IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

First Named Inventor	:	Matthew Sharifi	Art Unit	:	2658
Serial No.	:	14/675,932	Examiner	:	Michael C. Colucci
Filed	:	April 1, 2015	Conf. No.	:	4235
Title	:	HOTWORD DETECTION ON MULTIPLE DEVICES			

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

AMENDMENT

AND

INTERVIEW SUMMARY

AMENDMENTS TO THE CLAIMS

This Listing of Claims replaces all previous versions, and listings, of claims in the application.

Listing of Claims:

1. (Currently amended) A computer-implemented method comprising:

receiving, by a first computing device, audio data that corresponds to an utterance;

determining, based on an analysis of the audio data and without performing automated speech recognition processing on the audio data, a first value ~~corresponding to~~ that reflects a first likelihood that the utterance includes a particular hotword;

receiving a second value ~~corresponding to~~ that reflects a second likelihood that the utterance includes the particular hotword, ~~the second value being~~ as determined by a second computing device;

transmitting, to the second device, the first value that reflects the first likelihood that the utterance includes the particular hotword;

comparing the first value that reflects the first likelihood that the utterance includes the particular hotword and the second value that reflects the second likelihood that the utterance includes the particular hotword; and

based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword to the second value that reflects the second likelihood that the utterance includes the particular hotword, ~~initiating~~ determining whether to perform automated speech recognition processing on the audio data.

2. (Currently amended) The method of claim 1, comprising:

determining that the first value satisfies a hotword score threshold,

wherein the first value that reflects the first likelihood that the utterance includes the particular hotword is transmitted to the second device based on determining that the first value satisfies the hotword score threshold.

3. (Cancelled)

4. (Currently amended) The method of claim 1, comprising:

determining an activation state of the first computing device based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword and the second value that reflects the second likelihood that the utterance includes the particular hotword.

5. (Currently amended) The method of claim 4, wherein determining an activation state of the first computing device based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword and the second value that reflects the second likelihood that the utterance includes the particular hotword comprises:

determining that the activation state is an active state.

6. (Currently amended) The method of claim 1, comprising:

receiving, by the first computing device, additional audio data that corresponds to an additional utterance;

determining, based on an analysis of the additional audio data and without performing automated speech recognition processing on the additional audio data a third value ~~corresponding to~~ that reflects a third likelihood that the additional utterance includes the particular hotword;

receiving a fourth value ~~corresponding to~~ that reflects a fourth likelihood that the additional utterance includes the particular hotword, ~~the fourth value being~~ as determined by a third computing device;

comparing the third value that reflects the third likelihood that the additional utterance includes the particular hotword and the fourth value that reflects the fourth likelihood that the additional utterance includes the particular hotword; and

based on comparing the third value that reflects the third likelihood that the additional utterance includes the particular hotword and the fourth value that reflects the fourth likelihood

that the additional utterance includes the particular hotword, determining that the activation state of the first computing device is an inactive state.

7. (Currently amended) The method of ~~claim 3~~ claim 1, wherein:

transmitting, to the second computing device, the first value ~~to the second computing device~~ that reflects the first likelihood that the utterance includes the particular hotword comprises:

transmitting, to a server, through a local network, or through a short range radio, the first value that reflects the first likelihood that the utterance includes the particular hotword, and

receiving a second value ~~corresponding to~~ that reflects a second likelihood that the utterance includes the particular hotword, ~~the second value being~~ as determined by a second computing device comprises:

receiving, from the server, through the local network, or through the short range radio, a second value that ~~was determined by a second computing device~~ reflects the second likelihood that the utterance includes the particular hotword.

8. (Currently amended) The method of claim 1, comprising:

identifying the second computing device; and

determining that the second computing device is configured to respond to utterances that include the particular hotword.

9. (Currently amended) The method of ~~claim 3~~ claim 1, wherein:

transmitting, to the second computing device, the first value ~~to the second computing device~~ that reflects the first likelihood that the utterance includes the particular hotword comprises:

transmitting a first identifier ~~for the~~ of the first computing device, and

receiving a second value ~~corresponding to~~ that reflects a second likelihood that the utterance includes the particular hotword, ~~the second value being~~ as determined by a second computing device comprises:

receiving a second identifier ~~for the~~ of the second computing device.

10. (Original) The method of claim 5, wherein determining that the activation state is an active state comprises:

determining that a particular amount of time has elapsed since receiving the audio data that corresponds to the utterance.

11. (Currently amended) The method of claim 5, comprising:

continuing, for a particular amount of time, to transmit the first value that reflects the first likelihood that the utterance includes the particular hotword based on determining that the activation state is an active state.

12. (Currently amended) A computing device comprising:

one or more storage devices storing instructions that are operable, when executed by the computing device, to cause the computing device to perform operations comprising:

receiving, by a first computing device, audio data that corresponds to an utterance;

determining, based on an analysis of the audio data and without performing automated speech recognition processing on the audio data, a first value ~~corresponding to~~ that reflects a first likelihood that the utterance includes a particular hotword;

receiving a second value ~~corresponding to~~ that reflects a second likelihood that the utterance includes the particular hotword, ~~the second value being~~ as determined by a second computing device;

transmitting, to the second device, the first value that reflects the first likelihood that the utterance includes the particular hotword;

comparing the first value that reflects the first likelihood that the utterance includes the particular hotword and the second value that reflects the second likelihood that the utterance includes the particular hotword; and

based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword to the second value that reflects the second likelihood

that the utterance includes the particular hotword, ~~initiating~~ determining whether to perform automated speech recognition processing on the audio data.

13. (Currently amended) The ~~system~~ device of claim 12, wherein the operations further comprise:

determining that the first value satisfies a hotword score threshold,

wherein the first value that reflects the first likelihood that the utterance includes the particular hotword is transmitted to the second device based on determining that the first value satisfies the hotword score threshold.

