



(22) Date de dépôt/Filing Date: 2007/09/28

(41) Mise à la disp. pub./Open to Public Insp.: 2007/12/11

(45) Date de délivrance/Issue Date: 2011/07/19

(30) Priorités/Priorities: 2006/09/29 (US11/540,736);
2007/09/27 (USPCT/US07/79768)

(51) Cl.Int./Int.Cl. *G10L 11/02* (2006.01),
G10L 11/00 (2006.01), *G10L 21/00* (2006.01),
H04M 3/487 (2006.01), *H04M 3/493* (2006.01),
G06Q 10/00 (2006.01)

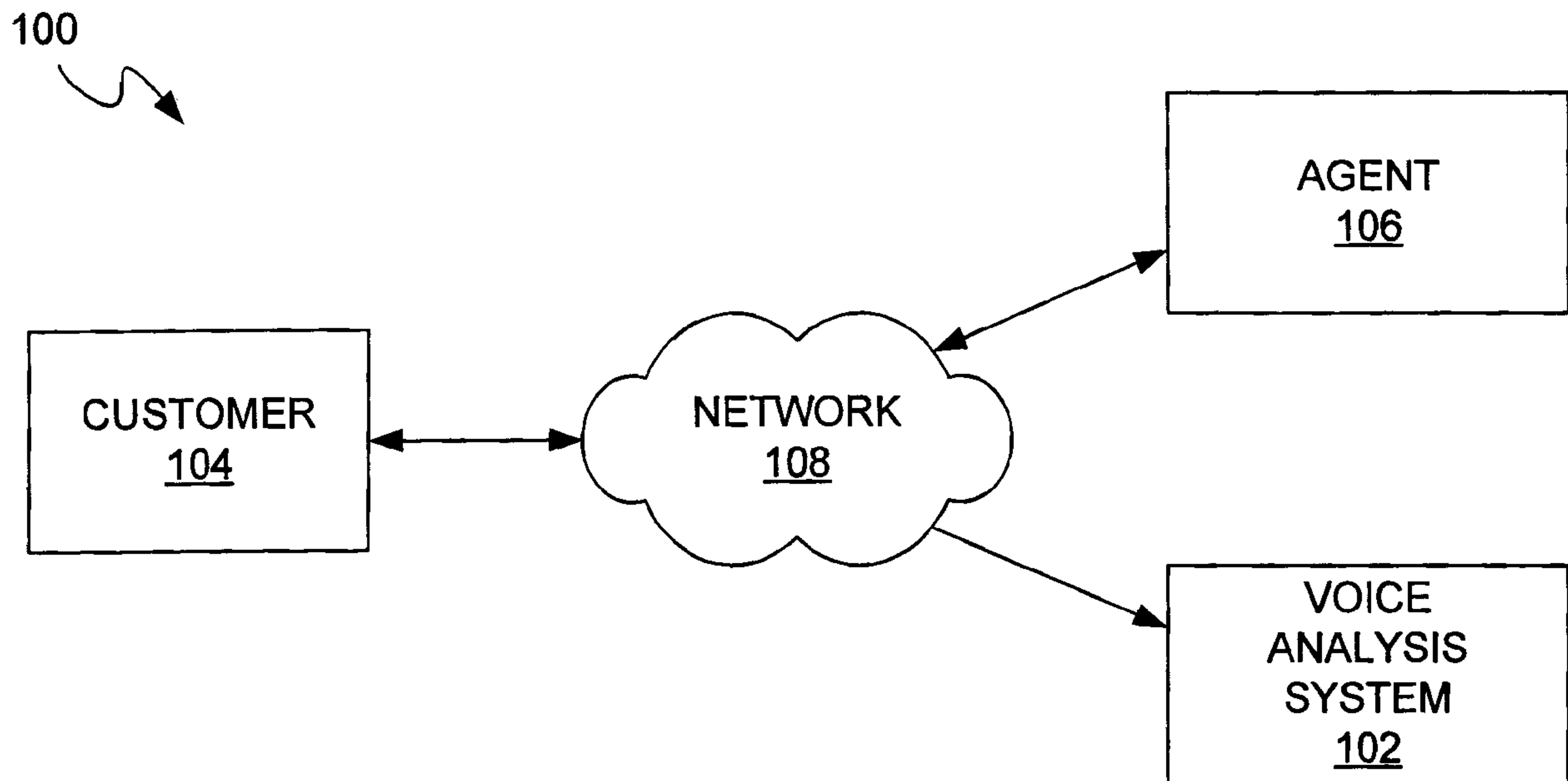
(72) Inventeurs/Inventors:
BLAIR, CHRISTOPHER D., GB;
WATSON, JOSEPH, US

(73) Propriétaire/Owner:
VERINT SYSTEMS INC., US

(74) Agent: SIM & MCBURNEY

(54) Titre : SYSTEMES ET METHODES D'ANALYSE DE SESSIONS DE COMMUNICATION

(54) Title: SYSTEMS AND METHODS FOR ANALYZING COMMUNICATION SESSIONS



(57) Abrégé/Abstract:

Systems and methods for analyzing communication sessions are provided. A representative method includes: recording the communication session; identifying those portions of the communication session not containing speech of the users party to the communications session; and performing post-recording processing on the recording of the communication session based, at least in part, on whether the portions contain speech of the users party to the communications session.



ABSTRACT OF THE DISCLOSURE

Systems and methods for analyzing communication sessions are provided. A representative method includes: recording the communication session; identifying those portions of the communication session not containing speech of the users party to the communications session; and performing post-recording processing on the recording of the communication session based, at least in part, on whether the portions contain speech of the users party to the communications session.

SYSTEMS AND METHODS FOR ANALYZING COMMUNICATION SESSIONS

TECHNICAL FIELD

[0001] The present disclosure generally relates to analysis of communication sessions.

DESCRIPTION OF THE RELATED ART

[0002] Contact centers are staffed by agents who are trained to interact with customers.

