



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.
16/036,345	07/16/2018	James David Meyers	P29742-US02CON	7622
136714	7590	03/14/2019	EXAMINER	
Pierce Atwood LLP (Attn: Amazon Docketing)			YEN, ERIC L	
Attn: Patent Docketing (Amazon)				
100 Summer Street			ART UNIT	
Suite 2250			PAPER NUMBER	
Boston, MA 02110			2658	
			NOTIFICATION DATE	
			DELIVERY MODE	
			03/14/2019	
			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):

patent@pierceatwood.com

DETAILED ACTION

Notice of Pre-AIA or AIA Status

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

Claim Objections

2. Claim 12 is objected to because of the following informalities:

Claim 12, line 3 recites “an receipt time” which should be a receipt time— (grammar).

Claim 12 recites, in the 3rd to last line “determining that the sound is detected is within” which should recite —determining that the sound is detected within—(i.e. delete “is” between “detected” and “within” [grammar])

Appropriate correction is required.

Claim Rejections - 35 USC § 112

For clarity of the record, NO new matter rejections are required for Claim 6's last limitation (the offset corresponding to a distance between a first device and a second device). Paragraph 21 of the Specification describes “echoing offset time may correspond to an amount of time between when the speech is outputted and when an echo of that audio data is detected by one or more... microphones” (Specification, paragraph 21). While this paragraph explicitly describes time, this also supports distance between the output device (speaker) and the input device (microphone) based on basic physics.

3. The following is a quotation of 35 U.S.C. 112(b):

(b) CONCLUSION.—The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the inventor or a joint inventor regards as the invention.

The following is a quotation of 35 U.S.C. 112 (pre-AIA), second paragraph:
The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

4. Claims 12-14 are rejected under 35 U.S.C. 112(b) or 35 U.S.C. 112 (pre-AIA), second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which the inventor or a joint inventor, or for pre-AIA the applicant regards as the invention.

As per Claim 12:

Claim 12 recites (in line 4) “a second time difference” when there is no first time difference claimed in claims 6 and 11 or any earlier portion of claim 12.

Claim 13 recites (lines 2-3) “a first device” when claim 6 already recites “a first device” and recites (line 4) “a second device” when claim 6 already recites “a second device”. At a minimum, it is not clear if the “first device” and “second device” in claim 13 is supposed to be the same “first device” and “second device”.

Claim 13, in line 5, recites “the second device” which is ambiguous because there are two recitations of “a second device” that this phrase can refer to (see previous paragraph) and it is not clear which one it is supposed to refer to.

As per Claim 14:

“the audio data received during the estimated time window” (end of the claim) lacks antecedent basis.

The “receiving...” step recites that the audio data is received, but not “during the estimated time window”. The “determining, based at least in part on the data tag...” limitation at best recites where the audio data is output during the “estimated time window” (not received).

Claim Rejections - 35 USC § 103

5. 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 patent for a claimed invention may not be obtained, notwithstanding that the claimed invention is not identically disclosed as set forth in section 102, if the differences between the claimed invention and the prior art are such that the claimed invention as a whole would have been obvious before the effective filing date of the claimed invention to a person having ordinary skill in the art to which the claimed invention pertains. Patentability shall not be negated by the manner in which the invention was made.

6. Claims 14, 19, 20, is/are rejected under 35 U.S.C. 103 as being unpatentable over Makovicka (US 6,725,193), in view of Klisch et al. (US 2002/0042709), hereafter Klisch.

As per Claim 14, Makovicka suggests **a computer-implemented method comprising: receiving, from a device, audio data, determining a data tag associated with a sound within the audio data, determining, based at least in part on the data tag, a time window when the sound will be output during output of the audio data, and disregarding the audio data when the sound is detected within the audio data received during the time window** (“remote system generates an audio

command such as... seven to delete... command is output on the near end speaker... 'seven' and 'delete' may well be picked up by the microphone and carried in the outgoing signal which, when received and processed... could cause a message to be deleted even though the near end user did nothing", col. 1, lines 52-61; "communication system... connects a far end... and near end... example, public telephone systems, private office systems... cellular systems... analog and digital", col. 3, lines 23-42; "incoming signals... including audio signals... output as an audio sound... microphone will pick up not only the speech... also pick up the audio sounds... from the speaker", col. 3, lines 42-52; "echo canceller... suppress audio sounds 70a from the speaker... picked up by the microphone... delay of the signal coming from the speaker... can be the delay or an extension of the delay required for the linear echo cancellation... time alignment of words in the audio sounds 70a output by the speaker... and when the words occur in the microphone... can be narrowed by the peak of the impulse response of the acoustic environment", col. 6, lines 21-40; "incoming signal... played out of the speaker... generated locally...so that the speech recognition echo canceller knows what words are said and when they are said...command words being output from the speaker... signaled to the speech recognition unit and then canceled from the outgoing signal by ignoring their detection at the time of use... during the time window when the command word is being played out of the speaker... or by removing the word from the words being searched for by the speech recognition unit", col. 7, lines 50-64; Figure 1;

