audio input obtained by the computing device. While the audio detection module may persistently monitor for audio input, the remaining components of the power management subsystem may remain in a low-power, inactive state until activated (either by the power management module or by a different module). If the audio detection module determines that an audio input meets a threshold energy level or volume, a speech detection module may be activated to determine whether the audio input includes speech. If the speech detection module determines that the audio input includes speech, a speech processing module included in the power management subsystem may be activated. The speech processing module may determine whether the speech includes a wakeword, and may optionally classify the speech to determine if a particular user spoke the wakeword. If the speech processing module determines that the speech includes the wakeword, an application processing module may be activated, which application processing module may implement a speech recognition application module stored in memory of the computing device. The speech recognition application may include, for example, an intelligent agent frontend, such as that described in "Intelligent Automated Assistant," which was filed on January 10, 2011 and published as U.S. Publication No. 2012/0016678 on January 19, 2012. The selectively activated network interface module may also be activated as discussed above, and the audio input may be transmitted to a remote computing device for processing. This example implementation is discussed in greater detail below with respect to FIG. 3. Alternately, the power management subsystem may, responsive to detecting the wakeword, activate a processing unit that implements any on-device speech recognition capabilities of the computing device.

[0010] By selectively activating modules of the computing device, the power management subsystem may advantageously improve the energy efficiency of the computing device. The power management subsystem may further improve the energy efficiency of the computing device by selectively activating one or more of its own modules. While such implementations are particularly advantageous for computing devices that rely on battery power, all computing devices for which power management may be desirable can benefit from the principles of the present disclosure.

[0011] Turning now to FIG. 1, an illustrative power management subsystem 100 that may be included in a computing device is shown. The power management subsystem 100 may include an analog/digital converter 102; a memory buffer module 104; an audio detection module 106; a speech detection module 108; a speech processing module 110; an application processing module 112; and a power management module 120. The memory buffer module 104 may be in communication with the audio detection module 106; the speech detection module 108; the speech processing module 110; the application processing module 112; and a network interface module 206. The power management module 120 may likewise be in communication with audio detection module 106; the speech detection module 108; the speech processing module 110; the application processing module 112; and the network interface module 206.

Petitioner Exhibit 1004, Page 714 of 1518

[0012] The analog/digital converter 102 may receive an audio input from an audio input module 208. The audio input module 208 is discussed in further detail below with respect to FIG. 2. The analog/digital converter 102 may be configured to convert analog audio input to digital audio input for processing by the other components of the power management subsystem 100. In embodiments in which the audio input module 208 obtains digital audio input (e.g., the audio input module 208 includes a digital microphone or other digital audio input device), the analog/digital converter 102 may optionally be omitted from the power management subsystem 100. Thus, the audio input module 208 may provide audio input directly to the other modules of the power management subsystem 100.

[0013] The memory buffer module 104 may include one or more memory buffers configured to store digital audio input. The audio input obtained by the audio input module 208 (and, if analog, converted to digital form by the analog/digital converter 102) may be recorded to the memory buffer module 104. The audio input recorded to the memory buffer module 104 may be accessed by other modules of the power management subsystem 100 for processing by those modules, as discussed further herein.

[0014] The one or more memory buffers of the memory buffer module 104 may include hardware memory buffers, software memory buffers, or both. The one or more memory buffers may have the same capacity, or different capacities. A memory buffer of the memory buffer module 104 may be selected to store an audio input depending on which other modules are activated. For example, if only the audio detection module 106 is active, an audio input may be stored to a hardware memory buffer with relatively small capacity. However, if other modules are activated, such as the speech detection module 108; the speech processing module 110; the application processing module 112; and/or the network interface module 206, the audio input may be stored to a software memory buffer of relatively larger capacity. In some embodiments, the memory buffer module 104 includes a ring buffer, in which audio input may be recorded and overwritten in the order that it is obtained by the audio input module 208.

[0015] The audio detection module 106 may process audio input to determine an energy level of the audio input. In some embodiments, the audio detection module 106 includes a low-power digital signal processor (or other type of processor) configured to determine an energy level (such as a volume, intensity, amplitude, etc.) of an obtained audio input and for comparing the energy level of the audio input to an energy level threshold. The energy level threshold may be set according to user input, or may be set automatically by the power management subsystem 100 as further discussed below with respect to FIG. 3. In some embodiments, the audio detection module 106 is further configured to determine that the audio input has an energy level satisfying a threshold for at least a threshold duration of time. In such embodiments, high-energy audio inputs of relatively short duration, which may correspond to sudden noises that are relatively unlikely to include speech, may be ignored and not processed by other components of the power management subsystem 100.

[0016] If the audio detection module 106 determines that the obtained audio input has an energy level satisfying an energy level threshold, it may communicate with the power management module 120 to direct the power management module 120 to activate the speech detection module 108. Alternately, the audio detection module 106 may communicate the energy level to the power management module 120, and the power management module 120 may compare the energy level to the energy level threshold (and optionally to the threshold duration) to determine whether to activate the speech detection module 108. In another alternative, the audio detection module 106 may communicate directly with the speech detection module 108 to activate it. Optionally, the power management module 120 (or audio detection module 106) may direct the audio input module 208 to increase its sampling rate (whether measured in frame rate or bit rate) responsive to the audio detection module 106 determining that the audio input has an energy level satisfying a threshold.

[0017] The speech detection module 108 may process audio input to determine whether the audio input includes speech. In some embodiments, the speech detection module 108 includes a low-power digital signal processor (or other type of processor) configured to implement one or more techniques to determine whether the audio input includes speech. In some embodiments, the speech detection module 108 applies voice activity detection (VAD) techniques. Such techniques may determine whether speech is present in an audio input based on various quantitative aspects of the audio input, such as the spectral slope between one or more frames of the audio input; the energy levels of the audio input in one or more spectral bands; the signal-to-noise ratios of the audio input in one or more spectral bands; or other quantitative aspects. In other embodiments, the speech detection module 108 implements a limited classifier configured to distinguish speech from background noise. The classifier may be implemented by techniques such as linear classifiers, support vector machines, and decision trees. In still other embodiments, the speech detection module 108 applies Hidden Markov Model (HMM) or Gaussian Mixture Model (GMM) techniques to compare the audio input to one or more acoustic models, which acoustic models may include models corresponding to speech, noise (such as environmental noise or background noise), or silence. Still other techniques may be used to determine whether speech is present in the audio input.

[0018] Using any of the techniques described above, the speech detection module 108 may determine a score or a confidence level whose value corresponds to a likelihood that speech is actually present in the audio input (as used herein, "likelihood" may refer to common usage,

Petitioner Exhibit 1004, Page 715 of 1518

whether something is likely, or the usage in statistics). If the score satisfies a threshold, the speech detection module 108 may determine that speech is present in the audio input. However, if the score does not satisfy the threshold, the speech detection module 108 may determine that there is no speech in the audio input.

[0019] The speech detection module 108 may communicate its determination as to whether speech is present in the audio input to the power management module 120. If speech is present in the audio input, the power management module 120 may activate the speech processing module 110 (alternately, the speech detection module 108 may communicate directly with the speech processing module 110). If speech is not present in the audio input, the power management module 120 may deactivate the speech detection module 108. Alternately, the speech detection module 108 may communicate the score to the power management module 120, whereupon the power management module 120 may determine whether to activate the speech processing module 110 or deactivate the speech detection module 108.

[0020] The speech processing module 110 may process the audio input to determine whether a keyword is included in the speech. In some embodiments, the speech processing module 110 includes a microprocessor configured to detect a keyword in the speech, such as a wakeword or sleepword. The speech processing module 110 may be configured to detect the keyword using HMM techniques, GMM techniques, or other speech recognition techniques.

[0021] The speech processing module 110 may be able to separate speech that incidentally includes a keyword from a deliberate utterance of the keyword by determining whether the keyword was spoken immediately before or after one or more other phonemes or words. For example, if the keyword is "ten," the speech processing module 110 may be able to distinguish the user saying "ten" by itself from the user saying "ten" incidentally as part of the word "Tennessee," the word "forgotten," the word "stent," or the phrase "ten bucks."

[0022] The speech processing module 110 may further be configured to determine whether the speech is associated with a particular user of a computing device in which the power management subsystem 100 is included, or whether the speech corresponds to background noise; audio from a television; music; or the speech of a person other than the user, among other classifications. This functionality may be implemented by techniques such as linear classifiers, support vector machines, and decision trees, among other techniques for classifying audio input.

[0023] Using any of the techniques described above, the speech processing module 110 may determine a score or confidence level whose value corresponds to a likelihood that a keyword is actually present in the speech. If the score satisfies a threshold, the speech processing module 110 may determine that the keyword is present in the speech. However, if the score does not satisfy the threshold, the speech processing module 110 may determine that there is no keyword in the speech.

[0024] The speech processing module 110 may communicate its determination as to whether a keyword is present in the speech to the power management module 120. If the keyword is present in the speech and the keyword is a wakeword, the power management module 120 may activate application processing module 112 and the network interface module 206 (alternately, the speech processing module 110 may communicate directly with these other modules). If the keyword is not present in the audio input (or the keyword is a sleepword), the power management module 120 may deactivate the speech processing module 110 and the speech detection module 108. Alternately, the speech processing module 110 may communicate the score to the power management module 120, whereupon the power management module 120 may determine whether to activate the application processing module 112 and the network interface module 206 or deactivate the speech processing module 110 and the speech detection module 108. In some embodiments, these activations and/or deactivations only occur if the speech processing module 110 determines that a particular user spoke the speech that includes the keyword.

[0025] The application processing module 112 may include a microprocessor configured to implement a speech recognition application provided with a computing device in which the power management subsystem is included. The speech recognition application may include any application for which speech recognition may be desirable, such as a dictation application, a messaging application, an intelligent agent frontend application, or any other application. The speech recognition application may also be configured to format the speech (e.g., by compressing the speech) for transmission over a network to a remote computing device, such as a speech recognition server.

[0026] In some embodiments, the application processing module 112 includes a dedicated microprocessor for implementing the speech recognition application. In other embodiments, the application processing module 112 includes a general-purpose microprocessor that may also implement other software provided with a computing device in which the power management subsystem 100 is included, such as the processing unit 202 shown in FIG. 2, which is discussed further below.

[0027] The network interface module 206, when activated, may provide connectivity over one or more wired or wireless networks. Upon its activation, the network interface module 206 may transmit the received audio input recorded to the memory buffer module 104 over a network to a remote computing device, such as a speech recognition server. The remote computing device may return recognition results (e.g., a transcription or response to an intelligent agent query) to the computing device in which the network interface module 206 is included, whereupon the network interface module 206 may provide the received recognition results to the ap-

Petitioner Exhibit 1004, Page 716 of 1518

plication processing module 112 for processing. The network interface module 206 is discussed further below with respect to FIG. 2.

[0028] The modules of the power management subsystem 100 may be combined or rearranged without departing from the scope of the present disclosure. The functionality of any module described above may be allocated among multiple modules, or combined with a different module. As discussed above, any or all of the modules may be embodied in one or more integrated circuits, one or more general-purpose microprocessors, or in one or more special-purpose digital signal processors or other dedicated microprocessing hardware. One or more modules may also be embodied in software implemented by a processing unit 202 included in a computing device, as discussed further below with respect to FIG. 2. Further, one or more of the modules may be omitted from the power management subsystem 100 entirely.

[0029] Turning now to FIG. 2, a user computing device 200 in which a power management subsystem 100 may be included is illustrated. The user computing device 200 includes a processing unit 202; a non-transitory computer-readable medium drive 204; a network interface module 206; the power management subsystem 100 as shown in FIG. 1; and an audio input module 208, all of which may communicate with one another by way of a communication bus. The user computing device 200 may also include a power supply 218, which may provide power to the various components of the user computing device 200, such as the processing unit 202; the non-transitory computer-readable medium drive 204; the network interface module 206; the power management subsystem 100 as shown in FIG. 1; and the audio input module 208.

[0030] The processing unit 202 may include one or more general-purpose microprocessors configured to communicate to and from the memory 210 to implement various software modules stored therein, such as a user interface module 212, operating system 214, and speech recognition application module 216. The processing unit 202 may also communicate with the power management subsystem 100 and may further implement any modules of the power management subsystem 100 embodied in software. Accordingly, the processing unit 202 may be configured to implement any or all of the audio detection module 106; the speech detection module 108; the speech processing module 110; the application processing module 112; and the power management module 120. Further, the processing unit 202 may be configured to implement on-device automatic speech recognition capabilities that may be provided with the user computing device 200.

[0031] The memory 210 generally includes RAM, ROM, and/or other persistent or non-transitory computer-readable storage media. The user interface module 212 may be configured to present a user interface via a display of the user computing device 200 (not shown). The user interface module 212 may be further configured to process user input received via a user input device (not shown), such as a mouse, keyboard, touchscreen, keypad, etc. The user interface presented by the user interface module 212 may provide a user with the opportunity to customize the operation of the power management subsystem 100 and/or other operations implemented by the user computing device 200. An example of a user interface is discussed further below with respect to FIG. 5. The memory 210 may additionally store an operating system 214 that provides computer program instructions for use by the processing unit 202 in the general administration and operation of the user computing device 200. The memory 210 can further include computer program instructions that the application processing module 112 and/or processing unit 202 executes in order to implement one or more embodiments of a speech recognition application module 216. As discussed above, the speech recognition application module 216 may be any application that may use speech recognition results, such as a dictation application; a messaging application; an intelligent agent application frontend; or any other application that may advantageously use speech recognition results. In some embodiments, the memory 210 may further include an automatic speech recognition engine (not shown) that may be implemented by the processing unit 202.

[0032] The non-transitory computer-readable medium drive 204 may include any electronic data storage known in the art. In some embodiments, the non-transitory computer-readable medium drive 204 stores one or more keyword models (e.g., wakeword models or sleepword models) to which an audio input may be compared by the power management subsystem 100. The non-transitory computer-readable medium drive 204 may also store one or more acoustic models and/or language models for implementing any on-device speech recognition capabilities of the user computing device 200. Further information regarding language models and acoustic models may be found in U.S. Patent App. No. 13/587,799, entitled "DISCRIMINATIVE LANGUAGE MODEL PRUNING," filed on August 16, 2012; and in U.S. Patent App. No. 13/592,157, entitled "UNSUPERVISED ACOUSTIC MODEL TRAINING," filed on August 22, 2012.

[0033] The network interface module 206 may provide the user computing device 200 with connectivity to one or more networks, such as a network 410, discussed further below with respect to FIG: 4A, FIG. 4B, and FIG. 4C. The processing unit 202 and the power management subsystem 100 may thus receive instructions and information from remote computing devices that may also communicate via the network 410, such as a speech recognition server 420, as also discussed further below. In some embodiments, the network interface module 206 comprises a wireless network interface that provides the user computing device 200 with connectivity over one or more wireless networks.

[0034] In some embodiments, the network interface

Petitioner Exhibit 1004, Page 717 of 1518

module 206 is selectively activated. While the network interface module 206 is in a deactivated or "sleeping" state, it may provide limited or no connectivity to networks or computing systems so as to conserve power. In some embodiments, the network interface module 206 is in a deactivated state by default, and becomes activated responsive to a signal from the power management subsystem 100. While the network interface module 206 in an activated state, it may provide a relatively greater amount of connectivity to networks or computing systems, such that the network interface module 206 enables the user computing device 200 to send audio input to a remote computing device and/or receive a keyword confirmation, speech recognition result, or deactivation instruction from the remote computing device, such as a speech recognition server 420.

[0035] In a particular, non-limiting example, the network interface module 206 may be activated responsive to the power management subsystem 100 determining that an audio input includes a wakeword. The power management subsystem 100 may cause transmission of the audio input to a remote computing device (such as a speech recognition server 420) via the activated network interface module 206. Optionally, the power management subsystem 100 may obtain a confirmation of a wakeword from a remote computing device before causing the transmission of subsequently received audio inputs to the remote computing device. The power management subsystem 100 may later deactivate the activated network interface module 206 in response to receiving a deactivation instruction from the remote computing device, in response to determining that at least a predetermined amount of time has passed since an audio input satisfying an energy level threshold has been obtained, or in response to receiving an audio input that includes a sleepword.

[0036] The audio input module 208 may include an audio input device, such as a microphone or array of microphones, whether analog or digital. The microphone or array of microphones may be implemented as a directional microphone or directional array of microphones. In some embodiments, the audio input module 208 receives audio and provides the audio to the power management subsystem 100 for processing, substantially as discussed above with respect to FIG. 1. The audio input module 208 may also receive instructions from the power management subsystem 100 to set a sampling rate (whether in frame rate or bitrate) for obtaining audio. The audio input module 208 may also (or instead) include one or more piezoelectric elements and/or micro-electrical-mechanical systems (MEMS) that can convert acoustic energy to an electrical signal for processing by the power management subsystem 100. The audio input module 208 may further be provided with amplifiers, rectifiers, and other audio processing components as desired.

[0037] One or more additional input devices such as light sensors, position sensors, image capture devices, or the like may be provided with the user computing device 200. Such additional input devices are not shown in FIG. 2 so as not to obscure the principles of the present disclosure. In some embodiments, an additional input device may detect the occurrence or non-occurrence of a condition. Information pertaining to such conditions may be provided to the power management subsystem 100 to determine whether one or more components of the user computing device 200 or the power management subsystem 100 should be activated or deactivated. In one embodiment, the additional input device includes a light sensor configured to detect a light level. The power management module 120 may only act network interface module 206 may only be activated if the light level detected by the light sensor does not satisfy a threshold. In another embodiment, the additional input device includes an image capture device configured with facial recognition capabilities. In this embodiment, the network interface module 206 may only be activated if the image capture device recognizes the face of a user associated with the user computing device 200. More information on controlling speech recognition capabilities with input devices may be found in U.S. Patent App. No. 10/058,730, entitled "AUTOMATIC SPEECH RECOGNITION SYSTEM AND METHOD," filed on January 30, 2002, which published as U.S. Patent Pub. No. 2003/0144844 on July 31, 2003. Further information on controlling speech recognition capabilities may be found in U.S. Patent No. 8,326,636, entitled "USING A PHYSICAL PHENOMENON DETECTOR TO CONTROL OPERATION OF A SPEECH RECOGNITION ENGINE," which issued on December 4, 2012.