Although capable of conducting these interactions using various media, the most common scenario involves voice communications using telephones. In this regard, when a customer contacts a contact center by phone, the call is typically provided to an automated call distributor (ACD) that is responsible for routing the call to an appropriate agent. Prior to an agent receiving the call, however, the call can be placed on hold by the ACD for a variety of reasons. By way of example, the ACD can enable an interactive voice response system (IVR) to query the user for information so that an appropriate queue for handling the call can be determined. As another example, the ACD can place the call on hold until an agent is available for handling the call. In such an on hold period, music (which is referred to as “music on hold”) and/or various announcements (which can be prerecorded or use synthetic human voices) can be provided to the customer.

[0003] For a number of reasons, such as compliance regulations, it is commonplace to record communication sessions. Notably, an entire call (including on hold periods) can be recorded. However, a significant portion of such a recording can be attributed to music on hold, announcements and/or IVR queries that do not tend to provide substantive information for analysis.

SUMMARY

- [0004] In this regard, systems and methods for analyzing communication sessions are provided. An exemplary embodiment of such a system comprises a voice analysis system that is operative to receive information corresponding to a communication session and perform post-recording processing on the information. The voice analysis system is configured to exclude a portion of the information corresponding to the communication session, that is not attributable to speech of the users party to the communication session, from post-recording processing.
- [0005] An exemplary embodiment of a method for analyzing communication sessions comprises excluding a portion of the communication session, not attributable to at least one party of the communication session, from post-recording processing.
- [0006] Another exemplary embodiment of a method for analyzing communication sessions comprises: recording the communication session; identifying those portions of the communication session not containing speech of the users party to the communication session; and performing post-recording processing on the recording of the communication session based, at least in part, on whether the portions contain speech of at least one of the users party to the communications session.
- [006a] In accordance with an aspect of the present invention, there is provided a method for analyzing communication sessions between an agent of a contact center and a customer, said method comprising: recording the communication session; identifying those portions of the communication session not containing speech of the users party to the communications session comprising identifying the presence of at least one of an announcement and audio from an interactive voice response (IVR) system; and performing post-recording processing on the recording of the communication session based, at least in part, on whether the portions contain speech of the users who are

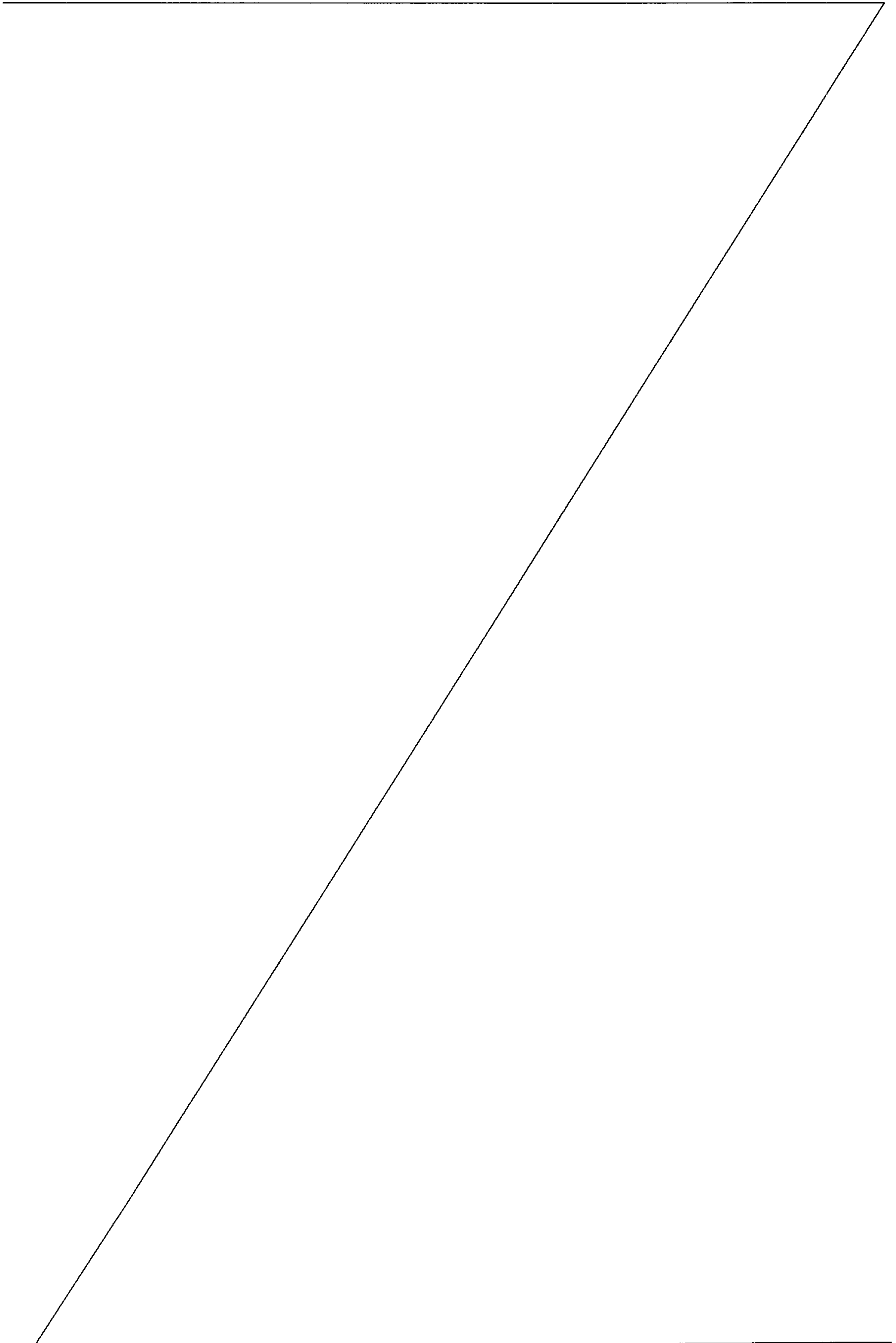
party to the communications session by providing access to information corresponding to a database of potential announcements and potential audio from the IVR system such that the post-recording processing can analyze the at least one of the announcement and the audio using the database.