14. (Cancelled)

15. (Currently amended) The ~~system~~ device of claim 12, wherein the operations further comprise:

determining an activation state of the first computing device based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword and the second value that reflects the second likelihood that the utterance includes the particular hotword.

16. (Currently amended) The ~~system~~ device of claim 15, wherein determining an activation state of the first computing device based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword and the second value that reflects the second likelihood that the utterance includes the particular hotword comprises:

determining that the activation state is an active state.

17. (Currently amended) The ~~system~~ device of claim 12, wherein the operations further comprise:

receiving, by the first computing device, additional audio data that corresponds to an additional utterance;

determining, based on an analysis of the additional audio data and without performing automated speech recognition processing on the additional audio data a third value ~~corresponding to~~ that reflects a third likelihood that the additional utterance includes the particular hotword;

receiving a fourth value ~~corresponding to~~ that reflects a fourth likelihood that the additional utterance includes the particular hotword, ~~the fourth value being~~ as determined by a third computing device;

comparing the third value that reflects the third likelihood that the additional utterance includes the particular hotword and the fourth value that reflects the fourth likelihood that the additional utterance includes the particular hotword; and

based on comparing the third value that reflects the third likelihood that the additional utterance includes the particular hotword and the fourth value that reflects the fourth likelihood that the additional utterance includes the particular hotword, determining that the activation state of the first computing device is an inactive state.

18. (Currently amended) The ~~system~~ device of claim 14, wherein:

transmitting, to the second computing device, the first value ~~to the second computing device~~ that reflects the first likelihood that the utterance includes the particular hotword comprises:

transmitting, to a server, through a local network, or through a short range radio, the first value that reflects the first likelihood that the utterance includes the particular hotword, and

receiving a second value ~~corresponding to~~ that reflects a second likelihood that the utterance includes the particular hotword, ~~the second value being~~ as determined by a second computing device comprises:

receiving, from the server, through the local network, or through the short range radio, a second value that ~~was determined by a second computing device~~ reflects the second likelihood that the utterance includes the particular hotword.

19. (Currently amended) The ~~system~~ device of claim 12, wherein the operations further comprise:

identifying the second computing device; and

determining that the second computing device is configured to respond to utterances that include the particular hotword.

20. (Currently amended) A non-transitory computer-readable medium storing software comprising instructions executable by one or more computers which, upon such execution, cause the one or more computers to perform operations comprising:

receiving, by a first computing device, audio data that corresponds to an utterance;

determining, based on an analysis of the audio data and without performing automated speech recognition processing on the audio data, a first value ~~corresponding to~~ that reflects a first likelihood that the utterance includes a particular hotword;

receiving a second value ~~corresponding to~~ that reflects a second likelihood that the utterance includes the particular hotword, ~~the second value being~~ as determined by a second computing device;

transmitting, to the second device, the first value that reflects the first likelihood that the utterance includes the particular hotword;

comparing the first value that reflects the first likelihood that the utterance includes the particular hotword and the second value that reflects the second likelihood that the utterance includes the particular hotword; and

based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword to the second value that reflects the second likelihood that the utterance includes the particular hotword, ~~initiating~~ determining whether to perform automated speech recognition processing on the audio data.

21. (New) The method of claim 1, comprising:

based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword to the second value that reflects the second likelihood that the

utterance includes the particular hotword, determining that the first value that reflects the first likelihood that the utterance includes the particular hotword is greater than the second value that reflects the second likelihood that the utterance includes the particular hotword,

wherein determining whether to perform automated speech recognition processing on the audio data comprises:

based on determining that the first value that reflects the first likelihood that the utterance includes the particular hotword is greater than the second value that reflects the second likelihood that the utterance includes the particular hotword, determining to perform automated speech recognition processing on the audio data.

22. (New) The method of claim 1, comprising:

based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword to the second value that reflects the second likelihood that the utterance includes the particular hotword, determining that the first value that reflects the first likelihood that the utterance includes the particular hotword is less than the second value that reflects the second likelihood that the utterance includes the particular hotword,

wherein determining whether to perform automated speech recognition processing on the audio data comprises:

based on determining that the first value that reflects the first likelihood that the utterance includes the particular hotword is less than the second value that reflects the second likelihood that the utterance includes the particular hotword, determining not to perform automated speech recognition processing on the audio data.

REMARKS

Claims 1-20 were pending as of the action of June 12, 2015. Claims 1, 12, and 20 are in independent form and have been amended. Claims 3 and 14 have been cancelled. New claims 21 and 22 have been added. No new matter has been added.

Interview Summary

The undersigned thanks Examiner Colucci for the courtesies extended during the telephonic interview on August 12, 2015. The undersigned participated on behalf of the Applicant. In the interview, the Applicant's representatives discussed an overview of the application and the proposed feature of "based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword to the second value that reflects the second likelihood that the utterance includes the particular hotword, determining whether to perform automated speech recognition processing on the audio data." An agreement was reached that the proposed amendment overcomes the current rejection.

Rejections under 35 U.S.C. § 103

Claims 1-20 were rejected under 35 U.S.C. § 103(a) over US 7,904,297 (Mirkovic) and further in view of US 2011/0060587 (Phillips). Withdrawal of the rejection is requested.

Claim 1 has been amended to recite "receiving a second value that reflects a second likelihood that the utterance includes the particular hotword, as determined by a second computing device," and "based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword to the second value that reflects the second likelihood that the utterance includes the particular hotword, determining whether to perform automated speech recognition processing on the audio data." Claims 12 and 20 recite similar features. The cited references do not describe or suggest these features.

Mirkovic describes that "the ASR 920 and parser 922 components may be combined into a single ASR/Parser unit that is coupled to, but functionally independent of DMV classifier 924 and slot classifier 926," (Mirkovic, 17:57-60) which is not the same as "based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword to the second value that reflects the second likelihood that the utterance includes the particular

hotword, determining whether to perform automated speech recognition processing on the audio data" where the "second value that reflects a second likelihood that the utterance includes the particular hotword" is "determined by a second computing device," as recited by claim 1.

