(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 November 2014 (13.11.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/180100 Al

(51) **International Patent Classification:**
H04M 3/56 (2006.01)

(21) **International Application Number:**
PCT/CN2013/085913

(22) **International Filing Date:**
24 October 2013 (24.10.2013)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**
2013 10172926.4 10 May 2013 (10.05.2013) CN

(71) **Applicant: TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED** [CN/CN]; Room 403, East Block 2, SEG Park, Zhenxing Road, Futian District, Shenzhen, Guangdong 518044 (CN).

(72) **Inventors: ZHANG, Xiaolong;** Room 403, East Block 2, SEG Park, Zhenxing Road, Futian District, Shenzhen, Guangdong 518044 (CN). **ZHAO, Yuan;** Room 403, East Block 2, SEG Park, Zhenxing Road, Futian District, Shenzhen, Guangdong 518044 (CN). **YANG, Ganrong;** Room 403, East Block 2, SEG Park, Zhenxing Road, Futian District, Shenzhen, Guangdong 518044 (CN).

(74) **Agent: ADVANCE CHINA LP. LAW OFFICE;** Suite 918-920, Dongshan Plaza, No. 69 Xianlie Central Road, Guangzhou, Guangdong 510095 (CN).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available):* AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

[Continued on nextpage]

(54) Title: METHOD, APPLICATION, AND DEVICE FOR AUDIO SIGNAL TRANSMISSION

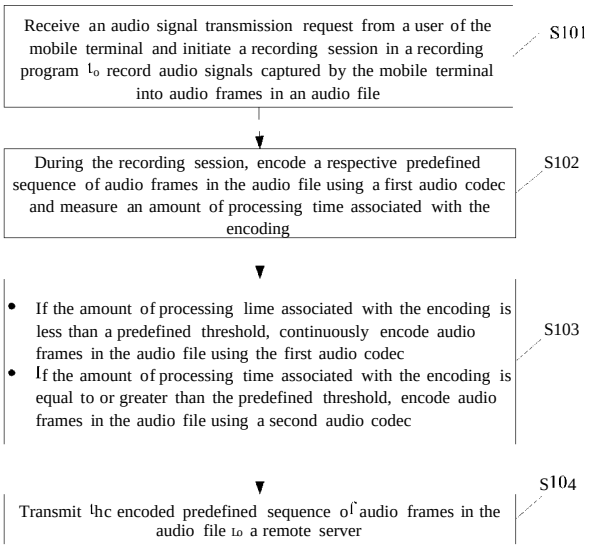


Fig. 1

(57) Abstract: The current invention discloses methods, applications, and devices for audio transmission from a mobile terminal. After receiving an audio signal transmission request from a user, the mobile terminal may initiate a recording session to record audio signals into audio frames. During the recording session, the terminal may adjust the audio codecs used for encoding the audio frames based on the workload and the performance of the terminal. By measuring and evaluating the encoding time, the terminal may change between using a floating-point AMR audio codec and a fixed-point AMR audio codec. The encoded audio frames are transmitted to a remote server. The current invention provides a flexible and efficient approach for audio signal encoding and transmission, balancing signal integrity and encoding speed at the same time.

WO 2014/180100 A1

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available):* ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

Description

METHOD, APPLICATION, AND DEVICE FOR AUDIO SIGNAL TRANSMISSION

RELATED APPLICATIONS

[0001] This application claims priority to Chinese Patent Application No. 2013 10172926.4, "Method, Application, and Terminal for Audio Signal Transmission," filed on May 10, 2013, which is hereby incorporated by reference in its entirety.

FIELD OF THE INVENTION

[0002] The present invention relates to the field of network technology, and in particular to audio technology involving methods, applications and devices for audio transmission.

BACKGROUND OF THE INVENTION

[0003] At present, recording technology is widely used in the client-side applications such as Instant Messenger (IM) and Social Networking Services (SNS). Recording technology is the basis of audio transmission, and it provides more convenient services for exchange and communication between users. In general, the conventional process of audio transmission includes: application provides a "Record" button in the chat interface; a user sends out an audio transmission request while pressing this button; the application initiates the recording program installed on the terminal, which starts recording; the recording program records audio input by users; the recording is ended when user releases the button; the recording program compiles the obtained complete audio signal data to form an audio file; the recording program loads the audio file to the application to conduct audio encoding and audio transmission. In the aforementioned process, since the compilation, encoding, and transmission of audio signals can only be conducted after the audio recording process is finished, the delay for audio transmission may be significant. Such a process cannot meet the needs of some scenarios that require swift audio transmission, making the user experience unsatisfactory regarding audio communication.

[0004] Accordingly, it is necessary and desirable to provide a new technology, so as to resolve the technical problem and improve the above-mentioned approach.

SUMMARY

[0005] The above deficiencies and other problems associated with audio encoding and transmission are reduced or eliminated by the invention disclosed below. In some embodiments, the invention is implemented in a computer system that has one or more processors, memory and one or more modules, programs or sets of instructions stored in the memory for performing multiple functions. Instructions for performing these functions may be included in a computer program product configured for execution by one or more processors.

[0006] One aspect of the invention involves a computer-implemented method performed by a mobile terminal. The mobile terminal may receive an audio signal transmission request from a user of the mobile terminal and initiate a recording session in a recording program to record audio signals captured by the mobile terminal into audio frames in an audio file. During the recording session, the mobile terminal may encode a respective predefined sequence of audio frames in the audio file using a first audio codec and measuring an amount of processing time associated with the encoding. If the amount of processing time associated with the encoding is less than a predefined threshold, the mobile terminal may continuously encode audio frames in the audio file using the first audio codec; and if the amount of processing time associated with the encoding is equal to or greater than the predefined threshold, the mobile terminal may encode audio frames in the audio file using a second audio codec. The mobile terminal may also transmit the encoded predefined sequence of audio frames in the audio file to a remote server.

[0007] Another aspect of the invention involves a computer system. The computer system may comprise a mobile terminal having one or more processors, memory, and one or more program modules stored in the memory and configured for execution by the one or more processors, the one or more program modules including: a receiving module configured to receive an audio signal transmission request from a user of the mobile terminal; a recording module configured to initiate a recording session in a recording program to record audio signals captured by the mobile terminal into audio frames in an audio file; an encoding module having a codec determination unit and an encoding unit, wherein: during the recording session: the encoding unit is configured to encode a respective predefined sequence of audio frames in the audio file using a first audio codec, the codec determination unit is configured to measure an amount of processing time associated with the encoding, if the amount of processing time associated with the encoding is less than a predefined threshold, the encoding unit is configured to continuously encode audio frames in the audio file using the first audio codec, if the amount of processing time associated with the encoding is equal to or greater than the predefined threshold, the encoding unit is configured to encode audio frames in the

audio file using a second audio codec; and a transmission module configured to transmit the encoded predefined sequence of audio frames in the audio file to a remote server.

[0008] Another aspect of the invention involves a non-transitory computer readable storage medium having stored therein instructions, which, when executed by a mobile terminal, cause the mobile terminal to: receive an audio signal transmission request from a user of the mobile terminal; initiate a recording session in a recording program to record audio signals captured by the mobile terminal into audio frames in an audio file; during the recording session: encode a respective predefined sequence of audio frames in the audio file using a first audio codec and measure an amount of processing time associated with the encoding; if the amount of processing time associated with the encoding is less than a predefined threshold, continuously encode audio frames in the audio file using the first audio codec; if the amount of processing time associated with the encoding is equal to or greater than the predefined threshold, encode audio frames in the audio file using a second audio codec; and transmit the encoded predefined sequence of audio frames in the audio file to the remote server.

[0009] Some embodiments may be implemented on either the terminal side or the server side of a terminal-server network environment.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] The aforementioned features and advantages of the invention as well as additional features and advantages thereof will be more clearly understood hereinafter as a result of a detailed description of preferred embodiments when taken in conjunction with the drawings.

[0011] Figure 1 is a flowchart illustrative of a method for audio encoding and transmission by a mobile terminal in accordance with some embodiments of the current invention.

[0012] Figure 2 is a flowchart illustrative of a method for audio encoding and transmission by a mobile terminal in accordance with some embodiments of the current invention, providing more details regarding the transmission process.

[0013] Figure 3 is a flowchart illustrative of a method for audio encoding and transmission by a mobile terminal in accordance with some embodiments of the current invention, providing more details regarding the encoding process.

[0014] Figure 4 is a flowchart illustrative of a method for audio encoding and transmission by a mobile terminal in accordance with some embodiments of the current invention, providing more details regarding the encoding process, particularly the codec determination process.

[0015] Figure 5 is a block diagram illustrative of a mobile terminal comprising modules configured to encode and transmit audio signals in accordance with some embodiments of the current invention.

[0016] Figure 6 is a block diagram illustrative of a mobile terminal comprising modules configured to encode and transmit audio signals in accordance with some embodiments of the current invention, providing more details for the modules.

[0017] Figure 7 is a block diagram illustrative of an encoding module in accordance with some embodiments of the current invention.

[0018] Figure 8 is a block diagram illustrative of a codec determination unit in accordance with some embodiments of the current invention.

[0019] Figure 9 is a block diagram illustrative of a transmission trigger module in accordance with some embodiments of the current invention.

[0020] Figure 10 is a block diagram of a mobile terminal in accordance with some embodiments of the current invention.

[0021] Like reference numerals refer to corresponding parts throughout the several views of the drawings.