[006b] In accordance with another aspect of the present invention, there is provided a method for analyzing communication sessions comprising: recording a communication session; identifying the portions of the communication session that include at least one of an announcement and audio from an interactive voice response (IVR) system; providing access to information corresponding to a database of potential announcements and potential audio from the IVR system; and post-recording processing the identified portions using the database.

[006c] In accordance with another aspect of the present invention, there is provided a system for analyzing communication sessions comprising: a voice analysis system operative to: receive identifiers of the portions of the communication session that include at least one of an announcement and audio from an interactive voice response (IVR) system; receive access to information corresponding to a database of potential announcements and potential audio from the IVR system; and post-record process the identified portions using the database to exclude a portion of the information corresponding to the communication session that is not attributable to speech of at least one party of the communication session, wherein the portion of the communication session comprises audio from the interactive voice response (IVR) system.

[0006] Other systems, methods, features and/or advantages will be or may become apparent to one with skill in the art upon examination of the following drawings and detailed description. It is intended that all such additional systems, methods, feature

and/or advantages be included within this description and be protected by the accompanying claims.



BRIEF DESCRIPTION OF THE DRAWINGS

[0008] The components in the drawings are not necessarily to scale relative to each other.

Like reference numerals designate corresponding parts throughout the several views.

[0009] FIG. 1 is a schematic diagram illustrating an embodiment of a system for analyzing communication sessions.

[0010] FIG. 2 is a flowchart depicting functionality (or method steps) associated with an embodiment of a system for analyzing communication sessions.

[0011] FIG. 3 is a schematic diagram illustrating another embodiment of a system for analyzing communication sessions.

[0012] FIG. 4 is a flowchart depicting functionality (or method steps) associated with an embodiment of a system for analyzing communication sessions.

[0013] FIG. 5 is a schematic diagram of an embodiment of a system for analyzing communication sessions that is implemented by a computer.

DETAILED DESCRIPTION

[0014] As will be described in detail here with reference to several exemplary embodiments, systems and methods for analyzing communication sessions can potentially enhance post-recording processing of communication sessions. In this regard, it is known that compliance recording and/or recording of communication sessions for other purposes involves recording various types of information that are of relatively limited substantive use. By way of example, music, announcements and/or queries by IVR systems commonly are recorded. Such information can cause problems during post-recording processing in that these types of information can make it difficult for accurate processing by speech recognition and phonetic analysis

systems. Additionally, since such information affords relatively little substantive value, inclusion of such information tends to use recording resources, i.e., the information takes up space in memory, thereby incurring cost without providing corresponding value.

[0015] Referring now to FIG. 1, FIG. 1 depicts an exemplary embodiment of a system 100 for analyzing communication sessions that incorporates a voice analysis system 102. Voice analysis system 102 receives information corresponding to a communication session, such as a session occurring between a customer 104 and an agent 106 via a communication network 108. As a non-limiting, example, communications network 108 can include a Wide Area Network (WAN), the Internet and/or a Local Area Network (LAN). In some embodiments, the voice analysis system can receive the information corresponding to the communication session from a data storage device, e.g., a hard drive, that is storing a recording of the communication session.

[0016] FIG. 2 depicts the functionality (or method) associated with an embodiment of a system for analyzing communications, such as the embodiment of FIG. 1. In this regard, the depicted functionality involves excluding a portion of a communication session from post-recording processing (block 202). That is, information that does not correspond to a voice component of a party to the communication session, e.g., the agent and the customer, can be excluded. Notably, various types of information, such as music, announcements and/or queries of an IVR system are not attributable to one of the parties. As such, these types of information can be excluded from post-recording processing (block 204), which can involve speech recognition and/or phonetic analysis.

[0017] In some embodiments, information that does not correspond to a voice component of any party to the communication session is deleted from the recording of the communication session. As another example, such information could be identified and any post-recording processing algorithms could ignore those portions, thereby enabling processing resources to be devoted to analyzing other portions of the recordings.

[0018] As a further example, at least with respect to announcements and queries from IVR systems that involve pre-recorded or synthetic human voices (i.e., computer generated voices), information regarding those audio components can be provided to the post-recording processing algorithms so that analysis can be accomplished efficiently. In particular, if the processing system has knowledge of the actual words that are being spoken in those audio components, the processing algorithm can more quickly and accurately convert those audio components to transcript form (as in the case of speech recognition) or to phoneme sequences (as in the case of phonetic analysis).

[0019] FIG. 3 depicts another exemplary embodiment of a system for analyzing communication sessions. In this regard, system 300 is implemented in a contact center environment that includes a voice analysis system 302. Voice analysis system 302 incorporates an identification system 304 and a post-recording processing system 306. The post-recording processing system incorporates a speech recognition system 310 and a phonetic analysis system 312.

[0020] The contact center also incorporates an automated call distributor (ACD) 314 that facilitates routing of a call between the customer and the agent. The communication session is recorded by a recording system 316 that is able to provide information corresponding to the communication session to the voice analysis system for analysis.

[0021] In operation, the voice analysis system receives information corresponding to a communication session that occurs between a customer 320 and an agent 322, with the session occurring via a communication network 324. Specifically, the ACD routes the call so that the customer and agent can interact and the recorder records the communication session.

[0022] With respect to the voice analysis system 302, the identification system 304 analyzes the communication session (e.g., from the recording) to determine whether post-recording processing should be conducted with respect to each of the recorded portions of the session. Based on the determinations, which can be performed in various manners (examples of which are described in detail later), processing can be performed by the post-recording processing system 306. By way of example, the embodiment of FIG. 3 includes both a speech recognition system and a phonetic analysis system that can be used either individually or in combination to process portions of the communication session.

[0023] Notably, the ACD 314 can be responsible for providing various announcements to the customer. In some embodiments, these announcements can be provided via synthetic human voices and/or recordings. It should be noted that other types of announcements can be present in recordings that are not provided by an ACD. By way of example, a telephone central office can introduce announcements that could be recorded. As another example, voice mail systems can provide announcements. The principles described herein relating to treatment of ACD announcements are equally applicable to such other forms of announcements regardless of the manner in which the announcements become associated with a recording.