Phillips describes that the "ASR client 118 may facilitate speech-enabled text entry to each of the multiple applications," (Phillips, ¶ [0061]) which is not the same as "based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword to the second value that reflects the second likelihood that the utterance includes the particular hotword, determining whether to perform automated speech recognition processing on the audio data" where the "second value that reflects a second likelihood that the utterance includes the particular hotword" is "determined by a second computing device," as recited by claim 1.

Accordingly, the Applicant requests reconsideration and withdrawal of the rejection of claims 1, 12, and 20 and their dependent claims.

Conclusion

The other claims in the application are each dependent on the independent claims, and are allowable for at least the above reasons. Because each claim is deemed to define additional aspects of the disclosure, however, the individual consideration of each claim on its own merits is respectfully requested.

By responding in the foregoing remarks only to particular positions taken in the action, Applicant does not acquiesce with other positions that have not been explicitly addressed. In addition, Applicant's selecting some particular arguments for the patentability of a claim should not be understood as implying that no other reasons for the patentability of that claim exist. Finally, Applicant's decision to amend or cancel any claim should not be understood as implying that Applicant agrees with any positions taken in the action with reference to that claim or other claims, or that Applicant adopts or agrees with any position except as specifically stated in this paper. In particular, by amending a claim, Applicant does not necessarily concede that the claim prior to amendment was unpatentable.

Please charge any fees that need to be paid and apply any credits to deposit account 06-1050.

Respectfully submitted,

Date: August 19, 2015

/Sam Williams/
Sam L. Williams
Reg. No. 68,071

Customer Number 26192
Fish & Richardson P.C.
Telephone: (202) 626-6404
Facsimile: (877) 769-7945

41126384.docx

UNITED STATES PATENT AND TRADEMARK OFFICE

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

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
14/675,932	04/01/2015	Matthew Sharifi	16113-6450001	4235

26192 7590 10/01/2015

FISH & RICHARDSON P.C.
PO BOX 1022
MINNEAPOLIS, MN 55440-1022

EXAMINER	
COLUCCI, MICHAEL C	

ART UNIT	PAPER NUMBER
2658	

NOTIFICATION DATE	DELIVERY MODE
10/01/2015	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):

PATDOCTC@fr.com

Office Action Summary	Application No. 14/675,932	Applicant(s) SHARIFI, MATTHEW	
	Examiner MICHAEL COLUCCI	Art Unit 2658	AIA (First Inventor to File) Status Yes

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

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTHS FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *08/20/2015*.

☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

4)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims*

5)☒ Claim(s) *1,2,4-13 and 15-22* is/are pending in the application.

5a) Of the above claim(s) _____ is/are withdrawn from consideration.

6)☐ Claim(s) _____ is/are allowed.

7)☒ Claim(s) *1,2,4-13 and 15-22* is/are rejected.

8)☐ Claim(s) _____ is/are objected to.

9)☐ Claim(s) _____ are subject to restriction and/or election requirement.

* If any claims have been determined allowable, you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see http://www.uspto.gov/patents/init_events/pph/index.jsp or send an inquiry to PPHfeedback@uspto.gov.

Application Papers

10)☐ The specification is objected to by the Examiner.

11)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

Certified copies:

a)☐ All b)☐ Some** c)☐ None of the:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

** See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Information Disclosure Statement(s) (PTO/SB/08a and/or PTO/SB/08b) Paper No(s)/Mail Date _____.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

4) ☐ Other: _____.

1. The present application, filed on or after March 16, 2013, is being examined under the first inventor to file provisions of the AIA.

DETAILED ACTION

Response to Arguments

1. Applicant's arguments with respect to claims 1, 2, 4-13, and 15-22 have been considered but are moot in view of the new ground(s) of rejection. Applicant's arguments are directed to the amended subject matter; new prior art is provided below. While the amendments appear to overcome the prior art of record, Rosenberger US 8340975 B1 (hereinafter Rosenberger) has been incorporated to address the communication of probabilistic weighted signals from each device to each other, and comparing probabilistic values between devices, to determine whether to handle the speech command i.e. continue processing speech.

NOTE: During current analysis of the present invention it is realized that contradictory to the claim language, the specification in fact states that speech recognition is performed, which from a technical standpoint is almost inherent given the nature of command and control. In order to spot a keyword some sort of speech recognition must be performed. There is no indication of waveform analysis that would suggest otherwise in the specification. Therefore being consistent with both the spec and technological field and in view of prior art Rosenberger, it is believed the claims mean that audio input is processed with ASR but segmented into two portions, where a

first portion has speech recognition performed on it, however the remaining portion of the audio is not processed until probabilistic weighted values are compared between devices. The remaining portion of audio therefore does not have ASR performed on it as consistent with the claims. For instance a user states “Hello thermostat” where SR is performed, the following command afterwards “turn up 2 degrees” is not processed until a decision is made between the devices to process the remaining command. This is precisely taught in newly incorporated Rosenberger.

See present invention spec 0005-0006 below:

“Consequently, each time the words "OK computer" are spoken, it is picked up by a microphone, conveyed to the system, which performs speech recognition techniques to determine whether the hotword was spoken and, if so, awaits an ensuing command or query. Accordingly, utterances directed at the system take the general form [HOTWORD] [QUERY], where "HOTWORD" in this example is "OK computer" and "QUERY" can be any question, command, declaration, or other request that can be speech recognized, parsed and acted on by the system, either alone or in conjunction with the server via the network.

[0006] According to one innovative aspect of the subject matter described in this specification, a user device receives an utterance that is spoken by a user. The user device determines whether the utterance includes a hotword and computes a hotword confidence score that indicates a likelihood that the utterance includes the hotword. The user device transmits this score to other user devices in the near vicinity. The other user devices likely received the same utterance. The other user devices compute a hotword

confidence score and transmit their scores to the user device. The user device compares the hotword confidence scores. If the user device has the highest hotword confidence score, then the user device remains active and prepares to process additional audio. If the user device does not have the highest hotword confidence score, then the user device does not process the additional audio."