[0038] Still further input devices may be provided, which may include user input devices such as mice, keyboards, touchscreens, keypads, etc. Likewise, output devices such as displays, speakers, headphones, etc. may be provided. In a particular example, one or more output devices configured to present speech recognition results in audio format (e.g., via text-to-speech) or in visual format (e.g., via a display) may be included with the user computing device 200. Such input and output devices are well known in the art and need not be discussed in further detail herein, and are not shown in FIG. 2 so as to avoid obscuring the principles of the present disclosure.

[0039] The power supply 218 may provide power to the various components of the user computing device 200. The power supply 218 may include a wireless or portable power supply, such as a disposable or rechargeable battery or battery pack; or may include a wired power supply, such as an alternating current (AC) power supply configured to be plugged into an electrical outlet. In some embodiments, the power supply 218 communicates the level of power that it can supply to the power management subsystem 100 (e.g., a percentage of battery life remaining, whether the power supply 218 is plugged into an electrical outlet, etc.). In some embodiments, the power management subsystem 100 selectively activates or deactivates one or more modules based at least in part

Petitioner Exhibit 1004, Page 718 of 1518

on the power level indicated by the power supply. For example, if the user computing device 200 is plugged in to an electrical outlet, the power management subsystem 100 may activate the network interface module 206 and leave it in an activated state. If the user computing device 200 is running on battery power, the power management subsystem 100 may selectively activate and deactivate the network interface module 206 as discussed above.

[0040] Turning now to FIG. 3, an illustrative routine 300 is shown in which modules of the power management subsystem 100 may be selectively activated for processing an audio input. The illustrative routine 300 represents an escalation of processing and/or power consumption, as modules that are activated later in the illustrative routine 300 may have relatively greater processing requirements and/or power consumption.

[0041] The illustrative routine 300 may begin at block 302 as the audio input module 208 monitors for audio input. The audio input module 208 may receive an audio input at block 304. At block 306, the received audio input may be recorded to the memory buffer module 104. At block 308, the audio detection module 106 may determine whether the audio input has an energy level that satisfies an energy level threshold (and, optionally, whether the audio input has an energy level that satisfies an energy level threshold for at least a threshold duration). If the audio input's energy level does not satisfy the energy level threshold, the audio input module 208 may continue to monitor for audio input in block 310 until another audio input is received.

[0042] Returning to block 308, if the audio detection module 106 determines that the audio input has an energy level satisfying a threshold, the power management module 120 may activate the speech detection module 108 at block 312 (alternately, the audio detection module 106 may directly activate the speech detection module 108, and the power management module 120 may be omitted as well in the following blocks). At block 314, the speech detection module 108 may determine whether speech is present in the obtained audio input, substantially as discussed above with respect to FIG. 1. If the speech detection module 108 determines that speech is not present (or not likely to be present) in the audio input, the power management module 120 may deactivate the speech detection module 108 at block 316. The audio input module 208 may then continue to monitor for audio input in block 310 until another audio input is received.

[0043] Returning to block 314, if the speech detection module 108 determines that the audio input includes speech, the power management module 120 may activate the speech processing module 110 at block 318. As discussed above, the speech processing module 110 may determine whether a wakeword is present in the speech at block 320. If the speech processing module 110 determines that the wakeword is not present in the speech (or not likely to be present in the speech), the speech processing module 110 may be deactivated at block 322. The speech detection module 108 may also be deactivated at block 316. The audio input device 208 may then continue to monitor for audio input in block 310 until another audio input is received.

[0044] Returning to block 320, if in some embodiments, the speech processing module 110 determines that the wakeword is present in the speech, user 401 the speech processing module 110 optionally determines in block 324 whether the speech is associated with a particular user (e.g., whether the wakeword was spoken by the user), substantially as discussed above with respect to FIG. 1. If the speech is not associated with the particular user, the speech processing module 110 may be deactivated at block 322. The speech detection module 108 may also be deactivated at block 316. The audio input device 208 may then continue to monitor for audio input in block 310 until another audio input is received. If the speech is associated with the particular user, the illustrative routine 300 may proceed to block 326. In other embodiments, block 324 may be omitted, and the illustrative routine 300 may proceed directly from block 320 to block 326 responsive to the speech processing module 110 determining that a wakeword is present in the speech.

[0045] At block 326, the power management module 120 may activate the application processing module 112, which may implement the speech recognition application module 216 shown in FIG. 2. The power management module 120 may also activate the network interface module 206 at block 328. With the network interface module 206 activated, the audio input recorded to the memory buffer module 104 may be transmitted over a network via the network interface module 206. In some embodiments, while the network interface module 206 is active, subsequently obtained audio inputs are provided from the audio input module 208 directly to the application processing module 112 and/or the network interface module 206 for transmission to the remote computing device. However, in other embodiments, any or all of the speech detection module 108, speech processing module 110, and application processing module 112 process the audio input before providing it to the network interface module 206 to be transmitted over the network 410 to a remote computing device.

[0046] In some embodiments, not shown, the power management subsystem 100 waits until the remote computing device returns a confirmation that the wakeword is present in the first audio input to transmit subsequent audio inputs for recognition. If no confirmation of the wakeword is provided by the remote computing device, or if a deactivation instruction is received via the network interface module 206, the network interface module 206 and one or more modules of the power management subsystem 100 may be deactivated.

[0047] As many of the operations of the power management subsystem 100 may generate probabilistic rather than exact determinations, errors may occur during the illustrative routine 300. In some instances, a particular module of the power management subsystem 100 may

Petitioner Exhibit 1004, Page 719 of 1518

provide a "false positive," causing one or more modules to be incorrectly activated. For example, the speech detection module 108 may incorrectly determine that speech is present at block 314, or the speech processing module 110 may incorrectly determine that the speech includes the wakeword at block 320or that the speech belongs to the user at block 324. Adaptive thresholding and cross-validation among the modules of the power management subsystem 100 may be advantageously used to reduce false positives. Two examples of adaptive thresholding are discussed herein, but other types of adaptive thresholding are possible. As discussed above, the speech detection module may determine that speech is present in an audio input at block 314. However, the speech processing module 110, which may recognize speech more precisely than the speech detection module 108 owing to its superior processing power, may determine that in fact no speech is present in the audio input. Accordingly, the speech processing module 110 may direct the speech detection module 108 to increase its score threshold for determining that speech is present in the audio input, so as to reduce future false positives. Likewise, if the remote computing device (such as a speech recognition server 420) includes speech recognition capabilities, the remote computing device may transmit to the user computing device 200 an indication that no wakeword was present in the speech, even though the speech processing module 110 may have indicated that the wakeword was present. Accordingly, the score threshold of the speech processing module 110 for determining that the wakeword is present in the speech may be increase, so as to reduce future false positives. Further, a user interface may be provided so that a user may increase one or more score thresholds to reduce false positives, as further described below with respect to FIG. 5.

[0048] In other instances, a particular component may provide a "false negative," such that components of the power management subsystem 100 are not activated and/or the network interface module 206 is not activated, even though the user has spoken the wakeword. For example, the speech detection module 108 may incorrectly determine that no speech is present at block 314, or the speech processing module 110 may incorrectly determine that the speech does not include the wakeword at block 320 or that the speech does not belong to the user at block 324. To reduce the likelihood of false negatives, the power management subsystem 100 may periodically lower the threshold scores, e.g., lower the score required to satisfy the thresholds in blocks 314, 320, and/or 324. The threshold may continue to be lowered until one or more false positives are obtained, as described above. Once one or more false positives are obtained, the threshold may not be lowered further, or may be slightly increased. Further, a user interface may accordingly be provided so that a user may decrease one or more score thresholds to reduce false negatives, as further described below with respect to FIG. 5.

[0049] In some embodiments, not all activated components are deactivated if a negative result is obtained at any of blocks 314, 320, or 324. For example, if a wakeword is not recognized at block 320, the speech processing module 110 may be deactivated at block 322, but the speech detection module 108 may remain activated. Additionally, blocks may be skipped in some implementations. In some embodiments, a score satisfying a threshold at either of blocks 314 or 320 prompts one or more subsequent blocks to be skipped. For example, if the speech processing module 110 determines with very high confidence that the wakeword is present in the speech at block 320, the illustrative routine 300 may skip directly to block 326

[0050] Further, in some embodiments, the user computing device 200 may include an automatic speech recognition engine configured to be executed by the processing unit 202. As such on-device speech recognition may have especially high power consumption, the processing unit 202 may only implement the automatic speech recognition engine to recognize speech responsive to the speech processing module 110 determining that the wakeword has been spoken by a user.

[0051] With reference now to FIG. 4A, FIG. 4B, and FIG. 4C, example operations of a distributed speech recognition service are shown in the illustrative environment 400. The environment 400 may include a user 401; a user computing device 200 as described above; a network 410; a speech recognition server 420; and a data store 430.

[0052] The network 410 may be any wired network, wireless network or combination thereof. In addition, the network 410 may be a personal area network, local area network, wide area network, cable network, satellite network, cellular telephone network, or combination thereof. Protocols and devices for communicating via the Internet or any of the other aforementioned types of communication networks are well known to those skilled in the art of computer communications, and thus need not be described in more detail herein.

[0053] The speech recognition server 420 may generally be any computing device capable of communicating over the network 410. In some embodiments, the speech recognition server 420 is implemented as one or more server computing devices, though other implementations are possible. The speech recognition server 420 may be capable of receiving audio input over the network 410 from the user computing device 200. This audio input may be processed in a number of ways, depending on the implementation of the speech recognition server 420. In some embodiments, the speech recognition server 420 processes the audio input received from the user computing device 200 to confirm that a wakeword is present (e.g., by comparing the audio input to a known model of the wakeword), and transmits the confirmation to the user computing device 200. The speech recognition server 420 may further be configured to identify a user 401 that spoke the wakeword using known speaker

Petitioner Exhibit 1004, Page 720 of 1518

identification techniques.

[0054] The speech recognition server 420 may process the audio input received from the user computing device 200 to determine speech recognition results from the audio input. For example, the audio input may include a spoken query for an intelligent agent to process; speech to be transcribed to text; or other speech suitable for a speech recognition application. The speech recognition server 420 may transmit the speech recognition results over the network 410 to the user computing device 200. Further information pertaining to distributed speech recognition applications may be found in U.S. Patent No. 8,117,268, entitled "Hosted voice recognition system for wireless devices" and issued on February 14, 2012, the disclosure of which is hereby incorporated by reference in its entirety.

[0055] The speech recognition server 420 may be in communication either locally or remotely with a data store 430. The data store 430 may be embodied in hard disk drives, solid state memories, and/or any other type of non-transitory, computer-readable storage medium accessible to the speech recognition server 420. The data store 430 may also be distributed or partitioned across multiple storage devices as is known in the art without departing from the spirit and scope of the present disclosure. Further, in some embodiments, the data store 430 is implemented as a network-based electronic storage service.

[0056] The data store 430 may include one or more models of wakewords. In some embodiments, a wakeword model is specific to a user 401, while in other embodiments, the Upon receiving an audio input determined by the user computing device 200 to include a wakeword, the speech recognition server may compare the audio input to a known model of the wakeword stored in the data store 430. If the audio input is sufficiently similar to the known model, the speech recognition server 420 may transmit a confirmation of the wakeword to the user computing device 200, whereupon the user computing device 200 may obtain further audio input to be processed by the speech recognition server 420.

[0057] The data store 430 may also include one or more acoustic and/or language models for use in speech recognition. These models may include general-purpose models as well as specific models. Models may be specific to a user 401; to a speech recognition application implemented by the user computing device 200 and/or the speech recognition server 420; or may have other specific purposes. Further information regarding language models and acoustic models may be found in U.S. Patent App. No. 13/587,799, entitled "DISCRIMINATIVE LANGUAGE MODEL PRUNING," filed on August 16, 2012; and in U.S. Patent App. No. 13/592,157, entitled "UNSUPERVISED ACOUSTIC MODEL TRAINING," filed on August 22, 2012.

[0058] The data store 430 may further include data that is responsive to a query contained in audio input received by the speech recognition server 420. The speech recognition server 420 may recognize speech included in the audio input, identify a query included in the speech, and process the query to identify responsive data in the data store 430. The speech recognition server 420 may then provide an intelligent agent response including the responsive data to the user computing device 200 via the network 410. Still further data may be included in the data store 430.

[0059] It will be recognized that many of the devices described above are optional and that embodiments of the environment 400 may or may not combine devices. Furthermore, devices need not be distinct or discrete. Devices may also be reorganized in the environment 400. For example, the speech recognition server 420 may be represented as a single physical server computing device, or, alternatively, may be split into multiple physical servers that achieve the functionality described herein. Further, the user computing device 200 may have some or all of the speech recognition functionality of the speech recognition server 420.

[0060] Additionally, it should be noted that in some embodiments, the user computing device 200 and/or speech recognition server 420 may be executed by one more virtual machines implemented in a hosted computing environment. The hosted computing environment may include one or more rapidly provisioned and released computing resources, which computing resources may include computing, networking and/or storage devices. A hosted computing environment may also be referred to as a cloud computing environment. One or more of the computing devices of the hosted computing environment may include a power management subsystem 100 as discussed above.

[0061] With specific reference to FIG. 4A, an illustrative operation by which a wakeword may be confirmed is shown. The user 401 may speak the wakeword 502. The user computing device 200 may obtain the audio input that may include the user's speech (1) and determine that the wakeword 402 is present in the speech (2), substantially as discussed above with respect to FIG. 3. The audio input may also include a voice command or query. Responsive to determining that the speech includes the wakeword, the application processing module 112 and the network interface module 206 of the user computing device 200 may be activated (3) and the audio input transmitted (4) over the network 410 to the speech recognition server 420. The speech recognition server 420 may confirm (5) that the wakeword is present in the audio input, and may transmit (6) a confirmation to the user computing device 200 over the network 410.

[0062] Turning now to FIG. 4B, responsive to receiving the confirmation of the wakeword from the speech recognition server 420, the user computing device 200 may continue to obtain audio input (7) to be provided to the speech recognition server 420 for processing. For example, the obtained audio input may include an intelligent agent query 404 for processing by the speech recognition server 420. Alternately, the obtained audio input may in-

Petitioner Exhibit 1004, Page 721 of 1518

clude speech to be transcribed by the speech recognition server 420 (e.g., for use with a dictation, word processing, or messaging application executed by the application processing module 112). The user computing device 200 may transmit the audio input (8) over the network 410 to the speech recognition server 420. Optionally, an identifier of the speech recognition application for which speech recognition results are to be generated may be provided to the speech recognition server 420, so that the speech recognition server 420 may generate results specifically for use with the speech recognition application implemented by the application processing module 112. The speech recognition server 420 may recognize speech (9) included in the audio input and generate speech recognition results (10) therefrom. The speech recognition results may include, for example, a transcription of the speech, an intelligent agent response to a query included in the speech, or any other type of result. These speech recognition results may be transmitted (11) from the speech recognition server 420 to the user computing device 200 over the network 410. In response to receiving the results, the application processing module 112 may cause presentation of the results (12) in audible format (e.g., via text-to-speech) or in visual format (e.g., via a display of the user computing device 200).

[0063] With reference now to FIG. 4C, the user computing device 200 may continue to obtain audio input (13) to be provided to the speech recognition server 420 for processing. The user computing device 200 may transmit the audio input (14) over the network 410 to the speech recognition server 420. The speech recognition server may recognize any speech (15) included in the audio input. Responsive to recognizing the speech, the speech recognition server 420 may determine that the user 401 is no longer speaking to the user computing device 200 and stop (16) any subsequent speech recognition. For example, the user 401 may speak words that do not correspond to a structured command or query, such as undirected natural language speech 406. The speech recognition server 420 may also analyze the speech's speed, carefulness, inflection, or clarity to determine that the speech is not directed to the user computing device 200 and should not be processed into speech recognition results.

[0064] Other types of audio inputs may also prompt the speech recognition server 420 to stop subsequent speech recognition. Alternately, the speech recognition server 420 may determine that the received audio input does not include speech. Responsive to receiving one or more audio inputs that do not include speech directed to the user computing device 200, the speech recognition server 420 may determine that speech recognition results should not be generated and that the speech recognition should stop. Further, the audio input may include a predetermined sleepword, which may be selected by the user 401. If the speech recognition server 420 detects the sleepword, the speech recognition server 420 may stop performing speech recognition on the audio input. Further, the speech recognition server 420 may determine that multiple users 401 are present in the vicinity of the user computing device 200 (e.g., by performing speaker identification on multiple audio inputs obtained by the user computing device 200). If the number of identified users 401 satisfies a threshold (which may be any number of users 401 greater than one), the speech recognition server 420 may determine that any audio inputs obtained by the user computing device 200 are not likely intended to be processed into speech recognition results.

[0065] Responsive to determining that the speech of the user 401 is not directed to the user computing device 200 (or determining that subsequent speech recognition should not be performed for any of the other reasons discussed above), the speech recognition server 420 may transmit a deactivation instruction (17) over the network 410 to the user computing device 200. In response to receiving the deactivation instruction, the user computing device 200 may deactivate (18) its network interface module 206 and one or more components of the power management subsystem 100, such as the application processing module 112, the speech processing module 110, and/or the speech detection module 108. Other conditions may also prompt the speech recognition server 420 to transmit the deactivation instruction to the user computing device 200. For example, returning to FIG. 4A, if the speech recognition server 420 determines that a wakeword is not present in the audio input received at state (1), the speech recognition server 420 may transmit a deactivation instruction to the user computing device 200. Alternately, the speech recognition server 420 may determine that a threshold amount of time has passed since it last received an audio input including speech from the user computing device 200, and may accordingly transmit a deactivation instruction to the user computing device 200. Still other criteria may be determined for transmitting a deactivation instruction to the user computing device 200.

[0066] Returning again to FIG. 4A, upon receiving a subsequent audio input determined to include a wakeword, the user computing device 200 may activate the components of the power management subsystem 100 and the network interface module 206 and transmit the audio input to the speech recognition server 420. The example operations shown herein may thus repeat themselves.