[0024] Additionally or alternatively, the ACD can facilitate interaction of the customer with an IVR system that queries the customer for various information. Additionally or

alternatively, the ACD can provide music on hold, such as when the call is queued awaiting pickup by an agent. It should be noted that other types of music can be present in recordings that are not provided by an ACD. By way of example, a customer could be speaking to an agent when music is being played in the background. The principles described herein relating to treatment of ACD music on hold are equally applicable to such other forms of music regardless of the manner in which the music becomes associated with a recording.

[0025] FIG. 4 is a flowchart depicting functionality of an embodiment of a system for analyzing communication sessions, such as the system depicted in FIG. 3. In this regard, the functionality (or method steps) may be construed as beginning at block 402, in which a communication session is recorded. In block 404, portions of the communication session are identified as containing music, announcements and/or IVR audio. Then, as depicted in block 406, a determination is made as to whether the music, announcements and/or IVR audio that were identified are to be deleted from the recording. If it is determined that the music, announcements and/or IVR audio are to be deleted, the process proceeds to block 408, in which deletion from the recording is performed. Then, the process proceeds to block 410. If, however, it is determined that the music, announcements and/or IVR audio are not to be deleted, the process also proceeds to block 410.

[0026] In block 410, information regarding the presence of the music, announcements and/or IVR audio is used to influence post-recording processing of a communication session. By way of example, the corresponding portions of the recording can be designated or otherwise flagged with information indicating that music, announcements and/or IVR audio is present. Other manners in which such a post-recording process can be influenced will be described in greater detail later.

[0027] Thereafter, the process proceeds to block 412, in which post-recording processing is performed. In particular, such post-recording processing can include at least one of speech recognition and phonetic analysis.

[0028] With respect to the identification of various portions of a communication session, a voice analysis system can be used to distinguish those portions of a communication session that include voice components of a party to the communication from other audio components. Depending upon the particular embodiment, such a voice analysis system could identify the voice components of the parties as being suitable for both post-recording analysis and/or could identify other portions as not being suitable for post-recording analysis.

[0029] In some embodiments, a voice analysis system is configured to identify dual tone multi-frequency (DTMF) tones, i.e., the sounds generated by a touch tone phone. In some of these embodiments, the tones can be removed from the recording. In removing such tones prior to speech recognition and/or phonetic analysis, such analysis may be more effective as the DTMF tones may no longer mask some of the recorded speech.

[0030] As an additional benefit, the desire for improved security of personal information may require in some circumstances that such DTMF tones not be stored or otherwise made available for later access. For instance, a customer responding to an IVR system query may input DTMF tones corresponding to a social security number or a bank account number. Clearly, recording such tones could increase the likelihood of this information being compromised. However, an embodiment of a voice analysis system that deletes these tones does not incur this potential liability.

[0031] In some embodiments, signaling tones, such as distant and local ring tones and busy equipment signals, can be identified. With respect to the identification of ring

tones, identification of regional tones can provide additional information about a call that may be useful. By way of example, such tones could identify the region to which an agent placed a call while a customer was on hold. Moreover, once identified, the signaling tones can be removed from the recording of the communication session.

[0032] Regional identification of audio components also can occur in some embodiments with respect to announcements. In this regard, some regions provide unique announcements, such as those originating from a central telephone office. For example, in the United States an announcement may be as follows, "I am sorry, all circuits are busy. Please try your call again later." Identifying such an audio component in a recording could then inform a user that a party to the communication session attempted to place a call to the United States.

[0033] Various techniques can be used for differentiating the various portions of a communication session. In this regard, energy envelope analysis, which involves graphically displaying the amplitude of audio of a communication session, can be used to distinguish music from voice components. This is because music tends to follow established tempo patterns and oftentimes exhibits higher energy levels than voice components.

[0034] In some embodiments, such identification can be accomplished manually, semi-automatically or automatically. By way of example, a semi-automatic mode of identification can include providing a user with a graphical user interface that depicts an energy envelope corresponding to a communication session. The graphical user interface could then provide the user with a sliding window that can be used to identify contiguous portions of the communication session. In this regard, the sliding window can be altered to surround a portion of the recording that is identified, such as by listening to that portion, as music. The portion of the communication session that

has been identified within such a sliding window as being attributable to music can then be automatically compared by the system to other portions of the recorded communication session. When a suitable match is automatically identified, each such portion also can be designated as being attributable to music.

[0035] Additionally or alternatively, some embodiments of a voice analyzer system can differentiate between announcements and tones that are regional in nature. This can be accomplished by comparing the recorded announcements and/or tones to a database of known announcements and tones to check for parity. Once designations are made about the portions of a communication sessions containing regional characteristics, the actual audio can be discarded or otherwise ignored during post-recording processing. In this manner, speech analysis does not need to be undertaken with respect to those portions of the audio, thereby allowing speech analysis systems to devote more time and resources to other portions of the communication session. Notably, however, the aforementioned designations can be retained in the records of the communication session so that information corresponding to the occurrence of such characteristics is not discarded.

[0036] In some embodiments, a database can be used for comparative purposes to identify variable announcements. That is an announcement that includes established fields, within which information can be changed. An example of such a variable announcement includes an airline reservation announcement that indicates current rate promotions. Such an announcement usually includes a fixed field identifying the airline and then variable fields identifying a destination and a fare. Knowledge of the first variable field involving a destination could be used to simplify post-recording processing in some embodiments, whereas other embodiments may avoid processing of that portion once a determination is made that the portion corresponds to an

announcement. Alternatively, a hybrid approach could involve not processing of audio corresponding to fixed fields and allowing post-recording processing on the audio corresponding to the variable fields.

[0037] Another form of variable announcements relates to voicemail systems. In this regard, voicemail systems use variable fields to inform a caller that a voice message can be recorded. In some embodiments, these announcements can be identified and handled such as described before. One notable distinction, however, involves the use of the actual voicemail message that is left by a caller. If such a caller indicates that the message is “private,” some embodiments can delete the message or otherwise avoid post-recording processing of the message.

