



US009270817B2

(12) **United States Patent**  
**Bigue et al.**

(10) **Patent No.:** **US 9,270,817 B2**  
(45) **Date of Patent:** **Feb. 23, 2016**

(54) **METHOD FOR DETERMINING THE ON-HOLD STATUS IN A CALL**

(75) Inventors: **Jason P. Bigue**, Toronto (CA); **Shai Berger**, Toronto (CA); **Michael J. Pultz**, Toronto (CA)

(73) Assignee: **FonCloud, Inc.**, Toronto (CA)

(\* ) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 480 days.

(21) Appl. No.: **12/276,621**

(22) Filed: **Nov. 24, 2008**

(65) **Prior Publication Data**

US 2009/0136014 A1 May 28, 2009

**Related U.S. Application Data**

(60) Provisional application No. 60/989,908, filed on Nov. 23, 2007.

(51) **Int. Cl.**  
**H04M 3/428** (2006.01)

(52) **U.S. Cl.**  
CPC ..... **H04M 3/428** (2013.01); **H04M 2201/40** (2013.01); **H04M 2203/2077** (2013.01)

(58) **Field of Classification Search**  
CPC ..... H04M 3/51; H04M 3/428  
USPC ..... 379/201.01, 207.02, 207.04, 215.01, 379/265.02, 67.1, 88.18  
See application file for complete search history.

(56) **References Cited**

**U.S. PATENT DOCUMENTS**

4,169,217 A 9/1979 Szanto et al.  
4,228,324 A 10/1980 Rasmussen et al.

4,425,479 A 1/1984 Dubner et al.  
4,731,822 A 3/1988 Berry, III et al.  
4,834,551 A 5/1989 Katz  
4,870,680 A 9/1989 Ohtsuka et al.  
5,627,884 A 5/1997 Williams et al.  
5,640,448 A 6/1997 Toyoshima  
5,737,393 A 4/1998 Wolf  
5,764,746 A 6/1998 Reichelt  
5,802,526 A 9/1998 Fawcett et al.  
5,822,405 A 10/1998 Astarabadi  
6,031,905 A 2/2000 Furman et al.  
6,049,600 A 4/2000 Nabkel et al.  
6,104,797 A 8/2000 Nabkel et al.  
6,122,346 A 9/2000 Grossman  
6,141,328 A 10/2000 Nabkel et al.  
6,195,417 B1 2/2001 Dans  
6,201,855 B1 3/2001 Kennedy  
6,501,750 B1 12/2002 Shaffer et al.  
6,512,825 B1 1/2003 Lindholm et al.

(Continued)

**FOREIGN PATENT DOCUMENTS**

EP 1156649 A 11/2001  
JP 2001285493 10/2001

(Continued)

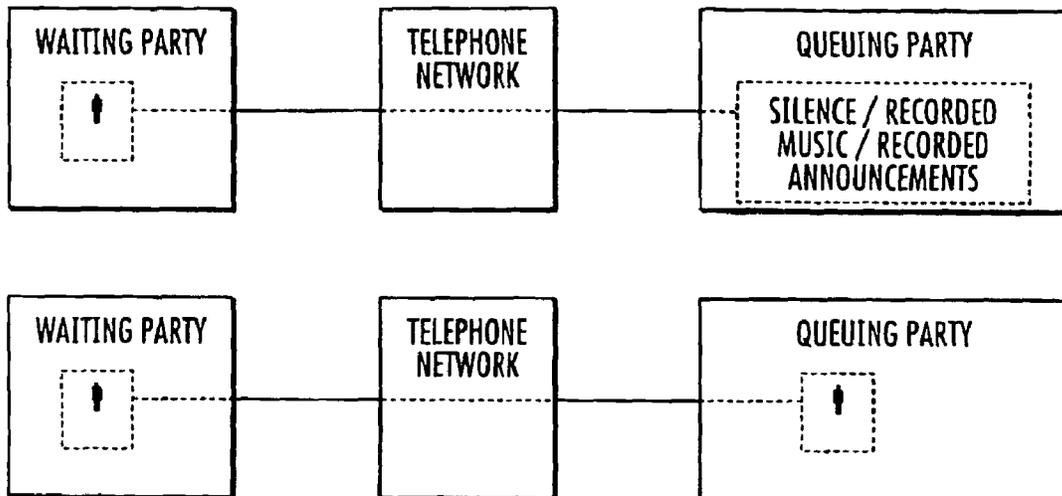
*Primary Examiner* — Sonia Gay

(74) *Attorney, Agent, or Firm* — Carl M. Napolitano; Gray Robinson, P.A.