DESCRIPTION OF EMBODIMENTS

[0022] Reference will now be made in detail to embodiments, examples of which are illustrated in the accompanying drawings. In the following detailed description, numerous specific details are set forth in order to provide a thorough understanding of the subject matter presented herein. But it will be apparent to one skilled in the art that the subject matter may be practiced without these specific details. In other instances, well-known methods, procedures, components, and circuits have not been described in detail so as not to unnecessarily obscure aspects of the embodiments.

[0023] Figure 1 is a flowchart illustrative of a method for audio encoding and transmission by a mobile terminal in accordance with some embodiments of the current invention. As shown by

step S101 in Figure 1, the mobile terminal may receive an audio signal transmission request from a user of the mobile terminal and initiate a recording session in a recording program to record audio signals captured by the mobile terminal into audio frames in an audio file. Step SI02 shows that during the recording session, the mobile terminal may encode a respective predefined sequence of audio frames in the audio file using a first audio codec and measure an amount of processing time associated with the encoding. As shown by Step 103, if the amount of processing time associated with the encoding is less than a predefined threshold, the mobile terminal may continuously encode audio frames in the audio file using the first audio codec; if the amount of processing time associated with the encoding is equal to or greater than the predefined threshold, the mobile terminal may encode audio frames in the audio file using a second audio codec. Step 104 shows that the mobile terminal may transmit the encoded predefined sequence of audio frames in the audio file to a remote server.

[0024] Figure 2 is a flowchart illustrative of a method for audio encoding and transmission by a mobile terminal in accordance with some embodiments of the current invention, providing more details regarding the transmission process.

[0025] As shown by step S201 of Figure 2, the mobile terminal may receive an audio signal transmission request from a user of the mobile terminal. The mobile terminal may be any device, such as but not limited to smart phones, tablet computers, personal digital assistants (PDAs), e-book readers, vehicle-associated terminals, and digital photo frames. It should also be noted that mobile terminal is used here as an example and the process and devices herein disclosed also apply to devices such as laptop and desktop computers. The audio signal transmission request refers to inputs, commands, and signals from the user of the mobile terminal to record, compile, encode, store, edit, and/or transmit audio signals. In addition to the user, the audio signal transmission request may be sent out by other computing devices such as a server.

[0026] The mobile terminal may receive the transmission request through the client-side of an application program installed on the mobile terminal, wherein the application program may include but not limited to: an instant messaging (IM) program, a live chat program, and a social networking program. A user of the mobile terminal may provide inputs through the user interfaces of the client-side of the program. As to the steps conducted by the mobile terminal as described below, the mobile terminal may carry out such steps through the client-side of the application program.

[0027] Referring to Figure 2, Step S202 shows that the mobile terminal may initiate a recording session. In some embodiments, the initiation is carried out by the client-side of an application program such as an IM program, which may be used as an example. The IM program may not incorporate a recording function, wherein the IM program may connect to a recording program and initiate the recording session in the recording program. For example, in a mobile terminal using the Android® operating system, the recording program may be a default Android recording program. In some embodiments, the client-side of the IM program may provide a user interface which allows the user to press a button to send an audio transmission request and initiate a recording session. For instance, the user interface may include a "Press-to-talk" button that may be used by the user. When the IM program receives the audio transmission request, the IM program may connect to a recording application programming interface (API) in the mobile terminal, which may be used to carry out the recording session. In some embodiments, the program (e.g. IM program) that receives the audio signal transmission request may also serve as the recording program and the initiation of the recording session is conducted within the same program.

[0028] The recording session is used to record audio signals captured by the mobile terminal into audio frames in an audio file. The recorder component of the mobile terminal is used to capture the original audio signals and the recording program may record, compile, and/or convert such audio signals into audio frames in an audio file. The audio file may be any format and the suitable format for the audio file and audio frames may vary. In some embodiments, the recording session may include: the recording program conducts digitization, sampling, and quantization of the analog audio data from the user or other sources, forming at least one audio frame; the least one audio frame may constitute an audio file having specific formats such as the PCM (pulse-code modulation) format.

[0029] An optional step may be included in the current process: the application program (e.g. IM program) receiving the audio transmission request may collect the audio frames from the recording program. Such a step is unneeded when the initiation and recording steps are conducted in the same program. The collection of audio frames from the recording program may be conducted in a frame-by-frame manner - whenever an audio frame is produced, it is collected, or by a batch-collection process - when certain number of audio frames are produced, these frames are collected together.

[0030] Referring to Figure 2, Step S203 shows that the mobile terminal may determine the codec to be used for encoding the audio frames. To balance encoding integrity and encoding speed, the codec that is used may be changed or adjusted from time to time. In some embodiments, the

encoding of the audio frames may be started instantaneously when the recording session begins. The encoding and transmission of the audio frame may be carried out almost entirely during the recording session. However, it should be noted that such simultaneity is not a requirement. Since the terminal may only encode the audio frames that have been created, there is likely to be a delay for the encoding and transmission and some parts of the encoding and transmission may be completed after the recording session is finished. In an alternatively approach, the encoding only starts after a delay that is set by the user or the program conducting the encoding.

[0031] The codec may be any kind of codec that may be used for audio frame encoding. In some embodiments, floating-point adaptive multi-rate (AMR) audio codec may be used; in some other embodiments, fixed-point AMR audio codec may be used. In general, floating-point AMR provides higher audio integrity during encoding but requires more system resources, such as central processing unit (CPU) capacity and memory occupancy. On the other hand, fixed-point AMR requires less system resources but is not as effective as the floating-point AMR audio codec in preserving audio integrity. In the current disclosure, floating-point AMR audio codec and the fixed-point AMR audio codec may be used as examples for two types of codec that may be selected by the mobile terminal to encode the audio frames and provide a balanced approach to maintain high audio quality and smooth operation. When there are enough system resources available, preferably the floating-point AMR audio codec is uses; and when system resources are limited, the fixed-point AMR audio codec is preferably used.

[0032] The mobile terminal may determine the codec to be used for encoding the audio frame by measuring the encoding time. The details of the determination process may be described below in Figures 3 and 4. Referring to Step S204 of Figure 2, the mobile terminal may encode the audio frames based on the determined codec. It should also be noted that there may be some variations as to which codec is used for the determination process, how many audio frames may be encoded after the codec is determined, and how often does the mobile terminal to change or adjust the codec.

[0033] Referring to Figure 2, Step S205 shows that the mobile terminal may cache the encoded audio frames.

[0034] As shown by Step S206 of Figure 2, the mobile terminal may determine whether transmission-triggering events occur. The transmission-triggering events may be any event or conditions that may or may not be satisfied. In some embodiments, the transmission-triggering event may not involve the cached audio frames from step S205. However, in some embodiments, the transmission-triggering event may be based on the cached audio frames. For example, the

transmission-triggering event occurs when the data size of the cached audio frames accumulate over a threshold, e.g. 2 kb, 3 kb, etc, or the transmission-triggering event occurs when the playtime of the cached audio frames are longer than a threshold, e.g. 1 second.

[0035] As shown in Figure 2, if the transmission-triggering event does not occur, the mobile terminal may not transmit the encoded audio frames. On the other hand, if the transmission-triggering event occurs, the mobile terminal may transmit the encoded audio frames in the audio file to a remote server. The thresholds may be set or scheduled by the programs and may be based on factors such as historical encoding performance, the system capacity (e.g. CPU speed and memory size) of the mobile terminal, network capacity, user requirement of audio integrity, and/or foreseeable workload. The remote server may be any computing device that provides processing and networking capabilities.

[0036] For example, the cached encoded audio frames may be measured by data size. Suppose every cached encoded audio frame is 1 kb and the threshold is set at 2 kb, the mobile terminal may detect that 2 kb of encoded audio frames are cached and may start the transmission. Alternatively, the cached encoded audio frames may be measured by the playing time of the frames. Suppose the playing time of the first three audio frames are 0.3 s, 0.5 s, and 0.4 s, respectively and the threshold is set at 1 second. Then after the first three audio frames are cached, the mobile terminal may detect that the accumulated playing time is greater than 1 second and transmit all three audio frames to the server.

[0037] In the present invention, when the mobile terminal receives the audio transmission request, it can initiate a recording session to record the audio signals captured by the mobile terminal into audio frames in an audio file. The audio frames are then encoded with the proper codec, cached, and transmitted to a remote server. The current invention thus realizes the process of recording the audio file while it sending out, reducing the audio transmission time delay, preserving audio integrity, effectively improving the audio transmission efficiency and satisfying the demands of users.

[0038] Figure 3 is a flowchart illustrative of a method for audio encoding and transmission by a mobile terminal in accordance with some embodiments of the current invention, providing more details regarding the encoding process. Here, the steps of S301 and S302 are similar to the steps S201 and S202 of Figure 2. The steps S303, S304, S305, and S306 provide an example for the process of determining the audio codec that may be used to encode the audio frames.

[0039] As shown in step S303, the mobile terminal may encode a respective predefined sequence of audio frames in the audio file using a first audio codec. The respective predefined sequence of audio frames may be the initial or subsequent sequence of audio frames of the audio file. The sequence may be defined by the program in the mobile terminal regarding the number of audio frames that are encoded and/or the start position of the sequence. The respective predefined sequence of audio frames is used to test the encoding speed and determine the codec for encoding the later sequences or the entire audio file. In some embodiments, the first audio codec may be the floating-point AMR audio codec, which may preserve more audio integrity but also consume more resources. In such cases, the goal is to conduct the test encoding with the codec that produces higher quality and continue with the same codec or switch to a codec that results in lower quality but consumes less resources, depending on the test encoding results. However, it should also be noted that in some embodiments, the first encoding codec may be any other codec such as the fixed-point AMR audio codec. In such cases, the basic goal is to save system resource and maintain smooth operation.