Claim Rejections - 35 USC § 103

1. The following is a quotation of 35 U.S.C. 103 which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

2. Claims 1, 2, 4-13, and 15-22 are rejected under 35 U.S.C. 103(a) as being unpatentable over Mirkovic; Danilo et al. US 7904297 B2 (hereinafter Mirkovic) in view of Phillips; Michael S. et al. US 20110060587 A1 (hereinafter Phillips) and further in view of Rosenberger US 8340975 B1 (hereinafter Rosenberger).

Re claims 1, 12, and 20, Mirkovic teaches

A computer-implemented method comprising:

receiving, by a first computing device, audio data that corresponds to an utterance;

determining, based on an analysis of the audio data and without performing automated speech recognition processing on the audio data, a first value that reflects a first likelihood that the utterance includes a particular hotword;

Phillips expressly teaches that other devices can be applications both local and external to produce results (Phillips fig. 1C).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Mirkovic to incorporate multiple devices as servers as taught by Phillips to allow for context based retrieval of best matches (Phillips fig. 1C), to associate the proper device with the best matching input.

However the combination fails to teach

receiving a second value that reflects a second likelihood that the utterance includes the particular hotword, as determined by a second computing device;

transmitting, to the second device, the first value that reflects the first likelihood that the utterance includes the particular hotword;

comparing the first value that reflects the first likelihood that the utterance includes the particular hotword and the second value that reflects the second likelihood that the utterance includes the particular hotword; and

based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword to the second value that reflects the second likelihood that the utterance includes the particular hotword, initiating determining whether to perform automated speech recognition processing on the audio data

Rosenberger expressly teaches handling an initial input speech to process, transmitting a probabilistic weighted value to other devices, comparing values of multiple devices, then determining whether to process the remaining portion of the input i.e. the command section based on whether an input value is higher than a threshold value associated with weighted probabilistic values of each device or one device depending on the number present (Rosenberger col 4 lines 5-53 with col 5 lines 1-50 with fig. 7 and claim 13).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Mirkovic and Phillips to incorporate the above limitations as taught by Rosenberger to allow for a modification and substitution of a resource saving system such as Rosenberger wherein speech recognition is not fully performed on the initial input such that each an every device is not compared, for instance A) “hello thermostat" versus B) "hello thermostat please raise the temperature 2 degrees” saves four times the amount of processing speed and resources per device, such that 8 devices can simultaneously handle A) wherein only 1 device can handle B) regardless of the amount of resources and speed available, wherein Mirkovic’s system can truncate input as a preamble i.e. A) and the rich content B) which is held until a threshold is compared to determine the best device to use.

Re claims 2, 4, 5, 8, 13, 15-17, and 19, Mirkovic teaches preserving n-best of devices with scores above a threshold, activation of the device with the best score, checking for more utterances, determining whether to respond or utilize a device if a keyword or hot word is found (receiving, producing confidence scores from several devices, comparing for best score above threshold, continuing processing on selected best device… claim 14 expressly and in fig. 9).

However the combination fails to teach likelihoods either greater or less than a second value to determine whether or not to perform ASR

in addition to

transmitting, to the second device, the first value that reflects the first likelihood that the utterance includes the particular hotword;

comparing the first value that reflects the first likelihood that the utterance includes the particular hotword and the second value that reflects the second likelihood that the utterance includes the particular hotword; and

based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword to the second value that reflects the second likelihood that the utterance includes the particular hotword, initiating determining whether to perform automated speech recognition processing on the audio data

Rosenberger expressly teaches handling an initial input speech to process, transmitting a probabilistic weighted value to other devices, comparing values of

multiple devices, then determining whether to process the remaining portion of the input i.e. the command section based on whether an input value is higher than a threshold value associated with weighted probabilistic values of each device or one device depending on the number present (Rosenberger col 4 lines 5-53 with col 5 lines 1-50 with fig. 7 and claim 13).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Mirkovic and Phillips to incorporate the above limitations as taught by Rosenberger to allow for a modification and substitution of a resource saving system such as Rosenberger wherein speech recognition is not fully performed on the initial input such that each an every device is not compared, for instance A) “hello thermostat" versus B) "hello thermostat please raise the temperature 2 degrees” saves four times the amount of processing speed and resources per device, such that 8 devices can simultaneously handle A) wherein only 1 device can handle B) regardless of the amount of resources and speed available, wherein Mirkovic’s system can truncate input as a preamble i.e. A) and the rich content B) which is held until a threshold is compared to determine the best device to use.

Re claims 6, 7, 9-11, and 18, Mirkovic teaches receiving, producing confidence scores from several devices, comparing for best score above threshold, continuing processing on selected best device…as in claim 14 expressly and in fig. 9, however

fails to teach multiple utterances in one input, durations of input, and sending first data to a second or third more devices, that is fails to teach:

3. The method of claim 1, comprising:

transmitting the first value to the second computing device

6. The method of claim 1, comprising:

receiving, by the first computing device, additional audio data that corresponds to an additional utterance

determining a third value corresponding to a likelihood that the additional utterance includes the hot word;

receiving a fourth value corresponding to a likelihood that the utterance includes the hot word, the fourth value being determined by a third computing device; comparing the third value and the fourth value; and

based on comparing the third value and the fourth value, determining that the activation state of the first computing device is an inactive state.

7. The method of claim 3, wherein:

transmitting the first value to the second computing device comprises:

transmitting, to a server, through a local network, or through a short range radio, the first value, and

receiving a second value corresponding to a likelihood that the utterance includes the hot word, the second value being determined by a second computing device comprises:

receiving, from the server, through the local network, or through the short range radio, a second value that was determined by a second computing device

9. The method of claim 3, wherein:

transmitting the first value to the second computing device comprises:

transmitting a first identifier for the first computing device, and

receiving a second value corresponding to a likelihood that the utterance includes the hotword, the second value being determined by a second computing device comprises:

receiving a second identifier for the second computing device

10. The method of claim 5, wherein determining that the activation state is an active state comprises:

determining that a particular amount of time has elapsed since receiving the audio data that corresponds to the utterance.