(57) **ABSTRACT**

A system and method is provided for detecting a hold status in a transaction between a waiting party and a queuing party. The system is adapted to use a preexisting cue profile database containing cue profile for a queuing party. A preexisting cue profile may be used for detecting a hold status in a call between a waiting party and a queuing party. The cue profile of the queuing party may include audio cues, text cues, and cue metadata. The transaction may be a telephone based, mobile-phone based, or internet based.

**16 Claims, 8 Drawing Sheets**



(56)

References Cited

U.S. PATENT DOCUMENTS

6,563,921 B1 5/2003 Williams et al.  
 6,584,184 B1 6/2003 Nabkel et al.  
 6,594,484 B1 7/2003 Hitchings, Jr.  
 6,643,641 B1\* 11/2003 Snyder ..... 707/709  
 6,674,725 B2 1/2004 Nabkel et al.  
 6,684,224 B2\* 1/2004 Meding et al.  
 6,694,008 B1 2/2004 Mukherji et al.  
 6,724,885 B1 4/2004 Deutsch et al.  
 6,754,334 B2 6/2004 Williams et al.  
 6,757,260 B2 6/2004 Pandit  
 6,763,090 B2 7/2004 Che et al.  
 6,788,770 B1 9/2004 Cook et al.  
 6,804,342 B1 10/2004 Gadant  
 6,807,274 B2 10/2004 Joseph et al.  
 6,813,636 B1 11/2004 Bean et al.  
 6,836,478 B1 12/2004 Huang et al.  
 6,850,602 B1\* 2/2005 Chou ..... 379/80  
 6,914,962 B2 7/2005 Neary  
 6,920,425 B1 7/2005 Will et al.  
 6,990,524 B1 1/2006 Hymel  
 6,999,944 B1 2/2006 Cook  
 7,027,408 B2 4/2006 Nabkel et al.  
 7,027,990 B2 4/2006 Sussman  
 7,065,203 B1 6/2006 Huart et al.  
 7,092,738 B2 8/2006 Creamer et al.  
 7,113,987 B2 9/2006 Nabkel et al.  
 7,120,244 B2 10/2006 Joseph et al.  
 7,130,411 B2 10/2006 Brown et al.  
 7,136,478 B1 11/2006 Brand et al.  
 7,174,011 B2 2/2007 Kortum et al.  
 7,215,759 B2 5/2007 Brown et al.  
 7,221,753 B2 5/2007 Hutton et al.  
 7,228,145 B2 6/2007 Burritt et al.  
 7,231,035 B2 6/2007 Walker et al.  
 7,251,252 B2 7/2007 Norby  
 7,315,617 B2 1/2008 Shaffer et al.  
 7,324,633 B2 1/2008 Gao et al.

7,349,534 B2 3/2008 Joseph et al.  
 7,386,101 B2\* 6/2008 Pugliese ..... 379/69  
 7,414,981 B2 8/2008 Jaramillo et al.  
 7,715,531 B1 5/2010 Golding et al.  
 8,155,276 B2 4/2012 Beauregard et al.  
 8,160,209 B2 4/2012 Wang et al.  
 8,223,929 B2 7/2012 Sylvain  
 2003/0043990 A1\* 3/2003 Gutta ..... H04M 3/428  
 2003/0112931 A1 6/2003 Brown et al. 379/215.01  
 2004/0202309 A1 10/2004 Baggenstoss et al.  
 2005/0069117 A1\* 3/2005 Fernandez ..... H04M 3/4286  
 2005/0147219 A1 7/2005 Comerford 379/210.01  
 2005/0278177 A1\* 12/2005 Gottesman ..... 704/270  
 2006/0095564 A1\* 5/2006 Gissel et al. .... 709/224  
 2006/0106613 A1 5/2006 Mills  
 2006/0126803 A1 6/2006 Patel et al.  
 2006/0245579 A1 11/2006 Bienfait et al.  
 2006/0256949 A1 11/2006 Noble  
 2007/0041564 A1 2/2007 Antilli  
 2007/0071223 A1\* 3/2007 Lee et al. .... 379/265.02  
 2007/0165608 A1 7/2007 Altberg et al.  
 2007/0280460 A1 12/2007 Harris et al.  
 2008/0039056 A1 2/2008 Mathews et al.  
 2008/0144786 A1 6/2008 Wang et al.  
 2008/0159495 A1\* 7/2008 Dahan ..... 379/93.01  
 2009/0149158 A1 6/2009 Goldfarb et al.  
 2009/0154578 A1 6/2009 Prakash  
 2009/0154678 A1 6/2009 Kewin et al.  
 2010/0057456 A1 3/2010 Grigsby et al.  
 2011/0103559 A1 5/2011 Andrews et al.