[0040] Referring to S303 of Figure 3, the mobile terminal may also measure an amount of processing time associated with the encoding of the respective predefined sequence of audio frames. In some embodiments, the measurement may be carried out by the program conducting the encoding.

[0041] As shown by step S304 of Figure 3, the mobile terminal may determine whether the amount of processing time associated with encoding the respective predefined sequence of audio frames is less than a predefined threshold. If the processing time is less than the threshold, the first audio codec is chosen and the mobile terminal continuously encode audio frames in the audio file using the first audio codec, e.g. the floating-point AMR audio codec, as indicated by step S305. On the other hand, if the processing time is equal or greater than the threshold, the mobile terminal encodes audio frames in the audio file using a second audio codec, e.g. the fixed-point AMR audio codec.

[0042] For example, suppose that audio file contains 10 audio frames and the threshold is one second. After the recording application starts to record the audio signals captured by the mobile terminal into audio frames, the first to third frames are collected and used as the respective predefined sequence by the program conducting the encoding, e.g. an instant message (IM) program. With a first audio codec, suppose the processing time for encoding the first frame is 0.2 s, the processing time for encoding the second frame is 0.8 s, and the processing time for encoding the third frame data is 0.2 s. Then the overall processing time may be (0.2 + 0.8 + 0.2) = 1.2 s, which is

more than the one second threshold, steering the IM program to switch to use the second audio codec. On the other hand, if the processing time for the three frames are: 0.1, 0.2, and 0.3 s, the overall time is 0.6 s, which is less than the threshold and the mobile terminal may continue encoding the audio frame with the first audio codec.

[0043] The threshold time may be set or schedule by the program conducting the encoding, based on factors such as historical encoding performance, the system capacity (e.g. CPU speed and memory size) of the mobile terminal, network capacity, user requirement of audio integrity, and/or foreseeable workload. The determination process may also be based on an average time for encoding each frame in the respective predefined sequence of audio frames, wherein the threshold is set to target the average time. In that case, suppose the threshold is 0.5 second and the processing time for the three frames are 0.2 s, 0.8 s, and 0.2 s, respectively, the average encoding is calculated as (0.2+0.8+0.2) /3 = 0.4s, less than the preset time threshold value 0.5 second, and the mobile terminal may determine that the first audio codec, e.g. the floating-point AMR audio codec, should be used.

[0044] As indicated above, the respective predefined sequence may be any sequence of audio frames in the audio file. In some embodiments, the respective predefined sequence may start from the first audio frame and after the encoding the respective predefined sequence, the mobile terminal may continue encoding the rest of the audio frames using the same codec for the respective predefined sequence or use the same audio codec for the entire audio file. Alternatively, the mobile terminal may switch to the second audio codec and use it for the rest of the audio frames or the all the audio frames in the audio file. On the other hand, the respective predefined sequence may start from positions other than the first audio frame of the audio file. The mobile terminal may use the first or second audio codec to encode the respective predefined sequence of audio frames, and then determines which audio codec should be used to encode the rest of the audio frames.

[0045] As shown in steps S307, S308, and S309 of Figure 3, the mobile terminal may cache the encoded audio frames, determine whether the transmission-triggering events exist, and transmit the encoded audio frames in the audio file to a remote server. These steps are largely similar to steps S205, S206, and S207 of Figure 2. In addition, the transmission of the encoded audio frames may include all the cached encoded audio frames or may be more limited. For example, the mobile terminal may transmit only the predefined sequence of encoded audio frames to the remote server.

[0046] Figure 4 is a flowchart illustrative of a method for audio encoding and transmission by a mobile terminal in accordance with some embodiments of the current invention, providing more details regarding the encoding process, particularly the codec determination process. Steps S401,

S402, S403, S404, S405, and S406 of Figure 4 are largely similar as the steps S301, S302, S303, S304, S305, and S306 of Figure 3. Referring to Figure 4, steps S407, S408, S409, and S410 provide examples for addition processes that may be used to further optimize codec selection and facilitate encoding.

[0047] As shown in step S407, in some embodiments, the mobile terminal may measure an amount of processing time associated with encoding a second set of respective predefined sequence of audio frames. Also as shown in step S407, in some embodiments, the mobile terminal may measure an average amount of processing time associated with encoding a frame in a preset number of frames. These approaches may be similar to what is illustrated in the step S403 and may be conducted after a preset time interval from the last codec selection or after encoding a preset number of audio frames after the last codec selection. In general, in steps S407-410, the mobile terminal make another attempt to adjust to the most suitable audio codec.

[0048] The measurement of encoding may vary, as indicated in step S407. For example, the mobile terminal may measure the total or average processing time of encoding a second set of respective predefined sequence of audio frames; the mobile terminal may measure the total or average processing time to encode each frame by the last used codec; the mobile terminal may also measure the total or average processing time to encode all the audio frames in the audio file that have been encoded. The key goal here is to provide a parameter that may be used to evaluate the encoding performance by the mobile terminal to decide what audio codec should be used for the rest of the audio frame in the audio file or until another determination process is carried out.

[0049] As shown in steps S408 to S410, the mobile terminal may determine if the amount of processing time is less than a second predefined threshold; if the processing time is less than the second predefined threshold, the mobile terminal may encode audio frames in the audio file using the first audio codec, e.g. the floating-point AMR audio codec; if the processing time is equal to or more than the second threshold, the mobile terminal may encode audio frames in the audio file using the second audio codec, e.g. the fixed-point AMR audio codec.

[0050] As indicated above, the particular measurement of processing time may vary. Accordingly, the threshold time should vary with the selection of the measurement. The threshold may be based on factors such as historical encoding performance, the system capacity (e.g. CPU speed and memory size) of the mobile terminal, network capacity, user requirement of audio integrity, and/or foreseeable workload.

[0051] Referring to Figure 4, steps S411, S412, and S414 are similar to the steps S307, S308, and S309 of Figure 3.

[0052] It should also be noted that during the encoding process, there may be one, two, or more determination processes (e.g. steps S403-S406 or steps S407-S410) regarding the audio codec. The number of determination processes or the intervals between the determination processes may be preset or scheduled by the encoding program.

[0053] The process shown in Figure 4 may be further illustrated by an example. After receiving an audio signal transmission request from the user - the user pressing the "press to talk" button on the user interface of an IM program - the IM program in the mobile terminal may initiate a recording session in a recording program to record audio signals captured by the mobile terminal into audio frames in an audio file. The IM program may then start to encode the audio frames after collecting the audio frames from the recording program. The IM program predefines the first three audio frames as the respective predefined sequence and uses the floating-point AMR audio codec to encode the first three frames. The processing time associated with encoding the first three audio frames is measured to be 0.2 s, 0.8 s, and 0.2 seconds, respectively. The predefined threshold time for the average processing time is set at 1 second, wherein the average time of encoding the respective predefined sequence of audio frames is (0.2 + 0.8 + 0.2)/3 = 0.4 s, which is less than the threshold. Thus, the mobile terminal continues to encode the audio frames using the floating-point AMR audio codec. After encoding an additional 12 audio frames (audio frames 4-15 of the audio file), the mobile terminal starts another determination process and encodes 3 audio frames (audio frames 16-18) using the floating-point AMR audio codec. The average processing time for these 3 audio frames is 0.8 s and the second threshold is set at 0.5 s. This time, the processing time is more than the second threshold and the mobile terminal switches to the fixed-point AMR audio codec for further encoding. The IM program at the same time caches all the encoded audio frames. The playing time for the first four audio frames are 0.2 s, 0.3 s, 0.4 s, and 0.3 second, while the threshold is set to be 1 second. After the first four encoded audio frames are cached, the transmission trigger event occurs and the four audio frames are transmitted by the mobile terminal to a remote server.

[0054] Figures 5-10 illustrate the computer systems that may be used to perform the methods described above. To avoid redundancy, not all the details and variations described for the method are herein included for the devices. Such details and variations should be considered included for the description of the devices as long as they are not in direct contradiction to the specific description provided for the methods.

[0055] Figure 5 is a block diagram illustrative of a mobile terminal 100 comprising modules configured to encode and transmit audio signals in accordance with some embodiments of the current invention. The terminal 100 may include one or more processors; memory; and one or more programs modules stored in the memory and configured for execution by the one or more processors, the one or more program modules including: a receiving module 501 configured to receive an audio signal transmission request from a user of the mobile terminal; a recording module 502 configured to initiate a recording session in a recording program to record audio signals captured by the mobile terminal into audio frames in an audio file; an encoding module 503 having a codec determination unit and an encoding unit, wherein: during the recording session: the encoding unit is configured to encode a respective predefined sequence of audio frames in the audio file using a first audio codec, the codec determination unit is configured to measure an amount of processing time associated with the encoding, if the amount of processing time associated with the encoding is less than a predefined threshold, the encoding unit is configured to continuously encode audio frames in the audio file using the first audio codec; if the amount of processing time associated with the encoding is equal to or greater than the predefined threshold, the encoding unit is configured to encode audio frames in the audio file using a second audio codec; and a transmission module 506 configured to transmit the encoded predefined sequence of audio frames in the audio file to a remote server.

[0056] Figure 6 is a block diagram illustrative of a mobile terminal 100 comprising modules configured to encode and transmit audio signals in accordance with some embodiments of the current invention, providing more details for the modules. In addition to the modules illustrated in Figure 5, the mobile terminal 100 may further comprise a caching module 504 and a transmission trigger module 505, wherein the caching module 504 is configured to cache the encoded predefined sequence of audio frames before transmission; the transmission trigger module 505 is configured to determine whether the cached audio frames exceed a threshold; and the transmission module 506 is further configured to transmit the encoded predefined sequence of audio frames starts when the cached audio frames exceed the threshold.