[0067] The example operations depicted in FIG. 4A, FIG. 4B, and FIG. 4C are provided for illustrative purposes. One or more states may be omitted from the example operations shown herein, or additional states may be added. In a particular example, the user computing device 200 need not obtain a confirmation of the wakeword from the speech recognition server 420 before transmitting an audio input for which speech recognition results are to be generated by the speech recognition server 420. Additionally, the user computing device 200 need not obtain a deactivation instruction before deactivating its network interface module 206 and/or one or more of

Petitioner Exhibit 1004, Page 722 of 1518

the components of its power management subsystem 100, such as the application processing module 112, speech processing module 110, or speech detection module 108. Rather, power management subsystem 100 may determine (via the audio detection module 106) that at least a threshold amount of time has passed since an audio input having an energy level satisfying an energy level threshold has been obtained by the user computing device 200. Alternately, the user computing device 200 may determine (via the speech detection module 108) that at least a threshold amount of time has passed since an audio input that includes speech has been obtained. Responsive to determining that the threshold amount of time has passed, the power management subsystem 100 may cause deactivation of the network interface module 206, and may deactivate one or more of its own components as described above with respect to FIG. 3.

[0068] Further, the power management subsystem 100 may be configured to recognize a sleepword selected and spoken by the user 401, in a manner substantially similar to how the wakeword is identified in FIG. 3. If the sleepword is detected by the power management subsystem 100 (e.g., by the speech processing module 110), the network interface module 206 and/or one or more of the components of the power management subsystem 100 may be deactivated. Likewise, if the user computing device 200 includes its own on-device speech recognition capabilities, they may be deactivated responsive to the sleepword being detected.

[0069] FIG. 5 depicts an illustrative user interface 500 that may be provided by a user computing device 200 for customizing operations of the power management subsystem 100 and of the user computing device 200. In one embodiment, the user interface module 212 processes user input made via the user interface 500 and provides it to the power management subsystem 100.

[0070] The energy level threshold element 502 may enable a user to specify a threshold energy level at which the speech detection module 108 should be activated, as shown in block 308 of FIG. 3. For example, if the user computing device 200 is in a relatively noisy environment or if the user computing device 200 is experiencing a significant number of "false positives" determined by the audio detection module 106, the user 401 may wish to increase the energy level threshold at which the speech processing module 108 is activated. If the user 401 is in a relatively quiet environment or if the user computing device 200 is experiencing a significant number of false negatives, the user 401 may wish to decrease the energy level threshold at which the speech detection module 108 is activated. As discussed above, the energy level threshold may correspond to a volume threshold, intensity threshold, amplitude threshold, or other threshold related to the audio input.

[0071] The keyword confidence threshold element 504 may enable a user to specify a threshold score at which the speech processing module 110 determines that a keyword is present. Likewise, the identification confidence threshold element may enable a user to specify a threshold score at which the speech processing module 110 determines that the user spoke the keyword. In one embodiment, the application processing module 112 and the network interface module 206 are activated responsive to the speech processing module 110 recognizing a wakeword (e.g., the speech processing module 110 determining a score that satisfies a threshold, which score corresponds to a likelihood that the wakeword is included in the speech).. In another embodiment, the application processing module 112 and the network interface module 206 are activated responsive to the speech processing module 110 determining that the wakeword is associated with the user 401 with at least the threshold score corresponding to a likelihood that the wakeword is associated with the user. In a further embodiment, the application processing module 112 and the network interface module 206 are activated responsive to the speech processing module 110 both recognizing the wakeword with at least the threshold score and determining that the wakeword is associated with the user 401 with at least the threshold score. Other threshold elements may be provided to enable the user 401 to set individual thresholds for activating any or all of the individual components of the power management subsystem 100. Further threshold elements may be provided to enable the user to specify scores at which one or more blocks of the illustrative routine 300 may be skipped, substantially as discussed above with respect to FIG. 3.

[0072] The user interface 500 may further include one or more timer elements 508A and 508B. Each timer element may be used to set a threshold time interval at which the network interface module 206 and/or one or more components of the power management subsystem 100 are automatically deactivated. With reference to timer element 508A, if the power management subsystem 100 determines that at least a threshold interval of time has passed since an audio input having an energy level satisfying an energy level threshold has been obtained by the user computing device 200, the network interface module 206 may be automatically deactivated, in addition to the application processing module 112, the speech processing module 110, and the speech detection module 108 of the power management subsystem 100. Further timer elements may also be used to set a threshold time interval after which the speech recognition server 420 automatically sends a deactivation instruction to the network interface module 206 and the power management subsystem 100, substantially as discussed above with respect to FIG. 4C. Timer elements for other modules of the power management subsystem 100 may also be provided.

[0073] With continued reference to FIG. 5, the user 401 can select whether the wakeword should be confirmed by the speech recognition server 420 with server confirmation element 510. In some embodiments, the application processing module 112 and network interface module 206 only remains activated after the speech process-

Petitioner Exhibit 1004, Page 723 of 1518

ing module 110 detects the wakeword if a confirmation of the wakeword is received from the speech recognition server 420. If the user 401 requires server confirmation of the wakeword, subsequently obtained audio inputs may not be transmitted to the speech recognition server 420 unless the wakeword is confirmed. However, as discussed above, confirmation is not necessarily required. If the user 401 does not require server confirmation of the wakeword, the user computing device 200 may transmit one or more audio inputs obtained subsequent to the speech processing module 110 detecting the wakeword in the speech and/or determining that the speech is associated with the user 401.

[0074] The user 401 may also select whether speaker identification is required with speaker identification element 512. If the user 401 requires speaker identification, the speech processing module 110 and/or the speech recognition server 420 may be used to determine whether an audio input including speech corresponding to a wakeword is associated with the user 401. The application processing module 112 and network interface module 206 may be activated responsive to the speech processing module 110 determining that the user 401 is the speaker of the speech. Likewise, the network interface module 206 may remain in an activated state responsive to receiving a confirmation from the speech recognition server 420 that the user 401 is indeed the speaker of the wakeword. If the user 401 does not require speaker identification, however, neither the speech processing module 110 nor the speech recognition server 420 need identify the speaker.

[0075] The user interface 500 may also include an on-device recognition selection element 514, wherein the user 401 may select whether the user computing device 200 generates speech recognition results by itself, or whether audio inputs are routed to the speech recognition server 420 for processing into speech recognition results. The on-device recognition selection element 514 may be optionally disabled or grayed out if the user computing device 200 does not include on-device speech recognition capabilities. Further, the on-device recognition selection element 514 may be automatically deselected (and on-device speech recognition capabilities automatically disabled) if the power supply 218 drops below a threshold power supply level (e.g., a battery charge percentage), as on-device speech recognition capabilities as implemented by the processing unit 202 and/or the application processing module 112 may require a relatively large power draw.

[0076] The wakeword pane 516 and sleepword pane 518 may include user interface elements whereby the user 401 may record and cause playback of a wakeword or sleepword spoken by the user 401. When the user 401 records a wakeword or sleepword, the network interface module 206 may be automatically activated so that the audio input including the user's speech may be provided to the speech recognition server 420. The speech recognition server 420 may return a transcription of the recorded wakeword or sleepword so that the user may determine whether the recorded wakeword or sleepword was understood correctly by the speech recognition server 420. Alternately, when the user 401 records a wakeword or sleepword, any on-device speech recognition capabilities of the user computing device 200 may be activated to transcribe the recorded speech of the user 401. A spectral representation of the spoken wakeword or sleepword may also be provided by the user interface 500. Optionally, the wakeword pane 516 and sleepword pane 518 may include suggestions for wakewords or sleepwords, and may also indicate a quality of a wakeword or sleepword provided by the user 401, which quality may reflect a likelihood that the wakeword or sleepword is to produce false positive or false negative. Further information regarding suggesting keywords may be found in U.S. Patent App. No. 13/670,316, entitled "WAKE WORD EVALUATION," which was filed on November 6, 2012.

[0077] Various aspects of the present disclosure have been discussed as hardware implementations for illustrative purposes. However, as discussed above, the power management subsystem 100 may be partially or wholly implemented by the processing unit 202. For example, some or all of the functionality of the power management subsystem 100 may be implemented as software instructions executed by the processing unit 202. In a particular, non-limiting example, the functionality of the speech processing module 110, the application processing module 112, and the power management module 120 may be implemented as software executed by the processing unit 202. The processing unit 202 may accordingly be configured to selectively activate and/or deactivate the network interface module 206 responsive to detecting a wakeword. Still further implementations are possible.

[0078] Depending on the embodiment, certain acts, events, or functions of any of the routines or algorithms described herein can be performed in a different sequence, can be added, merged, or left out altogether (e.g., not all described operations or events are necessary for the practice of the algorithm). Moreover, in certain embodiments, operations or events can be performed concurrently, e.g., through multi-threaded processing, interrupt processing, or multiple processors or processor cores or on other parallel architectures, rather than sequentially.

[0079] The various illustrative logical blocks, modules, routines, and algorithm steps described in connection with the embodiments disclosed herein can be implemented as electronic hardware, computer software, or combinations of both. To clearly illustrate this interchangeability of hardware and software, various illustrative components, blocks, modules, and steps have been described above generally in terms of their functionality. Whether such functionality is implemented as hardware or software depends upon the particular application and design constraints imposed on the overall system. The described functionality can be implemented in varying

Petitioner Exhibit 1004, Page 724 of 1518

ways for each particular application, but such implementation decisions should not be interpreted as causing a departure from the scope of the disclosure.

[0080] Conditional language used herein, such as, among others, "can," "could," "might," "may," "e.g.," and the like, unless specifically stated otherwise, or otherwise understood within the context as used, is generally intended to convey that certain embodiments include, while other embodiments do not include, certain features, elements and/or steps. Thus, such conditional language is not generally intended to imply that features, elements and/or steps are in any way required for one or more embodiments or that one or more embodiments necessarily include logic for deciding, with or without author input or prompting, whether these features, elements and/or steps are included or are to be performed in any particular embodiment. The terms "comprising," "including," "having," and the like are synonymous and are used inclusively, in an open-ended fashion, and do not exclude additional elements, features, acts, operations, and so forth. Also, the term "or" is used in its inclusive sense (and not in its exclusive sense) so that when used, for example, to connect a list of elements, the term "or" means one, some, or all of the elements in the list.

[0081] Conjunctive language such as the phrase "at least one of X, Y and Z," unless specifically stated otherwise, is to be understood with the context as used in general to convey that an item, term, etc. may be either X, Y, or Z, or a combination thereof. Thus, such conjunctive language is not generally intended to imply that certain embodiments require at least one of X, at least one of Y and at least one of Z to each is present.

[0082] While the above detailed description has shown, described, and pointed out novel features as applied to various embodiments, it can be understood that various omissions, substitutions, and changes in the form and details of the devices or algorithms illustrated can be made without departing from the disclosure. The claimed subject matter is defined by the appended claims.

Claims

1. A computer-implemented method of operating a first computing device (202), the method comprising:

   receiving an audio input;
   determining one or more values from the audio input, wherein the one or more values comprise at least one of:

      a first value indicating an energy level of the audio input; or
      a second value indicating a likelihood that the audio input comprises speech;

   activating a first module (110, 112; 202) of the first computing device based at least in part on the one or more values:
   performing an operation, by the first module, wherein the operation comprises:

      determining that the audio input comprises a wakeword and causing activation of a network interface module (206) in response to determining that the audio input comprises a wakeword; and
      causing transmission of at least a portion of the audio input to a second computing device via the activated network interface module.

2. The computer-implemented method of Claim 1, wherein:

   the first module comprises a processor that is switchable between a low-power state and a high-power state; and
   the processor only performs the operation when it is in the high-power state.

3. The computer-implemented method of Claim 2, wherein activating the first module comprises switching the processor from the low-power state to the high-power state.

4. The computer-implemented method of Claim 2 or Claim 3, further comprising deactivating the first module, wherein deactivating the first module comprises switching the processor from the high-power state to the low-power state.

5. The computer-implemented method of any one of Claims 2 to 4, wherein the processor comprises at least one of a digital signal processor or a microprocessor.

6. The computer-implemented method of any one of Claims 1 to 5, wherein the operation further comprises receiving speech recognition results from a second computing device.

7. The computer-implemented method of Claim 6 further comprising:

   activating a second module of the first computing device based at least in part on the one or more values, wherein the second module is configured to implement a speech recognition application; and
   processing the speech recognition results with the speech recognition application.

8. A device (200) comprising:

Petitioner Exhibit 1004, Page 725 of 1518

a first processor (106, 108) configured to:

determine one or more values, wherein the one or more values comprise at least one of a first value indicating an energy level of an audio input or a second value indicating a likelihood that the audio input comprises speech; and
cause activation of a second processor based at least in part on the one or more values;

the second processor (110,112; 202) configured to perform an operation, wherein the operation comprises:

determining that the audio input comprises a wakeword and causing activation of a network interface module (206) in response to determining that the audio input comprises a wakeword; and
causing transmission of at least a portion of the audio input to a second device via the activated network interface module.

9. The device of Claim 8, wherein the first processor comprises at least one of a digital signal processor or a microprocessor.

10. The device of Claim 8 or Claim 9, further comprising a memory buffer module configured to store the audio input.

11. The device of any one of Claims 8 to 10, further comprising an audio input module in communication with the first processor, wherein the audio input module is configured to obtain the audio input.

12. The device of any one of Claims 8 to 11, wherein:

the first value comprises a volume of the audio input;
the second value comprises a first score; and

determining that the audio input comprises a wakeword comprises determining that the audio input comprises a wakeword based at least in part on determining a second score indicating a likelihood that the audio input comprises a wakeword.

13. The device of Claim 12, wherein the first processor comprises an audio detection module and a speech detection module, and wherein the second processor comprises a wakeword recognition module.

14. The device of Claim 12 or Claim 13, wherein:

the speech detection module is configured to determine the first score using at least one of a hidden Markov model, a Gaussian mixture model, energies in a plurality of spectral bands, or signal to noise ratios in a plurality of spectral bands; and
the wakeword recognition module is further configured to determine the second score using at least one of an application processing module, a hidden Markov model, and a Gaussian mixture model.

15. The device of any one of Claims 12 to 14, wherein the wakeword recognition module is configured:

to cause deactivation of the audio detection module based at least in part on the first score; and
to cause deactivation of the speech detection module based at least in part on the second score.

Patentansprüche

1. Computerimplementiertes Verfahren zum Betreiben eines ersten Datenverarbeitungsgerätes (202), wobei das Verfahren Folgendes aufweist:

Empfangen einer Audioeingabe;
Bestimmen eines oder mehrerer Werte aus der Audioeingabe, wobei der eine oder die mehreren Werte mindestens eines der Folgenden aufweisen:

einen ersten Wert, der ein Energieniveau der Audioeingabe angibt; oder
einen zweiten Wert, der eine Wahrscheinlichkeit, dass die Audioeingabe Sprache aufweist, angibt;

Aktivieren eines ersten Moduls (110, 112; 202) des ersten Datenverarbeitungsgerätes mindestens teilweise basierend auf dem einen oder den mehreren Werten;
Durchführen einer Operation, durch das erste Modul, wobei die Operation Folgendes aufweist:

Bestimmen, dass die Audioeingabe ein Weckwort aufweist, und Veranlassen der Aktivierung eines Netzwerkschnittstellenmoduls (206) als Reaktion auf das Bestimmen, dass die Audioeingabe ein Weckwort aufweist; und
Veranlassen der Übertragung von mindestens einem Abschnitt der Audioeingabe an ein zweites Datenverarbeitungsgerät über das aktivierte Netzwerkschnittstellenmodul.

Petitioner Exhibit 1004, Page 726 of 1518

2. Computerimplementiertes Verfahren nach Anspruch 1, wobei:

das erste Modul einen Prozessor aufweist, der zwischen einem Zustand geringer Leistung und einem Zustand hoher Leistung schaltbar ist; und der Prozessor die Operation nur durchführt, wenn er sich in dem Zustand hoher Leistung befindet.

3. Computerimplementiertes Verfahren nach Anspruch 2, wobei das Aktivieren des ersten Moduls das Schalten des Prozessors aus dem Zustand geringer Leistung in den Zustand hoher Leistung aufweist.

4. Computerimplementiertes Verfahren nach Anspruch 2 oder 3, welches ferner das Deaktivieren des ersten Moduls aufweist, wobei das Deaktivieren des ersten Moduls das Schalten des Prozessors aus dem Zustand hoher Leistung in den Zustand geringer Leistung aufweist.

5. Computerimplementiertes Verfahren nach einem der Ansprüche 2 bis 4, wobei der Prozessor mindestens entweder einen digitalen Signalprozessor oder einen Mikroprozessor aufweist.

6. Computerimplementiertes Verfahren nach einem der Ansprüche 1 bis 5, wobei die Operation ferner das Empfangen von Spracherkennungsergebnissen von einem zweiten Datenverarbeitungsgerät aufweist.

7. Computerimplementiertes Verfahren nach Anspruch 6, welches ferner Folgendes aufweist:

Aktivieren eines zweiten Moduls des ersten Datenverarbeitungsgerätes mindestens teilweise basierend auf dem einen oder den mehreren Werten, wobei das zweite Modul zum Implementieren einer Spracherkennungsanwendung konfiguriert ist; und Verarbeiten der Spracherkennungsergebnisse mit der Spracherkennungsanwendung.

8. Gerät (200), welches Folgendes aufweist:

einen ersten Prozessor (106, 108), der zu Folgendem konfiguriert ist:

Bestimmen eines oder mehrerer Werte, wobei der eine oder die mehreren Werte mindestens entweder einen ersten Wert, der ein Energieniveau einer Audioeingabe angibt, oder einen zweiten Wert, der eine Wahrscheinlichkeit, dass die Audioeingabe Sprache aufweist, angibt, aufweisen; und

Veranlassen der Aktivierung eines zweiten Prozessors mindestens teilweise basierend auf dem einen oder den mehreren Werten;

den zweiten Prozessor (110, 112; 202), der zum Durchführen einer Operation konfiguriert ist, wobei die Operation Folgendes aufweist:

Bestimmen, dass die Audioeingabe ein Weckwort aufweist, und Veranlassen der Aktivierung eines Netzwerkschnittstellenmoduls (206) als Reaktion auf das Bestimmen, dass die Audioeingabe ein Weckwort aufweist; und
Veranlassen der Übertragung von mindestens einem Abschnitt der Audioeingabe an ein zweites Gerät über das aktivierte Netzwerkschnittstellenmodul.