[0038] FIG. 5 is a schematic diagram illustrating an embodiment of system for analyzing communication sessions that is implemented by a computer. Generally, in terms of hardware architecture, system 500 includes a processor 502, memory 504, and one or more input and/or output (I/O) devices interface(s) 506 that are communicatively coupled via a local interface 508. The local interface 508 can include, for example but not limited to, one or more buses or other wired or wireless connections. The local interface may have additional elements, which are omitted for simplicity, such as controllers, buffers (caches), drivers, repeaters, and receivers to enable communications.

[0039] Further, the local interface may include address, control, and/or data connections to enable appropriate communications among the aforementioned components. The processor may be a hardware device for executing software, particularly software stored in memory.

[0040] The memory can include any one or combination of volatile memory elements (*e.g.*, random access memory (RAM, such as DRAM, SRAM, SDRAM, *etc.*)) and

nonvolatile memory elements (*e.g.*, ROM, hard drive, tape, CDROM, *etc.*).

Moreover, the memory may incorporate electronic, magnetic, optical, and/or other types of storage media. Note that the memory can have a distributed architecture, where various components are situated remote from one another, but can be accessed by the processor. Additionally, the memory includes an operating system 510, as well as instructions associated with a voice analysis system 512, exemplary embodiments of which are described above.

[0041] One should note that the flowcharts included herein show the architecture, functionality and/or operation of a possible implementation of one or more embodiments that can be implemented in software and/or hardware. In this regard, each block can be interpreted to represent a module, segment, or portion of code, which comprises one or more executable instructions for implementing the specified logical functions. It should also be noted that in some alternative implementations, the functions noted in the blocks may occur out of the order in which depicted. For example, two blocks shown in succession may in fact be executed substantially concurrently or the blocks may sometimes be executed in the reverse order, depending upon the functionality involved.

[0042] One should note that any of the functions (such as depicted in the flowcharts) can be embodied in any computer-readable medium for use by or in connection with an instruction execution system, apparatus, or device, such as a computer-based system, processor-containing system, or other system that can fetch the instructions from the instruction execution system, apparatus, or device and execute the instructions. In the context of this document, a “computer-readable medium” can be any means that can contain, store, communicate, propagate, or transport the program for use by or in connection with the instruction execution system, apparatus, or device. The computer

readable medium can be, for example but not limited to, an electronic, magnetic, optical, electromagnetic, infrared, or semiconductor system, apparatus, or device. More specific examples (a nonexhaustive list) of the computer-readable medium could include an electrical connection (electronic) having one or more wires, a portable computer diskette (magnetic), a random access memory (RAM) (electronic), a read-only memory (ROM) (electronic), an erasable programmable read-only memory (EPROM or Flash memory) (electronic), an optical fiber (optical), and a portable compact disc read-only memory (CDROM) (optical). In addition, the scope of the certain embodiments of this disclosure can include embodying the functionality described in logic embodied in hardware or software-configured mediums.

[0043] It should be emphasized that many variations and modifications may be made to the above-described embodiments. All such modifications and variations are intended to be included herein within the scope of this disclosure and protected by the following claims.

CLAIMS

1. A method for analyzing communication sessions between an agent of a contact center and a customer, said method comprising:

recording the communication session;

identifying those portions of the communication session not containing speech of the users party to the communications session comprising identifying the presence of at least one of an announcement and audio from an interactive voice response (IVR) system; and

performing post-recording processing on the recording of the communication session based, at least in part, on whether the portions contain speech of the users who are party to the communications session by providing access to information corresponding to a database of potential announcements and potential audio from the IVR system such that the post-recording processing can analyze the at least one of the announcement and the audio using the database.

2. The method of claim 1, wherein the method further comprises:

deleting the portions not attributable to at least one of the agent and the customer from the recording; and wherein

performing post recording processing comprises performing post-recording processing on the remaining portions.

3. The method of claim 1, wherein identifying comprises identifying presence of music in the communication session.

4. The method of claim 1, further comprising deleting audio from the recording corresponding to a private voicemail message.
5. A method for analyzing communication sessions comprising:
 - recording a communication session;
 - identifying the portions of the communication session that include at least one of an announcement and audio from an interactive voice response (IVR) system;
 - providing access to information corresponding to a database of potential announcements and potential audio from the IVR system; and
 - post-recording processing the identified portions using the database.
6. The method of claim 5, wherein the post recording processing comprises speech recognition processing.
7. The method of claim 5, wherein the post-recording processing comprises phonetic analysis.
8. The method of claim 5, wherein the identified portions of the communication session comprise announcements.
9. The method of claim 8, wherein the announcement comprises synthetic human voices.
10. A system for analyzing communication sessions comprising:
 - a voice analysis system operative to:

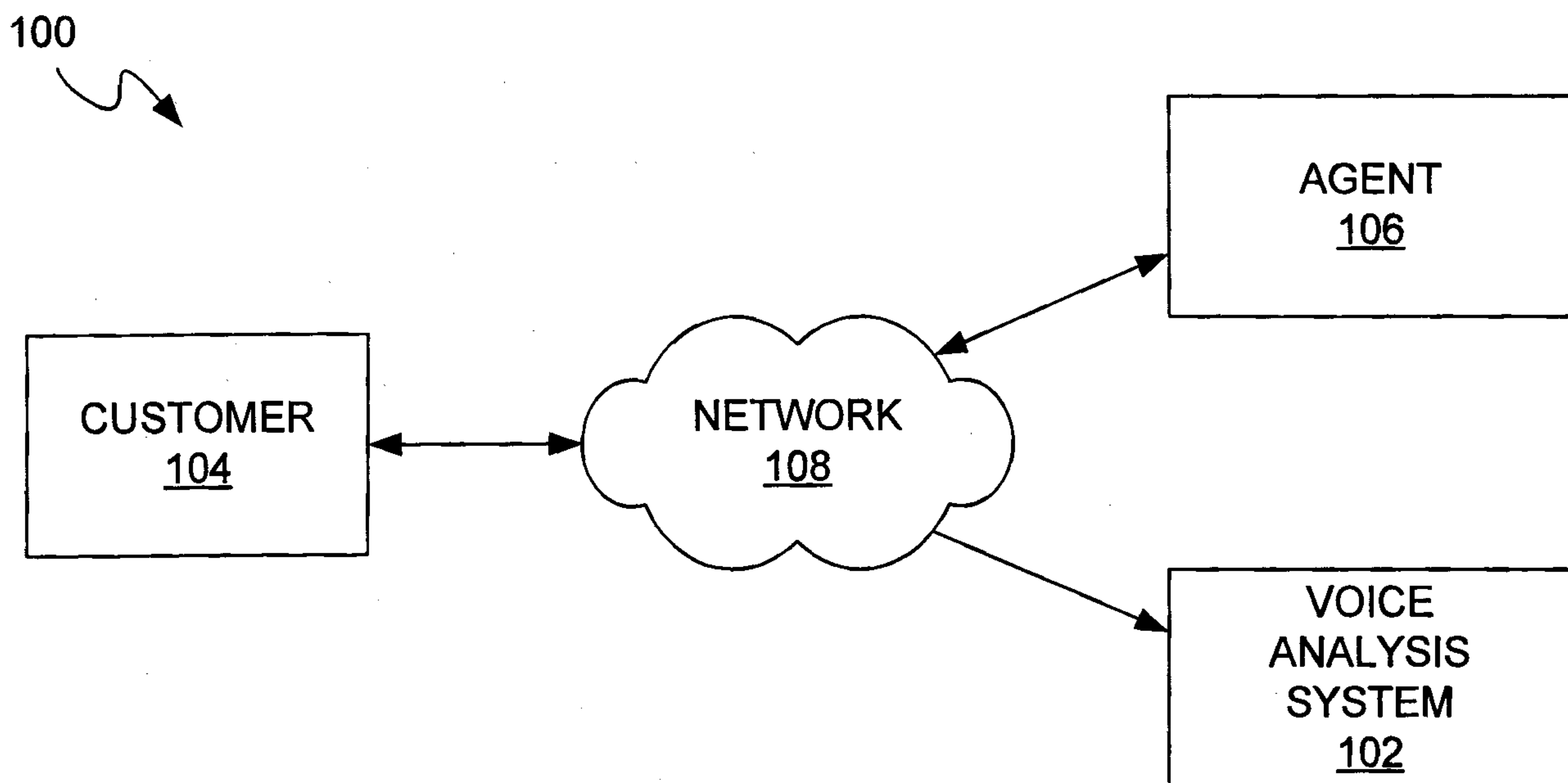
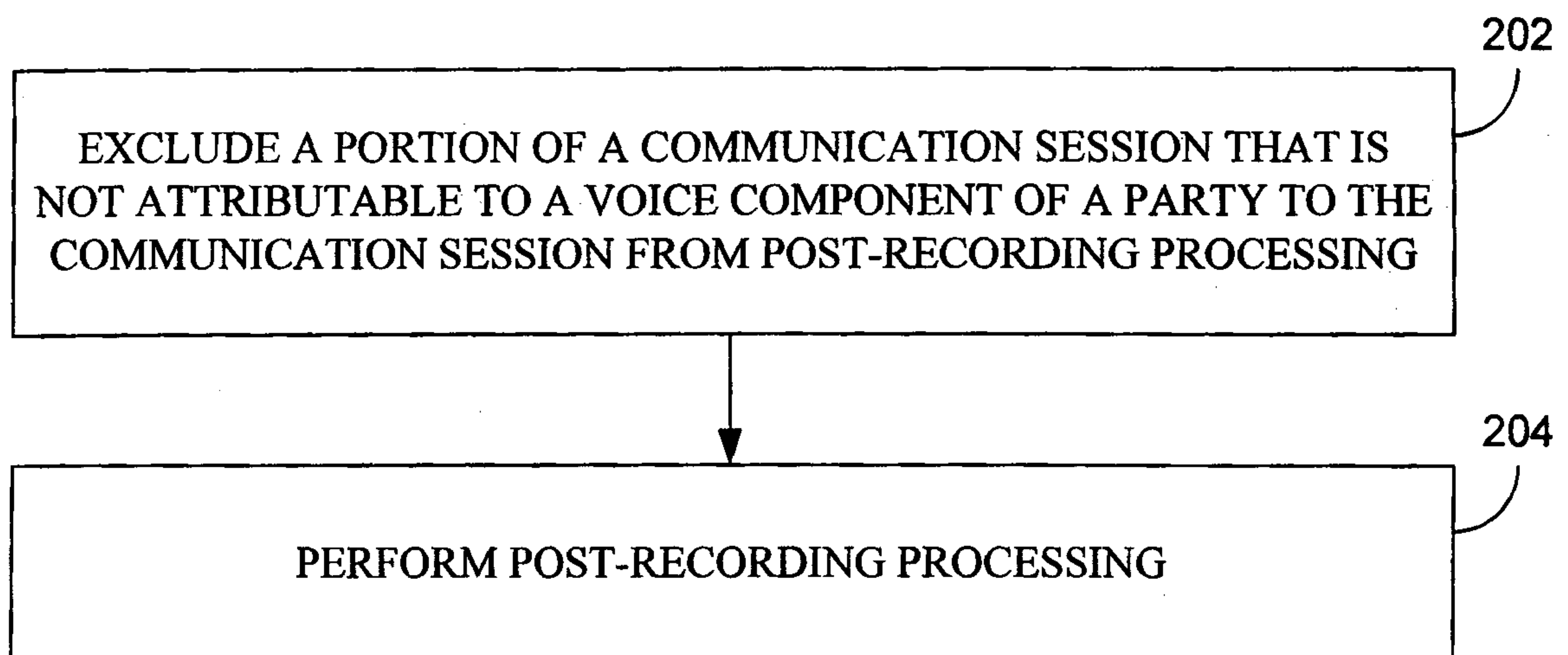
receive identifiers of the portions of the communication session that include at least one of an announcement and audio from an interactive voice response (IVR) system;

receive access to information corresponding to a database of potential announcements and potential audio from the IVR system; and

post-record process the identified portions using the database to exclude a portion of the information corresponding to the communication session that is not attributable to speech of at least one party of the communication session, wherein the portion of the communication session comprises audio from the interactive voice response (IVR) system.

