

(51) International Patent Classification:
H04M 9/08 (2006.01)(21) International Application Number:
PCT/IB2014/003164(22) International Filing Date:
29 December 2014 (29.12.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
201410005371.9 6 January 2014 (06.01.2014) CN(71) Applicant: **ALCATEL LUCENT** [FR/FR]; 148/152 route
de la Reine, F-92100 Boulogne-Billancourt (FR).(72) Inventors: **LI, Zhouzhou**; 388# Ningqiao Road, Pudong
Jinqiao, Shanghai 201206 (CN). **CAI, Yigang**; 23643
Royal Worlington Dr, Naperville, Illinois 60564 (US).(74) Agent: **SCIAUX, Edmond**; Alcatel-lucent International,
148/152 route de la Reine, F-92100 Boulogne-Billancourt
(FR).(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, 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.(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, ST, 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))
- before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments (Rule 48.2(h))

(54) Title: METHOD AND DEVICE FOR PACKET ACOUSTIC ECHO CANCELLATION

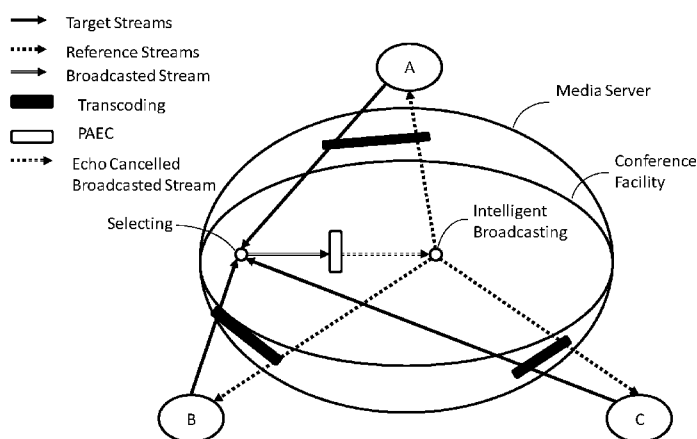


Fig. 7

(57) **Abstract:** An object of the invention is providing a method and device for packet acoustic echo cancellation. The echo canceling device obtains one or more source voice packet streams in a multi-party call, selects packets from the one or more source voice packet streams so as to obtain a broadcasted stream corresponding to the multi-party call, cancels echo in a target packet stream corresponding to the broadcasted stream based on a reference packet stream corresponding to the broadcasted stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream, and finally broadcasts the echo-cancelled packet stream to participants in the multi-party call. Compared with the prior art, the present invention realizes packet acoustic echo cancellation for a multi-party call, reduces the buffer consumption, channel allocation and corresponding overheads of maintenance and signaling for packet acoustic echo cancellation, enhances the performance of packet acoustic echo cancellation, and makes the performance of packet acoustic echo cancellation independent of the number of participants.



METHOD AND DEVICE FOR PACKET ACOUSTIC ECHO CANCELLATION

FIELD OF THE INVENTION

[0001] The present invention relates to the field of communication, and in particular to the technology of packet acoustic echo cancellation.

BACKGROUND OF THE INVENTATION

[0002] In packet-based telephony networks, media servers provide a variety of basic and enhanced services, such as conferencing, audio and video interactive voice response (IVR), transcoding, audio and video announcements, and other advanced speech services. Conference facility of the media server allows a number of participants to join a conference. Figure 1 shows a simplified N-party (N=3) audio conference mode. The media server providing conferencing service is supporting a 3-party audio conference. A, B, C are the 3 participants. Transcoding may be applied to each conference participant by the media server to guarantee the packet voice from each participant will be transcoded to a unique codec type, so that later the “mixing” function of the conference facility can work correctly. The “mixing” function is used to form unique broadcasted stream from all the target streams (incoming streams, from media server perspective). The broadcasted stream is then fed into “broadcasting” function to form all the reference streams (outgoing streams, from media server perspective).

[0003] The acoustic echo in telephony networks is due to badly designed handsets or other hands free equipment where the sound from the speakers is fed to the microphone of the receiver (and then back to voice sender). It greatly impairs the voice communication quality, especially in packet-based audio conferences.

[0004] In general, for an audio conference (including the audio session of a video conference), there are 3 types of acoustic echoes.

[0005] 1. Local echo. As shown in Figure 2, A's packet voice being fed back to A, this is called “local” echo.

[0006] 2. Far echo. As shown in Figure 3, A's voice played back by B's speaker to B's microphone if B's device does not have echo cancellation capability, then sent back to A, this is called “far” echo.

[0007] 3. Mate echo. As shown in Figure 4, A's voice sent to C and B, and played back by B's speaker to B's microphone if B's device does not have echo cancellation capability, then A's voice (i.e. the echo in B) sent to C again, thus C receives twice of A's voice, this is called “mate” echo.

[0008] All these types of acoustic echoes will impair the audio conference quality. Applying acoustic echo cancellation (AEC) against them is necessary to ensure the audio conference quality, thus satisfy customers.

[0009] In the prior art, there are 2 methods to cancel acoustic echo in a packet-based audio conference.

[0010] One method relies on AECs allocated at all party devices to guarantee no echo will be sent to the conference facility then broadcasted to other parties. Figure 5 depicts this method. Target stream (outgoing stream, from participant perspective) will be compared to history reference stream (incoming stream, from participant perspective) to detect acoustic echo, then target packets containing acoustic echo will be cancelled/suppressed, thus no acoustic echo will be sent to conference facility then broadcasted.

[0011] Its flaw is obvious – Firstly since this method can only guarantee that no echo is sent from the participant to the conference facilities, but cannot guarantee the echo cancellation from the broadcasting side to the participants, so it can not eliminate the local echo. Secondly, it requires all party devices equipped with AECs. If there is one party device isn't equipped with an AEC, the other parties may still receive far echo and mate echo generated from that party. For example, B's device is not equipped with AEC, it causes far echo to A and mate echo to C. AEC allocated at A party can only cancel the echo in the target stream from A, which can't cancel the echo in the reference stream to A (far echo). AEC allocated at C party can only cancel the echo in the target stream from C, which can't cancel the echo in the reference stream to C (mate echo). Even "Bidirectional PAEC" allocated at C's device can't cancel the mate echo to C because "Bidirectional PAEC" relies on 2 direction streams while both the original voice from A to C and the mate echo from B to C are presented in the same reference flow to C.