9. Gerät nach Anspruch 8, wobei der erste Prozessor mindestens entweder einen digitalen Signalprozessor oder einen Mikroprozessor aufweist.

10. Gerät nach Anspruch 8 oder 9, welches ferner ein Speicherpuffermodul aufweist, das zum Speichern der Audioeingabe konfiguriert ist.

11. Gerät nach einem der Ansprüche 8 bis 10, welches ferner ein Audioeingabemodul in Kommunikation mit dem ersten Prozessor aufweist, wobei das Audioeingabemodul zum Erhalten der Audioeingabe konfiguriert ist.

12. Gerät nach einem der Ansprüche 8 bis 11, wobei:

der erste Wert eine Lautstärke der Audioeingabe aufweist;
der zweite Wert eine erste Punktzahl aufweist; und
das Bestimmen, dass die Audioeingabe ein Weckwort aufweist, das Bestimmen, dass die Audioeingabe ein Weckwort aufweist, mindestens teilweise basierend auf dem Bestimmen einer zweiten Punktzahl, die eine Wahrscheinlichkeit, dass die Audioeingabe ein Weckwort aufweist, angibt, aufweist.

13. Gerät nach Anspruch 12, wobei der erste Prozessor ein Audioerkennungsmodul und ein Spracherkennungsmodul aufweist, und wobei der zweite Prozessor ein Weckworterkennungsmodul aufweist.

14. Gerät nach Anspruch 12 oder 13, wobei:

das Spracherkennungsmodul zum Bestimmen der ersten Punktzahl unter Verwendung von mindestens entweder einem Hidden-Markov-Modell, einem Gaußschen Mischverteilungs-

Petitioner Exhibit 1004, Page 727 of 1518

modell, Energien in mehreren Spektralbändern oder Signal-Rausch-Verhältnissen in mehreren Spektralbändern konfiguriert ist; und
das Weckworterkennungsmodul ferner zum Bestimmen der zweiten Punktzahl unter Verwendung von mindestens entweder einem Anwendungsverarbeitungsmodul, einem Hidden-Markov-Modell oder einem Gaußschen Mischverteilungsmodell konfiguriert ist.

15. Gerät nach einem der Ansprüche 12 bis 14, wobei das Weckworterkennungsmodul zu Folgendem konfiguriert ist:

Veranlassen der Deaktivierung des Audioerkennungsmoduls mindestens teilweise basierend auf der ersten Punktzahl; und
Veranlassen der Deaktivierung des Spracherkennungsmoduls mindestens teilweise basierend auf der zweiten Punktzahl.

Revendications

1. Procédé mis en oeuvre par ordinateur pour le fonctionnement d'un premier dispositif informatique (202), le procédé comprenant les étapes suivantes :

recevoir une entrée audio ;
déterminer une ou plusieurs valeurs provenant de l'entrée audio, la ou les valeurs comprenant au moins un élément parmi :

une première valeur indiquant un niveau d'énergie de l'entrée audio ; ou
une deuxième valeur indiquant une probabilité que l'entrée audio comprenne de la parole ;

activer un premier module (110, 112 ; 202) du premier dispositif informatique au moins en partie en fonction de la ou des valeurs ;
effectuer une opération, par le premier module, l'opération comprenant les étapes suivantes :

déterminer que l'entrée audio comprend un mot d'activation et provoquer l'activation d'un module d'interface réseau (206) en réponse à la détermination que l'entrée audio comprend un mot d'activation ; et
provoquer la transmission d'au moins une partie de l'entrée audio à un deuxième dispositif informatique par le biais du module d'interface réseau activé.

2. Procédé mis en oeuvre par ordinateur de la revendication 1, dans lequel :

le premier module comprend un processeur qui est commutable entre un état à basse puissance et un état à haute puissance ; et
le processeur effectue l'opération uniquement quand il se trouve dans l'état à haute puissance.

3. Procédé mis en oeuvre par ordinateur de la revendication 2, dans lequel l'activation du premier module comprend la commutation du processeur de l'état à basse puissance à l'état à haute puissance.

4. Procédé mis en oeuvre par ordinateur de la revendication 2 ou la revendication 3, comprenant en outre la désactivation du premier module, la désactivation du premier module comprenant la commutation du processeur de l'étape à haute puissance à l'état à basse puissance.

5. Procédé mis en oeuvre par ordinateur de l'une quelconque des revendications 2 à 4, dans lequel le processeur comprend un processeur de signal numérique et/ou un microprocesseur.

6. Procédé mis en oeuvre par ordinateur de l'une quelconque des revendications 1 à 5, dans lequel l'opération comprend en outre la réception de résultats de reconnaissance de la parole provenant d'un deuxième dispositif informatique.

7. Procédé mis en oeuvre par ordinateur de la revendication 6 comprenant en outre :

l'activation d'un deuxiéme module du premier dispositif informatique au moins en partie en fonction de la ou des valeurs, le deuxième module étant configuré pour mettre en oeuvre une application de reconnaissance de la parole ; et le traitement des résultats de reconnaissance de la parole avec l'application de reconnaissance de la

8. Dispositif (200) comprenant :

un premier processeur (106, 108) configuré pour :

déterminer une ou plusieurs valeurs, la ou les valeurs comprenant une première valeur indiquant un niveau d'énergie d'une entrée audio et/ou une deuxième valeur indiquant une probabilité que l'entrée audio comprenne de la parole ; et
provoquer l'activation d'un deuxième processeur au moins en partie en fonction de la ou des valeurs ;

le deuxième processeur (110, 112 ; 202) configuré pour effectuer une opération, l'opération

Petitioner Exhibit 1004, Page 728 of 1518

comprenant les étapes suivantes :

déterminer que l'entrée audio comprend un mot d'activation et provoquer l'activation d'un module d'interface réseau (206) en réponse à la détermination que l'entrée audio comprend un mot d'activation ; et
provoquer la transmission d'au moins une partie de l'entrée audio à un deuxième dispositif par le biais du module d'interface réseau activé.

9. Dispositif de la revendication 8, dans lequel le premier processeur comprend un processeur de signal numérique et/ou un microprocesseur.

10. Dispositif de la revendication 8 ou la revendication 9, comprenant en outre un module de mémoire tampon configuré pour stocker l'entrée audio.

11. Dispositif de l'une quelconque des revendications 8 à 10, comprenant en outre un module d'entrée audio en communication avec le premier processeur, le module d'entrée audio étant configuré pour obtenir l'entrée audio.

12. Dispositif de l'une quelconque des revendications 8 à 11, dans lequel :

la première valeur comprend un volume de l'entrée audio ;
la deuxième valeur comprend un premier score ; et

la détermination que l'entrée audio comprend un mot d'activation comprend la détermination que l'entrée audio comprend un mot d'activation au moins en partie en fonction de la détermination d'un deuxième score indiquant une probabilité que l'entrée audio comprenne un mot d'activation.

13. Dispositif de la revendication 12, dans lequel le premier processeur comprend un module de détection audio et un module de détection de la parole, et dans lequel le deuxième processeur comprend un module de reconnaissance de mot d'activation.

14. Dispositif de la revendication 12 ou la revendication 13, dans lequel :

le module de détection de la parole est configuré pour déterminer le premier score en utilisant au moins un élément parmi un modèle de Markov caché, un modèle de mélange gaussien, des énergies dans une pluralité de bandes spectrales, ou des rapports signal/bruit dans une pluralité de bandes spectrales ; et
le module de reconnaissance de mot d'activation est en outre configuré pour déterminer le deuxième score en utilisant au moins un élément parmi un module de traitement d'application, un modèle de Markov caché, et un modèle de mélange gaussien.

15. Dispositif de l'une quelconque des revendications 12 à 14, dans lequel le module de reconnaissance de mot d'activation est configuré :

pour provoquer la désactivation du module de détection audio au moins en partie en fonction du premier score ; et
pour provoquer la désactivation du module de détection de la parole au moins en partie en fonction du deuxième score.

Petitioner Exhibit 1004, Page 729 of 1518

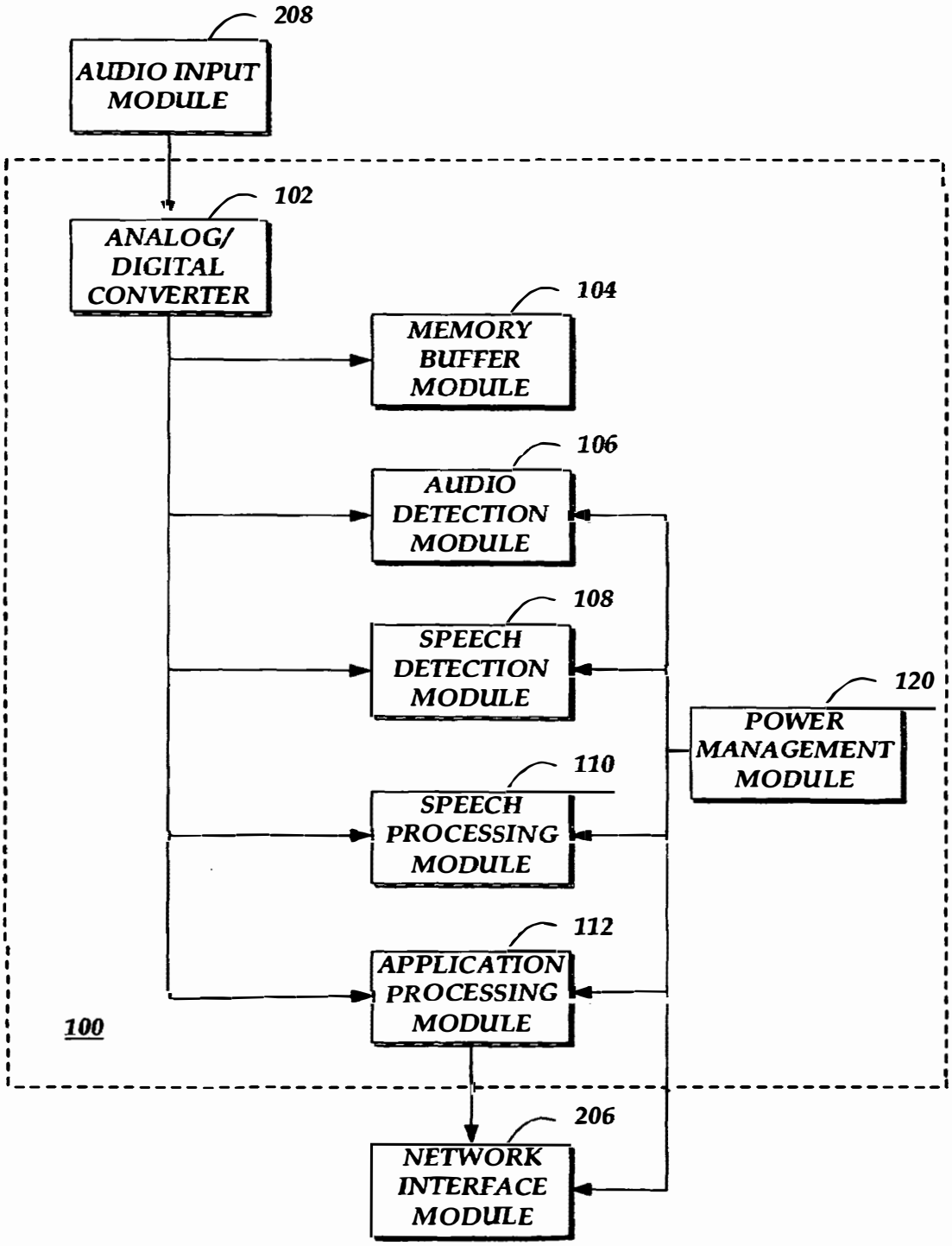


Fig. 1.

Petitioner Exhibit 1004, Page 730 of 1518

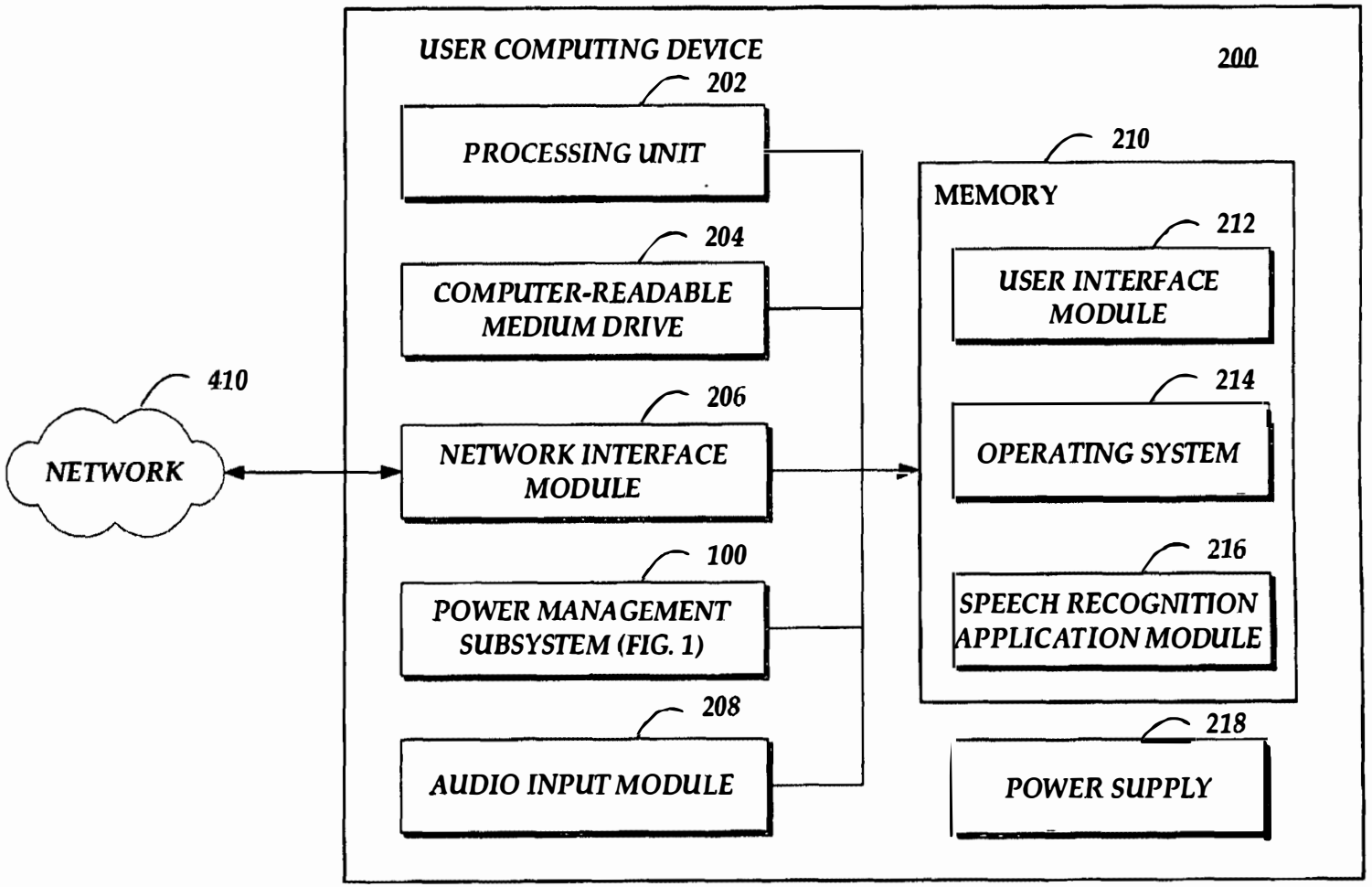

USER COMPUTING DEVICE
200
202
PROCESSING UNIT
204
COMPUTER-READABLE MEDIUM DRIVE
206
NETWORK INTERFACE MODULE
100
POWER MANAGEMENT SUBSYSTEM (FIG. 1)
208
AUDIO INPUT MODULE
210
MEMORY
212
USER INTERFACE MODULE
214
OPERATING SYSTEM
216
SPEECH RECOGNITION APPLICATION MODULE
218
POWER SUPPLY
410
NETWORK
20

Fig. 2.