FOREIGN PATENT DOCUMENTS

JP 2004304770 10/2004  
 KR 20040039586 A 5/2004  
 KR 20040106487 12/2004  
 KR 1020050002930 1/2005

\* cited by examiner

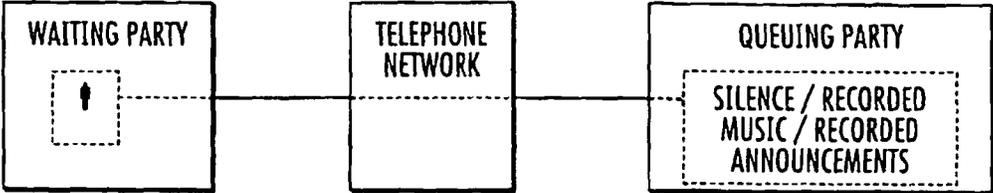


FIG. 1A

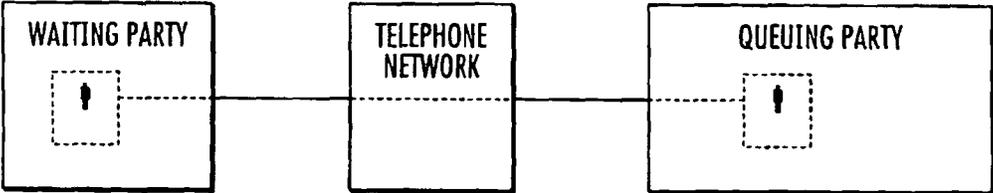


FIG. 1B

COMPANY A CUE METADATA	
AUDIO CUES	
HOLD STATUS AUDIO CUES	
1.	Ⓚ
2.	Ⓚ
3.	Ⓚ
TRANSITION AUDIO CUES	
1.	Ⓚ
2.	Ⓚ
3.	Ⓚ
TEXT CUES	
HOLD STATUS TEXT CUE	
1.	"Abc"
2.	"Abc"
3.	"Abc"
TRANSITION TEXT CUES	
1.	"Abc"
2.	"Abc"
3.	"Abc"

Ⓚ = AUDIO DATA BLOCK, TYPICALLY STORED AS PCM, 8 BIT, 8kHz SAMPLING  
"Abc" = TEXT DATA BLOCK, TYPICALLY STORED IN "UNICODE" FORMAT.