[0012] The second method relies on PAEC channels allocated on the conference facility, as shown in Figure 6, each target stream (incoming stream, from server perspective) will be compared to its history reference stream (outgoing stream, from server perspective) to detect acoustic echo. Target packets containing acoustic echo will be cancelled/suppressed, thus no acoustic echo will be sent to mixing function then broadcasting function. The broadcasting function needs a little bit intelligence to prevent backward broadcasting, with which voice packets from one party won't be sent back to that party, thus the local echo is avoided.

[0013] Its shortage is obvious – It requires allocating PAEC channels for all the participants. With more participants joining the conference, more resources on the conference facility need to be allocated and occupied for PAEC.

SUMMARY OF THE INVENTION

[0014] An object of the invention is providing a method and device for packet acoustic echo cancellation.

[0015] According to one aspect of the invention, a method for packet acoustic echo cancellation is provided, wherein the method comprises the following steps:

[0016] a. obtaining one or more source voice packet streams in a multi-party call, wherein acoustic echoes in the one or more source voice packet streams are to be cancelled and each source voice packet stream is from one participant in the multi-party call;

[0017] b. selecting packets from the one or more source voice packet streams so as

to obtain a broadcasted stream corresponding to the multi-party call;

[0018] c. cancelling echo in a target packet stream corresponding to the broadcasted stream based on a reference packet stream corresponding to the broadcasted stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream;

[0019] d. broadcasting the echo-cancelled packet stream to participants in the multi-party call.

[0020] According to another aspect of the invention, an echo cancelling device for packet acoustic echo cancellation is further provided, wherein the device comprises:

[0021] obtaining apparatus for obtaining one or more source voice packet streams in a multi-party call, wherein acoustic echoes in the one or more source voice packet streams are to be cancelled and each source voice packet stream is from one participant in the multi-party call;

[0022] selecting apparatus for selecting packets from the one or more source voice packet streams so as to obtain a broadcasted stream corresponding to the multi-party call;

[0023] cancelling apparatus for cancelling echo in a target packet stream corresponding to the broadcasted stream based on a reference packet stream corresponding to the broadcasted stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream;

[0024] broadcasting apparatus for broadcasting the echo-cancelled packet stream to participants in the multi-party call.

[0025] Compared with the prior art, the present invention realizes packet acoustic echo cancellation for a multi-party call, by obtaining one or more source voice packet streams in a multi-party call, selecting packets from the one or more source voice packet streams so as to obtain a broadcasted stream corresponding to the multi-party call, cancelling echo in a target packet stream corresponding to the broadcasted stream based on a reference packet stream corresponding to the broadcasted stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream, and finally broadcasting the echo-cancelled packet stream to participants in the multi-party call. The present invention reduces the buffer consumption, channel allocation and corresponding overheads of maintenance and signaling for packet acoustic echo cancellation, enhances the performance of packet acoustic echo cancellation, and makes the performance of packet acoustic echo cancellation independent of the number of participants.

[0026] Furthermore, the present invention could determine parameter information of the storage space based on parameters of the multi-party call, in conjunction with packet voice lengths of the target packet stream and the reference packet stream, wherein the storage space is used for storing the target packet stream and the reference packet stream. Thereby, it avoids traversing throughout the whole storage space by utilizing the parameter information, and enhances the echo detection efficiency.

[0027] Furthermore, the present invention could cancel echo in the target packet stream based on the reference packet stream corresponding to the broadcasted stream, in conjunction with source address information corresponding to each packet in a

corresponding target packet stream of the broadcasted stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream. Thereby, it detects the state of the participant corresponding to the source address information based on the source address information, and if the terminal is provided with a packet acoustic echo cancellation device, a comparison computation for new packets from the participant may be saved, which enhances the performance of packet acoustic echo cancellation.

[0028] Furthermore, the present invention could determine one or more target participants in the multi-party call based on source address information corresponding to each packet in the target packet stream, wherein the target participant is corresponding to the echo-cancelled packet stream, and broadcast the echo-cancelled packet stream to the one or more target participants. Thereby, it prevents backward broadcasting by determining the target participants before broadcasting, which may avoid local echo.

BRIEF DESCRIPTION OF THE DRAWINGS

[0029] Other features, purposes and advantages of the invention will become more explicit by means of reading the detailed statement of the non-restrictive embodiments made with reference to the accompanying drawings.

[0030] Fig. 1 shows a structure diagram of a simplified N-party (where $N=3$) audio conference model according to one aspect of the present invention;

[0031] Fig. 2 shows a diagram of forming a local echo according to one aspect of the present invention;

[0032] Fig. 3 shows a diagram of forming a far echo according to one aspect of the present invention;

[0033] Fig. 4 shows a diagram of forming a mate echo according to one aspect of the present invention;

[0034] Fig. 5 shows a diagram of a terminal-based multi-party packet acoustic echo cancellation according to the prior art;

[0035] Fig. 6 shows a diagram of a server-based multi-party packet acoustic echo cancellation according to the prior art;

[0036] Fig. 7 shows a diagram of a server-based multi-party packet acoustic echo cancellation independent of the number of conference participants according to one aspect of the present invention;

[0037] Fig. 8 shows a schematic diagram of an echo cancelling device for packet acoustic echo cancellation according to one aspect of the present invention;

[0038] Fig. 9 shows a schematic diagram of an echo cancelling device for packet acoustic echo cancellation according to a preferred embodiment of the present invention;

[0039] Fig. 10 shows a flow diagram of a method for packet acoustic echo cancellation according to another aspect of the present invention;

[0040] Fig. 11 shows a flow diagram of a method for packet acoustic echo cancellation according to a preferred embodiment of the present invention;

[0041] Fig. 12 shows a diagram of comparing packets for performing packet acoustic echo cancellation according to one aspect of the present invention.