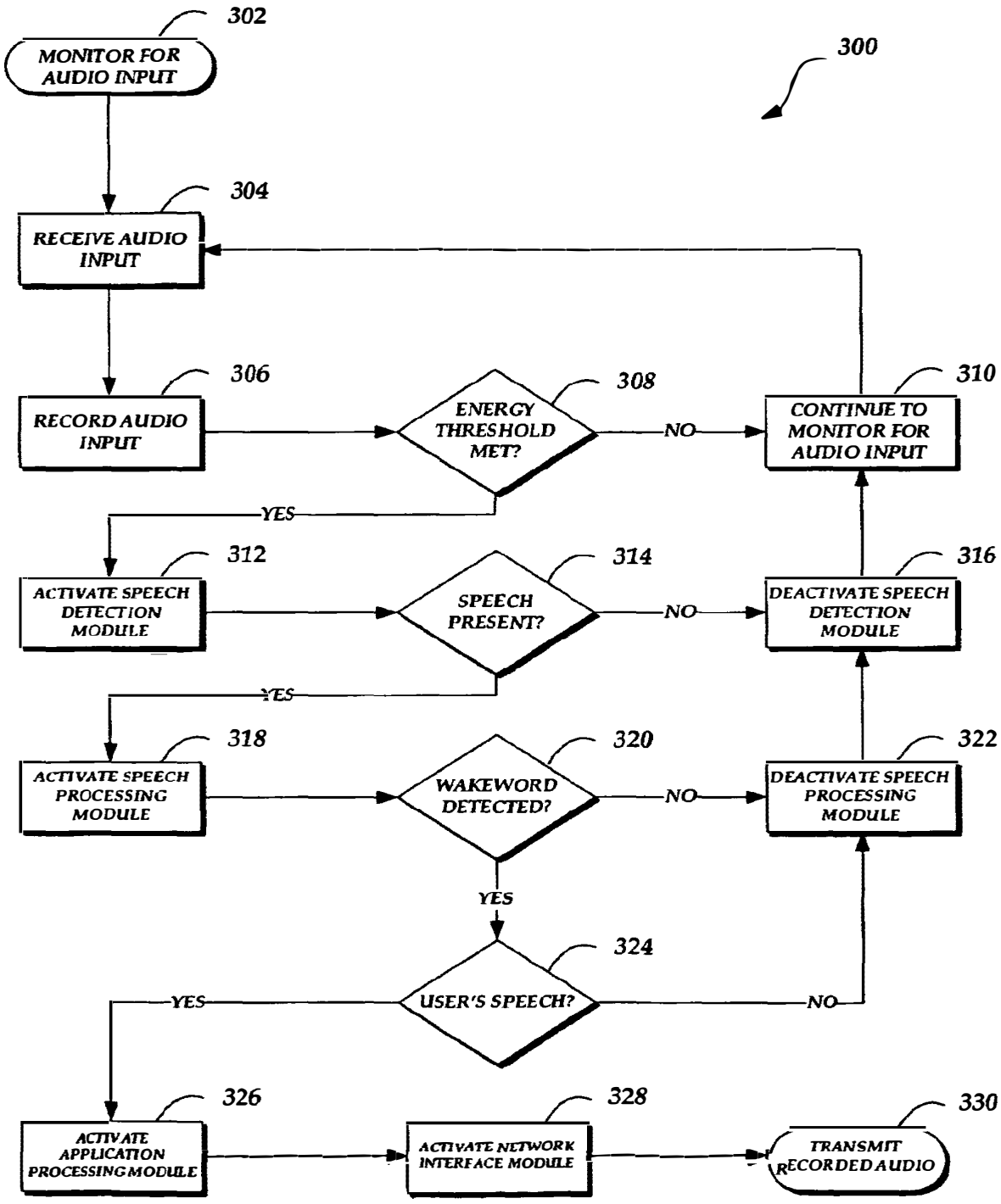


Fig.3.

Petitioner Exhibit 1004, Page 732 of 1518

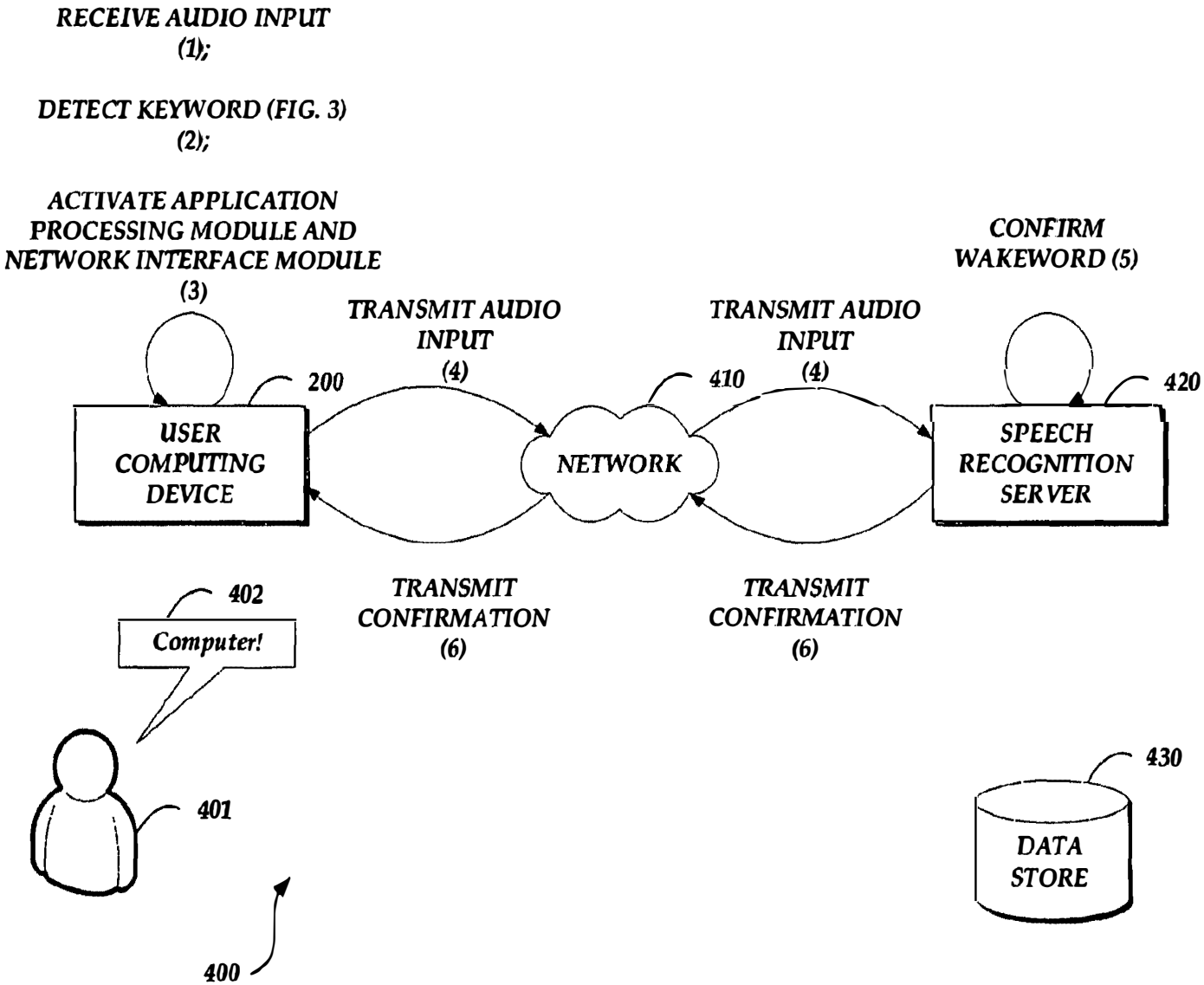

RECEIVE AUDIO INPUT (1);
DETECT KEYWORD (FIG. 3) (2);
ACTIVATE APPLICATION PROCESSING MODULE AND NETWORK INTERFACE MODULE (3)
TRANSMIT AUDIO INPUT (4)
TRANSMIT AUDIO INPUT (4)
CONFIRM WAKEWORD (5)
200
410
420
USER COMPUTING DEVICE
NETWORK
SPEECH RECOGNITION SERVER
TRANSMIT CONFIRMATION (6)
TRANSMIT CONFIRMATION (6)
402
Computer!
401
400
430
DATA STORE

Fig.4A.

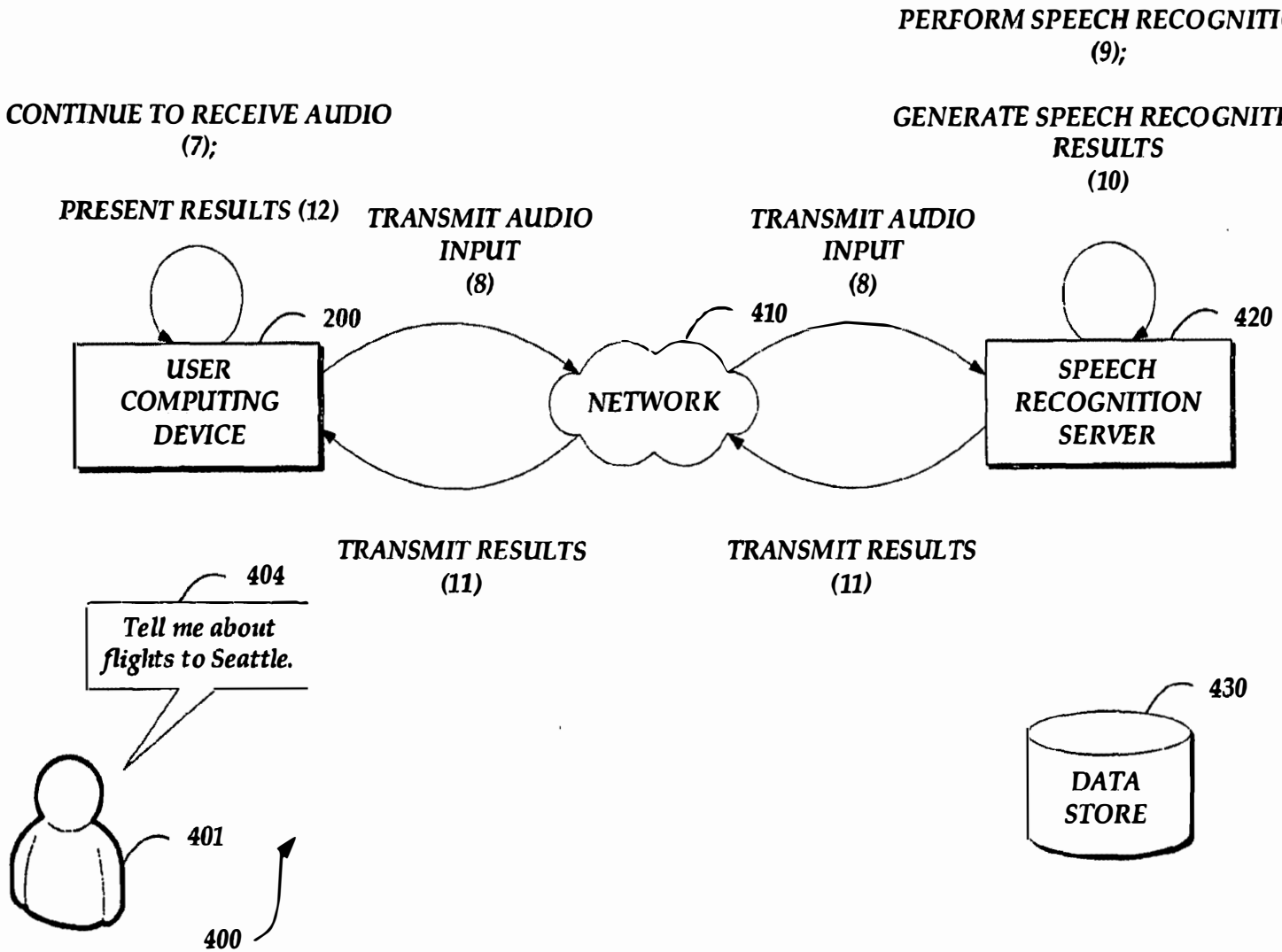

PERFORM SPEECH RECOGNITION (9);
CONTINUE TO RECEIVE AUDIO (7);
GENERATE SPEECH RECOGNITION RESULTS (10)
PRESENT RESULTS (12)
TRANSMIT AUDIO INPUT (8)
TRANSMIT AUDIO INPUT (8)
200
410
420
USER COMPUTING DEVICE
NETWORK
SPEECH RECOGNITION SERVER
TRANSMIT RESULTS (11)
TRANSMIT RESULTS (11)
404
Tell me about flights to Seattle.
430
DATA STORE
401
400

Fig.4B.

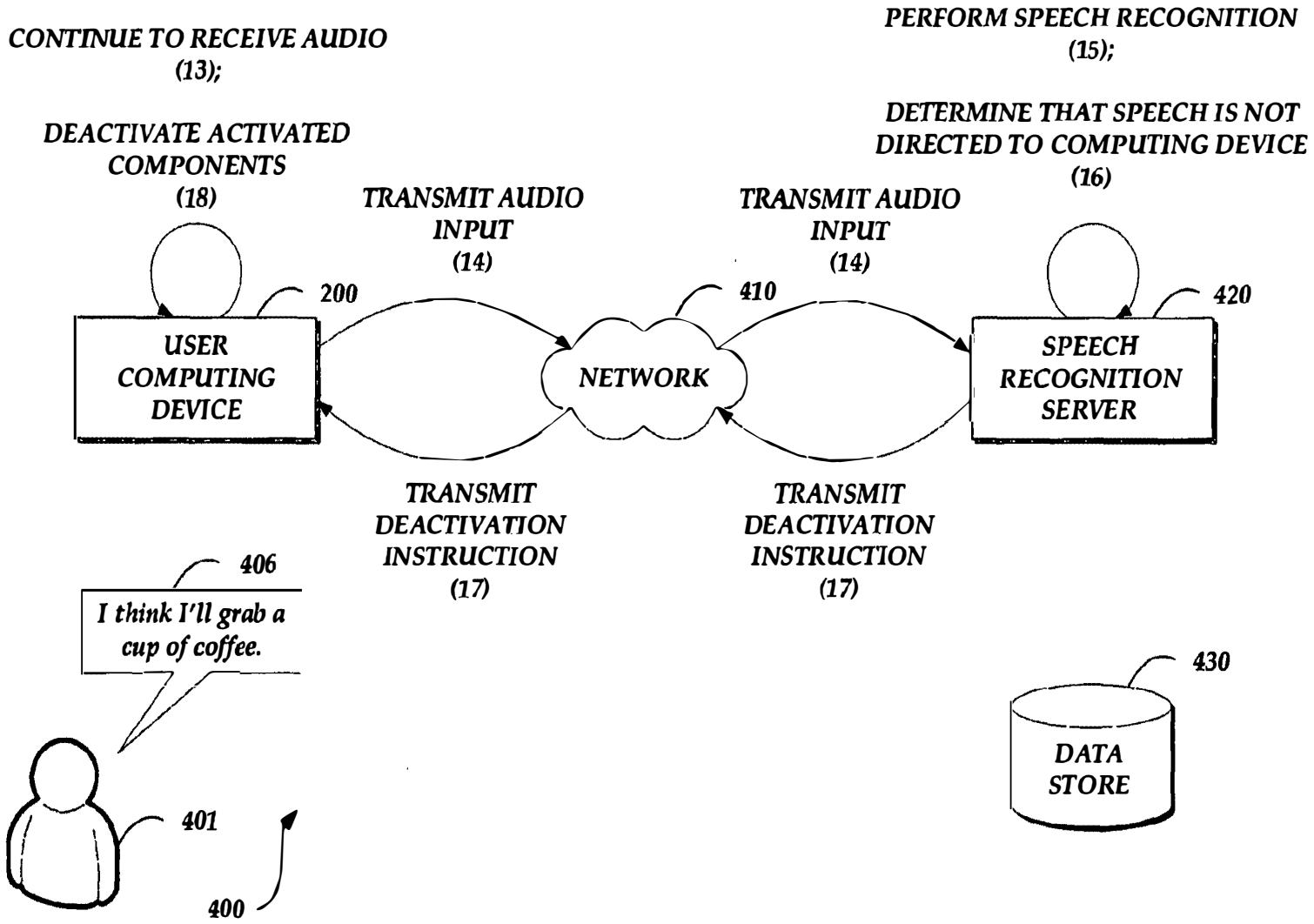


Fig.4C.

EP 2 932 500 B1

Petitioner Exhibit 1004, Page 735 of 1518

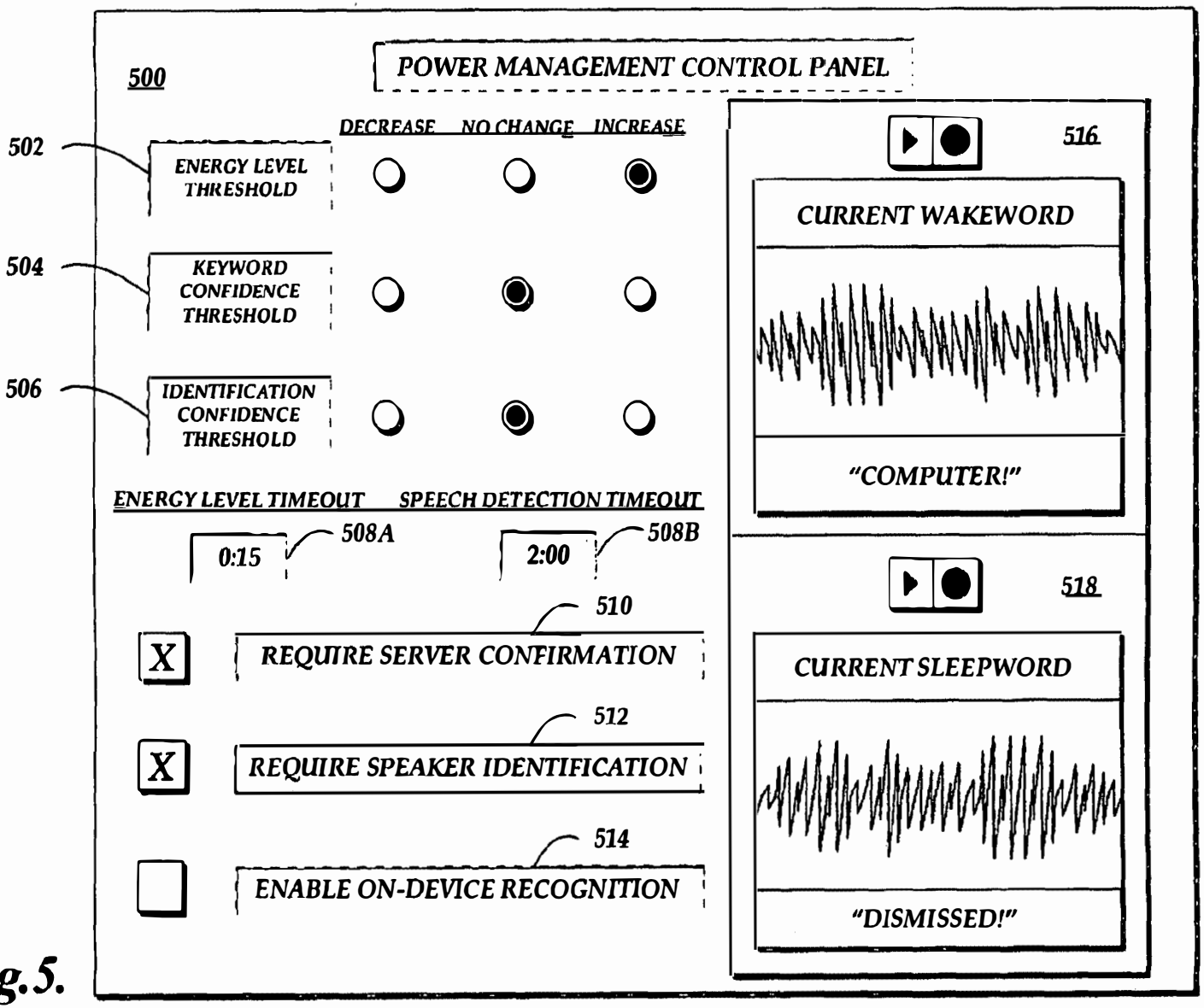


Fig.5.