[0057] Figure 7 is a block diagram illustrative of an encoding module 503 in accordance with some embodiments of the current invention, wherein the encoding module 503 may comprise a codec determination unit 701 and an encoding unit 702. Figure 8 is a block diagram illustrative of the codec determination unit 701 in accordance with some embodiments of the current invention, wherein the codec determination unit 701 may comprise a measuring subunit 801 and a codec determination subunit 802.

[0058] In some embodiments, the encoding unit 702 is configured to encode a respective predefined sequence of audio frames in the audio file using a first audio codec, the measuring subunit 801 of the codec determination unit 701 is configured to measure an amount of processing time associated with the encoding. If the codec determination subunit 802 determines that the amount of processing time associated with the encoding is less than a predefined threshold, the encoding unit 702 is configured to continuously encode audio frames in the audio file using the first audio codec. If the codec determination subunit 802 determines that the amount of processing time associated with the encoding is equal to or greater than the predefined threshold, the encoding unit 702 is configured to encode audio frames in the audio file using a second audio codec. As indicated above, the first audio codec may be a floating-point adaptive multi-rate (AMR) audio codec and the second audio codec may be a fixed-point AMR audio codec.

[0059] In some embodiments, during the recording session, after a preset time interval, the encoding unit 702 may be further configured to continuously encode a second set of respective predefined sequence of audio frames in the audio file. The measuring subunit 801 of the codec determination unit 701 may be further configured to measure an amount of processing time associated with encoding the second set of respective predefined sequence. If the codec determination subunit 802 of the codec determination unit 701 determines that the amount of processing time associated with encoding the second set of respective predefined sequence is less than a second predefined threshold, the encoding unit 702 is further configured to encode audio frames in the audio file using the first audio codec. If the codec determination subunit 802 of the codec determination unit 701 determines that the amount of processing time associated with encoding the second set of respective predefined sequence is equal to or greater than the second predefined threshold, the encoding unit 702 is further configured to encode audio frames in the audio file using the second audio codec.

[0060] In some embodiments, during the recording session, after a preset number of frames are encoded, , the measuring subunit 801 of the codec determination unit 701 is further configured to measure an average amount of processing time associated with encoding a frame in the preset number of frames. If the codec determination subunit 802 of the codec determination unit 701 determines that the average amount of processing time associated with encoding a frame in the preset number of frames is less than a second predefined threshold, the encoding unit 702 is further configured to continue to encode audio frames in the audio file using the first audio codec. If the codec determination subunit 802 of the codec determination unit 701 determines that the average amount of processing time associated with encoding a frame in the preset number of frames is equal

to or greater than the second predefined threshold, the encoding unit 702 is further configured to continue to encode audio frames in the audio file using the second audio codec.

[0061] Figure 9 is a block diagram illustrative of a transmission trigger module 505 in accordance with some embodiments of the current invention. As shown in Figure 9, the transmission trigger module 505 may comprise a data size detection unit 901, a first determination unit 902, a playing time detection unit 903, and a second determination unit 904. As indicated above, the transmission trigger module 505 is configured to determine whether the cached audio frames exceed a threshold. In some embodiments, the data size determination unit 901 is configured to measure the data size of the cached audio frames and the first determination unit 902 is configured to determine whether the measurement exceeds the threshold, which refers to data size, e.g. 10 kb. In some embodiments, the playing time detection unit 903 is configured to measure the playing time of the audio frames and the second determination unit 904 is configured to determine whether the measure exceeds the threshold, which refers to a time period, e.g. 2 seconds.

[0062] Figure 10 is a block diagram of a mobile terminal in accordance with some embodiments of the current invention. The exemplary mobile terminal 100 typically includes one or more processing units (CPU's) 1002, one or more network or other communications interfaces 1004, memory 1010, and one or more communication buses 1009 for interconnecting these components. The communication buses 1009 may include circuitry (sometimes called a chipset) that interconnects and controls communications between system components. The mobile terminal 100 may include a user interface 1005, for instance, a display 1006, a keyboard 1008, a speaker 1003, and a recorder 1007. When the mobile terminal 100 is a smart phone or tablet, the user interface 1005 may include a touch screen, which is both a display and an input device. Memory 1010 may include high speed random access memory and may also include non-volatile memory, such as one or more magnetic disk storage devices. Memory 1010 may include mass storage that is remotely located from the CPU's 1002. In some embodiments, memory 1010 stores the following programs, modules and data structures, or a subset or superset thereof:

- an operating system 1012 that includes procedures for handling various basic system services and for performing hardware dependent tasks;
- a network communication module 1014 that is used for connecting the mobile terminal 101 to the server, the mobile terminals, and/or other computers via one or more communication networks (wired or wireless), such as the Internet, other wide area networks, local area networks, metropolitan area networks, and so on;

- a user interface module 1016 configured to receive user inputs through the user interface 1005;
- and a number of terminal-side application modules 1018 including the following:
- a receiving module 501 configured to receive an audio signal transmission request from a user of the mobile terminal;
- a recording module 502 configured to initiate a recording session in a recording program to record audio signals captured by the mobile terminal into audio frames in an audio file 1040;
- an encoding module 503 having a codec determination unit and an encoding unit, wherein: during the recording session: the encoding unit is configured to encode a respective predefined sequence of audio frames in the audio file 1040 using a first audio codec, the codec determination unit is configured to measure an amount of processing time associated with the encoding, if the amount of processing time associated with the encoding is less than a predefined threshold, the encoding unit is configured to continuously encode audio frames in the audio file 1040 using the first audio codec; if the amount of processing time associated with the encoding is equal to or greater than the predefined threshold, the encoding unit is configured to encode audio frames in the audio file 1040 using a second audio codec;
- a caching module 504 configured to cache the encoded predefined sequence of audio frames before transmission, wherein the cached audio frames are stored in the cache 1030;
- a transmission trigger module 505 configured to determine whether the cached audio frames exceed a threshold; and
- a transmission module 506 configured to transmit the encoded predefined sequence of audio frames in the audio file 1040 to a remote server.

[0063] While particular embodiments are described above, it will be understood it is not intended to limit the invention to these particular embodiments. On the contrary, the invention includes alternatives, modifications and equivalents that are within the spirit and scope of the appended claims. Numerous specific details are set forth in order to provide a thorough understanding of the subject matter presented herein. But it will be apparent to one of ordinary skill in the art that the subject matter may be practiced without these specific details. In other instances, well-known methods, procedures, components, and circuits have not been described in detail so as not to unnecessarily obscure aspects of the embodiments.

[0064] The terminology used in the description of the invention herein is for the purpose of describing particular embodiments only and is not intended to be limiting of the invention. As used in the description of the invention and the appended claims, the singular forms "a," "an," and "the" are intended to include the plural forms as well, unless the context clearly indicates otherwise. It will also be understood that the term "and/or" as used herein refers to and encompasses any and all possible combinations of one or more of the associated listed items. It will be further understood that the terms "includes," "including," "comprises," and/or "comprising," when used in this specification, specify the presence of stated features, operations, elements, and/or components, but do not preclude the presence or addition of one or more other features, operations, elements, components, and/or groups thereof.

[0065] As used herein, the term "if may be construed to mean "when" or "upon" or "in response to determining" or "in accordance with a determination" or "in response to detecting," that a stated condition precedent is true, depending on the context. Similarly, the phrase "if it is determined [that a stated condition precedent is true]" or "if [a stated condition precedent is true]" or "when [a stated condition precedent is true]" may be construed to mean "upon determining" or "in response to determining" or "in accordance with a determination" or "upon detecting" or "in response to detecting" that the stated condition precedent is true, depending on the context.

[0066] Although some of the various drawings illustrate a number of logical stages in a particular order, stages that are not order dependent may be reordered and other stages may be combined or broken out. While some reordering or other groupings are specifically mentioned, others will be obvious to those of ordinary skill in the art and so do not present an exhaustive list of alternatives. Moreover, it should be recognized that the stages could be implemented in hardware, firmware, software or any combination thereof.

[0067] The foregoing description, for purpose of explanation, has been described with reference to specific embodiments. However, the illustrative discussions above are not intended to be exhaustive or to limit the invention to the precise forms disclosed. Many modifications and variations are possible in view of the above teachings. The embodiments were chosen and described in order to best explain the principles of the invention and its practical applications, to thereby enable others skilled in the art to best utilize the invention and various embodiments with various modifications as are suited to the particular use contemplated.

Claims

1. A method for transferring audio signals from a mobile terminal to a remote server, the method comprising:

at the mobile terminal having one or more processors and memory storing programs executed by the one or more processors,

receiving an audio signal transmission request from a user of the mobile terminal;

initiating a recording session in a recording program to record audio signals captured by the mobile terminal into audio frames in an audio file;

during the recording session:

encoding a respective predefined sequence of audio frames in the audio file using a first audio codec and measuring an amount of processing time associated with the encoding;

if the amount of processing time associated with the encoding is less than a predefined threshold, continuously encoding audio frames in the audio file using the first audio codec;

if the amount of processing time associated with the encoding is equal to or greater than the predefined threshold, encoding audio frames in the audio file using a second audio codec; and

transmitting the encoded predefined sequence of audio frames in the audio file to the remote server.

2. The method of claim 1, wherein:

the first audio codec is a floating-point adaptive multi-rate (AMR) audio codec.

3. The method of claim 1, wherein:

the second audio codec is a fixed-point AMR audio codec.