[0042] The same or similar reference signs in the drawings represent the same or similar component parts.

DETAILED DESCRIPTION OF THE INVENTION

[0043] Below, details of the invention will be further provided in combination with the accompanying drawings.

[0044] Fig. 8 shows a schematic diagram of an echo cancelling device for packet acoustic echo cancellation according to one aspect of the present invention; wherein the echo cancelling device comprises: a obtaining apparatus 1, a selecting apparatus 2, a cancelling apparatus 3, and a broadcasting apparatus 4. Specifically, the obtaining apparatus 1 obtains one or more source voice packet streams in a multi-party call, wherein acoustic echoes in the one or more source voice packet streams are to be cancelled and each source voice packet stream is from one participant in the multi-party call; the selecting apparatus 2 selects packets from the one or more source voice packet streams so as to obtain a broadcasted stream corresponding to the multi-party call; the cancelling apparatus 3 cancels echo in a target packet stream corresponding to the broadcasted stream based on a reference packet stream corresponding to the broadcasted stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream; the broadcasting apparatus 4 broadcasts the echo-cancelled packet stream to participants in the multi-party call.

[0045] Here, the echo cancelling device includes, but not limited to, an electronic hardware device or software device that automatically perform numerical value computation and information processing according to a pre-set or pre-stored instruction, wherein the hardware device includes, but not limited to, a microprocessor, an application-specific integrated circuit (ASIC), a field programmable gate array (FPGA), a digital processor (DSP), an embedded device, etc. Those skilled in the art should understand that other echo cancelling devices, if applicable to the present invention, should also be included within the protection scope of the present invention and are incorporated here by reference.

[0046] The echo cancelling device may be applied in any multi-party conference.

[0047] The above apparatuses work continuously therebetween. Here, those skilled in the art should understand that “continuously” means each of the above apparatuses obtains one or more source voice packet streams in a multi-party call, obtains a broadcasted stream, obtains the echo-cancelled packet stream, broadcasts the echo-cancelled packet stream, and the like in real time or according to a preset or real-time adjusted working pattern requirements, until the echo cancelling device stops obtaining one or more source voice packet streams in a multi-party call.

[0048] The obtaining apparatus 1 obtains one or more source voice packet streams in a multi-party call, wherein acoustic echoes in the one or more source voice packet streams are to be cancelled and each source voice packet stream is from one participant in the multi-party call.

[0049] Specifically, the obtaining apparatus 1 interacts with one or more participants in a multi-party call (e.g., in one conference) to obtain source voice packet streams to be cancelled packet acoustic echo from the participant to the broadcast end, wherein the source voice packet stream comprises one or more packets, and the packet of the source voice packet stream may contain an echo packet. Since each voice packet stream is from one participant in the multi-party call, the obtaining apparatus 1 may obtain one or more source voice packet streams in the multi-party call; for example, if the multi-party call includes participant A, participant B, and participant C, then the obtaining apparatus 1 may obtain a source voice packet stream from the participant A to the broadcast end, a source voice packet stream from the participant B to the broadcast end, and a source voice packet stream from the participant C to the broadcast end.

[0050] The selecting apparatus 2 selects packets from the one or more source voice packet streams so as to obtain a broadcasted stream corresponding to the multi-party call.

[0051] Specifically, the selecting apparatus 2 selects packets from the one or more source voice packet streams, based on voice intensity information of the packets in the source voice packet streams or continuity information (e.g., the person talking previously has a higher priority to be selected continuously) of the packets in the source voice packet streams, and generates a broadcasted stream corresponding to the multi-party call based on the information of the selected packets, so as to guarantee that no mixing will occur.

[0052] This is because in a packet-based audio conference, when more than one participant are talking at the same time, mixing their voices to broadcast won't help improving the voice quality and mixing is expensive operation, which includes decoding, synthesizing, encoding packets. Instead, selecting voice from one of the talking participants for some period is a simple and efficient method.

[0053] The cancelling apparatus 3 cancels echo in a target packet stream corresponding to the broadcasted stream based on a reference packet stream corresponding to the broadcasted stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream.

[0054] Specifically, the cancelling apparatus 3 regards, based on the broadcasted stream, one or more new packets in the broadcasted stream as the target packet stream corresponding to the broadcasted stream, wherein the new packet might contain an echo, and regards one or more historical packets in the same broadcasted stream as the reference packet stream corresponding to the broadcasted stream, wherein the historical packets are earlier than the packets of the new packets in time sequence, and in special cases, the historical packets may contain echo.

[0055] The cancelling apparatus 3 compares, in the buffer or other means that may perform high-speed read/write, one or more packets in the reference packet stream with one or more packets in the target packet stream, for example, determines whether they are similar based on a packet acoustic echo cancellation algorithm (PAEC algorithm), thereby detecting whether the target packet stream includes echo packets; if one or more similar packets exist, then the target packet stream includes echo

packet(s); wherein the means that may perform high-speed read/write may maintain a longer period of historical voice.

[0056] If it includes an echo packet, the cancelling apparatus 3 cancels echo in the target packet stream through deleting the echo packet or by substituting the detected echo packet with a replacement packet, or through other manners. Specifically, substitute the detected echo packet with a replacement packet to obtain an echo-cancelled packet stream corresponding to the target packet stream, wherein the replacement packet includes, but not limited to, a noise packet (for example, a packet including a certain type of noise, for example white noise, comfortable noise, etc.), a silent packet (for example, an empty packet), a 1/8 rate packet as finally buffered in the target packet stream, and the like or a combination thereof.

[0057] Here, the cancelling apparatus 3 may also regard the echo-cancelled packet stream as its corresponding reference packet stream of the same broadcasted stream. Here, the reference packet stream does not include any echo.

[0058] Here, since the far echo and the mate echo are similar packets occurring after the original packets, the above method may identify whether the new packet is similar to the historical packet, thereby the far echo and the mate echo may be cancelled.