In col. 7, lines 50-64, an obvious embodiment is where a locally generated incoming signal [obviously an audio signal based on col. 3, lines 42-52 and based on the fact that it is played out of a speaker] includes command words [also described as

played out of a speaker] and where there is data/knowledge that indicates “what words are said and when they are said” [obviously a different time for each command word to be played out since language words are conventionally communicated one word at a time] where the “when it is said” information for a command word is used to determine the time window when the command word is being played-out/”said” by the speaker [since “when they are said” is the only other time information described in col. 7, lines 50-64 and directly indicates the play-out time when interpreting play-out time as the time that the speaker “says” the command word] and then the command word is ignored/disregarded during the corresponding play-out time window [by canceling, ignoring detection, or removing the word from speech recognition search]

In the obvious embodiment described in the previous paragraph:

“audio data” [the locally generated incoming audio signal including command words] is “receiv[ed]” [at least for speaker output] “from a device” [from a physically adjacent far end that generates the audio signal, see col. 3, lines 23-42 which describes where incoming signals come from far end]

One of the audio command words in the “audio data” can be interpreted as “a sound within the audio data” and its corresponding “when it is said” play out time information can be interpreted as a “data tag” [time data assigned/tagged to the command word indicating the command word’s play out time]. The time “data tag” is “associated with a sound within the audio data” [since it indicates the play out time of the audio command word] and is obviously generated/”determined” at some point [it is illogical for the time data to exist prior to the local generation of the audio signal containing the time data’s command word].

As discussed earlier, the obvious embodiment includes where the “when it is said” information for a command word is used to determine the time window when the command word is being played-out/”said” by the speaker [since “when they are said” is the only other time information described in col. 7, lines 50-64 and directly indicates the play-out time when interpreting play-out time as the time that the speaker “says” the command word]. Col. 7, lines 50-64 thus suggests “a... time window when the sound will be output during output of the audio data” [the time window when the command word “sound” is being played-out/”output” while the speaker is playing out the locally generated incoming signal “audio data” including the command word “sound”] is “determin[ed], based at least in part on the data tag” [where the time “data tag” directly indicates “when” the command word is “said”/output-by-the-speaker and thus is obviously used to determine/identify the play-out “time window”].

Ignoring the detection during the time window can be interpreted as “disregarding the audio data” [i.e. ignoring the portion of the audio data containing the command word] “when the sound is detected within the audio data received during the... time window” [in order to ignore detection of a command word “sound”, there is obviously a detection of the command word “sound” during the time window to ignore]

Makovicka does not, but Klisch (in combination with Makovicka) suggests **determining a data tag associated with a sound within the audio data, and determining, based at least in part on the data tag, an estimated time window when the sound will be output during output of the audio data, and disregarding the audio data when the sound is detected within the audio data received during the estimated time window** (“During automatic speech recognition not only the words

themselves but also their respective start and end points on a time axis are computed”, paragraph 21; “Upon speaking... five hundred thirty... automatic speech recognizer... recognizes each of the spoken numbers as well as connecting words... During the recognition process, the starting and end points in time of the recognized numbers and connecting words are computed”, paragraph 26; Figure 2;

As discussed in the portion of this rejection of claim 14 based on Makovicka [alone], Makovicka suggests “determining a data tag associated with a sound within the audio data” in the sense that the time information for a command word is obviously determined/generated since it would be illogical for such time information to exist before the audio signal including the command word is generated, and Makovicka also suggests where the “time window when the sound will be output during output of the audio data” is “determin[ed], based at least in part on the data tag” [since the when it is said time information is the only other time information in Makovicka, col. 7, lines 50-64 and directly indicates “when the command word is being played out”].

Makovicka does not specifically describe how the time information “data tags” for the command words are “determined” [Makovicka only suggests that the “data tags” are “determined”], but Klisch [in paragraphs 21 and 26] describes where timing information for words in audio can be determined [computed, as described in paragraph 21] by performing speech recognition [recognized words of spoken numbers, see Figure 2] that recognizes not only the spoken words but also their respective start and end “points on a time axis” [obviously starting and ending times].

The references thus suggest a combination where the “when they are said” obviously-respective time “data tags” for each command word in the locally generated

audio signal [in the obvious embodiment of Makovicka discussed in the portion of this rejection based on Makovicka, alone, including where the “when they are said” time information is obviously used to determine the time window when the command word are output] are start time and end time “data tags” for each command word [one command word is the “sound within the audio data”] which are “determined”/computed by performing speech recognition to identify the command words in the locally generated audio signal and to compute each command word’s start and end times in the locally generated audio signal.

In the combination, recognizing the start and end times of one of the command words “determin[es] a data tag” [recognizes the start/end times, either of which can be interpreted as a data tag] “associated with a sound within the audio data” [they are start/end times of the one of the audio command words]

In the combination, the start and end times of the command word [“data tags” which are obviously “estimates” since they are “computed”, as described in paragraph 21, by recognition which is obviously not always perfectly accurate] are the “when it is said” time information of the command word that is obviously used to “determin[e]... [a]... time window when the sound will be output during output of the audio data” [as discussed in the portion of this rejection based on Makovicka, alone] where the time window is an “estimated” time window [the data tags are obviously estimates such that the window based on the time estimates is obviously an estimated window].