11. The method of claim 5, comprising:

continuing, for a particular amount of time, to transmit the first value based on determining that the activation state is an active state

Phillips expressly teaches that other devices can be applications both local and external to produce results (Phillips fig. 1C).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Mirkovic to incorporate the above limitations as taught by Phillips to allow for multiple servers as devices with different context based models attached including expected duration to avoid clipping an input of audio, wherein the highest scoring server can be output to handle consecutive utterances e.g. "Directions to Arthur Bryants in Kansas City Missouri", such that different servers can modified into the devices of Mirkovic or when servers are devices, utilize different context e.g. Maps, Directions, Google recognition performed once determination of highest score is made (Phillips fig. 7, 0163, fig. 5a-5f, fig. fig. 1C), to associate the proper device with the best matching input.

However the combination fails to teach likelihoods either greater or less than a second value to determine whether or not to perfrom ASR

in addition to

transmitting, to the second device, the first value that reflects the first likelihood that the utterance includes the particular hotword;

comparing the first value that reflects the first likelihood that the utterance includes the particular hotword and the second value that reflects the second likelihood that the utterance includes the particular hotword; and

based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword to the second value that reflects the second likelihood that the utterance includes the particular hotword, initiating determining whether to perform automated speech recognition processing on the audio data

Rosenberger expressly teaches handling an initial input speech to process, transmitting a probabilistic weighted value to other devices, comparing values of multiple devices, then determining whether to process the remaining portion of the input i.e. the command section based on whether an input value is higher than a threshold value associated with weighted probabilistic values of each device or one device depending on the number present (Rosenberger col 4 lines 5-53 with col 5 lines 1-50 with fig. 7 and claim 13).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Mirkovic and Phillips to incorporate the above limitations as taught by Rosenberger to allow for a modification and substitution of a resource saving system such as Rosenberger wherein speech recognition is not fully performed on the initial input such that each an every device is not compared, for instance A) “hello thermostat" versus B) "hello thermostat please raise the temperature 2 degrees” saves four times the amount of processing speed and resources per device, such that 8 devices can simultaneously handle A) wherein only 1 device can handle B) regardless of the amount of resources and speed available, wherein Mirkovic’s system can truncate input as a preamble i.e. A) and the rich content B) which is held until a threshold is compared to determine the best device to use.

Re claim 21, Mirkovic and Phillips teaches speech processing for several devices, servers, and applications, but fails to teach

21. (New) The method of claim 1, comprising:

based on comparing the first value that reflects the first likelihood that the utterance

includes the particular hotword to the second value that reflects the second likelihood that th

utterance includes the particular hotword, determining that the first value that reflects the first likelihood that the utterance includes the particular hotword is greater than the second value that reflects the second likelihood that the utterance includes the particular hotword,

wherein determining whether to perform automated speech recognition processing on the audio data comprises:

based on determining that the first value that reflects the first likelihood that the utterance includes the particular hotword is greater than the second value that reflects the second likelihood that the utterance includes the particular hotword, determining to perform automated speech recognition processing on the audio data.

Rosenberger expressly teaches handling an initial input speech to process, transmitting a probabilistic weighted value to other devices, comparing values of multiple devices, then determining whether to process the remaining portion of the input

i.e. the command section based on whether an input value is higher than a threshold value associated with weighted probabilistic values of each device or one device depending on the number present (Rosenberger col 4 lines 5-53 with col 5 lines 1-50 with fig. 7 and claim 13).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Mirkovic and Phillips to incorporate the above limitations as taught by Rosenberger to allow for a modification and substitution of a resource saving system such as Rosenberger wherein speech recognition is not fully performed on the initial input such that each an every device is not compared, for instance A) “hello thermostat" versus B) "hello thermostat please raise the temperature 2 degrees” saves four times the amount of processing speed and resources per device, such that 8 devices can simultaneously handle A) wherein only 1 device can handle B) regardless of the amount of resources and speed available, wherein Mirkovic’s system can truncate input as a preamble i.e. A) and the rich content B) which is held until a threshold is compared to determine the best device to use.

Re claim 22, Mirkovic and Phillips teaches speech processing for several devices, servers, and applications, but fails to teach

22. (New) The method of claim 1, comprising:

based on comparing the first value that reflects the first likelihood that the utterance

includes the particular hotword to the second value that reflects the second likelihood that the utterance includes the particular hotword, determining that the first value that reflects the first likelihood that the utterance includes the particular hotword is less than the second value that reflects the second likelihood that the utterance includes the particular hotword,

wherein determining whether to perform automated speech recognition processing on the audio data comprises:

based on determining that the first value that reflects the first likelihood that the utterance includes the particular hotword is less than the second value that reflects the second likelihood that the utterance includes the particular hotword, determining not to perform automated speech recognition processing on the audio data

Rosenberger expressly teaches handling an initial input speech to process, transmitting a probabilistic weighted value to other devices, comparing values of multiple devices, then determining whether to process the remaining portion of the input i.e. the command section based on whether an input value is higher than a threshold value associated with weighted probabilistic values of each device or one device depending on the number present (Rosenberger col 4 lines 5-53 with col 5 lines 1-50 with fig. 7 and claim 13).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Mirkovic and Phillips to incorporate the above limitations as taught by Rosenberger to allow for a modification and substitution of a resource saving system such as Rosenberger wherein speech recognition is not

fully performed on the initial input such that each an every device is not compared, for instance A) “hello thermostat" versus B) "hello thermostat please raise the temperature 2 degrees” saves four times the amount of processing speed and resources per device, such that 8 devices can simultaneously handle A) wherein only 1 device can handle B) regardless of the amount of resources and speed available, wherein Mirkovic’s system can truncate input as a preamble i.e. A) and the rich content B) which is held until a threshold is compared to determine the best device to use.