FIG. 2

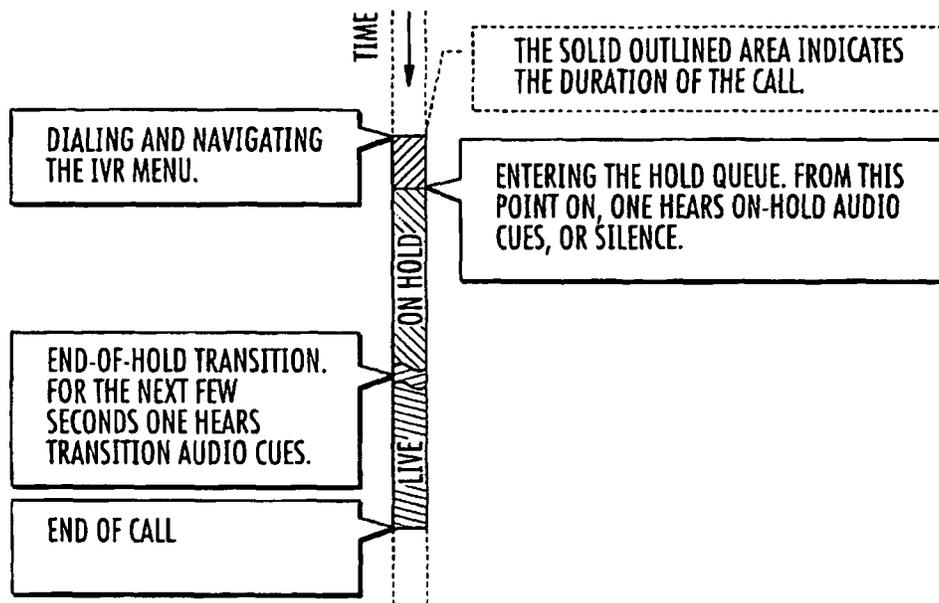


FIG. 3A

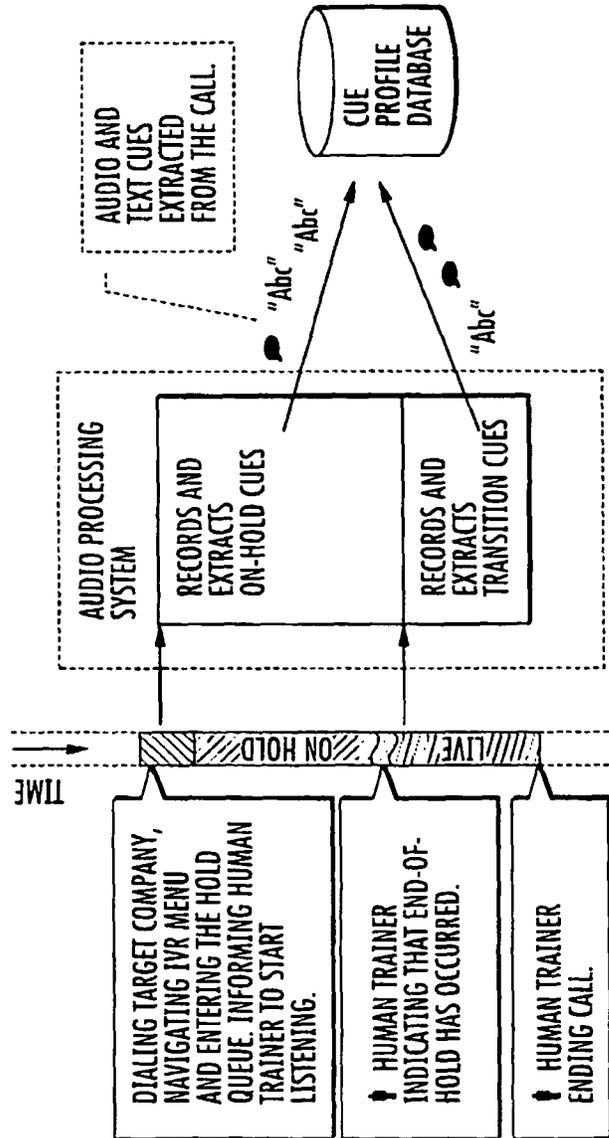


FIG. 3B

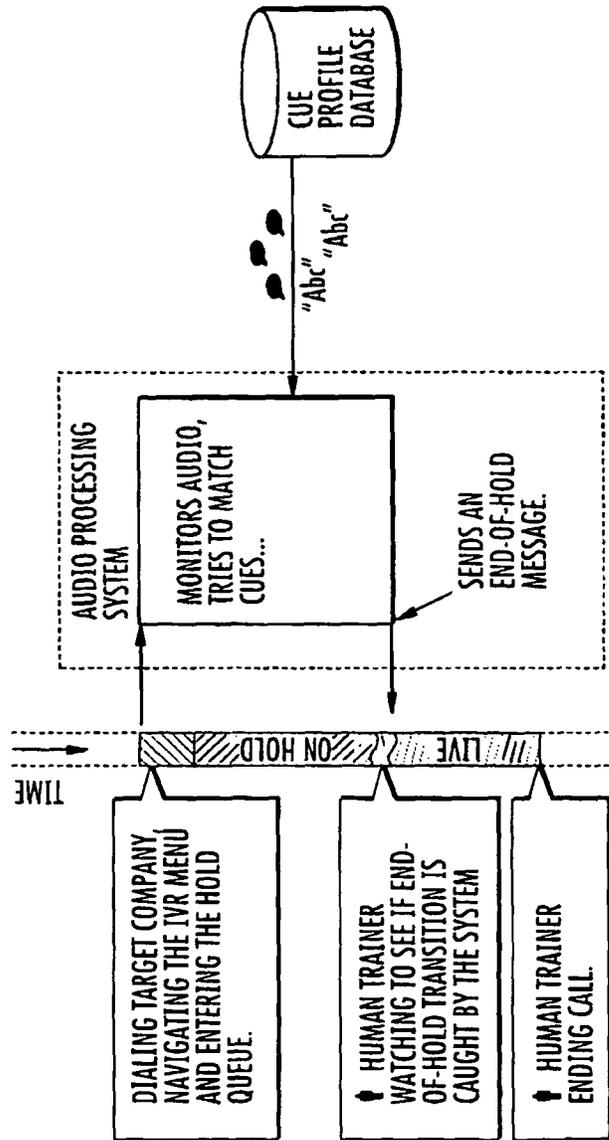


FIG. 3C

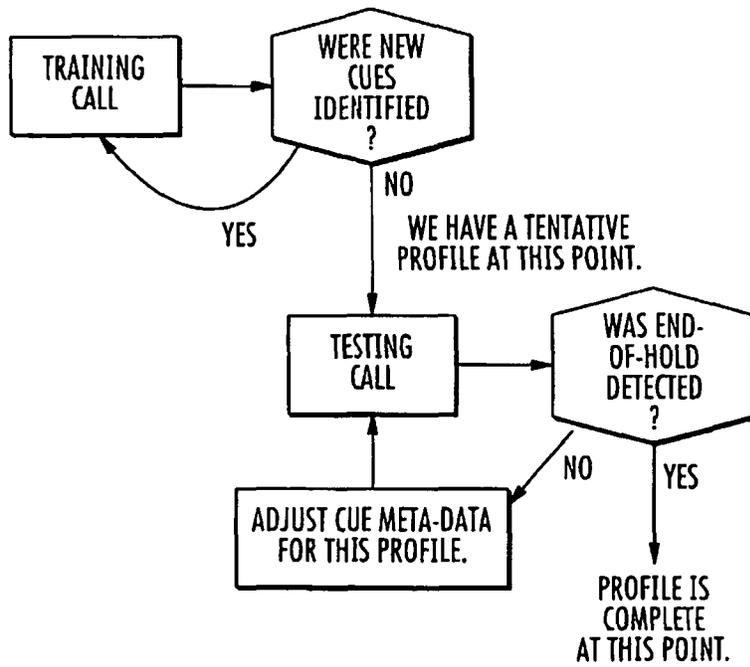


FIG. 3D

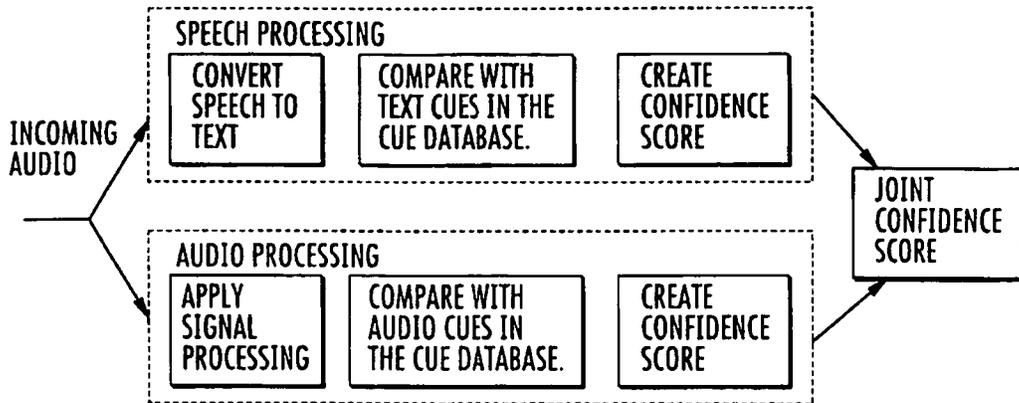


FIG. 4A

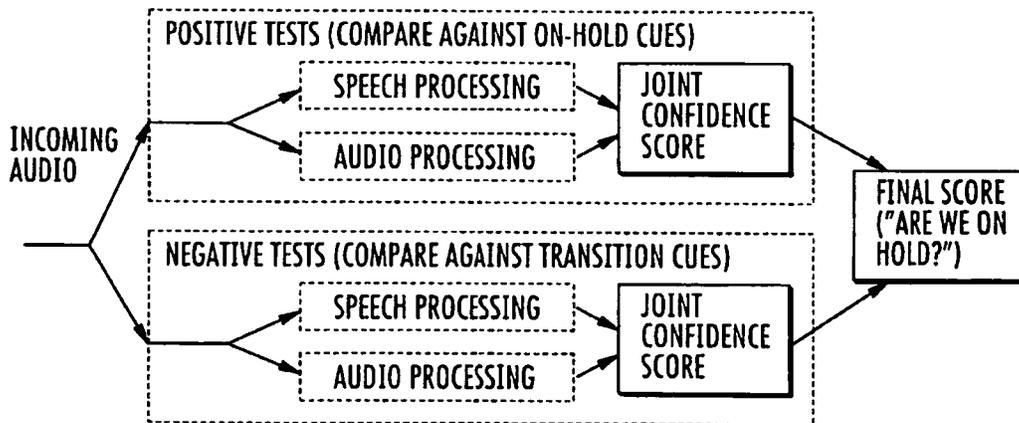


FIG. 4B

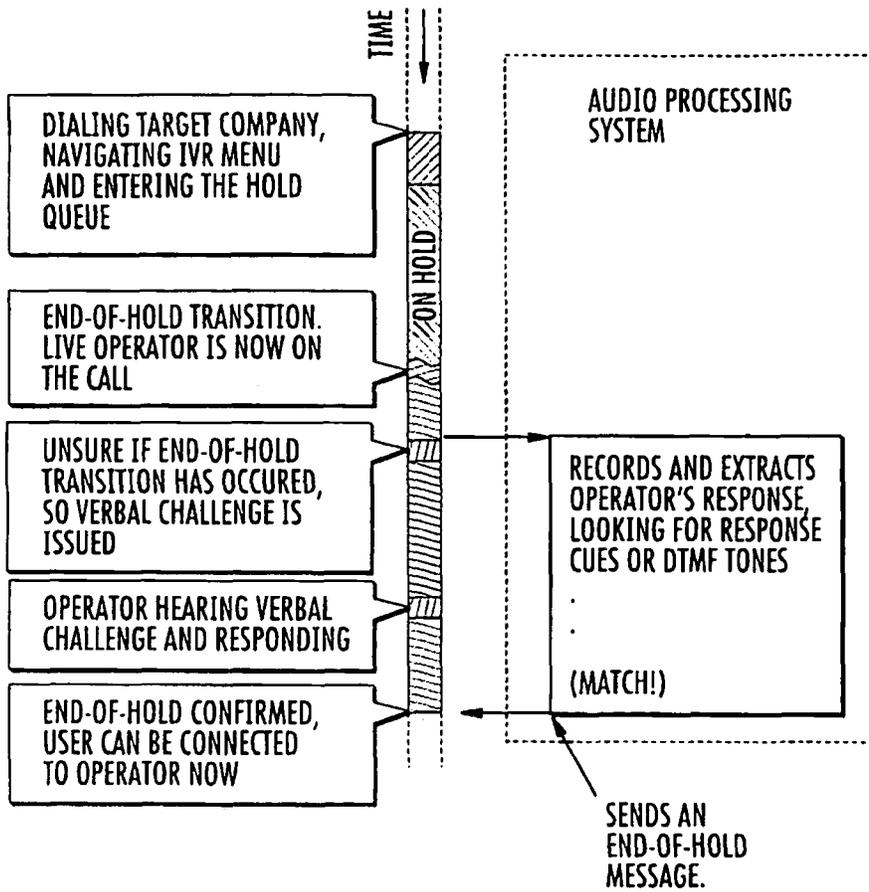


FIG. 5

1

## METHOD FOR DETERMINING THE ON-HOLD STATUS IN A CALL

### CROSS REFERENCE TO RELATED APPLICATIONS

This application claims priority from U.S. Provisional Patent Application Ser. No. 60/989,908 filed Nov. 23, 2007, the disclosure of which is herein incorporated by reference in its entirety.

### FIELD OF INVENTION

Various embodiments related to telephone-based or internet-based call transactions are presented.

### BACKGROUND

In telephone-based or internet-based communication, data, voice or sound (or a combination) is exchanged between parties on a call (typically two parties). Traditionally, businesses have utilized people to participate in telephone-based transactions with their clients. However, recently there are an increasing number of transactions that use automated services and do not engage a person until a certain stage of the call. The embodiments presented herein, relate to such transactions.