4. The method of claim 1, further comprising:

during the recording session:

after a preset time interval, continuously encoding a second set of respective predefined sequence of audio frames in the audio file and measuring an amount of processing time associated with encoding the second set of respective predefined sequence;

if the amount of processing time associated with encoding the second set of respective predefined sequence is less than a second predefined threshold, encoding audio frames in the audio file using the first audio codec; and

if the amount of processing time associated with encoding the second set of respective predefined sequence is equal to or greater than the second predefined threshold, encoding audio frames in the audio file using the second audio codec.

5. The method of claim 1, further comprising:

during the recording session:

after a preset number of frames are encoded, measuring an average amount of processing time associated with encoding a frame in the preset number of frames;

if the average amount of processing time associated with encoding a frame in the preset number of frames is less than a second predefined threshold, continuing to encode audio frames in the audio file using the first audio codec; and

if the average amount of processing time associated with encoding a frame in the preset number of frames is equal to or greater than the second predefined threshold, continuing to encode audio frames in the audio file using the second audio codec.

6. The method of claim 1, further comprising:

caching the encoded predefined sequence of audio frames before transmission.

7. The method of claim 6, wherein:

the step of transmitting the encoded predefined sequence of audio frames starts when the cached audio frames exceed a threshold.

8. The method of claim 7, wherein:

the threshold is a time period and the cached audio frames are measured by a playing time of the audio frames.

9. A computer system comprising a mobile terminal, the mobile terminal having:

one or more processors;

memory; and

one or more programs modules stored in the memory and configured for execution by the one or more processors, the one or more program modules including:

a receiving module configured to receive an audio signal transmission request from a user of the mobile terminal;

a recording module configured to initiate a recording session in a recording program to record audio signals captured by the mobile terminal into audio frames in an audio file;

an encoding module having a codec determination unit and an encoding unit, wherein:

during the recording session:

the encoding unit is configured to encode a respective predefined sequence of audio frames in the audio file using a first audio codec,

the codec determination unit is configured to measure an amount of processing time associated with the encoding,

if the amount of processing time associated with the encoding is less than a predefined threshold, the encoding unit is configured to continuously encode audio frames in the audio file using the first audio codec,

if the amount of processing time associated with the encoding is equal to or greater than the predefined threshold, the encoding unit is configured to encode audio frames in the audio file using a second audio codec; and

a transmission module configured to transmit the encoded predefined sequence of audio frames in the audio file to a remote server.

10. The computer system of claim 9, wherein:

the first audio codec is a floating-point adaptive multi-rate (AMR) audio codec.

11. The computer system of claim 9, wherein:

the second audio codec is a fixed-point AMR audio codec.

12. The computer system of claim 9, wherein:

during the recording session:

after a preset time interval, the encoding unit is further configured to continuously encode a second set of respective predefined sequence of audio frames in the audio file;

the codec determination unit is further configured to measure an amount of processing time associated with encoding the second set of respective predefined sequence;

if the amount of processing time associated with encoding the second set of respective predefined sequence is less than a second predefined threshold, the encoding unit is further configured to encode audio frames in the audio file using the first audio codec; and

if the amount of processing time associated with encoding the second set of respective predefined sequence is equal to or greater than the second predefined threshold, the encoding unit is further configured to encode audio frames in the audio file using the second audio codec.

13. The computer system of claim 9, wherein:

during the recording session:

after a preset number of frames are encoded, the codec determination unit is further configured to measure an average amount of processing time associated with encoding a frame in the preset number of frames;

if the average amount of processing time associated with encoding a frame in the preset number of frames is less than a second predefined threshold, the encoding unit is further configured to continue to encode audio frames in the audio file using the first audio codec; and

if the average amount of processing time associated with encoding a frame in the preset number of frames is equal to or greater than the second predefined threshold, the encoding unit is further configured to continue to encode audio frames in the audio file using the second audio codec.

14. The computer system of claim 9, further comprising a caching module, wherein:

the caching module is configured to cache the encoded predefined sequence of audio frames before transmission.

15. The computer system of claim 14, wherein:

the caching module is configured to determine whether the cached audio frames exceed a threshold; and

the transmission module is configured to transmit the encoded predefined sequence of audio frames starts when the cached audio frames exceed the threshold.

16. The computer system of claim 15, wherein:

the threshold is a time period and the cached audio frames are measured by a playing time of the audio frames.

17. A non-transitory computer readable storage medium having stored therein one or more instructions, which, when executed by a mobile terminal, cause the mobile terminal to:

receive an audio signal transmission request from a user of the mobile terminal;

initiate a recording session in a recording program to record audio signals captured by the mobile terminal into audio frames in an audio file;

during the recording session:

encode a respective predefined sequence of audio frames in the audio file using a first audio codec and measure an amount of processing time associated with the encoding;

if the amount of processing time associated with the encoding is less than a predefined threshold, continuously encode audio frames in the audio file using the first audio codec;

if the amount of processing time associated with the encoding is equal to or greater than the predefined threshold, encode audio frames in the audio file using a second audio codec; and

transmit the encoded predefined sequence of audio frames in the audio file to the remote server.

18. The non-transitory computer readable storage medium of claim 17, wherein:

the first audio codec is a floating-point adaptive multi-rate (AMR) audio codec.

19. The non-transitory computer readable storage medium of claim 17, wherein:

the second audio codec is a fixed-point AMR audio codec.

20. The non-transitory computer readable storage medium of claim 17, wherein:

during the recording session, the instruction further causes the mobile terminal to:

after a preset time interval, continuously encode a second set of respective predefined sequence of audio frames in the audio file and measure an amount of processing time associated with encoding the second set of respective predefined sequence;

if the amount of processing time associated with encoding the second set of respective predefined sequence is less than a second predefined threshold, encode audio frames in the audio file using the first audio codec; and

if the amount of processing time associated with encoding the second set of respective predefined sequence is equal to or greater than the second predefined threshold, encode audio frames in the audio file using the second audio codec.

21. The non-transitory computer readable storage medium of claim 17, wherein:

during the recording session, the instruction further causes the mobile terminal to:

after a preset number of frames are encoded, measure an average amount of processing time associated with encoding a frame in the preset number of frames;

if the average amount of processing time associated with encoding a frame in the preset number of frames is less than a second predefined threshold, continue to encode audio frames in the audio file using the first audio codec; and

if the average amount of processing time associated with encoding a frame in the preset number of frames is equal to or greater than the second predefined threshold, continue to encode audio frames in the audio file using the second audio codec.

22. The non-transitory computer readable storage medium of claim 17, wherein the instruction further causes the mobile terminal to:

cache the encoded predefined sequence of audio frames before transmission.

23. The non-transitory computer readable storage medium of claim 22, wherein:

the mobile terminal starts to transmit the encoded predefined sequence of audio frames when the cached audio frames exceed a threshold.

24. The non-transitory computer readable storage medium of claim 23, wherein:

the threshold is a time period and the cached audio frames are measured by a playing time of the audio frames.

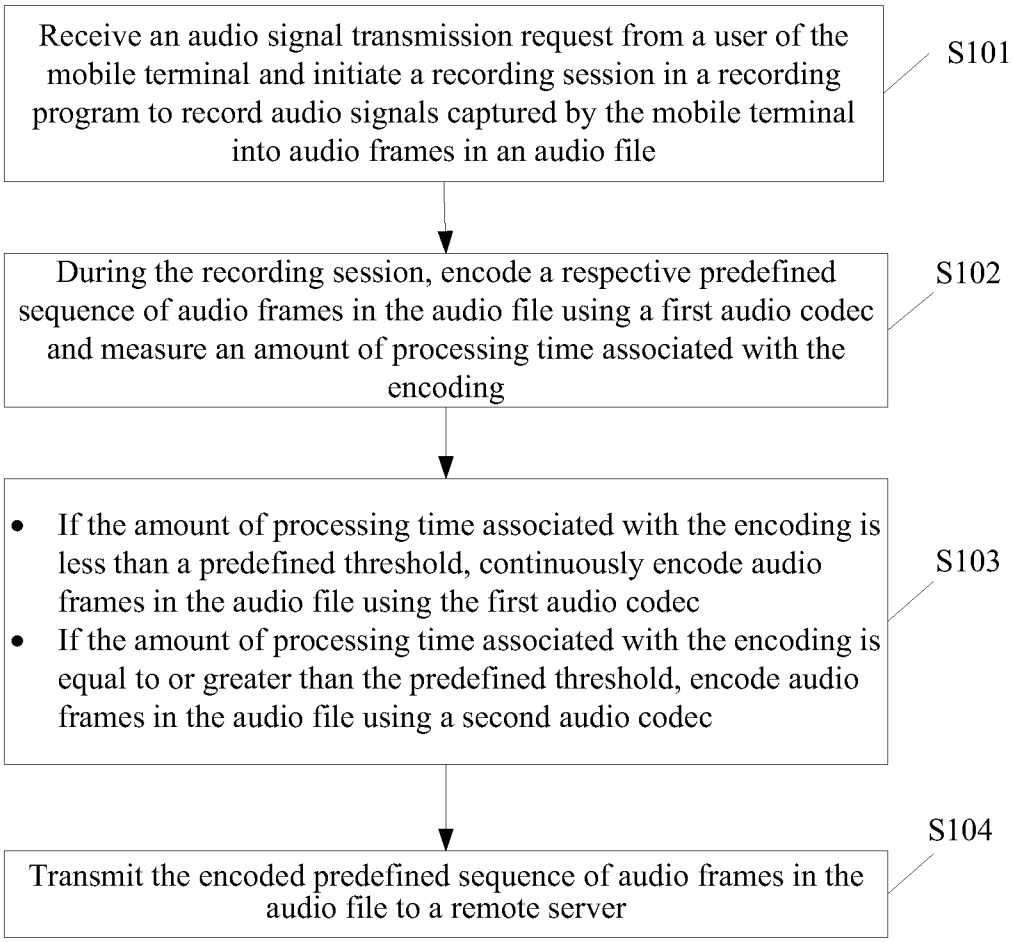

Receive an audio signal transmission request from a user of the mobile terminal and initiate a recording session in a recording program to record audio signals captured by the mobile terminal into audio frames in an audio file
S101
During the recording session, encode a respective predefined sequence of audio frames in the audio file using a first audio codec and measure an amount of processing time associated with the encoding
S102
If the amount of processing time associated with the encoding is less than a predefined threshold, continuously encode audio frames in the audio file using the first audio codec
If the amount of processing time associated with the encoding is equal to or greater than the predefined threshold, encode audio frames in the audio file using a second audio codec
S103
Transmit the encoded predefined sequence of audio frames in the audio file to a remote server
S104