In the combination, the detection of the audio command word is ignored/disregarded “during the estimated time window” [as discussed in the portion of this rejection of claim 14 based on Makovicka, alone, except, in the combination, the

time window is an estimated window since it is based on obviously-estimated start/end times]).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to perform a simple substitution of one type of audio word time information with another because the prior art teaches the claimed invention except for the substitution of audio word time information that is not necessarily computed from an audio signal with audio word time information that is. Klisch teaches that audio word time information that is computed from an audio signal was known in the art. One of ordinary skill in the art could have substituted one type of audio word time information with another to obtain the predictable results of a system that ignores command words during a time window when those command words are played out and which knows what words are said and when those words are said (as per Makovicka) where when those words are said is determined by recognizing start and end times of each word from the audio signal (as per Klisch).

As per Claim 19, Makovicka (in combination with Klisch) suggests **generating an indication to delete a portion of the audio data corresponding to the estimated time window** (Figure 4; “echo canceller... suppress audio sounds 70a from the speaker... picked up by the microphone... delay of the signal coming from the speaker... can be the delay or an extension of the delay required for the linear echo cancellation... time alignment of words in the audio sounds 70a output by the speaker... and when the words occur in the microphone... can be narrowed by the peak of the impulse response of the acoustic environment”, col. 6, lines 21-40; “incoming signal...

played out of the speaker... generated locally...so that the speech recognition echo canceller knows what words are said and when they are said...command words being output from the speaker... signaled to the speech recognition unit and then canceled from the outgoing signal by ignoring their detection at the time of use... during the time window when the command word is being played out of the speaker... or by removing the word from the words being searched for by the speech recognition unit”, col. 7, lines 50-64; Figure 1; col. 1, lines 52-61; col. 3, lines 23-42; col. 3, lines 42-52;

The combination [thus far] is as discussed in the rejection of claim 14.

Col. 7, lines 50-64 describes “cancell[ing]” command words being output from the speaker from the “outgoing signal” by ignoring detection and removing the command words from consideration. Col. 6, lines 21-40 describes an alternative form of canceling which is where audio from the speaker [i.e. the command words played out by the speaker] are “suppressed” and Figure 4 more specifically depicts where incoming signal data is subtracted [i.e. “deleted”] from the outgoing signal.

These portions of Makovicka thus at least suggest an alternative embodiment of the combination applied to reject claim 14 where the disregarding/ignoring of the command word during the estimated time window is, instead, subtracting the command word audio from the audio signal portion that exists in the window when the command word is played out [i.e. “generating an indication” to tell the system to “delete a” command word “portion of the audio data corresponding to the estimated time window”])

As per Claim 20, Makovicka (in combination with Klisch) suggests **receiving additional audio input data, determining that the additional audio input data**

includes an occurrence of the sound, determining that the occurrence of the sound was captured outside the estimated time window, and recording the additional audio input data (“comparator... does not match... valid detected word... matches... no signal... cancel out any feedback received... comparison of detected words is done during a selected time period or with a selected delay... outgoing word which is sufficiently delayed after a comparable word... when a... user/speaker... responds to a prompt... not be feedback... properly output”, col. 4, lines 17-37; “incoming signal... played out of the speaker... generated locally... so that the speech recognition echo canceller knows what words are said and when they are said... command words being output from the speaker... signaled to the speech recognition unit and then canceled from the outgoing signal by ignoring their detection at the time of use... during the time window when the command word is being played out of the speaker... or by removing the word from the words being searched for by the speech recognition unit”, col. 7, lines 50-64; col. 1, lines 52-61; col. 3, lines 23-42; col. 3, lines 42-52; col. 6, lines 21-40; Figure 1;

The combination [thus far] is as discussed in the rejection of claim 14.

Makovicka, in Col. 4, lines 17-37, describes where a process to cancel out feedback [speaker audio picked up by a microphone] is “done during a selected time period” [just like the ignoring/canceling of a command word in col. 7, lines 50-64 is during a particular time window]. Col. 4, lines 17-37 further describes where an outgoing word [i.e. a word picked up by the microphone and sent out] can be “sufficiently delayed” [obviously delayed beyond the selected time period] and properly output. Col. 4, lines 17-37 thus further suggests an embodiment of what is described in

col. 7, lines 50-64 where a command word is spoken with “sufficient delay” outside the ignore/disregard time window and processing the command word properly [instead of ignoring it].

Figure 1 also depicts where outgoing signals can be speech-recognized to detect words.

Col. 4, lines 17-37 and Col. 7, lines 50-64 of Makovicka thus suggest where the combination applied to reject claim 14 further “receiv[es] additional audio input data” and “record[s] the additional audio input data” [receiving, outside the time window, user speech including the same command word that is played out, where the received user speech is obviously received/picked up via microphone and is obviously recorded somewhere on the system as data after being picked up/received by the microphone].

Figure 1 further suggests where, in the combination, the command word spoken by the user outside the time window is speech-recognized such that the combination “determine[s] that the additional audio input data includes an occurrence of the sound” [recognizing that the user-spoken outside-window same command word is an occurrence of the same command word].