### SUMMARY

The present embodiments provides in one aspect, a system for detecting a hold status in a transaction between a waiting party and a queuing party, said system comprising a device adapted to use a preexisting cue profile database containing cue profile for at least one queuing party.

In another aspect, the present embodiments provide for the use of a preexisting cue profile for detecting a hold status in a call between a waiting party and a queuing party.

In another aspect, the present embodiments provide a method for detecting a hold status in a transaction between a waiting party and a queuing party, said method comprising using a preexisting cue profile database containing cue profile for at least one queuing party.

### BRIEF DESCRIPTION OF THE DRAWINGS

For a fuller understanding of the invention, reference is made to the following detailed description, taken in connection with the accompanying drawings illustrating various embodiments of the present invention, in which:

FIG. 1A is an illustration of “on hold” and “Live” states in a call in which the human at the waiting party is “on hold”.

FIG. 1B is an illustration of the “on hold” and “Live” states in a call in which the human at the waiting party is connected “Live” to a human at the queuing party.

FIG. 2 is an illustration of an exemplary cue profile from a cue profile database.

FIG. 3A is an illustration of an exemplary call timeline of a call involving an on-hold state and a live state.

FIG. 3B is an illustration of an exemplary training call in creating an audio cue profile for a queuing party.

FIG. 3C is an illustration of an exemplary testing call in testing an exemplary audio cue profile for a queuing party.

FIG. 3D is an illustration of an exemplary call flow in creating an audio cue profile for a queuing party.

FIG. 4A is an illustration of an exemplary testing of audio clips with two channels of processing.

2

FIG. 4B is an illustration of an exemplary testing of audio clips in which both channels are used for real-time positive and negative testing.

FIG. 5 is an illustration of an exemplary verbal challenge.