Fig. 1

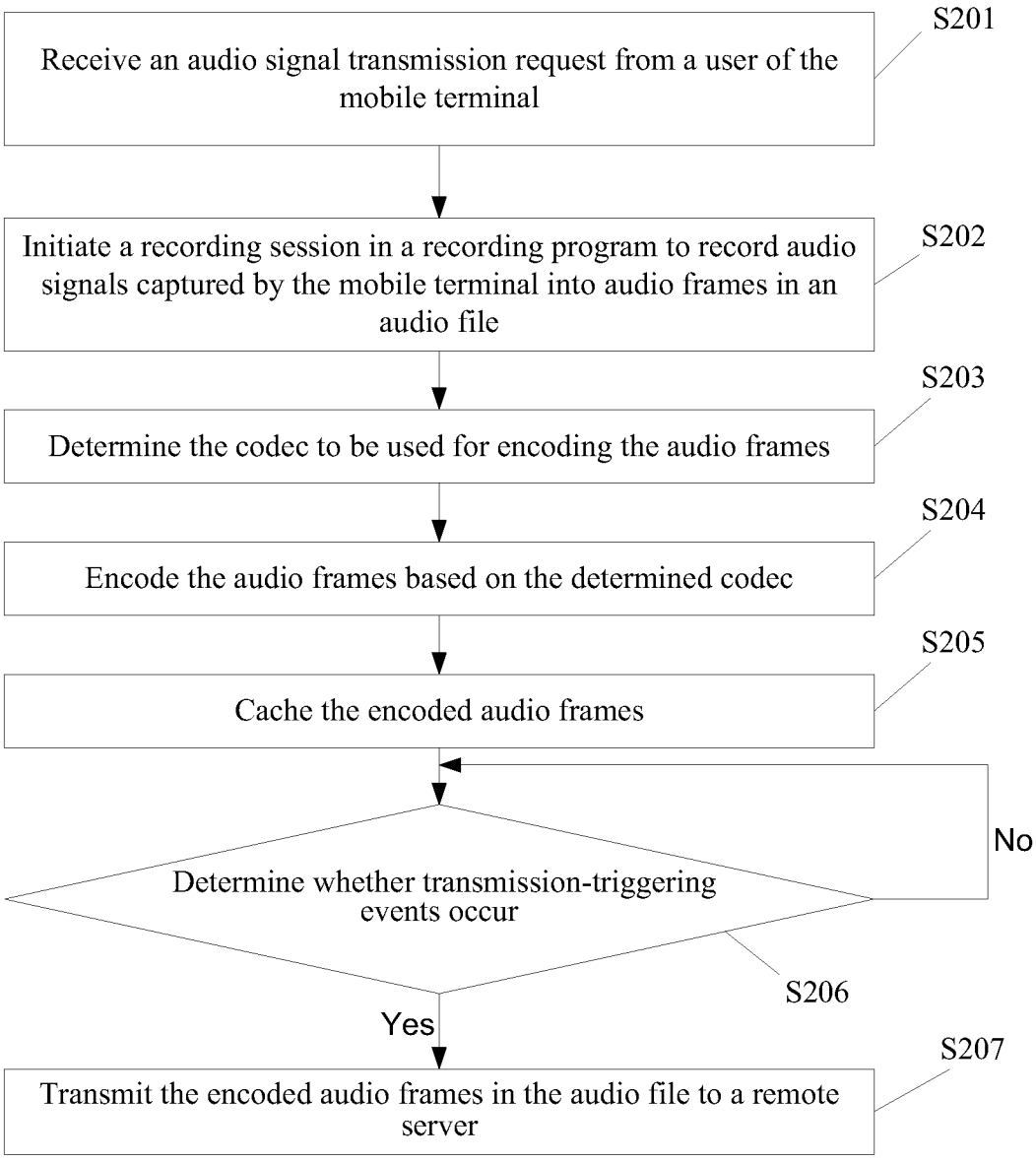

Receive an audio signal transmission request from a user of the mobile terminal
S201
Initiate a recording session in a recording program to record audio signals captured by the mobile terminal into audio frames in an audio file
S202
Determine the codec to be used for encoding the audio frames
S203
Encode the audio frames based on the determined codec
S204
Cache the encoded audio frames
S205
Determine whether transmission-triggering events occur
No
S206
Yes
Transmit the encoded audio frames in the audio file to a remote server
S207

Fig. 2

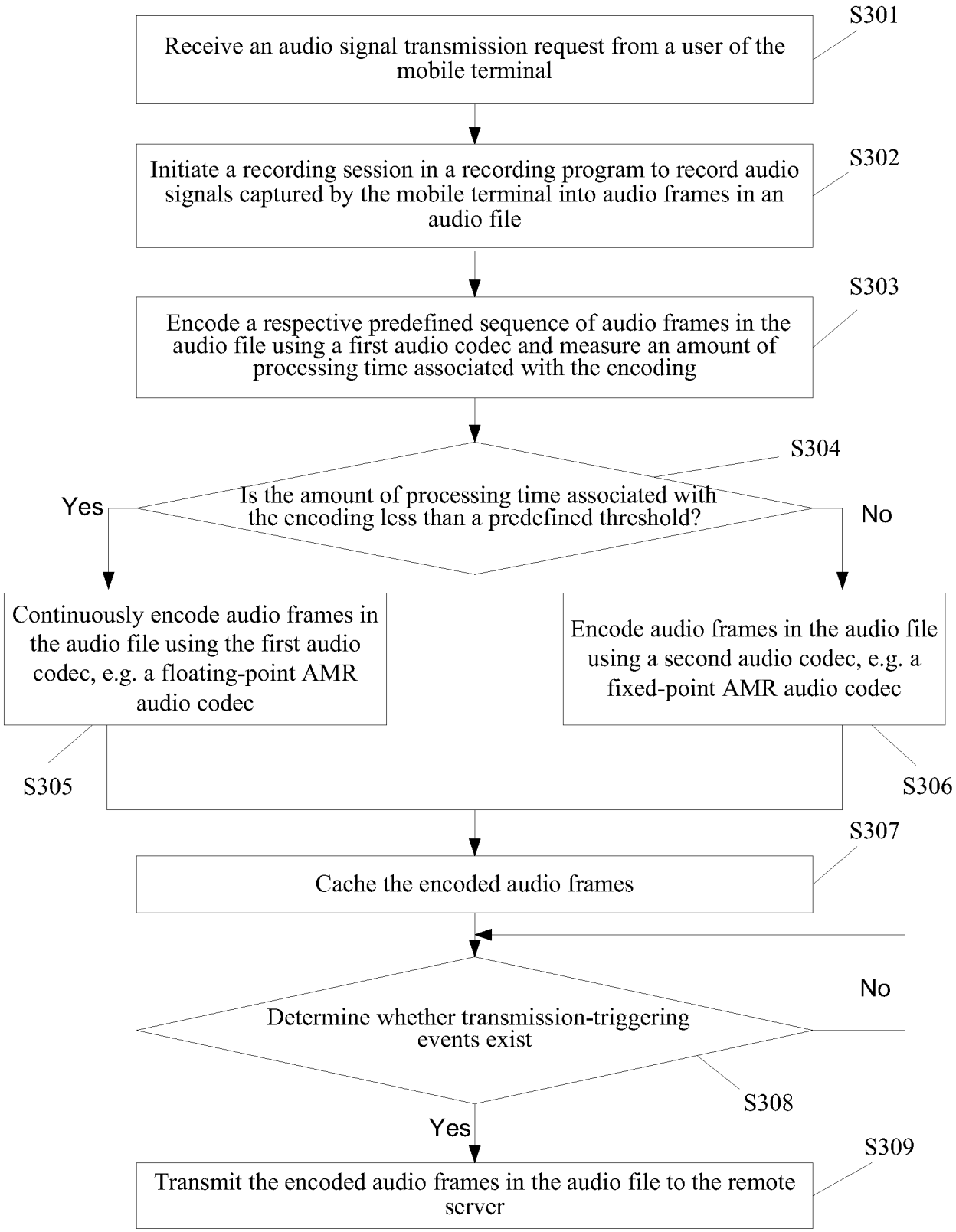

S301
Receive an audio signal transmission request from a user of the mobile terminal
S302
Initiate a recording session in a recording program to record audio signals captured by the mobile terminal into audio frames in an audio file
S303
Encode a respective predefined sequence of audio frames in the audio file using a first audio codec and measure an amount of processing time associated with the encoding
S304
Is the amount of processing time associated with the encoding less than a predefined threshold?
Yes
No
Continuously encode audio frames in the audio file using the first audio codec, e.g. a floating-point AMR audio codec
Encode audio frames in the audio file using a second audio codec, e.g. a fixed-point AMR audio codec
S305
S306
S307
Cache the encoded audio frames
Determine whether transmission-triggering events exist
No
S308
Yes
S309
Transmit the encoded audio frames in the audio file to the remote server