Col. 4, lines 17-37 and col. 7, lines 50-64 further suggest “determining that the occurrence of the sound was captured outside the estimated time window” [i.e. determining that the user-spoken command word was spoken outside the command word’s window and thus does not need to be ignored and should be properly processed])

7. Claims 1, 5, is/are rejected under 35 U.S.C. 103 as being unpatentable over Makovicka (US 6,725,193), in view of Klisch et al. (US 2002/0042709), hereafter Klisch, and Millman et al. (US 2008/0140412), hereafter Millman.

As per Claim 1, Makovicka suggests **a system comprising: at least one processor and a memory including instructions operable to be executed by the at least one processor to configure the system to: determine audio data, determine a data tag associated with a sound represented within the audio data; determine, based at least in part on the data tag, a time window when the sound will be output during output of the audio data; and generate instructions to disregard the sound within the time window** ("remote system generates an audio command such as... seven to delete... command is output on the near end speaker... 'seven' and 'delete' may well be picked up by the microphone and carried in the outgoing signal which, when received and processed... could cause a message to be deleted even though the near end user did nothing", col. 1, lines 52-61; "communication system... connects a far end... and near end... example, public telephone systems, private office systems... cellular systems... analog and digital", col. 3, lines 23-42; "incoming signals... including audio signals... output as an audio sound... microphone will pick up not only the speech... also pick up the audio sounds... from the speaker", col. 3, lines 42-52; "echo canceller... suppress audio sounds 70a from the speaker... picked up by the microphone... delay of the signal coming from the speaker... can be the delay or an extension of the delay required for the linear echo cancellation... time alignment of words in the audio sounds 70a output by the speaker... and when the words occur in

the microphone... can be narrowed by the peak of the impulse response of the acoustic environment”, col. 6, lines 21-40; “incoming signal... played out of the speaker... generated locally... so that the speech recognition echo canceller knows what words are said and when they are said... command words being output from the speaker... signaled to the speech recognition unit and then canceled from the outgoing signal by ignoring their detection at the time of use... during the time window when the command word is being played out of the speaker... or by removing the word from the words being searched for by the speech recognition unit”, col. 7, lines 50-64; Figure 1;

For the same reasons as those discussed above in the portion of the rejection of claim 14 based on Makovicka, alone, Makovicka suggests “determin[ing] a data tag associated with a sound represented within the audio data”, “determin[ing], based at least in part on the data tag... time window when the sound will be output during output of the audio data” and “generate instructions to disregard the sound within the... time window”.

Also, locally generating the incoming audio “audio data” [col. 7, lines 50-64] suggests “determin[ing] audio data” [i.e. determining what the audio data is by generated it].

Also, based on Figure 1 and col. 3, lines 23-42, an obvious embodiment is where Makovicka’s system is implemented using a generic computer [a system including a processor and memory and programming instructions])

Makovicka does not, but Klisch (in combination with Makovicka) suggests
determine a data tag associated with a sound represented within the audio data;
determine that the sound corresponds to a word; determine, based at least in part

on the data tag, an estimated time window when the sound will be output during output of the audio data, and generate instructions to disregard the sound within the estimated time window (“During automatic speech recognition not only the words themselves but also their respective start and end points on a time axis are computed”, paragraph 21; “Upon speaking... five hundred thirty... automatic speech recognizer... recognizes each of the spoken numbers as well as connecting words... During the recognition process, the starting and end points in time of the recognized numbers and connecting words are computed”, paragraph 26; Figure 2;

The combination is the same as described in the portion of the rejection of claim 14 based on Klisch [in combination with Makovicka], and the combination suggests “determine a data tag associated with a sound represented within the audio data” [determining start/end times of a command word], “determine, based at least in part on the data tag, an estimated time window when the sound will be output during output of the audio data” [obviously using the obviously-estimated/computed start and end times of the command word in the audio signal to determine the time window during which the command word is disregarded/ignored] and “generate instructions to disregard the sound within the estimated time window” [ignoring detection of the command word during the time window based on the computed start/end times obviously disregards the command word during the time window, and instructions to perform the ignoring are obviously generated and executed such that the system actually performs the ignoring function]

In the combination, words are also recognized in addition to the start and end times [as described in Klisch, paragraphs 21 and 26 and Figure 1] such that the

references further suggest where, in the combination, “what words are said” is determined by speech recognition performed on the locally generated audio signal including the command word, such that the combination further suggests “determin[ing] that the sound corresponds to a” word [i.e. the command word sound in the locally generated audio signal is recognized as the command word, in addition to recognizing the command word’s start and end times].)

Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to perform a simple substitution of one type of audio word information including what word is said and when the word is said with another because the prior art teaches the claimed invention except for the substitution of audio word information including what word is said and when the word is said which is not necessarily recognized from an audio signal with audio word information including what word is said and when the word is said that is. Klisch teaches that audio word information including what word is said and when the word is said and that is recognized from an audio signal was known in the art. One of ordinary skill in the art could have substituted one type of audio word information with another to obtain the predictable results of a system that ignores command words during a time window when those command words are played out and which knows what words are said and when those words are said (as per Makovicka) where what words are said and when those words are said is determined by recognizing words and the start and end times of each word from the audio signal (as per Klisch).