[0059] Preferably, the cancelling apparatus 3 may cancel echo in the target packet stream based on the reference packet stream corresponding to the broadcasted stream, in conjunction with source address information corresponding to each packet in a corresponding target packet stream of the broadcasted stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream.

[0060] Specifically, the cancelling apparatus 3 regards, based on the broadcasted stream, one or more new packets in the broadcasted stream as the target packet stream corresponding to the broadcasted stream, wherein the new packet might contain an echo, and regards one or more historical packets in the same broadcasted stream as the reference packet stream corresponding to the broadcasted stream, wherein the historical packets are earlier than the packets of the new packets in time sequence, and in special cases, the historical packets may contain echo.

[0061] Then, the cancelling apparatus 3 determines, based on the source address information in header information of each packet in the target packet stream, the participant corresponding to the packet. The cancelling apparatus 3 interacts with the participant or with other third-party device to determine whether the participant is provided with a packet acoustic echo cancellation (PAEC) apparatus. If the participant is provided with the PAEC apparatus, it indicates that the echo in the packet from the participant has been cancelled before the packet is sent to the echo cancelling device, such that the cancelling apparatus 3 will not perform echo detection to the packets corresponding to the participant, thereby saving the comparison computation for new packets from the participant, reducing the computation complexity, and enhancing the performance of packet acoustic echo cancellation. Here, those skilled in the art should understand that the cancelling apparatus 3 may not determine the source address information of the packet in the reference packet stream, because the far echo or mate echo comes from other communication terminal, not from the communication terminal corresponding to the original packet.

[0062] Otherwise, if the participant corresponding to the packet in the target packet stream is not provided with a packet acoustic echo cancellation (PAEC) apparatus, then the cancelling apparatus 3 compares the packets in the target packet stream with the ones in the corresponding reference packet stream to detect whether the packets in the target packet stream contain any echo packet; if it does, then the cancelling apparatus 3 cancels echo in the target packet stream by deleting the echo packet or by substituting the detected echo packet using a substitution packet.

[0063] Here, those skilled in the art should understand that the operation of the packet acoustic echo cancellation in the cancelling apparatus 3 is identical or similar to the one in the cancelling apparatus 3 in Fig. 8, which will not be detailed here but is incorporated here by reference.

[0064] Preferably, the echo cancelling device further comprises a parameter determining apparatus(not shown), wherein, the parameter determining apparatus determines parameter information of the storage space for storing the target packet stream and the reference packet stream based on parameters of the multi-party call, in conjunction with packet voice lengths of the target packet stream and the reference packet stream.

[0065] Specifically, the storage space includes, but not limited to, a buffer or other means that may perform high-speed read/ write, and the parameter information of the storage space may be determined based on information such as the call parameter of the multi-party call (e.g., echo delay information of each participant of the conference) and voice packet lengths (e.g., each packet duration is 20 ms) of the target packet stream and the reference packet stream. Here, the parameter information of the storage space includes, but not limited to, size of the storage space (e.g., buffer size) and/or window size of the reference packet stream.

[0066] Here, the design for the storage space of the broadcasted stream (for the buffer size of the reference packet stream as an example) mainly comprises three parameters:

[0067] The first is delay_{th} in millisecond, i.e., the maximum echo delay the conference facility can cancel, this is usually an engineering decision, but it should be greater than most echo delays. Then the buffer size of the reference packet can be decided based on the delay_{th} , in conjunction with packet voice lengths of the target packet stream and the reference packet stream, given each packet presents 20ms voice: the buffer size = $\text{delay}_{th} / 20 + M + 1$; wherein, M denotes the number of target packets or reference packets in the target packet sets or reference packet sets for comparison.

[0068] The second is delay_{min} in millisecond, i.e., the minimum echo delay of all the participants of one audio conference.

[0069] The third is delay_{max} in millisecond, i.e., the maximum echo delay of all

the participants of one audio conference.

[0070] $\text{delay}_{\min}$ and $\text{delay}_{\max}$ are dynamical per conference call parameters, which can be derived from statistics data. $\text{delay}_{\min}$ can be used to estimate the interval between the target packet window and the reference packet window, i.e. offset of the reference packet window. With both $\text{delay}_{\min}$ and $\text{delay}_{\max}$, the reference packet window size can be estimated, i.e.: the reference packet window size = $\text{delay}_{\max} / 20 + M + 1 - \text{delay}_{\min} / 20$.

[0071] Here, the $\text{delay}_{\max}$ and $\text{delay}_{\min}$, i.e., the maximum delay and the minimum delay of the multi-party call as computed by the parameter determining apparatus based on the echo delay information of each participant of the conference, may be used to determine the size of the reference packet window, so as to adjust the size of the reference stream to avoid unnecessary comparison and enhance the echo detection efficiency.

[0072] The broadcasting apparatus 4 broadcasts the echo-cancelled packet stream to participants in the multi-party call.

[0073] Specifically, the broadcasting apparatus 4 sends the echo-cancelled packet stream as obtained by the cancelling apparatus 3 to one or more participants in the multi-party call; wherein the echo-cancelled packet stream may be broadcasted to all participants in the multi-party call, or the echo-cancelled packet stream may be broadcasted to a part of participants in the multi-party call.

[0074] Preferably, the broadcasting apparatus 4 determines one or more target participants in the multi-party call corresponding to the echo-cancelled packet stream based on source address information corresponding to each packet in the target packet stream; and broadcasts the echo-cancelled packet stream to the one or more target participants.

[0075] Specifically, the broadcasting apparatus 4 determines, based on source address information in the header information of each packet in the echo-cancelled packet stream, one or more participants except the participant corresponding to the source address information in the multi-party call as the target participant; finally, the broadcasting apparatus 4 broadcasts the echo-cancelled packet stream to the one or more target participants.

[0076] Here, the broadcasting apparatus 4 determines the participant(s) corresponding to the packet based on the source address information of the packet so as to avoid sending the packet to its corresponding participant, i.e., if the source address information of the packet corresponds to participant A, the packet will not be sent to A anymore when broadcasting, thereby realizing intelligent broadcast and preventing backward broadcasting to avoid local echo.