Fig. 3

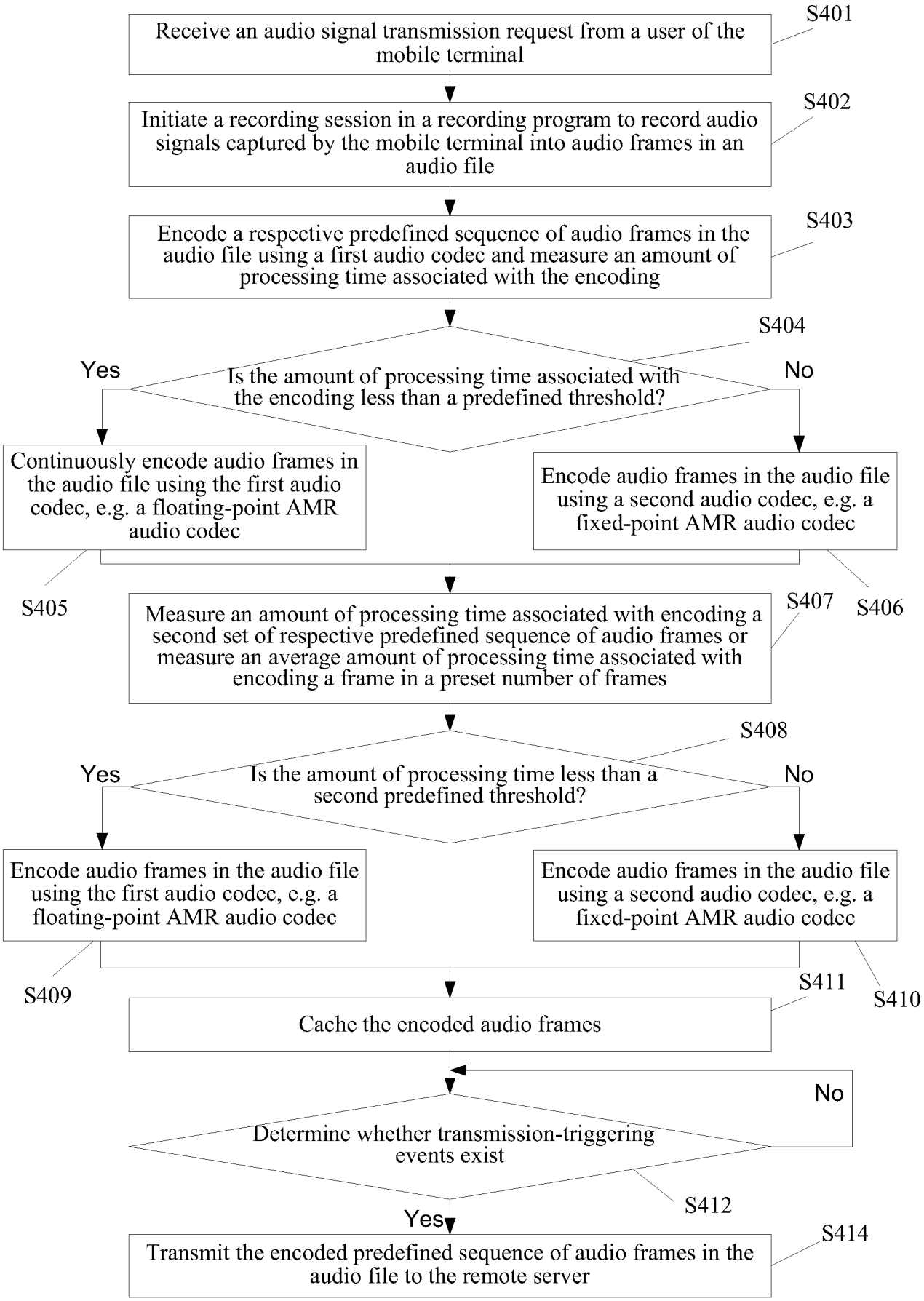

Receive an audio signal transmission request from a user of the mobile terminal
S401
Initiate a recording session in a recording program to record audio signals captured by the mobile terminal into audio frames in an audio file
S402
Encode a respective predefined sequence of audio frames in the audio file using a first audio codec and measure an amount of processing time associated with the encoding
S403
Is the amount of processing time associated with the encoding less than a predefined threshold?
S404
Yes
No
Continuously encode audio frames in the audio file using the first audio codec, e.g. a floating-point AMR audio codec
Encode audio frames in the audio file using a second audio codec, e.g. a fixed-point AMR audio codec
S405
S406
Measure an amount of processing time associated with encoding a second set of respective predefined sequence of audio frames or measure an average amount of processing time associated with encoding a frame in a preset number of frames
S407
Is the amount of processing time less than a second predefined threshold?
S408
Yes
No
Encode audio frames in the audio file using the first audio codec, e.g. a floating-point AMR audio codec
Encode audio frames in the audio file using a second audio codec, e.g. a fixed-point AMR audio codec
S409
S410
Cache the encoded audio frames
S411
Determine whether transmission-triggering events exist
No
S412
Yes
Transmit the encoded predefined sequence of audio frames in the audio file to the remote server
S414

Fig. 4

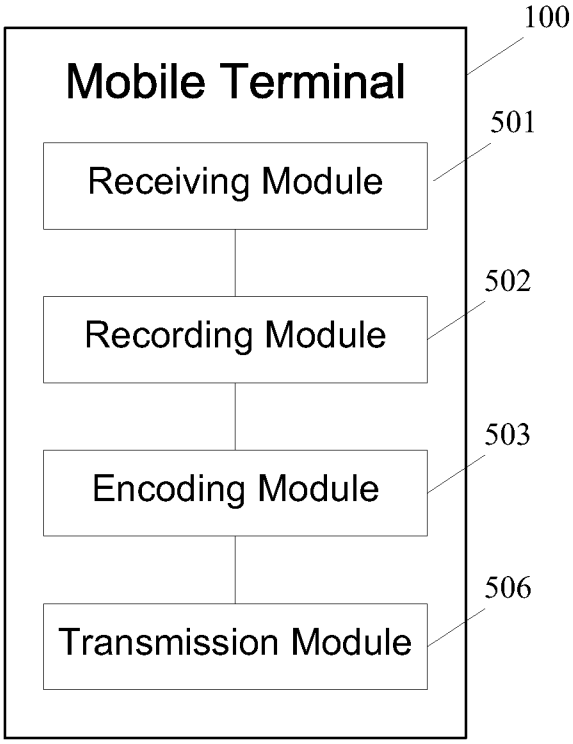

100
Mobile Terminal
501
Receiving Module
502
Recording Module
503
Encoding Module
506
Transmission Module

Fig. 5

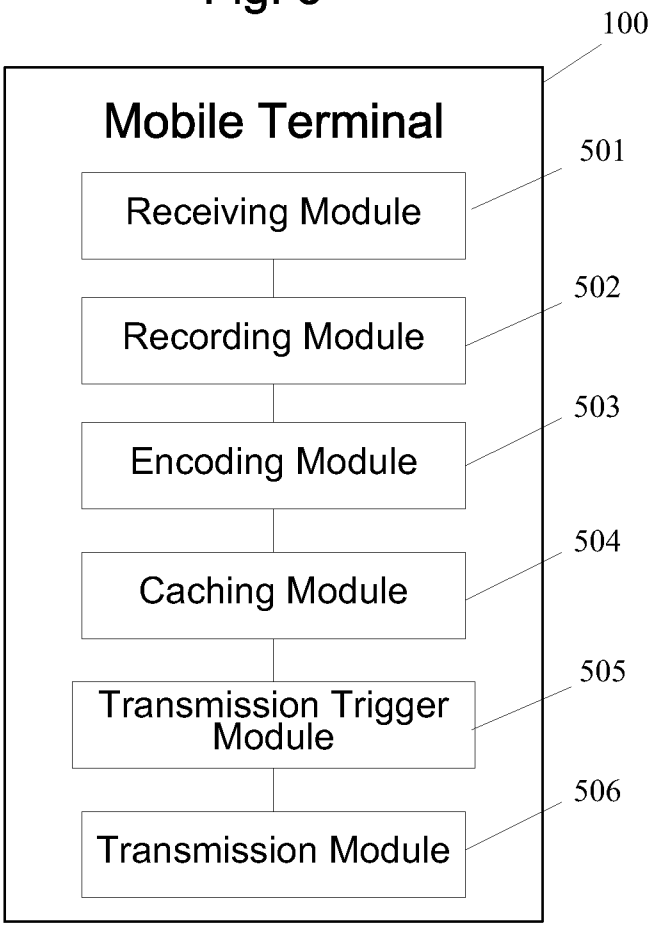

100
Mobile Terminal
501
Receiving Module
502
Recording Module
503
Encoding Module
504
Caching Module
505
Transmission Trigger Module
506
Transmission Module