Makovicka, in view of Klisch, do not, but Millman (in combination with Makovicka and Klisch) suggests **where the word that the sound is determined to correspond**

to is a wakeword (“audio input from the user... determines... if the input includes the wake-up word... checks for the presence of a wake-up word or command”, paragraph 132; “speech recognition... determine when audio input received from the user corresponds to the first word”, paragraph 38;

The combination [thus far] is as discussed in the portion of this rejection of claim 1 based on Klisch [in combination with Makovicka]

As discussed above in the portion of this rejection of claim 1 based on Klisch [in combination with Makovicka] recognizing the command word in the audio signal as part of determining what words are said and when they are said can be interpreted as “determining that the sound corresponds to a” word.

Millman describes where audio input from a user includes a wake-up word or command [paragraph 132], and using speech recognition to determine whether user audio input corresponds to a word [paragraph 38].

The references further suggest a combination where the command word mapped to the claimed “sound” which is recognized from the locally generated incoming audio signal [along with its start and end times] is, instead, a wake-up word/”wakeword” [such that recognizing the wakeword sound in the audio signal “determines that the sound corresponds to a wakeword”])

Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to perform a simple substitution of one type of recognizable word with another, because the prior art teaches the claimed invention except for the substitution of a recognizable word which is not necessarily a wakeword with a recognizable word that is. Millman teaches that a recognizable word that is a wakeword was known in the

art. One of ordinary skill in the art could have substituted one type of recognizable word with another to obtain the predictable results of a system that ignores command words during a time window when those command words are played out and which knows what words are said and when those words are said (as per Makovicka) where what words are said and when those words are said is determined by recognizing words and the start and end times of each word from the audio signal (as per Klisch) where the recognized words include wake-up words (as per Millman).

As per Claim 5, Makovicka (in combination with Klisch and Millman) suggests **wherein the memory further comprises instructions that further configure the system to: generating an indication to delete a portion of the audio data corresponding to the estimated time window** (Figure 4; “echo canceller... suppress audio sounds 70a from the speaker... picked up by the microphone... delay of the signal coming from the speaker... can be the delay or an extension of the delay required for the linear echo cancellation... time alignment of words in the audio sounds 70a output by the speaker... and when the words occur in the microphone... can be narrowed by the peak of the impulse response of the acoustic environment”, col. 6, lines 21-40; “incoming signal... played out of the speaker... generated locally... so that the speech recognition echo canceller knows what words are said and when they are said... command words being output from the speaker... signaled to the speech recognition unit and then canceled from the outgoing signal by ignoring their detection at the time of use... during the time window when the command word is being played out of the speaker... or by removing the word from the words being searched for by the

speech recognition unit”, col. 7, lines 50-64; Figure 1; col. 1, lines 52-61; col. 3, lines 23-42; col. 3, lines 42-52;

The combination [thus far] is as discussed in the rejection of claim 1.

Col. 7, lines 50-64 describes “cancell[ing]” command words being output from the speaker from the “outgoing signal” by ignoring detection and removing the command words from consideration. Col. 6, lines 21-40 describes an alternative form of canceling which is where audio from the speaker [i.e. the command words played out by the speaker] are “suppressed” and Figure 4 more specifically depicts where incoming signal data is subtracted [i.e. “deleted”] from the outgoing signal.

These portions of Makovicka thus at least suggest an alternative embodiment of the combination applied to reject claim 1 where the disregarding/ignoring of the wakeword word during the estimated time window is, instead, subtracting the wakeword audio from the audio signal portion that exists in the window when the wakeword is played out [i.e. “generating an indication” to tell the system to “delete a” wakeword “portion of the audio data corresponding to the estimated time window”])

Allowable Subject Matter

8. The following is a statement of reasons for the indication of allowable subject matter:

Claims 2-4 and 16-18 are not reasonably suggested by the prior art because Makovicka teaches away from disabling an input device while the loudspeaker is outputting audio (col. 2, lines 11-19)

As per Claims 6 (and consequently claims 7-13) and 15, the prior art does not teach or reasonably suggest modifying, based on an offset, an estimated time window when the sound (within the audio data and associated with the data tag) will be output during output of the audio data.

Prior art echo cancellation teaches delaying signals that are to be used to cancel echoes, where the delay is the time between speaker output and microphone input (logically, based on physics, dependent on the distance between the speaker and microphone).

2014/0003611 “calibration... determine acoustic delay between one or more loudspeakers coupled to the audio processing device... one or more microphones... acoustic delay may be used... to update tunable delay parameters... used to delay providing a reference signal... to the echo cancellation... increase an effective echo cancellation time range of the echo processing components”, paragraph 118; “audio processing... determine a delay between the signal output by the electronic device... and the signal received by the microphone”, paragraph 199;

2005/0203749 “delay section for giving an audio signal a delay corresponding to a period during which a voice outputted from the sound producing section... inputted to the sound picking section”, paragraph 85;

The prior art (including Makovicka) however, does not teach where the word play out window is the echo cancellation data being delayed.