Conclusion

1. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to MICHAEL COLUCCI whose telephone number is (571)270-1847. The examiner can normally be reached on 9 am - 6:00 pm , Monday - Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (571)-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

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

/MICHAEL COLUCCI/

Primary Examiner, Art Unit 2658

(571)-270-1847
Examiner FAX: (571)-270-2847
Michael.Colucci@uspto.gov

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

First Named Inventor	:	Matthew Sharifi	Art Unit	:	2658
Serial No.	:	14/675,932	Examiner	:	Michael C. Colucci
Filed	:	April 1, 2015	Conf. No.	:	4235
Title	:	HOTWORD DETECTION ON MULTIPLE DEVICES			

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

AMENDMENT AFTER FINAL ACTION

AND

INTERVIEW SUMMARY

AMENDMENTS TO THE CLAIMS

This Listing of Claims replaces all previous versions, and listings, of claims in the application.

Listing of Claims:

1. (Currently amended) A computer-implemented method comprising:

receiving, by a first computing device, audio data that corresponds to an utterance;

before beginning automated speech recognition processing on the audio data, processing the audio data using a classifier that classifies audio data as including a particular hotword or as not including the particular hotword;

determining, based on ~~an analysis of the audio data and without performing automated speech recognition processing on the audio data~~ the processing of the audio data using the classifier that classifies audio data as including a particular hotword or as not including the particular hotword, a first value that reflects a first likelihood that the utterance includes ~~a particular~~ the particular hotword;

receiving a second value that reflects a second likelihood that the utterance includes the particular hotword, as determined by a second computing device;

~~transmitting, to the second device, the first value that reflects the first likelihood that the utterance includes the particular hotword;~~

comparing the first value that reflects the first likelihood that the utterance includes the particular hotword and the second value that reflects the second likelihood that the utterance includes the particular hotword; and

based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword to the second value that reflects the second likelihood that the utterance includes the particular hotword, determining whether to ~~perform~~ begin performing automated speech recognition processing on the audio data.

2. (Currently amended) The method of claim 1, comprising:

determining that the first value satisfies a hotword score ~~threshold,~~; and

~~wherein the first value that reflects the first likelihood that the utterance includes the particular hotword is transmitted to the second device~~ based on determining that the first value satisfies the hotword score threshold, transmitting, to the second device, the first value that reflects the first likelihood that the utterance includes the particular hotword.

3. (Cancelled)

4. (Currently amended) The method of claim 1, comprising:

determining an activation state of the first computing device based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword ~~and the~~ to the second value that reflects the second likelihood that the utterance includes the particular hotword.

5. (Currently amended) The method of claim 4, wherein determining an activation state of the first computing device based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword ~~and the~~ to the second value that reflects the second likelihood that the utterance includes the particular hotword comprises:

determining that the activation state of the first computing device is an active state.

6. (Currently amended) The method of claim 1, comprising:

receiving, by the first computing device, additional audio data that corresponds to an additional utterance;

before beginning automated speech recognition processing on the audio data, processing the audio data using a classifier that classifies audio data as including a particular hotword or as not including the particular hotword;

determining, based on ~~an analysis of the additional audio data and without performing automated speech recognition processing on the additional audio data~~ the processing of the audio data using the classifier that classifies audio data as including a particular hotword or as not including the particular hotword, a third value that reflects a third likelihood that the additional utterance includes the particular hotword;

receiving a fourth value that reflects a fourth likelihood that the additional utterance includes the particular hotword, as determined by a third computing device;

comparing the third value that reflects the third likelihood that the additional utterance includes the particular hotword and the fourth value that reflects the fourth likelihood that the additional utterance includes the particular hotword; and

based on comparing the third value that reflects the third likelihood that the additional utterance includes the particular hotword and the fourth value that reflects the fourth likelihood that the additional utterance includes the particular hotword, determining ~~that the activation state of the first computing device is an inactive state~~ whether to begin performing automated speech recognition processing on the additional audio data.

7. (Currently amended) The method of claim 1, wherein:

~~transmitting, to the second computing device, the first value that reflects the first likelihood that the utterance includes the particular hotword comprises:~~

~~transmitting, to a server, through a local network, or through a short range radio, the first value that reflects the first likelihood that the utterance includes the particular hotword, and~~

receiving a second value that reflects a second likelihood that the utterance includes the particular hotword, as determined by a second computing device comprises:

receiving, from ~~the server~~ a server, through ~~the local~~ a local network, or through ~~the short~~ a short range radio communication channel, a second value that reflects the second likelihood that the utterance includes the particular hotword.

8. (Currently amended) The method of claim 1, comprising:

~~identifying the second computing device; and~~

determining that the second computing device is configured to respond to utterances that include the particular hotword,

wherein comparing the first value that reflects the first likelihood that the utterance includes the particular hotword and the second value that reflects the second likelihood that the

utterance includes the particular hotword is performed in response to determining that the second computing device is configured to respond to utterances that include the particular hotword.

9. (Currently amended) The method of claim 1, wherein:

~~transmitting, to the second computing device, the first value that reflects the first likelihood that the utterance includes the particular hotword comprises:~~

~~transmitting a first identifier of the first computing device, and~~

receiving a second value that reflects a second likelihood that the utterance includes the particular hotword, as determined by a second computing device comprises:

receiving a second identifier of the second computing device.

10. (Currently amended) The method of claim 5, wherein determining ~~that the activation state is an active state comprises:~~

whether to begin performing automated speech recognition processing on the audio data is further based on determining that a particular amount of time has elapsed since receiving the audio data that corresponds to the utterance.

11. (Currently amended) The method of claim 5, comprising:

~~continuing~~ transmitting, for a particular amount of time, ~~to transmit~~ the first value that reflects the first likelihood that the utterance includes the particular hotword based on determining that the activation state is an active state.