[0077] For example, Fig. 7 shows a diagram of a server-based multi-party packet

acoustic echo cancellation independent of the number of conference participants according to one aspect of the present invention.

[0078] Specifically, the target stream sent from each participant forms a single broadcasted stream by selection; the single broadcasted stream simultaneously includes target packets (selected new packets from each participant) and reference packets (according to Fig. 7, it is historical packets without any echo packet, because the echo packet has been cancelled during the previous AEC process). The target packets and the reference packets are compared according to the PAEC (Packet Acoustic Echo Cancellation) algorithm so as to remove the echo packet; the echo-cancelled packet is then broadcasted to the conference participants. Therefore, it transforms the universal method for detecting an echo in the packet field into a method of comparing new packets and historical packets in the broadcasted stream. It effectively cancels the packet echo in a packet-based audio conference independent of the number of conference participants. For a conference with N participants, only one PAEC resource (e.g., channel) is needed, which saves N-1 channels.

[0079] Fig. 9 shows a schematic diagram of an echo cancelling device for packet acoustic echo cancellation according to a preferred embodiment of the present invention; wherein the echo cancelling device comprises: a obtaining apparatus 1', a selecting apparatus 2', a cancelling apparatus 3', and a broadcasting apparatus 4', the cancelling apparatus 3' comprises an updating unit 31' and a cancelling unit 32'. Specifically, the obtaining apparatus 1' obtains one or more source voice packet streams in a multi-party call, wherein acoustic echoes in the one or more source voice packet streams are to be cancelled and each source voice packet stream is from one participant in the multi-party call; the selecting apparatus 2' selects packets from the one or more source voice packet streams so as to obtain a broadcasted stream corresponding to the multi-party call; the updating unit 31' updates the target packet stream and the reference packet stream corresponding to the multi-party call based on the broadcasted stream; the cancelling unit 32' cancels echo in the target packet stream based on the reference packet stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream; the broadcasting apparatus 4' broadcasts the echo-cancelled packet stream to participants in the multi-party call.

[0080] Herein, the obtaining apparatus 1', the selecting apparatus 2' and the broadcasting apparatus 4' as comprised in the echo cancelling device are identical or substantially identical to corresponding apparatus shown in Fig. 8, which are thus not detailed here, but incorporated here by reference.

[0081] The above apparatuses work continuously therebetween. Here, those skilled in the art should understand that "continuously" means each of the above apparatuses obtains one or more source voice packet streams in a multi-party call, obtains a broadcasted stream, updates the target packet stream and the reference packet stream, obtains the echo-cancelled packet stream, broadcasts the echo-cancelled packet stream, and the like in real time or according to a preset or real-time adjusted working pattern requirements, until the echo cancelling device stops obtaining one or more source voice packet streams in a multi-party call.

[0082] The updating unit 31' updates the target packet stream and the reference

packet stream corresponding to the multi-party call based on the broadcasted stream.

[0083] Specifically, the updating unit 31' regards, based on the broadcasted stream, one or more new packets in the broadcasted stream as the target packet stream corresponding to the broadcasted stream, wherein the new packet might contain an echo, and regards one or more historical packets in the same broadcasted stream as the reference packet stream corresponding to the broadcasted stream, wherein the historical packets are earlier than the packets of the new packets in time sequence, and the historical packets may contain echo.

[0084] Or, the updating unit 31' may also regard, based on the broadcasted stream, the same broadcasted stream which having been echo cancelled (i.e., the echo-cancelled packet stream earlier than the current new packets in time sequence) as the reference packet stream corresponding to the target packet stream. Here, the reference packet stream does not include any echo.

[0085] Preferably, the updating unit 31' could update the target packet stream corresponding to the multi-party call based on the broadcasted stream; and determine the reference packet stream corresponding to the target packet stream based on the broadcasted stream and the target packet stream.

[0086] Specifically, the updating unit 31' regards, based on the broadcasted stream, one or more new packets in the broadcasted stream as the target packet stream corresponding to the broadcasted stream, wherein the new packet might contain an echo, and regards one or more historical packets in the same broadcasted stream corresponding to the target packet stream as the reference packet stream corresponding to the broadcasted stream, wherein the historical packets are earlier than the packets of the new packets in time sequence, and the historical packets may contain echo.

[0087] Preferably, the echo cancelling device further comprises a reference updating apparatus (not shown), wherein, the reference updating apparatus updates the reference packet stream based on the echo-cancelled packet stream.

[0088] Specifically, the reference updating apparatus obtains the echo-cancelled packet stream by interacting with the cancelling apparatus 3', and then updates the reference packet stream based on the echo-cancelled packet stream. In this way, by using the echo-cancelled packet stream as a reference packet stream for comparing with the target packet stream, it has a better reference effect, thereby further enhancing the PAEC accuracy.

[0089] The cancelling unit 32' cancels echo in the target packet stream based on the reference packet stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream.

[0090] Specifically, the cancelling unit 32', based on the reference packet stream, cancels echo in the target packet stream by comparing the reference packet stream with the target packet stream, such as computing according to the PAEC algorithm, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream.

[0091] Here, those skilled in the art should understand that the operation of the packet acoustic echo cancellation in the cancelling unit 32' is identical or similar to

the one in the cancelling apparatus 3 in Fig. 8, which will not be detailed here but is incorporated here by reference.

[0092] Fig. 12 shows a diagram of comparing packets for performing packet acoustic echo cancellation according to one aspect of the present invention.

[0093] Specifically, when a new packet (which may contain an acoustic echo) in the broadcasted stream is added to the broadcasted stream, the new packet may act as a target packet for comparing with a historical packet which acts as the reference packet stream, so as to detect the acoustic echo. As shown in Fig. 12, a packet set of the target packet stream (packet j to packet j+M, i.e., new packets) are compared with set 1 (packet 1 to packet q+M, i.e., historical packets), set 2 (packet q+1 to packet q+M+1, i.e., historical packets), ..., set Q (packet q+Q-1 to packet q+Q+M-1, i.e., historical packets), respectively, to determine whether the target packet stream include any echo packet.