EP 2 932 500 B1

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 2003216909 A **[0001]**
- US 20120016678 A **[0009]**
- US 58779912 A **[0032] [0057]**
- US 59215712 A **[0032] [0057]**
- US 05873002 A **[0037]**
- US 20030144844 A **[0037]**
- US 8326636 B **[0037]**
- US 8117268 B **[0054]**
- US 67031612 A **[0076]**

Petitioner Exhibit 1004, Page 737 of 1518

DOCKET NO.: N0484.71090US01

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

First Named Inventor:	William F. Ganong
Application No.:	13/795,933
Confirmation No.:	2321
Filed:	March 12, 2013
For:	METHODS AND APPARATUS FOR DETECTING A VOICE COMMAND
Examiner:	Huy C. Ho
Art Unit:	2644

CERTIFICATE OF ELECTRONIC FILING UNDER 37 C.F.R. § 1.8

The undersigned hereby certifies that this paper, along with any paper referred to as being attached or enclosed, is being transmitted via the Office electronic filing system in accordance with § 1.6(a)(4), on the 5th day of October, 2017.

Electronic Signature: /Trish McDonald/

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

STATEMENT FILED PURSUANT TO THE DUTY OF DISCLOSURE UNDER 37 C.F.R. §§ 1.56, 1.97 AND 1.98

Sir:

Pursuant to the duty of disclosure under 37 C.F.R. §§ 1.56, 1.97 and 1.98, the undersigned requests consideration of this Information Disclosure Statement.

PART I: Compliance with 37 C.F.R. § 1.97

This Information Disclosure Statement has been filed before the mailing of a first Office Action after the filing of a Request for Continued Examination under 37 C.F.R. § 1.114.

No fee or certification is required.

PART II: Information Cited

The undersigned hereby makes of record in the above-identified application the information listed on the attached form PTO-1449 (modified PTO/SB/08). The order of

Petitioner Exhibit 1004, Page 738 of 1518

presentation of the references should not be construed as an indication of the importance of the references.

The undersigned hereby makes the following additional information of record in the above-identified application.

The undersigned would like to bring to the Examiner's attention the enclosed search report or other communication from a corresponding or related International or Foreign National Application:

Serial No.	Date of Mailing	Type(s) of Communication	Docket No.
CN 201480013903.1	**07-28-2017**	**Office Action**	**N0484.71090CN00**

PART III: Remarks

Documents cited anywhere in the Information Disclosure Statement are enclosed unless otherwise indicated. It is respectfully requested that:

1. The Examiner consider completely the cited information, along with any other information, in reaching a determination concerning the patentability of the present claims;

2. The enclosed form PTO-1449 (modified PTO/SB/08) be signed by the Examiner to evidence that the cited information has been fully considered by the United States Patent and Trademark Office during the examination of this application;

3. The citations for the information be printed on any patent which issues from this application.

By submitting this Information Disclosure Statement, the undersigned makes no representation that a search has been performed, of the extent of any search performed, or that more relevant information does not exist.

By submitting this Information Disclosure Statement, the undersigned makes no representation that the information cited in the Statement is, or is considered to be, material to patentability as defined in 37 C.F.R. § 1.56(b).

Petitioner Exhibit 1004, Page 739 of 1518

By submitting this Information Disclosure Statement, the undersigned makes no representation that the information cited in the Statement is, or is considered to be, in fact, prior art as defined by 35 U.S.C. § 102.

Notwithstanding any statements by the undersigned, the Examiner is urged to form his or her own conclusion regarding the relevance of the cited information.

An early and favorable action is hereby requested.

The Director is hereby authorized to charge any deficiency or credit any overpayment in the fees occasioned by the filing of this Information Disclosure Statement to our Deposit Account No. 23/2825 under Docket No. N0484.71090US01 from which the undersigned is authorized to draw.

Respectfully submitted,

By: /Thomas S. Chlebeck/
Thomas S. Chlebeck, Reg. No. 70,616
Wolf, Greenfield & Sacks, P.C.
600 Atlantic Avenue
Boston, Massachusetts 02210-2206
Telephone: (617) 646-8000

Docket No.: N0484.71090US01
Date: 2017-10-05
xNDDx

Electronic Acknowledgement Receipt

EFS ID:	30582253
Application Number:	13795933
International Application Number:	
Confirmation Number:	2321
Title of Invention:	METHODS AND APPARATUS FOR DETECTING A VOICE COMMAND
First Named Inventor/Applicant Name:	William F. Ganong
Customer Number:	91337
Filer:	Thomas Stephen Chlebeck/Trish McDonald
Filer Authorized By:	Thomas Stephen Chlebeck
Attorney Docket Number:	N0484.71090US01
Receipt Date:	05-OCT-2017
Filing Date:	12-MAR-2013
Time Stamp:	21:16:01
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	Miscellaneous Incoming Letter	N048471090US01-TRN-TSC.pdf	25034 43e1acb304537e81fdfc632f139d8da5ebfa37cf	no	1

Warnings:

Information:					
2		N048471090US01-IDS-TSC.pdf	30842 a9d6151504c0f17829b766a09c4ee1d82f6dbc86	yes	4
	Multipart Description/PDF files in .zip description				
	Document Description		**Start**	**End**	
	Transmittal Letter		1	3	
	Information Disclosure Statement (IDS) Form (SB08)		4	4	
Warnings:					
Information:					
3	Other Reference-Patent/App/Search documents	N048471090CN00_OA_Dated_07-28-2017.pdf	545151 c2521d594df48fcc6aa01ce7056e317ab6484868	no	15
Warnings:					
Information:					
4	Foreign Reference	CN105009204A.pdf	2177666 a96d10c13ceb22f2f3de266f9b559fc80d787524	no	29
Warnings:					
Information:					
5	Foreign Reference	EP2932500B1.pdf	1732727 f162ed64eff763feedf956bca415dc860c4e5d4f	no	26
Warnings:					
Information:					
		Total Files Size (in bytes):	4511420		

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.

Espacenet

Bibliographic data: CN105009204 (A) — 2015-10-28

SPEECH RECOGNITION POWER MANAGEMENT

Inventor(s): BASYE KENNETH JOHN; SECKER-WALKER HUGH EVAN; DAVID TONY; KNESER REINHARD; ADAMS JEFFREY PENROD; SALVADOR STAN WEIDNER; KRISHNAMOORTHY MAHESH ± (BASYE KENNETH JOHN, ; SECKER-WALKER HUGH EVAN, ; DAVID TONY, ; KNESER REINHARD, ; ADAMS JEFFREY PENROD, ; SALVADOR STAN WEIDNER, ; KRISHNAMOORTHY MAHESH)

Applicant(s): AMAZON TECH INC ± (AMAZON TECHNOLOGIES, INC)

Classification: **- international:** ***G10L15/26; G10L25/78***
- cooperative: **G10L15/28;** G10L15/30; G10L2015/088; G10L25/78

Application number: CN2013864898 20131209 Global Dossier

Priority number(s): WO2013US73913 20131209 ; US201213711510 20121211

Also published as: EP2932500 (A1) EP2932500 (B1) JP2016505888 (A) US2014163978 (A1) US9704486 (B2) more

Abstract of CN105009204 (A)

Power consumption for a computing device may be managed by one or more keywords. For example, if an audio input obtained by the computing device includes a keyword, a network interface module and/or an application processing module of the computing device may be activated. The audio input may then be transmitted via the network interface module to a remote computing device, such as a speech recognition server. Alternately, the computing device may be provided with a speech recognition engine configured to process the audio input for on-device speech recognition.

(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 105009204 A
(43)申请公布日 2015.10.28

(21)申请号 201380064898.2
(22)申请日 2013.12.09
(30)优先权数据
13/711,510 2012.12.11 US
(85)PCT国际申请进入国家阶段日
2015.06.11
(86)PCT国际申请的申请数据
PCT/US2013/073913 2013.12.09
(87)PCT国际申请的公布数据
WO2014/093238 EN 2014.06.19
(71)申请人 亚马逊技术有限公司
地址 美国内华达州
(72)发明人 肯尼思·约翰·巴斯耶
休·埃文·塞克-瓦尔克
托尼·大卫 赖因哈德·克内泽尔
杰弗瑞·彭罗德·亚当斯
斯坦·魏德纳·萨尔瓦多
马赫什·克里希纳穆尔蒂
(74)专利代理机构 北京天昊联合知识产权代理有限公司 11112
代理人 顾丽波 李荣胜
(51)Int.Cl.
G10L 15/26(2006.01)
G10L 25/78(2006.01)

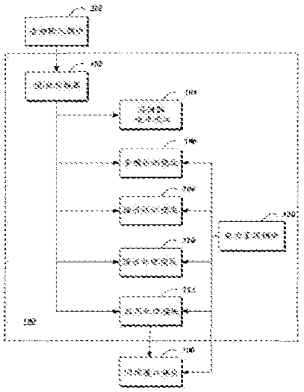
权利要求书2页 说明书18页 附图7页

(54)发明名称

语音识别功率管理

(57)摘要

计算装置的功耗可通过一个或多个关键字进行管理。例如，如果所述计算装置获得的音频输入包含关键字，那么所述计算装置的网络接口模块和/或应用处理模块可被启用。所述音频输入然后可经由所述网络接口模块传输到远程计算装置，诸如，语音识别服务器。或者，所述计算装置可配有语音识别引擎，所述语音识别引擎配置来处理所述音频输入以便进行装置上的语音识别。



CN 105009204 A

Petitioner Exhibit 1004, Page 745 of 1518

1. 一种系统，其包括：

音频输入模块，其配置来接收音频输入；

与所述音频输入模块通信的音频检测模块，所述音频检测模块配置来确定所述音频输入的音量，并且至少部分基于所确定的音量，使语音检测模块启用；

所述语音检测模块配置来确定指示所述音频输入包括语音的可能性的第一分数，并且至少部分基于所述分数，使唤醒字识别模块启用；

所述唤醒字识别模块配置来确定指示所述音频输入包括唤醒字的可能性的第二分数；以及

网络接口模块，其配置来将所获得的音频输入的至少一部分传输到远程计算装置。

2. 根据权利要求 1 所述的系统，其中所述音频输入装置包括麦克风，所述音频检测模块包括第一数字信号处理器，所述语音检测模块包括第二数字信号处理器，并且所述唤醒字识别模块包括微处理器。

3. 根据权利要求 1 所述的系统，其中：

所述语音检测模块还配置来使用以下各者中的至少一个来确定所述第一分数：隐马尔可夫模型、高斯混合模型、多个频谱带中的能量或者多个频谱带中的信噪比；并且

所述唤醒字识别模块还配置来使用以下各者中的至少一个来确定所述第二分数：应用处理模块、隐马尔可夫模型以及高斯混合模型。

4. 根据权利要求 1 所述的系统，其中：

所述唤醒字识别模块还配置来至少部分基于所述第一分数来使所述音频检测模块停用；并且

所述唤醒字识别模块还配置来至少部分基于所述第二分数来使所述语音检测模块停用。

5. 一种操作第一计算装置的计算机实施的方法，所述方法包括：

接收音频输入；

从所述音频输入确定一个或多个值，其中所述一个或多个值包括以下各者中的至少一个：

指示所述音频输入的能量水平的第一值；或

指示所述音频输入包括语音的可能性的第二值；

至少部分基于所述一个或多个值，启用所述第一计算装置的第一模块；

通过所述第一模块来执行操作，其中所述操作包括以下各者中的至少一个：

确定所述音频输入包括唤醒字；

对所述音频输入的至少一部分执行语音识别，以获得语音识别结果；或

使所述音频输入的至少一部分传输到第二计算装置。

6. 根据权利要求 5 所述的计算机实施的方法，其中：

所述第一模块包括处理器，所述处理器可在低功率状态与高功率状态之间切换；并且

所述处理器只有在处于所述高功率状态时才执行所述操作。

7. 根据权利要求 6 所述的计算机实施的方法，其中启用所述第一模块包括将所述处理器从所述低功率状态切换到所述高功率状态。

8. 根据权利要求 6 所述的计算机实施的方法，其还包括停用所述第一模块，其中停用

Petitioner Exhibit 1004, Page 746 of 1518

所述第一模块包括将所述处理器从所述高功率状态切换到所述低功率状态。

9. 根据权利要求 6 所述的计算机实施的方法，其中所述处理器包括数字信号处理器或微处理器中的至少一个。

10. 根据权利要求 5 所述的计算机实施的方法，其中所述操作包括使所述音频输入的至少一部分传输到所述第二计算装置，并且还包括从所述第二计算装置接收语音识别结果。

11. 根据权利要求 10 所述的计算机实施的方法，其还包括：

至少部分基于所述一个或多个值，启用所述第一计算装置的第二模块，其中所述第二模块配置来实施语音识别应用；以及

通过所述语音识别应用来处理所述语音识别结果。

12. 一种装置，其包括：

第一处理器，其配置来：

确定一个或多个值，其中所述一个或多个值包括以下各者中的至少一个：指示音频输入的能量水平的第一值，或指示所述音频输入包括语音的可能性的第二值；以及

至少部分基于所述一个或多个值，使第二处理器启用；

所述第二处理器，其配置来执行操作，其中所述操作包括以下各者中的至少一个：

确定所述音频输入包括唤醒字；

对所述音频输入的至少一部分执行语音识别，以获得语音识别结果；或

使所述音频输入的至少一部分传输到第二装置。

13. 根据权利要求 12 所述的装置，其中所述第一处理器包括数字信号处理器或微处理器中的至少一个。

14. 根据权利要求 12 所述的装置，其还包括配置来存储所述音频输入的存储器缓冲模块。

15. 根据权利要求 12 所述的装置，其还包括与所述第一处理器通信的音频输入模块，其中所述音频输入模块配置来获得所述音频输入。

Petitioner Exhibit 1004, Page 747 of 1518

语音识别功率管理

背景技术

[0001] 计算装置可包含语音识别能力。例如，计算装置可捕获音频输入并使用声学模型和语言模型来识别语音。声学模型用来生成假设，其中基于语音的声学特征，声音子字单元（例如，音素等）对应于语音。语言模型基于所说语音所用语言的词汇特征来确定使用声学模型生成的哪个假设最有可能是语音的抄本。计算装置也能够针对具体的语音识别应用来处理识别的语音。例如，有限的语法或自然语言处理技术可用来处理语音。

[0002] 附图简述

[0003] 参考以下详细说明，结合附图，将更易于了解、同时更好地理解本公开的各方面以及随之而来的很多优点，在附图中：

[0004] 图 1 为描绘说明性功率管理子系统的示意图。

[0005] 图 2 为描绘包括功率管理子系统的说明性用户计算装置的示意图。

[0006] 图 3 为描绘用于语音识别功率管理的说明性例行程序的流程图，所述语音识别功率管理可由图 1 的功率管理子系统实施。

[0007] 图 4A、图 4B 和图 4C 为描绘分布式语音识别系统的说明性操作的状态图。

[0008] 图 5 为描绘说明性用户接口的直观图，所述用户接口可由包括功率管理子系统的用户计算装置提供。

[0009] 详细说明

[0010] 在语音识别的一些当前方法中，语音识别能力被分配到分布式计算环境中的一个或多个计算装置上。在这些方法的一个具体实例中，第一计算装置可配置来捕获音频输入，并且可通过网络将音频输入传输到第二计算装置。第二计算装置可对音频输入执行语音识别，并且生成音频输入中包含的语音的抄本。语音的抄本随后可通过网络从第二计算装置传输回到第一计算装置。在其他当前方法中，第一计算装置可配置来捕获音频输入并独立地转录音频输入。

[0011] 在这些和其他当前方法中，第一计算装置可配置来保持在持续活动状态。在这种持续活动状态下，第一计算装置可连续不断地维持到第二计算装置的网络连接。第一计算装置也可继续为用来实施自身的语音识别能力的任何硬件提供电力。这些方法的一个缺点尤其在于，第一计算装置用来维持持续活动状态而消耗的能量可能是无法接受的。对于依赖电池电力的移动计算装置而言，这种能量需求特别成问题。当前方法中还有其他问题。

[0012] 因此，本公开的方面涉及语音识别的功率管理。计算装置可配有功率管理子系统，所述功率管理子系统选择性地启用或停用计算装置的一个或多个模块。此启用可响应于音频输入，所述音频输入包括一个或多个预先指定的所说字词，本文中有时称为“关键字”。促使启用一个或多个部件的关键字在本文中有时称为“唤醒字”，而促使停用一个或多个部件的关键字在本文中有时称为“休眠字”。在一个特定实例中，计算装置可包括选择性地启用的网络接口模块，所述网络接口模块在启用时会消耗能量以使得计算装置连接到第二计算装置，诸如，语音识别服务器或其他计算装置。功率管理子系统可处理音频输入以确定音频输入是否包含唤醒字，并且响应于确定音频输入包括唤醒字而启用网络接口模块。启用网

Petitioner Exhibit 1004, Page 748 of 1518

络接口模块之后，功率管理子系统可使音频输入传输到语音识别服务器以进行处理。

[0013] 功率管理子系统自身可包括一个或多个选择性地启用的模块。在一些实施例中，一个或多个选择性地启用的模块被实施为专用硬件（诸如集成电路、数字信号处理器或其他类型的处理器），其可从功能相对较少的低功率停用状态切换到功能相对较多的高功率启用状态，并且反之亦然。在其他实施例中，一个或多个模块被实施为软件，其包含由一个或多个通用处理器执行的计算机可执行代码。通过启用（或停用）通用处理器可启用（或停用）软件模块，所述通用处理器配置来或能够执行软件中包含的计算机可执行代码。在另外的实施例中，功率管理系统包括一个或多个硬件模块以及一个或多个软件模块。