12. (Currently amended) A computing device comprising:

one or more storage devices storing instructions that are operable, when executed by the computing device, to cause the computing device to perform operations comprising:

receiving, by a first computing device, audio data that corresponds to an utterance;

before beginning automated speech recognition processing on the audio data, processing the audio data using a classifier that classifies audio data as including a particular hotword or as not including the particular hotword;

determining, based on ~~an analysis of the audio data and without performing automated speech recognition processing on the audio data~~ the processing of the audio data using the classifier that classifies audio data as including a particular hotword or as not including the particular hotword, a first value that reflects a first likelihood that the utterance includes ~~a particular~~ the particular hotword;

receiving a second value that reflects a second likelihood that the utterance includes the particular hotword, as determined by a second computing device;

~~transmitting, to the second device, the first value that reflects the first likelihood that the utterance includes the particular hotword;~~

comparing the first value that reflects the first likelihood that the utterance includes the particular hotword and the second value that reflects the second likelihood that the utterance includes the particular hotword; and

based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword to the second value that reflects the second likelihood that the utterance includes the particular hotword, determining whether to ~~perform~~ begin performing automated speech recognition processing on the audio data.

13. (Currently amended) The device of claim 12, wherein the operations further comprise:

determining that the first value satisfies a hotword score ~~threshold,~~; and

~~wherein the first value that reflects the first likelihood that the utterance includes the particular hotword is transmitted to the second device~~ based on determining that the first value satisfies the hotword score threshold, transmitting, to the second device, the first value that reflects the first likelihood that the utterance includes the particular hotword.

14. (Cancelled)

15. (Currently amended) The device of claim 12, wherein the operations further comprise:

determining an activation state of the first computing device based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword ~~and the~~ to

the second value that reflects the second likelihood that the utterance includes the particular hotword.

16. (Currently amended) The device of claim 15, wherein determining an activation state of the first computing device based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword ~~and the~~ to the second value that reflects the second likelihood that the utterance includes the particular hotword comprises:

determining that the activation state of the first computing device is an active state.

17. (Currently amended) The device of claim 12, wherein the operations further comprise:

receiving, by the first computing device, additional audio data that corresponds to an additional utterance;

before beginning automated speech recognition processing on the audio data, processing the audio data using a classifier that classifies audio data as including a particular hotword or as not including the particular hotword;

determining, based on ~~an analysis of the additional audio data and without performing automated speech recognition processing on the additional audio data~~ the processing of the audio data using the classifier that classifies audio data as including a particular hotword or as not including the particular hotword, a third value that reflects a third likelihood that the additional utterance includes the particular hotword;

receiving a fourth value that reflects a fourth likelihood that the additional utterance includes the particular hotword, as determined by a third computing device;

comparing the third value that reflects the third likelihood that the additional utterance includes the particular hotword and the fourth value that reflects the fourth likelihood that the additional utterance includes the particular hotword; and

based on comparing the third value that reflects the third likelihood that the additional utterance includes the particular hotword and the fourth value that reflects the fourth likelihood that the additional utterance includes the particular hotword, determining ~~that the activation state of the first computing device is an inactive state~~ whether to begin performing automated speech recognition processing on the additional audio data.

18. (Currently amended) The device of claim 14, wherein:

~~transmitting, to the second computing device, the first value that reflects the first likelihood that the utterance includes the particular hotword comprises:~~

~~transmitting, to a server, through a local network, or through a short range radio, the first value that reflects the first likelihood that the utterance includes the particular hotword, and~~

receiving a second value that reflects a second likelihood that the utterance includes the particular hotword, as determined by a second computing device comprises:

receiving, from ~~the server~~ a server, through ~~the local~~ a local network, or through ~~the short~~ a short range radio communication channel, a second value that reflects the second likelihood that the utterance includes the particular hotword.

19. (Currently amended) The device of claim 12, wherein the operations further comprise:

~~identifying the second computing device; and~~

determining that the second computing device is configured to respond to utterances that include the particular hotword,

wherein comparing the first value that reflects the first likelihood that the utterance includes the particular hotword and the second value that reflects the second likelihood that the utterance includes the particular hotword is performed in response to determining that the second computing device is configured to respond to utterances that include the particular hotword.

20. (Currently amended) A non-transitory computer-readable medium storing software comprising instructions executable by one or more computers which, upon such execution, cause the one or more computers to perform operations comprising:

receiving, by a first computing device, audio data that corresponds to an utterance;

before beginning automated speech recognition processing on the audio data, processing the audio data using a classifier that classifies audio data as including a particular hotword or as not including the particular hotword;

determining, based on ~~an analysis of the audio data and without performing automated speech recognition processing on the audio data~~ the processing of the audio data using the classifier that classifies audio data as including a particular hotword or as not including the particular hotword, a first value that reflects a first likelihood that the utterance includes ~~a particular~~ the particular hotword;

receiving a second value that reflects a second likelihood that the utterance includes the particular hotword, as determined by a second computing device;

~~transmitting, to the second device, the first value that reflects the first likelihood that the utterance includes the particular hotword;~~

comparing the first value that reflects the first likelihood that the utterance includes the particular hotword and the second value that reflects the second likelihood that the utterance includes the particular hotword; and

based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword to the second value that reflects the second likelihood that the utterance includes the particular hotword, determining whether to ~~perform~~ begin performing automated speech recognition processing on the audio data.

21. (Currently amended) The method of claim 1, comprising:

based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword to the second value that reflects the second likelihood that the utterance includes the particular hotword, determining that the first value that reflects the first likelihood that the utterance includes the particular hotword is greater than the second value that reflects the second likelihood that the utterance includes the particular hotword,

wherein determining whether to perform automated speech recognition processing on the audio data comprises:

based on determining that the first value that reflects the first likelihood that the utterance includes the particular hotword is greater than the second value that reflects the second likelihood that the utterance includes the particular hotword, determining to ~~perform~~ begin performing automated speech recognition processing on the audio data.