### DETAILED DESCRIPTION

The embodiments and implementations described here are only exemplary. It will be appreciated by those skilled in the art that these embodiments may be practiced without certain specific details. In some instances however, certain obvious details have been eliminated to avoid obscuring inventive aspects the embodiments.

Embodiments presented herein relate to telephone-based (land or mobile) and internet-based call transactions. The words “transaction” and “call” are used throughout this application to indicate any type of telephone-based or internet based communication. It is also envisioned that such transactions could be made with a combination of telephone and internet-connected device.

In all such transactions, the client (normally, but not necessarily, the dialing party) is the waiting party or on-hold party who interacts with an automated telephone-based service (normally, but not necessarily, the receiver of the call) which is the queuing party or holding party (different from the on-hold party). The terms “waiting party” and “queuing party” are used throughout this application to indicate these parties, however, it could be appreciated by those skilled in the art that the scope of the embodiments given herein applies to any two parties engaged in such transactions.

During a typical transaction between a waiting party and a queuing party, the waiting party needs to take certain measures like pressing different buttons or saying certain phrases to proceed to different levels of the transaction. In addition, the waiting party may have to wait “on hold” for a duration, before being able to talk to an actual person. Any combination of the two is possible and is addressed in the embodiments given herein.

To understand one example, as shown in FIG. 1, two states during a transaction are considered. The state during which a waiting party is dealing with the automated system and has not reached an actual person is called the “on-hold state”. The state during which the waiting party is talking to an actual person is called the “live state”. Accordingly, the phrase “hold status” is used to refer to either the on-hold state or the live state, depending on whether or not the waiting party is on hold or talking to an actual person, respectively.

It is desirable for the waiting party to find out when the hold status changes from an on-hold state to a live state by a method other than constantly listening and paying attention. Accordingly, different embodiments presented herein address the issue of “hold status detection”.

A “cue profile” of a company, in this disclosure, is referred to as all the information available about the queuing party hold status. In some embodiments presented herein, the pre-existing cue profiles of different queuing parties are used to determine the hold status.

In some embodiments, the cue profile may contain the hold status “audio cues” which are used to detect the hold status for a particular queuing party. Audio cues are any audible cues that could bear information about the hold status. For instance, music, pre-recorded voice, silence, or any combination thereof could indicate an on-hold state. On the other hand, the voice of an actual person could indicate a live state. The event of transition from an on-hold state to a live state could be very subtle. For instance, the transition from a recorded message to a live agent speaking may not be accom-

panied by any distinguished audio message like a standard greeting. Nevertheless there are audio cues indicating the transition from an on-hold state to a live state. Such audio cues are called “transition audio cues”.

In some embodiments, certain preexisting data about a queuing party is used to determine the hold status. Such preexisting data is referred as “cue metadata”. For example, the cue metadata may indicate the sensitivity required for each cue in order to dependably identify it in the audio stream while avoiding false-positives. In these particular embodiments, combinations of hold status audio cues in combination with cue metadata are referred to as the cue profile.

Some embodiments described herein relate to finding the cue profile of a particular queuing party. In certain embodiments, the queuing party itself is used, at least partially, to provide cue metadata to create a cue profile. However, in other embodiments, the cooperation of the queuing party is not necessary.

In some embodiments, “dial-in profiling” is used to create a cue profile of a queuing party accessible through PSTN. The method used in these embodiments is an ordinary telephone connection as used by a typical waiting party.

Dial-in profiling is an iterative process that is done in order to figure out the hold status of a queuing party. FIGS. 3A, 3B, 3C, and 3D are exemplary illustrations of dial-in profiling according to one embodiment. Seen in these figures are different layers and branches of hold status. Once the profile of a certain queuing party is configured, it is entered into a cue profile database as seen in the figures.

In certain cases, dial-in profiling, as described herein, could be the only means for creating a cue profile of a queuing party. In addition, dial-in profiling, according to some embodiments, could also be used to update, expand, or edit a previously created cue profile.

Audio cues could be stored in a standardized format (for example, MP3) and are of fixed time length, for instance two seconds. Another type of cue used in some embodiments is a text cue, which is stored in a standard format (for example ASCII) and is of fixed length (for example two syllables).

In some embodiments these two cues are used create a confidence score. Shown in FIGS. 4A and 4B, certain sections of audio are extracted from a call. These sections, called audio samples, are then compared with audio cues of a given queuing party in what is called an audio test, to create a confidence score. A speech recognition engine in an audio processing system is then used to process the audio samples. The output of the speech recognition engine is compared with text cues to create a text-based confidence score in what is called a text test. The results of audio tests and text tests are then combined to create a final confidence score. The final confidence score is used to determine the hold status. The audio tests and text tests may happen in parallel or they may happen sequentially.