11. The system of claim 10, wherein the voice analysis system is operative to perform at least one of speech recognition and phonetic analysis during the post-record processing.

12. The system of claim 10, wherein the voice analysis system comprises an identification system operative to identify portions of the communication session containing music, announcements and synthetic human voices.

**FIG. 1****FIG. 2**

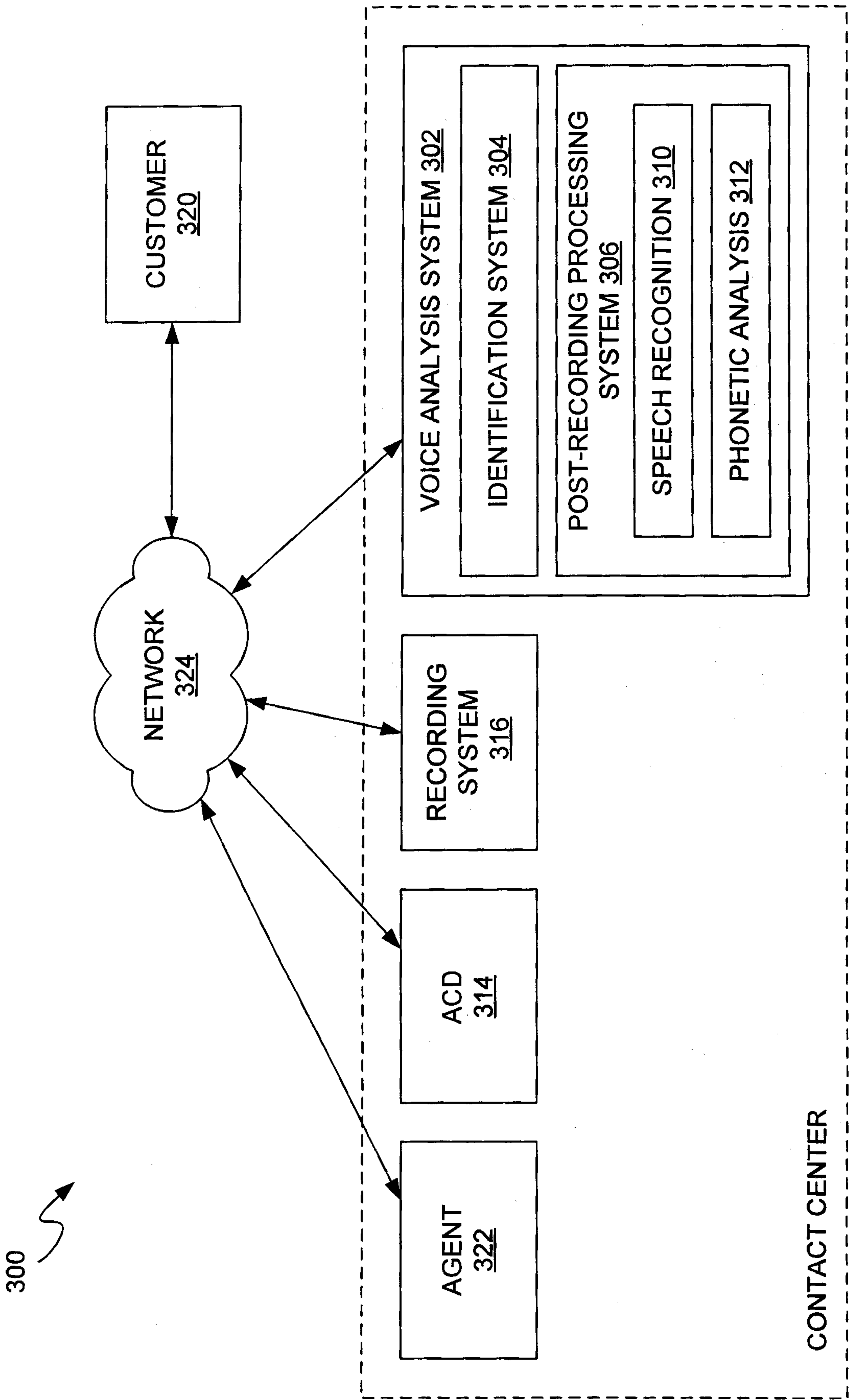
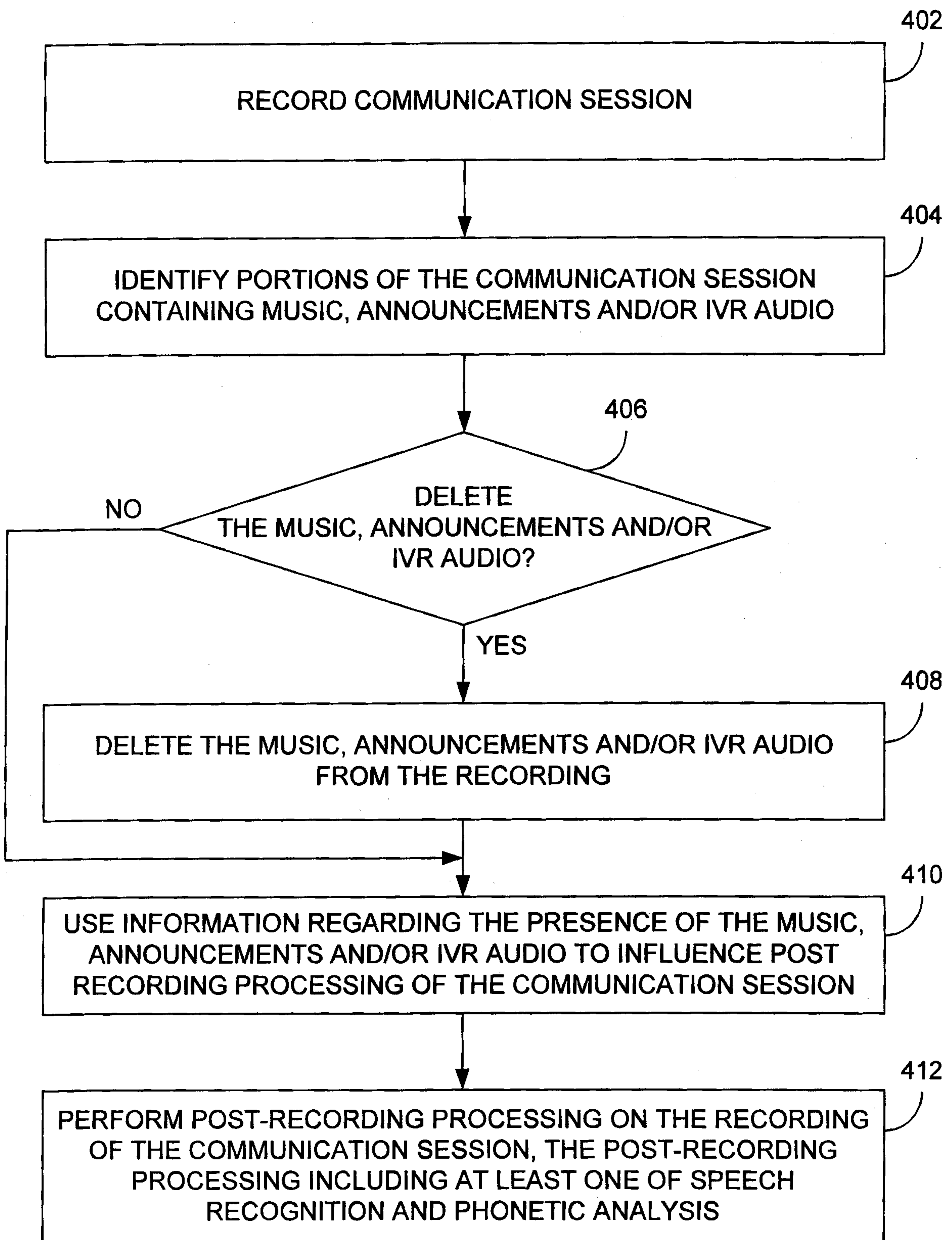