[0094] Equations 1 to 3 show the PAEC algorithm for cancelling packet acoustic echo.

$$[0095] \quad \mathbf{E}_j^B = [\mathbf{e}_{j,q}^B, \mathbf{e}_{j,q+1}^B, \dots, \mathbf{e}_{j,q+Q-1}^B] \quad (\text{equation 1})$$

$$[0096] \quad \mathbf{e}_{j,q}^B = \sqrt{\sum_{m=0}^M \sum_{\partial=1}^P (\ell_{j+m,\partial}^B - \ell_{q+m,\partial}^B)^2} \quad (\text{equation 2})$$

$$[0097] \quad \min \mathbf{e}_{j,q}^B \leq \mathbf{e}_{th} \quad (\text{equation 3})$$

[0098] Specifically, equation 1 is an equation for computing a distance vector, while equation 2 is an equation for computing a distance value between each new packet and each historical packet. “M+1” is the new packet window size. “M+Q” is the corresponding history packet window size. $\mathbf{e}_{j,q}^B$ represents the comparison result of M+1 packets from new window (j, j+1, ..., j+M) to history window (q, q+1, ..., q+M).

The minimum value of $\mathbf{e}_{j,q}^B$ (q=q, q+1, q+Q-1) will be compared to a similarity

threshold $\mathbf{e}_{TH}$, to decide if echo exists. If the result of the equation 2 satisfies equation 3, it shows that similarities exist in the target packet stream and the reference packet stream, and the target packet stream has echo.

[0099] Here, P denotes the value of LSP (Line Spectral Pair).

[00100] In this way, the present invention realizes a method for packet acoustic echo cancellation, which has the following advantages over the prior art:

[00101] - The prior art needs buffering reference streams for each participant, and N-party audio conference will need N sets of reference buffers, such that when a new packet (which might contain an echo) occurs in the target packet stream, it can be compared to the buffered reference packets to detect acoustic echo; while by using the present invention, only one (broadcasted) stream is buffered for PAEC comparison, no matter how big N is. The present invention can greatly reduce the buffer consumption

as well as the corresponding maintenance cost for PAEC.

[00102] - The prior art needs allocating PAEC channels for each participant. N-Party audio conference will need N PAEC channels. By using the present invention, only one PAEC channel is needed per conference, no matter how big N is. The present invention can greatly reduce the channel allocation as well as the corresponding signaling overhead for PAEC.

[00103] - The prior art usually ignores some call participants of the conference (by some algorithm) when their quantity exceeds the threshold. However this is a type of service degradation. By using the present invention, the PAEC capacity of the conference facility can be greatly enhanced, and its performance is participant quantity independent.

[00104] Fig. 10 shows a flow diagram of a method for packet acoustic echo cancellation according to another aspect of the present invention. Specifically, in the step s1, the echo cancelling device obtains one or more source voice packet streams in a multi-party call, wherein acoustic echoes in the one or more source voice packet streams are to be cancelled and each source voice packet stream is from one participant in the multi-party call; in the step s2, the echo cancelling device selects packets from the one or more source voice packet streams so as to obtain a broadcasted stream corresponding to the multi-party call; in the step s3, the echo cancelling device cancels echo in a target packet stream corresponding to the broadcasted stream based on a reference packet stream corresponding to the broadcasted stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream; in the step s4, the echo cancelling device broadcasts the echo-cancelled packet stream to participants in the multi-party call.

[00105] The above steps work continuously therebetween. Here, those skilled in the art should understand that “continuously” means each of the above steps obtains one or more source voice packet streams in a multi-party call, obtains a broadcasted stream, obtains the echo-cancelled packet stream, broadcasts the echo-cancelled packet stream, and the like in real time or according to a preset or real-time adjusted working pattern requirements, until the echo cancelling device stops obtaining one or more source voice packet streams in a multi-party call.

[00106] In the step s1, the echo cancelling device obtains one or more source voice packet streams in a multi-party call, wherein acoustic echoes in the one or more source voice packet streams are to be cancelled and each source voice packet stream is from one participant in the multi-party call.

[00107] Specifically, in the step s1, the echo cancelling device interacts with one or more participants in a multi-party call (e.g., in one conference) to obtain source voice packet streams to be cancelled packet acoustic echo from the participant to the broadcast end, wherein the source voice packet stream comprises one or more packets, and the packet of the source voice packet stream may contain an echo packet. Since each voice packet stream is from one participant in the multi-party call, in the step s1, the echo cancelling device may obtain one or more source voice packet streams in the multi-party call; for example, if the multi-party call includes participant A, participant B, and participant C, then in the step s1, the echo cancelling device may obtain a

source voice packet stream from the participant A to the broadcast end, a source voice packet stream from the participant B to the broadcast end, and a source voice packet stream from the participant C to the broadcast end.

[00108] In the step s2, the echo cancelling device selects packets from the one or more source voice packet streams so as to obtain a broadcasted stream corresponding to the multi-party call.

[00109] Specifically, in the step s2, the echo cancelling device selects packets from the one or more source voice packet streams, based on voice intensity information of the packets in the source voice packet streams or continuity information (e.g., the person talking previously has a higher priority to be selected continuously) of the packets in the source voice packet streams, and generates a broadcasted stream corresponding to the multi-party call based on the information of the selected packets, so as to guarantee that no mixing will occur.

[00110] This is because in a packet-based audio conference, when more than one participant are talking at the same time, mixing their voices to broadcast won't help improving the voice quality and mixing is expensive operation, which includes decoding, synthesizing, encoding packets. Instead, selecting voice from one of the talking participants for some period is a simple and efficient method.

[00111] In the step s3, the echo cancelling device cancels echo in a target packet stream corresponding to the broadcasted stream based on a reference packet stream corresponding to the broadcasted stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream.

[00112] Specifically, in the step s3, the echo cancelling device regards, based on the broadcasted stream, one or more new packets in the broadcasted stream as the target packet stream corresponding to the broadcasted stream, wherein the new packet might contain an echo, and regards one or more historical packets in the same broadcasted stream as the reference packet stream corresponding to the broadcasted stream, wherein the historical packets are earlier than the packets of the new packets in time sequence, and in special cases, the historical packets may contain echo.