Fig. 6

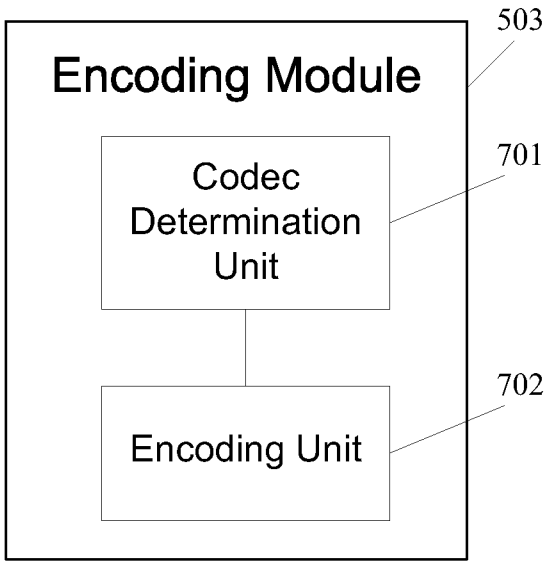


Fig. 7

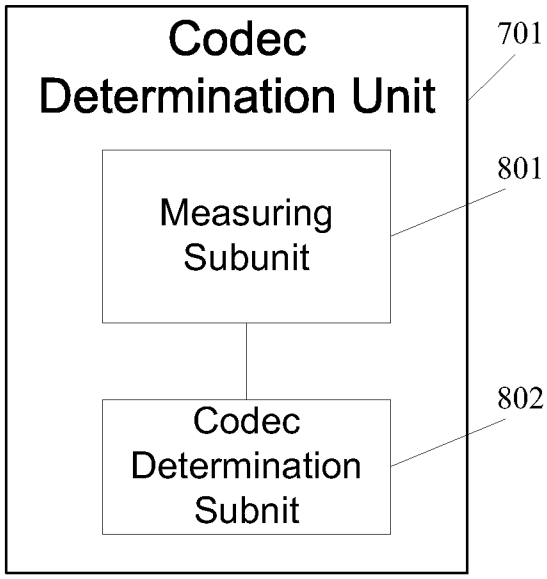


Fig. 8

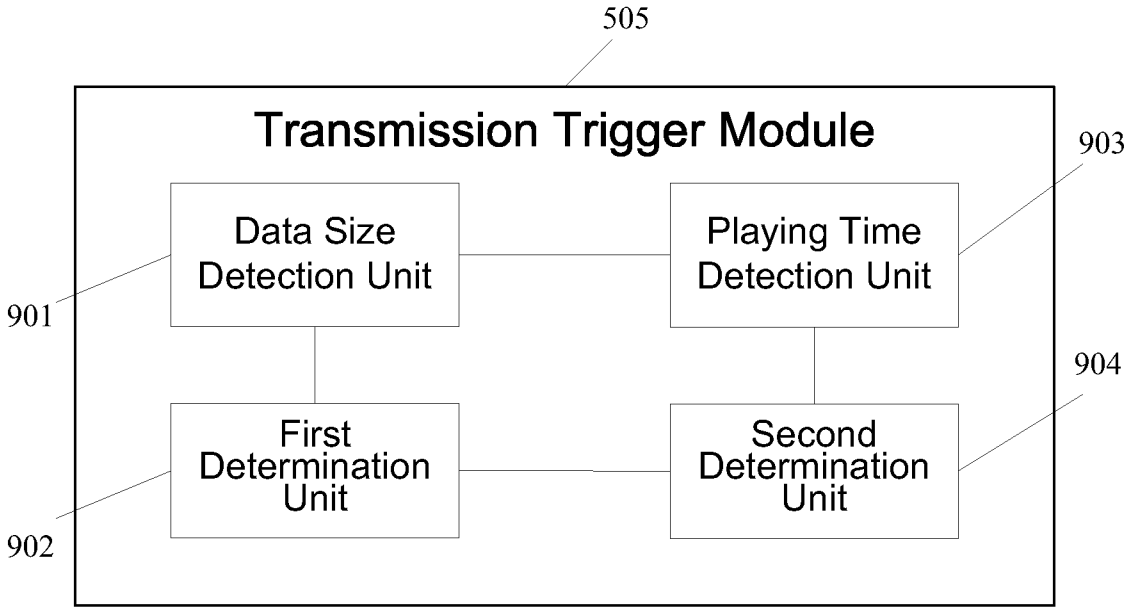


Fig. 9

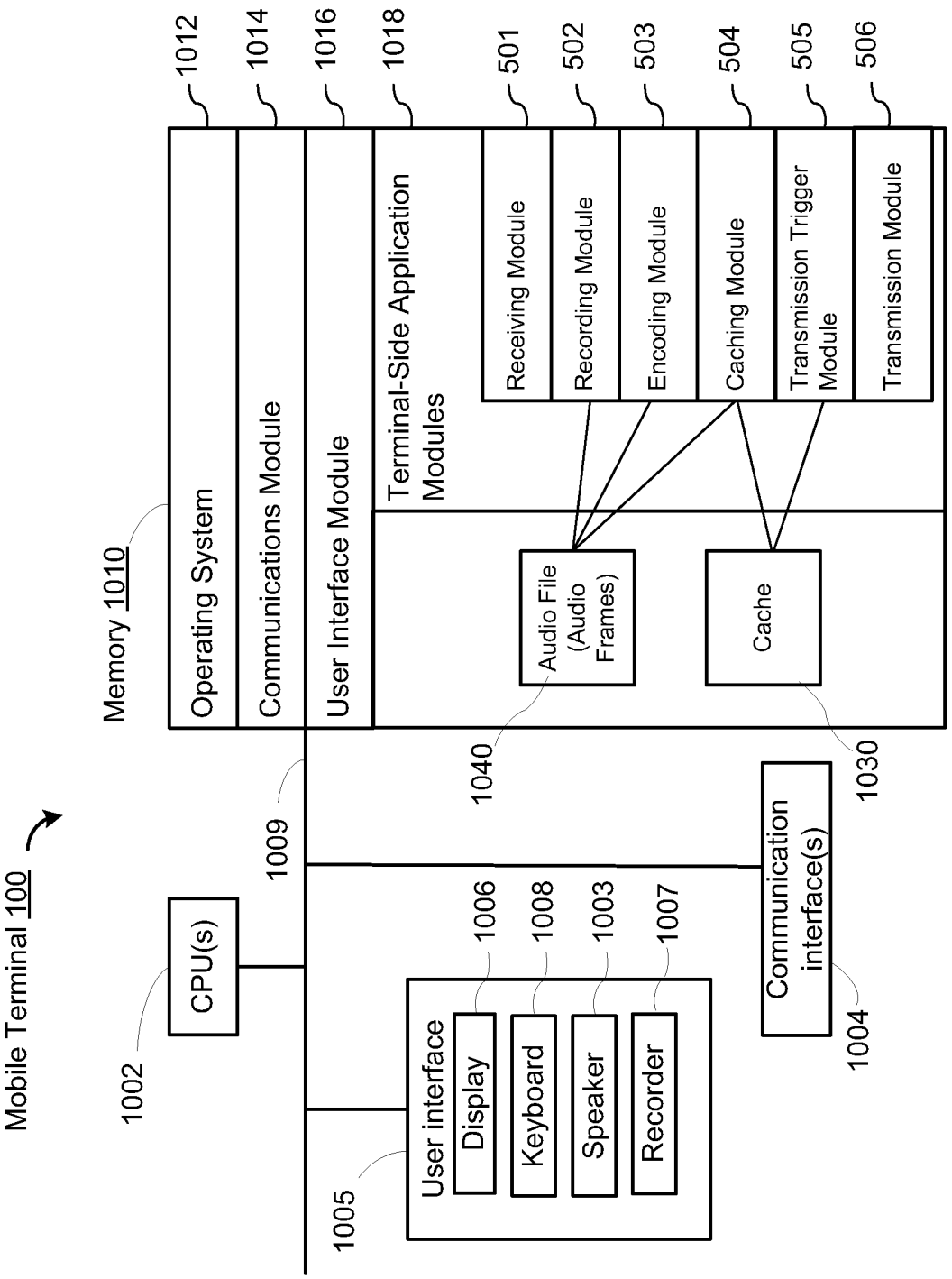

Mobile Terminal 100
1002
CPU(s)
1009
1005
User interface
Display
Keyboard
Speaker
Recorder
1006
1008
1003
1007
Communication interface(s)
1004
Memory 1010
Operating System
Communications Module
User Interface Module
Terminal-Side Application Modules
1012
1014
1016
1018
Receiving Module
Recording Module
Encoding Module
Caching Module
Transmission Trigger Module
Transmission Module
501
502
503
504
505
506
Audio File (Audio Frames)
Cache
1040
1030

Fig. 10

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2013/085913

A. CLASSIFICATION OF SUBJECT MATTER

H04M 3/56 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: H04M; H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT,CNKI, CNABS, GOOGLE, CNTXT, DWPI, WPI, EPODOC: audio, signal, transmit, transfer, record, frame, encode, processing time, threshold, terminal, server, AMR, audio codec, first, second

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2003063574 A1 (NOKIA CORPORATION) 03 April 2003(03.04.2003) description, paragraphs[0003]-[0016], [0060]-[0063]	1-3,6-11,14-19,22-24
A	CN 101588284 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.) 25 November 2009(25.11.2009) the whole document	1-24
A	CN 101394585 A (ZTE CORPORATION) 25 March 2009(25.03.2009) the whole document	1-24

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 January 2014 (10.01.2014)	**27 Feb. 2014 (27.02.2014)**
Name and mailing address of the ISA/CN The State Intellectual Property Office, the P.R.China 6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China 100088 Facsimile No. 86-10-62019451	Authorized officer HU, Yajuan Telephone No. (86-10)82245497

Form PCT/ISA /210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/CN2013/085913

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
US 2003063574 A1	03.04.2003	FI20011917 A	29.03.2003
		DE60209637 D1	04.05.2006
		CN1411256 A	16.04.2003
		EP1298906 A1	02.04.2003
		DE60209637 T2	21.12.2006
		AT320137 T	15.03.2006
CN 101588284 A	25.11.2009	None	
CN 101394585 A	25.03.2009	None	