FIG. 3

**FIG. 4**

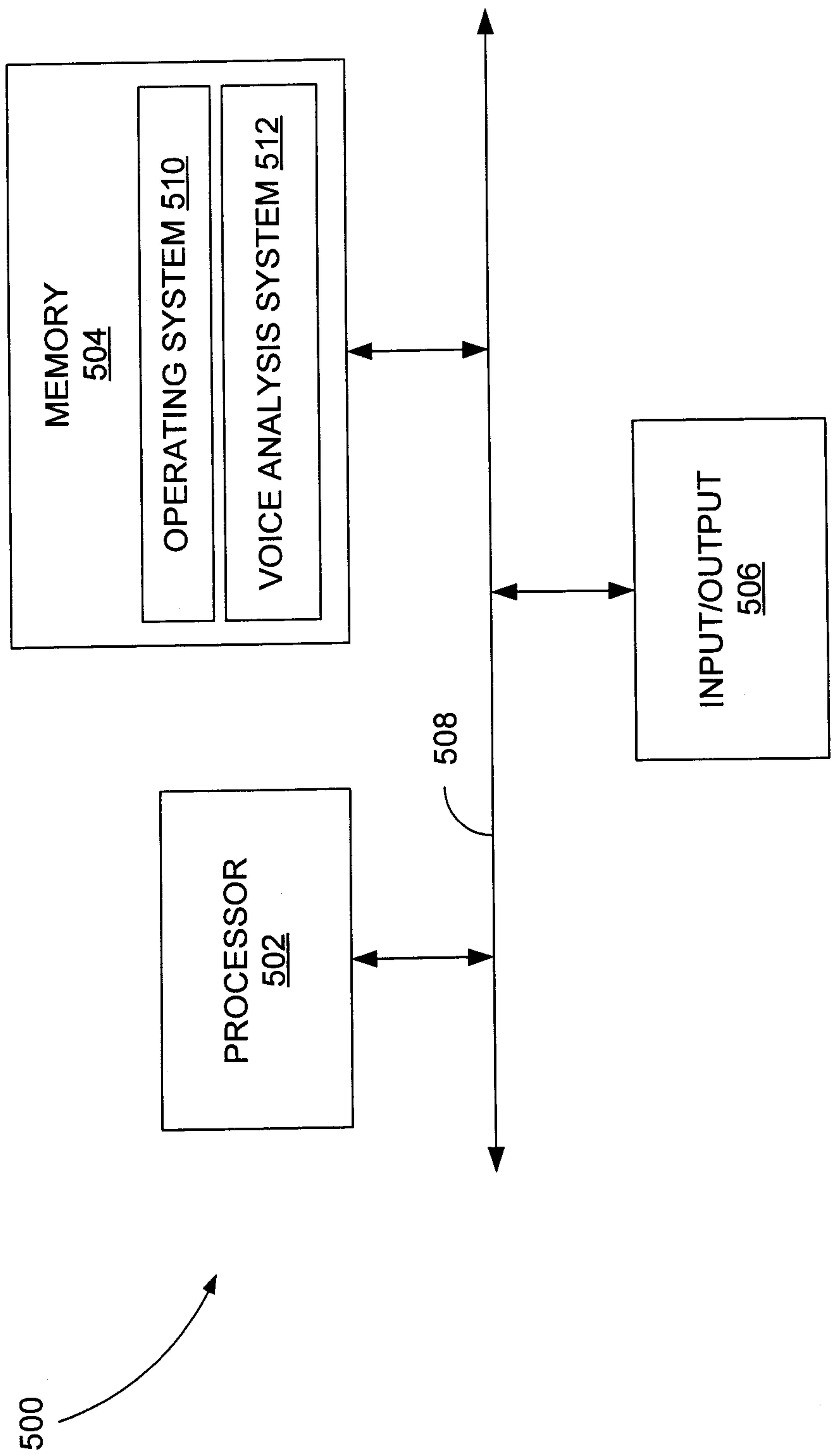


FIG. 5

100