[00113] In the step s3, the echo cancelling device compares, in the buffer or other means that may perform high-speed read/write, one or more packets in the reference packet stream with one or more packets in the target packet stream, for example, determines whether they are similar based on a packet acoustic echo cancellation algorithm (PAEC algorithm), thereby detecting whether the target packet stream includes echo packets; if one or more similar packets exist, then the target packet stream includes echo packet(s); wherein the means that may perform high-speed read/write may maintain a longer period of historical voice.

[00114] If it includes an echo packet, in the step s3, the echo cancelling device cancels echo in the target packet stream through deleting the echo packet or by substituting the detected echo packet with a replacement packet, or through other manners. Specifically, substitute the detected echo packet with a replacement packet to obtain an echo-cancelled packet stream corresponding to the target packet stream, wherein the replacement packet includes, but not limited to, a noise packet (for example, a packet including a certain type of noise, for example white noise,

comfortable noise, etc.), a silent packet (for example, an empty packet), a 1/8 rate packet as finally buffered in the target packet stream, and the like or a combination thereof.

[00115] Here, in the step s3, the echo cancelling device may also regard the echo-cancelled packet stream as its corresponding reference packet stream of the same broadcasted stream. Here, the reference packet stream does not include any echo.

[00116] Here, since the far echo and the mate echo are similar packets occurring after the original packets, the above method may identify whether the new packet is similar to the historical packet, thereby the far echo and the mate echo may be cancelled.

[00117] Preferably, in the step s3, the echo cancelling device may cancel echo in the target packet stream based on the reference packet stream corresponding to the broadcasted stream, in conjunction with source address information corresponding to each packet in a corresponding target packet stream of the broadcasted stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream.

[00118] Specifically, in the step s3, the echo cancelling device regards, based on the broadcasted stream, one or more new packets in the broadcasted stream as the target packet stream corresponding to the broadcasted stream, wherein the new packet might contain an echo, and regards one or more historical packets in the same broadcasted stream as the reference packet stream corresponding to the broadcasted stream, wherein the historical packets are earlier than the packets of the new packets in time sequence, and in special cases, the historical packets may contain echo.

[00119] Then, in the step s3, the echo cancelling device determines, based on the source address information in header information of each packet in the target packet stream, the participant corresponding to the packet. In the step s3, the echo cancelling device interacts with the participant or with other third-party device to determine whether the participant is provided with a packet acoustic echo cancellation (PAEC) apparatus. If the participant is provided with the PAEC apparatus, it indicates that the echo in the packet from the participant has been cancelled before the packet is sent to the echo cancelling device, such that in the step s3, the echo cancelling device will not perform echo detection to the packets corresponding to the participant, thereby saving the comparison computation for new packets from the participant, reducing the computation complexity, and enhancing the performance of packet acoustic echo cancellation. Here, those skilled in the art should understand that in the step s3, the echo cancelling device may not determine the source address information of the packet in the reference packet stream, because the far echo or mate echo comes from other communication terminal, not from the communication terminal corresponding to the original packet.

[00120] Otherwise, if the participant corresponding to the packet in the target packet stream is not provided with a packet acoustic echo cancellation (PAEC) apparatus, then in the step s3, the echo cancelling device compares the packets in the target packet stream with the ones in the corresponding reference packet stream to detect whether the packets in the target packet stream contain any echo packet; if it does, then in the step s3, the echo cancelling device cancels echo in the target packet stream by deleting the echo packet or by substituting the detected echo packet using a

substitution packet.

[00121] Here, those skilled in the art should understand that the operation of the packet acoustic echo cancellation in the step s3 is identical or similar to the one in the step s3 in Fig. 10, which will not be detailed here but is incorporated here by reference.

[00122] Preferably, the method further comprises a step s5(not shown), wherein, in the step s5, the echo cancelling device determines parameter information of the storage space for storing the target packet stream and the reference packet stream based on parameters of the multi-party call, in conjunction with packet voice lengths of the target packet stream and the reference packet stream.

[00123] Specifically, the storage space includes, but not limited to, a buffer or other means that may perform high-speed read/ write, and the parameter information of the storage space may be determined based on information such as the call parameter of the multi-party call (e.g., echo delay information of each participant of the conference) and voice packet lengths (e.g., each packet duration is 20 ms) of the target packet stream and the reference packet stream. Here, the parameter information of the storage space includes, but not limited to, size of the storage space (e.g., buffer size) and/or window size of the reference packet stream.

[00124] Here, the design for the storage space of the broadcasted stream (for the buffer size of the reference packet stream as an example) mainly comprises three parameters:

[00125] The first is delay_{th} in millisecond, i.e., the maximum echo delay the conference facility can cancel, this is usually an engineering decision, but it should be greater than most echo delays. Then the buffer size of the reference packet can be decided based on the delay_{th} , in conjunction with packet voice lengths of the target packet stream and the reference packet stream, given each packet presents 20ms voice: the buffer size = $\text{delay}_{th} / 20 + M + 1$; wherein, M denotes the number of target packets or reference packets in the target packet sets or reference packet sets for comparison.

[00126] The second is delay_{min} in millisecond, i.e., the minimum echo delay of all the participants of one audio conference.

[00127] The third is delay_{max} in millisecond, i.e., the maximum echo delay of all the participants of one audio conference.

[00128] delay_{min} and delay_{max} are dynamical per conference call parameters, which can be derived from statistics data. delay_{min} can be used to estimate the interval between the target packet window and the reference packet window, i.e. offset of the reference packet window. With both delay_{min} and delay_{max} , the reference packet

window size can be estimated, i.e.: the reference packet window size = $\text{delay}_{\max} / 20 + M + 1 - \text{delay}_{\min} / 20$.