22. (Currently amended) The method of claim 1, comprising:

based on comparing the first value that reflects the first likelihood that the utterance includes the particular hotword to the second value that reflects the second likelihood that the utterance includes the particular hotword, determining that the first value that reflects the first likelihood that the utterance includes the particular hotword is less than the second value that reflects the second likelihood that the utterance includes the particular hotword,

wherein determining whether to perform automated speech recognition processing on the audio data comprises:

based on determining that the first value that reflects the first likelihood that the utterance includes the particular hotword is less than the second value that reflects the second likelihood that the utterance includes the particular hotword, determining not to ~~perform~~ begin performing automated speech recognition processing on the audio data.

23. (New) The method of claim 1, wherein processing the audio data using a classifier that classifies audio data as including a particular hotword or as not including the particular hotword comprises:

extracting filterbank energies or mel-frequency cepstral coefficients from the audio data.

24. (New) The method of claim 1, wherein processing the audio data using a classifier that classifies audio data as including a particular hotword or as not including the particular hotword comprises:

processing the audio data using a support vector machine or a neural network.

REMARKS

Claims 1, 2, 4-13, and 15-22 were pending as of the action of October 1, 2015. Claims 1, 12, and 20 are in independent form and have been amended. New claims 23 and 24 have been added. Support for the amendments can be found throughout the application, including, for example, paragraphs [0022], [0028], and [0042] of the specification. No new matter has been added.

Interview Summary

The undersigned thanks Examiner Colucci for the courtesies extended during the telephonic interview on December 4, 2015. The undersigned participated on behalf of the Applicant. In the interview, the Applicant's representatives discussed an overview of the application and the proposed feature of "before beginning automated speech recognition processing on the audio data, processing the audio data using a classifier that classifies audio data as including a particular hotword or as not including the particular hotword." An agreement was reached that the proposed amendment overcomes the current rejection.

Rejections under 35 U.S.C. § 103

Claims 1, 2, 4-13, and 15-22 were rejected under 35 U.S.C. § 103(a) as over Mirkovic (US 7904297) in view of Phillips (US 2011/0060587) and further in view of Rosenberger (US 8340975). Withdrawal of the rejection is requested.

Claim 1 has been amended to recite "before beginning automated speech recognition processing on the audio data, processing the audio data using a classifier that classifies audio data as including a particular hotword or as not including the particular hotword," and "determining, based on the processing of the audio data using the classifier that classifies audio data as including a particular hotword or as not including the particular hotword, a first value that reflects a first likelihood that the utterance includes the particular hotword." Claims 12 and 20 recite similar features. The cited references do not describe or suggest these features.

Mirkovic describes that "the ASR 920 and parser 922 components may be combined into a single ASR/Parser unit that is coupled to, but functionally independent of DMV classifier 924

and slot classifier 926," (Mirkovic, 17:57-60) which is not the same as "before beginning automated speech recognition processing on the audio data, processing the audio data using a classifier that classifies audio data as including a particular hotword or as not including the particular hotword," and "determining, based on the processing of the audio data using the classifier that classifies audio data as including a particular hotword or as not including the particular hotword, a first value that reflects a first likelihood that the utterance includes the particular hotword," as recited by claim 1.

Phillips describes that the "ASR client 118 may facilitate speech-enabled text entry to each of the multiple applications," (Phillips, ¶ [0061]) which is not the same as "before beginning automated speech recognition processing on the audio data, processing the audio data using a classifier that classifies audio data as including a particular hotword or as not including the particular hotword," and "determining, based on the processing of the audio data using the classifier that classifies audio data as including a particular hotword or as not including the particular hotword, a first value that reflects a first likelihood that the utterance includes the particular hotword," as recited by claim 1.

Rosenberger describes that "[e]mbedded speech recognition IC modules and software (e.g., the above-noted Model LD3320 from IC Route and Model NLP-5x from Sensory, Inc.) calculate and return these measures each time they successfully match (i.e. recognize) a spoken word or phrase with their active vocabulary or grammar," (Rosenberger, 10:47-52) which is not the same as "before beginning automated speech recognition processing on the audio data, processing the audio data using a classifier that classifies audio data as including a particular hotword or as not including the particular hotword," and "determining, based on the processing of the audio data using the classifier that classifies audio data as including a particular hotword or as not including the particular hotword, a first value that reflects a first likelihood that the utterance includes the particular hotword," as recited by claim 1.

Accordingly, the Applicant requests reconsideration and withdrawal of the rejection of claims 1, 12, and 20 and their dependent claims.

Request for After-Final Consideration

The Applicant respectfully requests that Examiner Colucci consider this Amendment under the After-Final Consideration Pilot program. Accordingly, the Applicant has attached a completed form SB-434.

Conclusion

The other claims in the application are each dependent on the independent claims, and are allowable for at least the above reasons. Because each claim is deemed to define additional aspects of the disclosure, however, the individual consideration of each claim on its own merits is respectfully requested.

By responding in the foregoing remarks only to particular positions taken in the action, Applicant does not acquiesce with other positions that have not been explicitly addressed. In addition, Applicant's selecting some particular arguments for the patentability of a claim should not be understood as implying that no other reasons for the patentability of that claim exist. Finally, Applicant's decision to amend or cancel any claim should not be understood as implying that Applicant agrees with any positions taken in the action with reference to that claim or other claims, or that Applicant adopts or agrees with any position except as specifically stated in this paper. In particular, by amending a claim, Applicant does not necessarily concede that the claim prior to amendment was unpatentable.

Fees in the amount of $160 for two excess dependent claims are being paid concurrently herewith on the Electronic Filing System (EFS) by way of Deposit Account authorization. Please apply any charges or credits to Deposit Account No. 06-1050.

Respectfully submitted,

Date: December 7, 2015

/Sam Williams/
Sam L. Williams
Reg. No. 68,071

Customer Number 26192
Fish & Richardson P.C.
Telephone: (202) 626-6404
Facsimile: (877) 769-7945

41151508.docx