[0014] 功率管理子系统可还包括与一个或多个选择性地启用的模块通信的控制模块。此控制模块在本文中有时称为“功率管理模块”，并且可包括上述硬件或软件中的任一个。功率管理模块可使启用或停用功率管理子系统的模块。在一些实施例中，至少部分基于计算装置中包括的音频输入模块获得的音频输入的特征，功率管理模块启用或停用一个或多个模块。例如，功率管理子系统的模块可确定一个或多个值，例如，所述值可包括音频输入的能量水平或音量；对应于语音出现在音频输入中的可能性的分数；对应于关键字出现在语音中的可能性的分数；以及其他值。模块可将一个或多个值传送到功率管理模块，所述功率管理模块可与另一模块通信以使该模块启用，或者与从中接收一个或多个值的模块通信以使所述模块和/或其他模块停用。然而，在其他实施例中，第一选择性地启用的模块可直接与第二选择性地启用的模块通信，以使该模块启用。在此类实施例中，无需功率管理模块。在另外的实施例中，功率管理子系统可配有一个或多个模块，其中一个或多个模块中的至少一些彼此通信，但不与功率管理模块通信。

[0015] 在一个示例性实施方案中，功率管理子系统可包括音频检测模块，所述音频检测模块可配置来确定计算装置获得的音频输入的能量水平或音量。尽管音频检测模块可持续监测音频输入，但功率管理子系统的其余部件可保持在低功率非活动状态，直到被启用为止（被功率管理模块或另一不同模块启用）。如果音频检测模块确定音频输入满足阈值能量水平或音量，则语音检测模块可被启用以确定音频输入是否包含语音。如果语音检测模块确定音频输入包含语音，则功率管理子系统中包括的语音处理模块可被启用。语音处理模块可确定语音是否包含唤醒字，并且可任选地对语音分类，以确定特定的用户是否说出所述唤醒字。如果语音处理模块确定语音包含唤醒字，那么应用处理模块可被启用，所述应用处理模块可实施存储在计算装置的存储器中的语音识别应用模块。例如，语音识别应用可包括智能代理前端，诸如，如2011年1月10日申请并于2012年1月19日公开为美国公布号2012/0016678的“智能自动助理（Intelligent Automated Assistant）”中所描述的智能代理前端。该专利申请的公开内容以全文引用的方式并入本文中。选择性地启用的网络接口模块也可被启用，如上文所述，并且音频输入可被传输到远程计算装置，以便进行处理。下文相对于图3更详细地论述了此示例性实施方案。或者，功率管理子系统可响应于检测到唤醒字而启用处理单元，所述处理单元实施计算装置的任何装置上的语音识别能力。

[0016] 通过选择性地启用计算装置的模块，功率管理子系统可有利地提高计算装置的能量效率。通过选择性地启用一个或多个自身的模块，功率管理子系统还可提高计算装置的能量效率。尽管对于依赖电池电力的计算装置而言，此类实施方案尤其有利，但可能需要功率管理的所有计算装置均可受益于本公开的原理。

Petitioner Exhibit 1004, Page 749 of 1518

[0017] 现在转到图 1，示出可包括在计算装置中的说明性功率管理子系统 100。功率管理子系统 100 可包括模 / 数转换器 102；存储器缓冲模块 104；音频检测模块 106；语音检测模块 108；语音处理模块 110；应用处理模块 112；以及功率管理模块 120。存储器缓冲模块 104 可与音频检测模块 106、语音检测模块 108、语音处理模块 110、应用处理模块 112 以及网络接口模块 206 通信。功率管理模块 120 同样可与音频检测模块 106、语音检测模块 108、语音处理模块 110、应用处理模块 112 以及网络接口模块 206 通信。

[0018] 模 / 数转换器 102 可接收来自音频输入模块 208 的音频输入。下文相对于图 2 还详细地论述音频输入模块 208。模 / 数转换器 102 可配置来将模拟音频输入转换成数字音频输入，以便由功率管理子系统 100 的其他部件处理。在音频输入模块 208 获得数字音频输入（例如，音频输入模块 208 包括数字麦克风或其他数字音频输入装置）的实施例中，模 / 数转换器 102 可任选地从功率管理子系统 100 中省去。因此，音频输入模块 208 可将音频输入直接提供到功率管理子系统 100 的其他模块。

[0019] 存储器缓冲模块 104 可包含配置来存储数字音频输入的一个或多个存储器缓冲区。音频输入模块 208 获得的音频输入（并且如果是模拟音频输入，则由模 / 数转换器 102 转换成数字形式）可被记录到存储器缓冲模块 104。记录到存储器缓冲模块 104 的音频输入可由功率管理子系统 100 的其他模块存取，以便由这些模块处理，如本文中进一步论述。

[0020] 存储器缓冲模块 104 的一个或多个存储器缓冲区可包括硬件存储器缓冲区、软件存储器缓冲区或者两者。一个或多个存储器缓冲区可具有相同的容量或不同的容量。存储器缓冲模块 104 的存储器缓冲区可经选择以存储，其他模块可取决于所述音频输入而启用。例如，如果只有音频检测模块 106 活动，那么音频输入可被存储到容量相对较小的硬件存储器缓冲区。然而，如果其他模块被启用，诸如，语音检测模块 108、语音处理模块 110、应用处理模块 112 和 / 或网络接口模块 206，那么音频输入可存储到容量相对较大的软件存储器缓冲区。在一些实施例中，存储器缓冲模块 104 包含环形缓冲区，其中音频输入可按照音频输入模块 208 获得其的顺序进行记录和重写。

[0021] 音频检测模块 106 可处理音频输入，以确定音频输入的能量水平。在一些实施例中，音频检测模块 106 包括低功率数字信号处理器（或其他类型的处理器），所述处理器配置来确定所获得的音频输入的能量水平（诸如，音量、强度、振幅等）并且将音频输入的能量水平与能量水平阈值相比较。能量水平阈值可根据用户输入来设置，或者可由功率管理子系统 100 自动设置，如下文相对于图 3 进一步论述。在一些实施例中，音频检测模块 106 还配置来确定音频输入的能量水平至少在阈值持续时间内满足阈值。在此类实施例中，持续时间相对较短的高能量音频输入可对应于相对不可能包含语音的突发噪音，它们可被忽略并且不被功率管理子系统 100 的其他部件处理。

[0022] 如果音频检测模块 106 确定所获得的音频输入的能量水平满足能量水平阈值，那么它可与功率管理模块 120 通信，以指导功率管理模块 120 启用语音检测模块 108。或者，音频检测模块 106 可将能量水平传送到功率管理模块 120，并且功率管理模块 120 可将所述能量水平与能量水平阈值相比较（且任选地与阈值持续时间相比较），以确定是否启用语音检测模块 108。在另一替代方案中，音频检测模块 106 可直接与语音检测模块 108 通信以将其启用。任选地，响应于音频检测模块 106 确定音频输入的能量水平满足阈值，功率管理模块 120（或音频检测模块 106）可指导音频输入模块 208 提高其采样率（无论是以帧速率

Petitioner Exhibit 1004, Page 750 of 1518

还是位速率测量）。

[0023] 语音检测模块 108 可处理音频输入，以确定音频输入是否包含语音。在一些实施例中，语音检测模块 108 包括低功率数字信号处理器（或其他类型的处理器），所述处理器配置来实施一个或多个技术，以便确定音频输入是否包含语音。在一些实施例中，语音检测模块 108 应用话音活动检测 (VAD) 技术。此类技术可基于音频输入的各个量化方面确定语音是否存在于音频输入中，所述量化方诸如，音频输入的一个或多个帧之间的频谱斜率；一个或多个频谱带中的音频输入的能量水平；一个或多个频谱带中的音频输入的信噪比；或其他量化方面。在其他实施例中，语音检测模块 108 实施受限分类器，所述分类器配置来将语音与背景噪音区分开来。所述分类器可由诸如线性分类器、支持向量机以及决策树的技术实施。在另外的实施例中，语音检测模块 108 应用隐马尔可夫模型 (Hidden Markov Model，HMM) 或高斯混合模型 (Gaussian Mixture Model，GMM) 技术，以将音频输入与一个或多个声学模型相比较，所述声学模型可包括对应于语音、噪音（诸如环境噪音或背景噪音）或静默的模型。其他技术可用来确定语音是否存在于音频输入中。

[0024] 使用上述任一技术，语音检测模块 108 均可确定一个分数或置信度水平，其值对应于语音实际存在于音频输入中的可能性（本文中所用的“可能性”可指常见用法，不管某件事是可能的还是统计上的用法）。如果分数满足阈值，那么语音检测模块 108 可确定语音存在于音频输入中。然而，如果分数不满足阈值，那么语音检测模块 108 可确定音频输入中没有语音。

[0025] 语音检测模块 108 可将其关于语音是否存在于音频输入中的确定传送到功率管理模块 120。如果语音存在于音频输入中，那么功率管理模块 120 可启用语音处理模块 110（或者，语音检测模块 108 可直接与语音处理模块 110 通信）。如果语音不存在于音频输入中，那么功率管理模块 120 可停用语音检测模块 108。或者，语音检测模块 108 可将分数传送到功率管理模块 120，于是功率管理模块 120 可确定是启用语音处理模块 110 还是停用语音检测模块 108。

[0026] 语音处理模块 110 可处理音频输入，以确定语音中是否包含关键字。在一些实施例中，语音处理模块 110 包括微处理器，所述微处理器配置来检测语音中的关键字，诸如，唤醒字或休眠字。语音处理模块 110 可配置来使用 HMM 技术、GMM 技术或其他语音识别技术来检测关键字。

[0027] 通过确定关键字是否是在一个或多个音素或词语之前或之后立即说出来的，语音处理模块 110 能够将意外包含关键字的语音与有意说出的关键字分开。例如，如果关键字是“ten”，那么语音处理模块 110 能够将独立地说出“ten”的用户与意外说出“ten”（作为词语“Tennessee”、词语“forgotten”、词语“stent”或短语“ten bucks”的一部分）的用户区分开。

[0028] 语音处理模块 110 还可配置来确定语音是否与包括功率管理子系统 100 的计算装置的特定用户相关联，或者语音是否对应于背景噪音；电视音频；音乐；或者用户以外的人的语音，以及其他分类。这一功能可用以下技术实施：诸如线性分类器、支持向量机和决策树，以及用于对音频输入进行分类的其他技术。

[0029] 使用上述任一技术，语音处理模块 110 均可确定一个分数或置信度水平，其值对应于关键字实际存在于语音中的可能性。如果分数满足阈值，那么语音处理模块 110 可确

Petitioner Exhibit 1004, Page 751 of 1518

定关键字存在于语音中。然而，如果分数不满足阈值，那么语音处理模块 110 可确定语音中没有关键字。

[0030] 语音处理模块 110 可将其关于关键字是否存在于语音中的确定传送到功率管理模块 120。如果关键字存在于语音中并且该关键字是唤醒字，那么功率管理模块 120 可启用应用处理模块 112 和网络接口模块 206（或者，语音处理模块 110 可直接与这些其他模块通信）。如果音频输入中不存在关键字（或者关键字是休眠字），那么功率管理模块 120 可停用语音处理模块 110 和语音检测模块 108。或者，语音处理模块 110 可将分数传送到功率管理模块 120，于是功率管理模块 120 可确定是启用应用处理模块 112 和网络接口模块 206，还是停用语音处理模块 110 和语音检测模块 108。在一些实施例中，只有在语音处理模块 110 确定特定用户说出了包含关键字的语音时，才会发生这些启用和 / 或停用。

[0031] 应用处理模块 112 可包括配置来实施语音识别应用的微处理器，所述语音识别应用是由包含功率管理子系统的计算装置提供的。语音识别应用可以包括可能需要语音识别的任何应用，诸如听写应用、消息应用、智能代理前端应用或任何其他应用。语音识别应用也可配置来将语音格式化（例如，将语音压缩），以便通过网络传输到远程计算装置，诸如语音识别服务器。

[0032] 在一些实施例中，应用处理模块 112 包括专用微处理器，用于实施语音识别应用。在其他实施例中，应用处理模块 112 包括通用微处理器，所述微处理器也可实施由包括功率管理子系统 100 的计算装置提供的其他软件，诸如图 2 所示的处理单元 202，将在下文进一步论述。

[0033] 网络接口模块 206 在被启用时可通过一个或多个有线或无线网络提供连接。启用之后，网络接口模块 206 可通过网络将记录到存储器缓冲模块 104 的所接收音频输入传输到远程计算装置，诸如语音识别服务器。远程计算装置可将识别结果（例如，智能代理查询的抄本或响应）返回到包括网络接口模块 206 的计算装置，于是网络接口模块 206 可将接收的识别结果提供给应用处理模块 112 以便处理。下文相对于图 2 进一步论述网络接口模块 206。

[0034] 在不脱离本公开范围的情况下，功率管理子系统 100 的模块可进行组合或重新布置。上述任何模块的功能均可分配到多个模块，或与不同的模块相组合。如上所述，任何或所有模块均可实现在一个或多个集成电路中、一个或多个通用微处理器中，或者实现在一个或多个专用数字信号处理器或其他专用微处理硬件中。一个或多个模块也可实现在由计算装置中包括的处理单元 202 实施的软件中，如下文相对于图 2 进一步论述。此外，一个或多个模块可从功率管理子系统 100 中整个省去。

[0035] 现在转到图 2，示出了其中可包括功率管理子系统 100 的用户计算装置 200。用户计算装置 200 包括处理单元 202；非瞬时计算机可读媒体驱动器 204；网络接口模块 206；如图 1 所示的功率管理子系统 100；以及音频输入模块 208，它们均可借助通信总线而互相通信。用户计算装置 200 也可包括电源 218，所述电源可为用户计算装置 200 的各种部件提供电力，诸如处理单元 202；非瞬时计算机可读媒体驱动器 204；网络接口模块 206；如图 1 所示的功率管理子系统 100；以及音频输入模块 208。

[0036] 处理单元 202 可包括一个或多个通用微处理器，所述微处理器配置来与存储器 210 来回通信，以实施存储于其中的各种软件模块，诸如用户接口模块 212，操作系统 214 以

Petitioner Exhibit 1004, Page 752 of 1518

及语音识别应用模块216。处理单元202还可与功率管理子系统100通信，并且可还实施在软件中实现的功率管理子系统100的任何模块。因此，处理单元202可配置来实施任何或所有以下模块：音频检测模块106；语音检测模块108；语音处理模块110；应用处理模块112以及功率管理模块120。此外，处理单元202可配置来实施装置上的自动语音识别能力，用户计算装置200可具有所述能力。

[0037] 存储器210通常包括RAM、ROM和/或其他持续或非瞬时计算机可读存储媒体。用户接口模块212可配置来借助用户计算装置200的显示器（未示出）来呈现用户接口。用户接口模块212还可配置来处理经由用户输入装置（未示出）接收的用户输入，诸如鼠标、键盘、触摸屏、小键盘等。由用户接口模块212呈现的用户接口可为用户提供机会来定制功率管理子系统100的操作和/或用户计算装置200实施的其他操作。下文将相对于图5进一步论述用户接口的实例。存储器210可另外存储操作系统214，所述操作系统提供计算机程序指令，以便由处理单元202用在用户计算装置200的一般管理和操作中。存储器210还可包含计算机程序指令，应用处理模块112和/或处理单元202执行所述计算机程序指令，以便实施语音识别应用模块216的一个或多个实施例。如上所述，语音识别应用模块216可以是可使用语音识别结果的任何应用，诸如听写应用、消息应用、智能代理应用前端；或可有利地使用语音识别结果的任何其他应用。在一些实施例中，存储器210还可包含可由处理单元202实施的自动语音识别引擎（未示出）。

[0038] 非瞬时计算机可读媒体驱动器204可包括本领域已知的任何电子数据存储器。在一些实施例中，非瞬时计算机可读媒体驱动器204存储一个或多个关键字模型（例如，唤醒字模型或休眠字模型），其中功率管理子系统100可将音频输入与所述关键字模型相比较。非瞬时计算机可读媒体驱动器204还可存储一个或多个声学模型和/或语言模型，用于实施用户计算装置200的任何装置上的语音识别能力。关于语言模型和声学模型的其他信息可见2012年8月16日提交的标题为“判别式语言模型修剪(DISCRIMINATIVE LANGUAGE MODEL PRUNING)”美国专利申请号13/587,799，以及2012年8月22日提交的标题为“无监督声学模型训练(UNSUPERVISED ACOUSTIC MODEL TRAINING)”的美国专利申请号13/592,157。这两个申请的公开内容均以全文引用的方式并入本文中。

[0039] 网络接口模块206可将用户计算装置200连接到一个或多个网络，诸如网络410，将在下文相对于图4A、图4B和图4C进一步论述。处理单元202和功率管理子系统100因而可接收来自远程计算装置的指令和信息，所述远程计算装置也可通过网络410进行通信，诸如语音识别服务器420，如下文也进一步论述。在一些实施例中，网络接口模块206包括无线网络接口，所述无线网络接口通过一个或多个无线网络为用户计算装置200提供连接。

[0040] 在一些实施例中，网络接口模块206被选择性地启用。当网络接口模块206处于停用或“休眠”状态时，它可受限制地连接到或不连接到网络或计算系统，以便省电。在一些实施例中，网络接口模块206默认处于停用状态，并且响应于来自功率管理子系统100的信号而变得启用。当网络接口模块206处于启用状态时，它可相对大量地连接到网络或计算系统，因此网络接口模块206使得用户计算装置200能够将音频输入发送到远程计算装置和/或从远程计算装置（例如，语音识别服务器420）接收关键字确认、语音识别结果或者停用指令。

Petitioner Exhibit 1004, Page 753 of 1518

[0041] 在一个特定的非限制性实例中，响应于功率管理子系统 100 确定音频输入包含唤醒字，网络接口模块 206 可被启用。功率管理子系统 100 可使音频输入经由启用的网络接口模块 206 而传输到远程计算装置（诸如语音识别服务器 420）。任选地，在使随后接收的音频输入传输到远程计算装置之前，功率管理子系统 100 可获得来自远程计算装置的唤醒字的确认。响应于接收到来自远程计算装置的停用指令，响应于确定自从获得满足能量水平阈值的音频输入之后已过了至少预定量的时间，或者响应于接收到包含休眠字的音频输入，功率管理子系统 100 之后可将启用的网络接口模块 206 停用。