Double Patenting

9. The nonstatutory double patenting rejection is based on a judicially created doctrine grounded in public policy (a policy reflected in the statute) so as to prevent the

unjustified or improper timewise extension of the “right to exclude” granted by a patent and to prevent possible harassment by multiple assignees. A nonstatutory double patenting rejection is appropriate where the conflicting claims are not identical, but at least one examined application claim is not patentably distinct from the reference claim(s) because the examined application claim is either anticipated by, or would have been obvious over, the reference claim(s). See, e.g., *In re Berg*, 140 F.3d 1428, 46 USPQ2d 1226 (Fed. Cir. 1998); *In re Goodman*, 11 F.3d 1046, 29 USPQ2d 2010 (Fed. Cir. 1993); *In re Longi*, 759 F.2d 887, 225 USPQ 645 (Fed. Cir. 1985); *In re Van Ornum*, 686 F.2d 937, 214 USPQ 761 (CCPA 1982); *In re Vogel*, 422 F.2d 438, 164 USPQ 619 (CCPA 1970); *In re Thorington*, 418 F.2d 528, 163 USPQ 644 (CCPA 1969).

A timely filed terminal disclaimer in compliance with 37 CFR 1.321(c) or 1.321(d) may be used to overcome an actual or provisional rejection based on nonstatutory double patenting provided the reference application or patent either is shown to be commonly owned with the examined application, or claims an invention made as a result of activities undertaken within the scope of a joint research agreement. See MPEP § 717.02 for applications subject to examination under the first inventor to file provisions of the AIA as explained in MPEP § 2159. See MPEP §§ 706.02(I)(1) - 706.02(I)(3) for applications not subject to examination under the first inventor to file provisions of the AIA. A terminal disclaimer must be signed in compliance with 37 CFR 1.321(b).

The USPTO Internet website contains terminal disclaimer forms which may be used. Please visit www.uspto.gov/patent/patents-forms. The filing date of the application in which the form is filed determines what form (e.g., PTO/SB/25, PTO/SB/26,

PTO/AIA/25, or PTO/AIA/26) should be used. A web-based eTerminal Disclaimer may be filled out completely online using web-screens. An eTerminal Disclaimer that meets all requirements is auto-processed and approved immediately upon submission. For more information about eTerminal Disclaimers, refer to www.uspto.gov/patents/process/file/efs/guidance/eTD-info-I.jsp.

10. Claims 1-20 are rejected on the ground of nonstatutory double patenting as being unpatentable over claims 1, 4, 5, 8, 9, 11, 12, of U.S. Patent No. 9,691,378, hereafter Parent Patent. Although the claims at issue are not identical, they are not patentably distinct from each other because the claims of this application are rendered obvious by the claims of the Parent Patent.

As per Claim 1, Claims 1 and 5 of the Parent Patent suggests:

A system comprising: at least one processor; and a memory including instructions operable to be executed by the at least one processor to configured the system to: (Claim 5, "by an electronic device", where generic computers [including at least one processor and memory with instructions] are common/conventional and thus obvious embodiments of the electronic device)

determine audio data (Claim 5, "receiving... audio data" where receiving audio data generates/determines data content that previously did not exist prior to data receipt)

determine a data tag associated with a sound represented within the audio data (Claim 5, "receiving... a data tag associated with a sound to be output based, at least in

part, on the audio data” [in order for sound to be output based on the audio data, the sound is logically in the audio data])

determine that the sound corresponds to a wakeword (Claim 5, “determining that the sound is a trigger for the electronic device”, Claim 1, “a wakeword that is a series of temporally-related sounds that, when received by a microphone of the electronic device, causes functionality of the electronic device to be activated”; “determining that a portion of the audio input data represents the wakeword”)

determine, based at least in part on the data tag, an estimated time window when the sound will be output during output of the audio data (Claim 5, “determining that the sound is a trigger for the electronic device”; “determining, based at least in part on the data tag, a time window that the trigger is to be outputted by the audio data” [if the sound is a trigger and the trigger is to be output, then the sound logically is to be output])

generate instructions to disregard the sound within the estimated time window (Claim 1, “determining that a portion of the audio input data represents the wakeword”; “ignoring the portion such that functionality triggered by the wakeword remains inactive” [the window is when the trigger/wakeword/sound, as a portion of audio, is to be output, and claim 1 teaches where that audio portion is ignored])

As per Claims 2-4, Claims 8 and 1 of the Parent Patent suggests their limitations (an audio input device is both an audio data processing component and [particularly in microphone hardware embodiments, see Claim 1 which recites “audio input data is captured by a microphone”] a hardware component, and removing power logically

terminates power and logically disables the audio input device, and it is obvious to terminate power/disable audio input during the estimated window when that is the only window available [when claim 5 of the Parent Patent is, in an obvious manner, broadened to no longer require modification of the time window])

As per Claim 5, Claim 9 of the Parent Patent suggests its limitations (substantively identical except where, in claim 19, the window is not modified [obvious to delete data in the unmodified window when that is the only window available when claim 5 of the Parent Patent is, in an obvious manner, broadened to no longer require modification of the time window])

As per Claim 6, Claim 5 of the Parent Patent teach a computer-implemented method (Claim 5 is a method and the first two steps recite that the receiving is done “by an electronic device” where a computer is a common/conventional and thus obvious form of electronic device)

Claims 1 and 5 teach the first 4 limitations of Claim 6 for the same reasons as those discussed in the rejection of claim 1, above.