In one embodiment related to the case when the audio cues are not sufficient to detect the hold status, a verbal challenge is issued to the queuing party. A verbal challenge consists of a prerecorded message which is asked of the queuing party at specific instances. For example, one verbal challenge may be “is this a live person?” After a verbal challenge has been issued, a speech recognition engine determines whether there is any response from a live person to the verbal challenge. Based on this, a judgment is made as to the hold status. FIG. 5 is an illustration showing the function of the verbal challenge in the system.

Verbal challenges can also make use of DTMF tones. For example, the challenge could be “press 1 if you are a real human”. In this case, the audio processing system will be

searching for the DTMF tones instead of an audio cue. If the queuing party is in a live state, it may send an unprompted DTMF tone down the line in order to send preemptive notification of the end-of-hold transition. In an order to handle this case the audio system is always listening to and detecting DTMF tones.

A typical apparatus built in accordance with some embodiments presented herein, is referred to as a “hold detection system” and it could comprise, inter alia, some of the following components:

Audio processing system—for extracting audio clips from the phone call and preparing them for analysis by either the speech recognition engine or the audio pattern matching component.

Speech recognition engine—for taking an audio sample and converting human speech to text.

Audio pattern matching component—for taking an audio sample and comparing it to the relevant audio cues contained in a cue database.

Cue processor component—for taking results from the speech recognition engine and audio pattern matching component and computing a confidence score for the hold status.

Audio playback component—for playing pre-recorded audio for the verbal challenge.

Cue profile database—for containing the cue profiles for one or more companies.

It should be noted that any number of the components mentioned above could be integrated into a single component, device. And it should be noted that any device capable of using preexisting cue profile database to determine the hold status in a call or transaction falls within the scope of the embodiments presented herein.

The embodiments presented herein address, inter alia, the following difficulties:

Lack of formal signaling of the hold status in the telephone network.

Hold status cues vary widely between companies.

Hold status cues for a given company can change over time. Cues may not be sufficient to determine the end-of-hold transition.

Companies do not make available any information about their cues.

It will be obvious to those skilled in the art that one may be able to envision alternative embodiments without departing from the scope and spirit of the embodiments presented herein.

As will be apparent to those skilled in the art, various modifications and adaptations of the structure described above are possible without departing from the present invention, the scope of which is defined in the appended claims.

That which is claimed is:

1. A system for detecting a hold status in a transaction between a waiting party and a queuing party, the system comprising:

a cue profile database containing at least one cue profile for at least one queuing party, the at least one cue profile including on-hold cues and transition audio cues of the queuing party; and

a processor adapted to detect a hold status at least partially based on the at least one cue profile of the queuing party, wherein the system is independent of the queuing party.

2. The system of claim 1, wherein the cue profile of the queuing party comprises at least one of audio cues, cue metadata and text cues.

5

3. The system of claim 1, wherein the transaction is at least one of a telephone based, mobile-phone based, and internet based transaction.

4. The system of claim 1, wherein at least part of the cue profile is provided by the queuing party.

5. The system of claim 1, wherein the processor comprises, in combination, at least one of an audio processing system, a speech recognition engine, an audio pattern matching component and a cue processor component.

6. The system of claim 5, further comprising an audio playback component for playing pre-recorded audio used to perform a verbal challenge to detect a live person.

7. The system of claim 1, further comprising means to update the cue profile database after at least one of a certain period and a change in the cue profile.

8. The system of claim 1, further comprising means to use a verbal challenge to determine the hold status.

9. A method for detecting a hold status in a transaction between a waiting party and a queuing party, the method comprising:

using a cue profile database containing at least one cue profile for at least one queuing party, the cue profile containing on-hold cues and transition audio cues; and

6

detecting, by a processor, the hold status at least partially based on the cue profile, wherein the method is independent of the queuing party.

10. The method of claim 9, wherein the cue profile of the queuing party comprises at least one of audio cues, cue meta-data and text cues.

11. The method of claim 9, wherein the transaction is at least one of a telephone based, mobile-phone based, and internet based transaction.

12. The method of claim 9, wherein at least part of the cue profile is provided by the queuing party.

13. The method of claim 9, wherein the method comprises, in combination, at least one of audio processing, speech recognition, audio pattern matching, and cue processing.

14. The method of claim 13, further comprising playing pre-recorded audio used to perform a verbal challenge to detect a live person.

15. The method of claim 9, wherein the method updates the cue profile database after at least one of a certain period and a change in the cue profile.

16. The method of claim 9, wherein the method uses a verbal challenge to determine the hold status.

\* \* \* \* \*