[0042] 音频输入模块 208 可包括音频输入装置，诸如麦克风或麦克风阵列，不管是模拟的还是数字的。麦克风或麦克风阵列可实施为定向麦克风或麦克风定向阵列。在一些实施例中，音频输入模块 208 接收音频并且将音频提供到功率管理子系统 100 进行处理，大体上如上文相对于图 1 所论述。音频输入模块 208 还可接收来自功率管理子系统 100 的指令，以设置用于获得音频的采样率（无论是帧速率还是位速率）。音频输入模块 208 也可（或改为）包括一个或多个压电元件和 / 或微机电系统 (MEMS)，其可将声能转换成电信号，以便由功率管理子系统 100 处理。音频输入模块 208 还可配有放大器、整流器以及所需的其他音频处理部件。

[0043] 用户计算装置 200 可配有一个或多个另外的输入装置，诸如光传感器、位置传感器、图像捕获装置等。图 2 中并未示出此类另外的输入装置，以免模糊本公开的原理。在一些实施例中，另外的输入装置可检测条件的发生或不发生。关于此类条件的信息可被提供到功率管理子系统 100，以确定用户计算装置 200 或功率管理子系统 100 的一个或多个部件应被启用还是停用。在一个实施例中，另外的输入装置包括配置来检测光水平的光传感器。在光传感器检测到的光水平不满足阈值时，功率管理模块 120 才可充当网络接口模块 206、才可被启用。在另一实施例中，另外的输入装置包括配置有面部识别能力的图像捕获装置。在此实施例中，只有在图像捕获装置识别与用户计算装置 200 相关联的用户的面部时，网络接口模块 206 才可被启用。关于用输入装置控制语音识别能力的更多信息可见 2002 年 1 月 30 日提交的标题为“自动语音识别系统和方法 (AUTOMATIC SPEECH RECOGNITION SYSTEM AND METHOD)”的美国专利申请号 10/058,730，该申请在 2003 年 7 月 31 日被公开为美国专利公开号 2003/0144844，其公开内容以全文引用的方式并入本文中。关于控制语音识别能力的进一步信息可见 2012 年 12 月 4 日颁布的标题为“使用物理现象检测器来控制语音识别引擎的操作 (USING A PHYSICAL PHENOMENON DETECTOR TO CONTROL OPERATION OF A SPEECH RECOGNITION ENGINE)”的美国专利号 8,326,636。该专利的公开内容以全文引用的方式并入本文中。

[0044] 可提供另外的输入装置，其可包含用户输入装置，诸如，鼠标、键盘、触摸屏、小键盘等。同样，可提供输出装置，诸如，显示器、扬声器、耳机等。在一个特定实例中，用户计算装置 200 可包括一个或多个输出装置，其配置来采用音频形式（例如，通过文本到语音）或视觉形式（例如，通过显示器）呈现语音识别结果。此类输入和输出装置已为本领域所熟知，并且无需在本文中进一步详细论述，且未在图 2 中示出，以避免模糊本发明的原理。

[0045] 电源 218 可向用户计算装置 200 的各个部件提供电力。电源 218 可包括无线或便携式电源，诸如，一次性或可充电电池或电池组；或者可包括有线电源，诸如，配置来插入插座中的交流 (AC) 电源。在一些实施例中，电源 218 将其可供应的功率水平传送到功率管理

Petitioner Exhibit 1004, Page 754 of 1518

子系统 100（例如，剩余电池寿命的百分数、电源 218 是否插入插座等）。在一些实施例中，至少部分基于电源所指示的功率水平，功率管理子系统 100 选择性地启用或停用一个或多个模块。例如，如果用户计算装置 200 插入插座中，那么功率管理子系统 100 可启用网络接口模块 206 并让其处于启用状态。如果用户计算装置 200 使用电池电力，那么功率管理子系统 100 可选择性地启用和停用网络接口模块 206，如上所述。

[0046] 现在转到图 3，示出了说明性例行程序 300，其中功率管理子系统 100 的模块可被选择性地启用以处理音频输入。当说明性例行程序 300 中随后启用的模块可相对大量地处理需求和 / 或功率消耗时，说明性例行程序 300 代表处理和 / 或功率消耗的升级。

[0047] 说明性例行程序 300 可在框 302 处开始，此时，音频输入模块 208 监测音频输入。在框 304 处，音频输入模块 208 可接收音频输入。在框 306 处，接收的音频输入可被记录到存储器缓冲模块 104。在框 308 处，音频检测模块 106 可确定音频输入的能量水平是否满足能量水平阈值（以及任选地，音频输入的能量水平至少是否在阈值持续时间内满足能量水平阈值）。如果音频输入的能量水平并不满足能量水平阈值，那么在框 310 中，音频输入模块 208 可继续监测音频输入，直至接收到另一音频输入为止。

[0048] 返回到框 308，如果音频检测模块 106 确定音频输入的能量水平满足阈值，那么在框 312 处，功率管理模块 120 可启用语音检测模块 108（或者，音频检测模块 106 可直接启用语音检测模块 108，并且在接下来的框中，功率管理模块 120 也可省去）。在框 314 处，语音检测模块 108 可确定语音是否存在于所获得的音频输入中，大体如上文相对于图 1 所论述。如果语音检测模块 108 确定语音不存在于（或不可能存在于）音频输入中，那么在框 316 处，功率管理模块 120 可停用语音检测模块 108。在框 310 中，音频输入模块 208 随后可继续监测音频输入，直至接收到另一音频输入为止。

[0049] 返回到框 314，如果语音检测模块 108 确定音频输入包含语音，那么在框 318 处，功率管理模块 120 可启用语音处理模块 110。如上所述，在框 320 处，语音处理模块 110 可确定唤醒字是否存在于语音中。如果语音处理模块 110 确定唤醒字不存在于语音中（或不可能存在于语音中），那么在框 322 处，语音处理模块 110 可被停用。语音检测模块 108 也可在框 316 处被停用。在框 310 中，音频输入装置 208 随后可继续监测音频输入，直至接收到另一音频输入为止。

[0050] 返回到框 320，如果在一些实施例中，语音处理模块 110 确定唤醒字存在于语音中，那么用户 401，在框 324 中，语音处理模块 110 任选地确定语音是否与特定用户相关联（例如，用户是否说出唤醒字），大体如上文相对于图 1 所论述。如果语音并不与特定用户相关联，那么在框 322 处，语音处理模块 110 可被停用。语音检测模块 108 也可在框 316 处被停用。在框 310 中，音频输入装置 208 随后可继续监测音频输入，直至接收到另一音频输入为止。如果语音与特定用户相关联，那么说明性例行程序 300 可进行到框 326。在其他实施例中，框 324 可被省去，并且响应于语音处理模块 110 确定唤醒字存在于语音中，说明性例行程序 300 可直接从框 320 进行到框 326。

[0051] 在框 326 处，功率管理模块 120 可启用应用处理模块 112，所述应用处理模块可实施图 2 所示的语音识别应用模块 216。在框 328 处，功率管理模块 120 也可启用网络接口模块 206。网络接口模块 206 被启用之后，记录到存储器缓冲模块 104 的音频输入可经由网络接口模块 206 通过网络进行传输。在一些实施例中，当网络接口模块 206 活动时，随后获得

Petitioner Exhibit 1004, Page 755 of 1518

的音频输入从音频输入模块 208 直接提供到应用处理模块 112 和 / 或网络接口模块 206，以用于传输到远程计算装置。然而，在其他实施例中，在将音频输入提供到网络接口模块 206 以便通过网络 410 传输到远程计算装置之前，语音检测模块 108、语音处理模块 110 以及应用处理模块 112 中的任一个或全部均对所述音频输入进行处理。

[0052] 在未示出的一些实施例中，功率管理子系统 100 等待，直到远程计算装置返回唤醒字存在于第一音频输入中的确认，以便传输随后的音频输入用于识别。如果远程计算装置没有提供唤醒字的确认，或者如果经由网络接口模块 206 接收到停用指令，那么网络接口模块 206 以及功率管理子系统 100 的一个或多个模块可被停用。

[0053] 由于功率管理子系统 100 的许多操作可生成概率确定而非准确确定，因此，在说明性例行程序 300 期间可能出现误差。在一些情况下，功率管理子系统 100 的特定模块可提供“误报”，从而导致一个或多个模块被错误地启用。例如，在框 314 处，语音检测模块 108 可错误地确定存在语音，或者在框 320 处，语音处理模块 110 可错误地确定语音包含唤醒字，或在框 324 处，错误地确定语音属于用户。功率管理子系统 100 的模块之间的自适应阈值和交叉验证可有利地用来减少误报。本文中论述了自适应阈值的两个实例，但其他类型的自适应阈值是可能的。如上所述，在框 314 处，语音检测模块可确定语音存在于音频输入中。然而，语音处理模块 110 由于其较好的处理能力而可比语音检测模块 108 更准确地识别语音，它可确定实际上音频输入中没有语音。因此，语音处理模块 110 可指导语音检测模块 108 提高用于确定语音存在于音频输入中的分数阈值，以便减少将来的误报。同样地，如果远程计算装置（诸如语音识别服务器 420）包含语音识别能力，那么远程计算装置可将语音中不存在唤醒字的指示传输到用户计算装置 200，即使语音处理模块 110 可能已经指示存在唤醒字。因此，用于确定唤醒字存在于语音中的语音处理模块 110 的分数阈值可以提高，以便减少将来的误报。此外，可提供用户接口，从而使得用户可提高一个或多个分数阈值来减少误报，如下文相对于图 5 进一步描述。

[0054] 在其他情况下，特定部件可提供“漏报”，从而功率管理子系统 100 的部件不被启用和 / 或网络接口模块 206 不被启用，即使用户已经说出唤醒字也是如此。例如，在框 314 处，语音检测模块 108 可错误地确定不存在语音，或者在框 320 处，语音处理模块 110 可错误地确定语音不包含唤醒字，或在框 324 处，错误地确定语音不属于用户。为了降低漏报的可能性，功率管理子系统 100 可定期降低阈值分数，例如，在框 314、320 和 / 或 324 中，降低满足阈值所需的分数。阈值可继续降低，直至获得一个或多个误报为止，如上所述。一旦获得一个或多个误报，阈值可不再进一步降低，或者可稍微提高。此外，可相应地提供用户接口，从而使得用户可降低一个或多个分数阈值来减少漏报，如下文相对于图 5 进一步描述。

[0055] 在一些实施例中，如果在框 314、320 或 324 中的任一框处获得否定结果，那么并非所有启用的部件均被停用。例如，如果唤醒字在框 320 处未被识别出，那么语音处理模块 110 可在框 322 处被停用，但语音检测模块 108 可保持启用。此外，在一些实施方案中，可跳过某些框。在一些实施例中，在框 314 或 320 处满足阈值的分数会促使一个或多个随后的框被跳过。例如，如果在框 320 处，语音处理模块 110 非常肯定地确定唤醒字存在于语音中，那么说明性例行程序 300 可直接跳到框 326

[0056] 此外，在一些实施例中，用户计算装置 200 可包括被配置成由处理单元 202 执行的自动语音识别引擎。由于此类装置上的语音识别的功率消耗可能尤其高，因此，响应于语音

Petitioner Exhibit 1004, Page 756 of 1518

处理模块 110 确定用户已经说出唤醒字，处理单元 202 可只实施自动语音识别引擎来识别语音。

[0057] 现在参考图 4A、图 4B 和图 4C，说明性环境 400 中示出了分布式语音识别服务的示例性操作。环境 400 可包括用户 401；如上所述的用户计算装置 200；网络 410；语音识别服务器 420；以及数据暂存器 430。

[0058] 网络 410 可以是任何有线网络、无线网络或其组合。此外，网络 410 可以是个人区域网、局域网、广域网、电缆网络、卫星网络、蜂窝电话网或其组合。计算机通信领域的技术人员已熟知借由因特网或上述任何其他类型的通信网络进行通信的协议和装置，并且因此，无需在本文中更详细地描述。

[0059] 语音识别服务器 420 通常可以是能够通过网络 410 进行通信的任何计算装置。在一些实施例中，语音识别服务器 420 被实施为一个或多个服务器计算装置，但其他实施方式是可能的。语音识别服务器 420 可能能够通过网络 410 从用户计算装置 200 接收音频输入。此音频输入可用多种方式进行处理，取决于语音识别服务器 420 的实施方式。在一些实施例中，语音识别服务器 420 处理从用户计算装置 200 接收的音频输入，以确认唤醒字存在（例如，通过将音频输入与唤醒字的已知模型相比较），并且将确认传输到用户计算装置 200。语音识别服务器 420 还可配置来使用已知的说话人鉴别技术来鉴别说出唤醒字的用户 401。

[0060] 语音识别服务器 420 可处理从用户计算装置 200 接收的音频输入，以根据音频输入来确定语音识别结果。例如，音频输入可包含用于智能代理来处理的口头查询；将被转录成文本的语音；或者适用于语音识别应用的其他语音。语音识别服务器 420 可通过网络 410 将语音识别结果传输到用户计算装置 200。关于分布式语音识别应用的进一步信息可见于 2012 年 2 月 14 日颁布的标题为“用于无线装置的托管话音识别系统 (Hosted voice recognition system for wireless devices)”的美国专利号 8,117,268，该专利的公开内容以全文引用的方式并入本文中。

[0061] 语音识别服务器 420 可与数据暂存器 430 本地或远程通信。数据暂存器 430 可实现为以下形式：硬盘驱动器、固态存储器，和 / 或语音识别服务器 420 可存取的任何其他类型的非瞬时计算机可读存储媒体。在不脱离本公开的精神和范围的情况下，数据存储器 430 也可以是本领域已知的分布式或跨分区的多个存储装置。此外，在一些实施例中，数据暂存器 430 被实施为基于网络的电子存储服务。

[0062] 数据暂存器 430 可包括一个或多个唤醒字模型。在一些实施例中，唤醒字模型专用于用户 401，而在其他实施例中，接收到用户计算装置 200 确定包含唤醒字的音频输入之后，语音识别服务器可将音频输入与数据暂存器 430 中存储的唤醒字的已知模型相比较。如果音频输入极其类似于已知模型，那么语音识别服务器 420 可将唤醒字的确认传输到用户计算装置 200，于是用户计算装置 200 可获得待由语音识别服务器 420 处理的其他音频输入。

[0063] 数据暂存器 430 还可包含用于语音识别的一个或多个声学和 / 或语言模型。这些模型可包括通用模型，以及专用模型。模型可专用于用户 401；专用于用户计算装置 200 和 / 或语音识别服务器 420 实施的语音识别应用；或者可具有其他特定用途。关于语言模型和声学模型的其他信息可见 2012 年 8 月 16 日提交的标题为“判别式语言模型

Petitioner Exhibit 1004, Page 757 of 1518

修剪(DISCRIMINATIVE LANGUAGE MODEL PRUNING)”的美国专利申请号13/587,799,以及2012年8月22日提交的标题为“无监督声学模型训练(UNSUPERVISED ACOUSTIC MODEL TRAINING)”的美国专利申请号13/592,157。这两个申请案的公开内容先前以引用的方式并入。

[0064] 数据暂存器430还可包含数据,该数据响应于语音识别服务器420接收的音频输入中含有的查询。语音识别服务器420可识别音频输入中包含的语音,鉴别语音中包含的查询,并且处理查询以鉴别数据暂存器430中的响应数据。语音识别服务器420随后可通过网络410将包含响应数据的智能代理响应提供到用户计算装置200。数据暂存器430中还可包含其他数据。

[0065] 将认识到,很多上述装置都是任选的,而且环境400的实施例可以或可不结合装置。此外,装置无需不同或离散。装置也可以在环境400中重新组织。例如,语音识别服务器420可表示为单个物理服务器计算装置,或者可以分成实现本文所述功能的多个物理服务器。此外,用户计算装置200可具有语音识别服务器420的一些或所有语音识别功能。

[0066] 另外,应注意,在一些实施例中,用户计算装置200和/或语音识别服务器420可由实施在托管计算环境中的一个或多个虚拟机执行。托管计算环境可包含一个或多个快速配置和释放的计算资源,所述计算资源可包括计算装置、网络装置和/或存储装置。托管计算环境也可称为云计算环境。托管计算环境的一个或多个计算装置可包括如上所述的功率管理子系统100。

[0067] 具体参考图4A,示出了可确认唤醒字的说明性操作。用户401可能说出唤醒字502。用户计算装置200可获得可能包含用户语音的音频输入(1),并且确定唤醒字402存在于语音中(2),大体如上文相对于图3所论述。音频输入还可包含话音命令或查询。响应于确定语音包含唤醒字,应用处理模块112和用户计算装置200的网络接口模块206可被启用(3),并且音频输入通过网络410传输到语音识别服务器420(4)。语音识别服务器420可确认唤醒字存在于音频输入中(5),并且可通过网络410将确认传输到用户计算装置200(6)。

[0068] 现转到图4B,响应于接收来自语音识别服务器420的唤醒字的确认,用户计算装置200可继续获得音频输入(7),所述音频输入将被提供到语音识别服务器420以用于处理。例如,获得的音频输入可包含智能代理查询404,用于由语音识别服务器420处理。或者,获得的音频输入可包含将被语音识别服务器420转录的语音(例如,用于由应用处理模块112执行的听写、文字处理或消息应用)。用户计算装置200可通过网络410将音频输入传输到语音识别服务器420(8)。任选地,将生成语音识别结果的语音识别应用的识别符可被提供给语音识别服务器420,因此,语音识别服务器420可生成结果,所述结果具体用于应用处理模块112实施的语音识别应用。语音识别服务器420可识别音频输入中包含的语音(9),并且由此生成语音识别结果(10)。例如,语音识别结果可包含语音的抄本、语音中包含的查询的智能代理响应,或任何其他类型的结果。这些语音识别结果可通过网络410从语音识别服务器420传输到用户计算装置200(11)。响应于接收到结果,应用处理模块112可使以声音形式(例如,通过文本到语音)或以视觉形式(例如,通过用户计算装置200的显示器)来呈现结果(12)。

[0069] 现在参考图4C,用户计算装置200可继续获得音频输入(13),所述音频输入将被

Petitioner Exhibit 1004, Page 758 of 1518