Claims 1 and 5 further teach calculating, based on an offset and the estimated time window, a modified time window when the sound will be output during output of the audio data (Claim 5, “generating a modified time window based at least in part on at least one offset and the time window”; “determining... a time window that the trigger is to be outputted by the audio data”; “determining that the sound is a trigger”)

the offset corresponding to a distance between a first device and a second device (Claim 1, “determining a modified time window by applying... the echoing offset time to the time window”; “determining the echoing offset time comprises: determining an audio receipt time that audio input data is captured by the microphone... calculating... time difference between the output time and the audio receipt time”; “determining an output time that the audio data begins to be outputted by the speaker” [difference in time between when the audio data is outputted by a first speaker device and when the audio data is received by a second microphone device is, according to basic physics, directly dependent on the distance between the microphone and speaker])

As per Claim 7, Claim 5 of the Parent Patent teaches causing the audio data to be sent to the first device for output (“causing the audio data to be outputted from at least one speaker”) and Claim 8 of the Parent Patent teaches causing an audio data processing component of the second device to be disabled during the modified time window (removing power to an audio input device logically disables it)

As per Claim 8, Claims 8 and 1 of the Parent Patent teaches causing a hardware component of the second device to be disabled during the modified time window (Claim 8 describes removing power from, thus logically disabling, an audio input, Claim 1 recites “receiving audio input data using the microphone” which teaches where the audio input device can be a microphone [hardware])

As per Claim 9, Claim 11 of the Parent Patent suggests its limitations (alternatively interpreting the electronic device [including the speaker component] in Claim 11 of the Parent Patent as a first device and the microphone as the second device)

As per Claim 10, Claim 12 of the Parent Patent teaches its limitations.

As per Claims 11-12, the last 7 steps of Claim 1 of the Parent Patent suggest all of the limitations (where echo is, by definition, a feedback of audio output from a speaker to a microphone, which by basic physics, naturally requires an amount of time to travel from the speaker to the microphone) and Claim 8 of the Parent Patent more particularly teaches where an audio data processing component (specifically an audio input device) is disabled (power is removed).

As per Claim 13, Claims 1 and 8 of the Parent Patent suggests wherein determining the audio data comprises receiving the audio data from a first device (Claim 1, “receiving, by the electronic device, audio data representing a word”; “receiving audio input data using the microphone”; i.e. microphone is the first device)

sending, to a second device, a first instruction to output the audio data (Claim 1, “outputting the audio data using the speaker” [second device is the electronic device performing the method, not just the speaker])

sending, to the second device, a second instruction to disable an audio data processing component when an occurrence of the sound is detected during the

modified time window (last 3 steps of Claim 1 teach detecting a sound within the modified window and ignoring, and Claim 8 teaches disabling of an audio input/data-processing component [by removing power]).

As per Claim 14, Claims 1 and 5 of the Parent Patent suggest the first 3 limitations of Claim 14 for the same reasons as those discussed above pertaining to Claim 1 (“receiving” in claim 5 of the Parent Patent read on “determining” in claim 1 and more closely reads on “receiving, from a device, audio data” in claim 14)

Claim 1 further suggests disregarding the audio data when the sound is detected within the audio data received during the estimated time window (Claim 1 of the Parent Patent recites “determining that a portion of the audio input data represents the wakeword”, “determining that a detected time that the portion is detected by the microphone is within the... time window” and “ignoring the portion” [even though Claim 1 of the Parent Patent is directed to ignoring an audio wakeword portion in the modified window, it is further obvious to do the same for an unmodified window when that is the only window available, for example when the claims are broadened to no longer require modifying of a window])

As per Claim 15, Claim 5 of the Parent Patent teaches “generating a modified time window based on an offset and the estimated time window when the sound will be output during output of the audio data” (Claim 5, “generating a modified time window based at least in part on at least one offset and the time window”; “time window that the trigger is to be outputted by the audio data”; “determining that the sound is a trigger”)

As per Claims 16-17, Claims 8 and 1 of the Parent Patent teaches their limitations (an audio input device is both an audio data processing component and [particularly in microphone hardware embodiments, see Claim 1 which recites “audio input data is captured by a microphone”] a hardware component, and removing power logically disables the audio input device)

As per Claim 18, Claim 8 of the Parent Patent suggests its limitations (removing power can be interpreted as “terminating power”, audio input device includes hardware microphone embodiments, and it is obvious to terminate power during the estimated window when that is the only window available [when claim 5 of the Parent Patent is, in an obvious manner, broadened to no longer require modification of the time window])

As per Claim 19, Claim 9 of the Parent Patent suggests its limitations (substantively identical except where, in claim 19, the window is not modified [obvious to delete data in the unmodified window when that is the only window available when claim 5 of the Parent Patent is, in an obvious manner, broadened to no longer require modification of the time window])

As per Claim 20, Claim 4 of the Parent Patent suggests its limitations (Claim 20 is broader than Claim 4 of the Parent Patent and [as discussed above] it is obvious to use the unmodified window when that is the only window available)

11. Claims 1-20 are provisionally rejected on the ground of nonstatutory double patenting as being unpatentable over claims 1, 4, 5, 8, 9, 11, 12, of copending Application No. 15/633,529 (reference application), hereafter Parent Application. Although the claims at issue are not identical, they are not patentably distinct from each other because the claims of this application are rendered obvious by the claims of the Parent Patent.

This is a provisional nonstatutory double patenting rejection because the patentably indistinct claims have not in fact been patented.

Claims 1, 4, 5, 8, 9, 11, and 12 of the Parent Application are substantively identical to those of the Parent Patent, and therefore render obvious claims 1-20 for the same reasons as those discussed above in the rejections of claims 1-20 based on the claims of the Parent Patent.

Conclusion

12. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. See PTO-892.

13. Any inquiry concerning this communication or earlier communications from the examiner should be directed to ERIC YEN whose telephone number is (571)272-4249. The examiner can normally be reached on M-F 9:00-5:30PM.

Examiner interviews are available via telephone, in-person, and video conferencing using a USPTO supplied web-based collaboration tool. To schedule an

interview, applicant is encouraged to use the USPTO Automated Interview Request (AIR) at <http://www.uspto.gov/interviewpractice>.

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.

EY 3/8/19

/ERIC YEN/
Primary Examiner, Art Unit 2658

Notice of References Cited	Application/Control No. 16/036,345		Applicant(s)/Patent Under Reexamination Meyers et al.	
	Examiner ERIC YEN		Art Unit 2658	Page 1 of 2

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	CPC Classification	US Classification
*	A	US-6725193-B1	04-2004	Makovicka; Thomas J.	G10L21/02	704/233
*	B	US-20020042709-A1	04-2002	Klisch, Rainer	G10L15/04	704/231
*	C	US-20080140412-A1	06-2008	Millman; Jonathan Travis	G09B7/02	704/270
*	D	US-20140003611-A1	01-2014	Mohammad; Asif I.	H04R3/005	381/66
*	E	US-20050203749-A1	09-2005	Miyata, Kazuhiko	G10L15/22	704/275
*	F	US-6845356-B1	01-2005	Epstein; Mark E.	G10L25/78	704/235
*	G	US-5832440-A	11-1998	Woodbridge; Donald A.	B63H20/007	704/275
*	H	US-20090086934-A1	04-2009	Thomas; Trevor	G10L15/07	379/88.01
*	I	US-8972262-B1	03-2015	Buryak; Kirill	G10L15/28	704/251
*	J	US-20020055844-A1	05-2002	L'Esperance, Lauren	H04M1/271	704/260
*	K	US-20100081487-A1	04-2010	Chen; Shaohai	G10L21/0208	455/575.1
*	L	US-20110238417-A1	09-2011	YAMAMOTO; Koichi	G10L25/78	704/233
*	M	US-4658426-A	04-1987	Chabries; Douglas M.	G10L21/0208	381/71.11

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.

U.S. Patent and Trademark Office
PTO-892 (Rev. 01-2001)

Notice of References Cited

Part of Paper No. 20190305

Notice of References Cited	Application/Control No. 16/036,345		Applicant(s)/Patent Under Reexamination Meyers et al.	
	Examiner ERIC YEN		Art Unit 2658	Page 2 of 2

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	CPC Classification	US Classification
*	A	US-20080091415-A1	04-2008	Schafer; Ronald W.	G10L19/0208	704/200.1
*	B	US-20030065441-A1	04-2003	Funk, Karsten	G01C21/3608	701/468
*	C	US-20010028713-A1	10-2001	Walker, Michael	G10L21/0208	379/416
*	D	US-20100299149-A1	11-2010	Kurzweil; Raymond C.	G10L13/00	704/260
*	E	US-20130238326-A1	09-2013	KIM; Yongsin	G06F3/167	704/231
*	F	US-20100023333-A1	01-2010	Sato; Yasushi	G10L21/038	704/500
*	G	US-20030158732-A1	08-2003	Pi, Xiaobo	G01S13/66	704/251
*	H	US-20050027527-A1	02-2005	Junkawitsch, Jochen	G10L15/20	704/243
*	I	US-8738382-B1	05-2014	Herz; William Samuel	G06F3/167	704/233
*	J	US-20150281162-A1	10-2015	Shen; Jianqiang	H04L51/26	709/206
*	K	US-20150127346-A1	05-2015	Gruenstein; Alexander H.	G10L15/08	704/254
*	L	US-9047857-B1	06-2015	Barton; William F.	G10L15/22	1/1
	M					

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	CPC Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	
	V	
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

U.S. Patent and Trademark Office
PTO-892 (Rev. 01-2001)

Notice of References Cited

Part of Paper No. 20190305