[00129] Here, the $\text{delay}_{\max}$ and $\text{delay}_{\min}$, i.e., the maximum delay and the minimum delay of the multi-party call as computed by the echo cancelling device in the step s5 based on the echo delay information of each participant of the conference, may be used to determine the size of the reference packet window, so as to adjust the size of the reference stream to avoid unnecessary comparison and enhance the echo detection efficiency.

[00130] In the step s4, the echo cancelling device broadcasts the echo-cancelled packet stream to participants in the multi-party call.

[00131] Specifically, in the step s4, the echo cancelling device sends the echo-cancelled packet stream as obtained by the step s3 to one or more participants in the multi-party call; wherein the echo-cancelled packet stream may be broadcasted to all participants in the multi-party call, or the echo-cancelled packet stream may be broadcasted to a part of participants in the multi-party call.

[00132] Preferably, in the step s4, the echo cancelling device determines one or more target participants in the multi-party call corresponding to the echo-cancelled packet stream based on source address information corresponding to each packet in the target packet stream; and broadcasts the echo-cancelled packet stream to the one or more target participants.

[00133] Specifically, in the step s4, the echo cancelling device determines, based on source address information in the header information of each packet in the echo-cancelled packet stream, one or more participants except the participant corresponding to the source address information in the multi-party call as the target participant; finally, in the step s4, the echo cancelling device broadcasts the echo-cancelled packet stream to the one or more target participants.

[00134] Here, in the step s4, the echo cancelling device determines the participant(s) corresponding to the packet based on the source address information of the packet so as to avoid sending the packet to its corresponding participant, i.e., if the source address information of the packet corresponds to participant A, the packet will not be sent to A anymore when broadcasting, thereby realizing intelligent broadcast and preventing backward broadcasting to avoid local echo.

[00135] Fig. 11 shows a flow diagram of a method for packet acoustic echo cancellation according to a preferred embodiment of the present invention. Specifically, in the step s1', the echo cancelling device obtains one or more source voice packet streams in a multi-party call, wherein acoustic echoes in the one or more source voice packet streams are to be cancelled and each source voice packet stream is from one participant in the multi-party call; in the step s2', the echo cancelling device selects packets from the one or more source voice packet streams so as to obtain a broadcasted stream corresponding to the multi-party call; in the step s31', the echo cancelling device updates the target packet stream and the reference packet

stream corresponding to the multi-party call based on the broadcasted stream; in the step s32', the echo cancelling device cancels echo in the target packet stream based on the reference packet stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream; in the step s4', the echo cancelling device broadcasts the echo-cancelled packet stream to participants in the multi-party call.

[00136] Herein, the step s1', the step s2' and the step s4' of the method are identical or substantially identical to corresponding steps shown in Fig. 10, which are thus not detailed here, but incorporated here by reference.

[00137] The above steps work continuously therebetween. Here, those skilled in the art should understand that "continuously" means each of the above steps obtains one or more source voice packet streams in a multi-party call, obtains a broadcasted stream, updates the target packet stream and the reference packet stream, obtains the echo-cancelled packet stream, broadcasts the echo-cancelled packet stream, and the like in real time or according to a preset or real-time adjusted working pattern requirements, until the echo cancelling device stops obtaining one or more source voice packet streams in a multi-party call.

[00138] In the step s31', the echo cancelling device updates the target packet stream and the reference packet stream corresponding to the multi-party call based on the broadcasted stream.

[00139] Specifically, in the step s31', the echo cancelling device regards, based on the broadcasted stream, one or more new packets in the broadcasted stream as the target packet stream corresponding to the broadcasted stream, wherein the new packet might contain an echo, and regards one or more historical packets in the same broadcasted stream as the reference packet stream corresponding to the broadcasted stream, wherein the historical packets are earlier than the packets of the new packets in time sequence, and the historical packets may contain echo.

[00140] Or, in the step s31', the echo cancelling device may also regard, based on the broadcasted stream, the same broadcasted stream which having been echo cancelled (i.e., the echo-cancelled packet stream earlier than the current new packets in time sequence) as the reference packet stream corresponding to the target packet stream. Here, the reference packet stream does not include any echo.

[00141] Preferably, in the step s31', the echo cancelling device could update the target packet stream corresponding to the multi-party call based on the broadcasted stream; and determine the reference packet stream corresponding to the target packet stream based on the broadcasted stream and the target packet stream.

[00142] Specifically, in the step s31', the echo cancelling device regards, based on the broadcasted stream, one or more new packets in the broadcasted stream as the target packet stream corresponding to the broadcasted stream, wherein the new packet might contain an echo, and regards one or more historical packets in the same broadcasted stream corresponding to the target packet stream as the reference packet stream corresponding to the broadcasted stream, wherein the historical packets are earlier than the packets of the new packets in time sequence, and the historical packets may contain echo.

[00143] Preferably, the method further comprises a step s6' (not shown), wherein, in

the step s6', the echo cancelling device updates the reference packet stream based on the echo-cancelled packet stream.

[00144] Specifically, in the step s6', the echo cancelling device obtains the echo-cancelled packet stream by interacting with the step s3', and then updates the reference packet stream based on the echo-cancelled packet stream. In this way, by using the echo-cancelled packet stream as a reference packet stream for comparing with the target packet stream, it has a better reference effect, thereby further enhancing the PAEC accuracy.

[00145] In the step s32', the echo cancelling device cancels echo in the target packet stream based on the reference packet stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream.

[00146] Specifically, in the step s32', the echo cancelling device, based on the reference packet stream, cancels echo in the target packet stream by comparing the reference packet stream with the target packet stream, such as computing according to the PAEC algorithm, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream.

[00147] Here, those skilled in the art should understand that the operation of the packet acoustic echo cancellation in the step s32' is identical or similar to the one in the step s3 in Fig. 10, which will not be detailed here but is incorporated here by reference.

[00148] To those skilled in the art, apparently the present invention is not limited to the details of the aforementioned exemplary embodiments; moreover, under the premise of not deviating from the spirit or fundamental characteristics of the invention, this invention can be accomplished in other specific forms. Therefore, the embodiments should be considered exemplary and non-restrictive no matter from which point, the scope of the invention is defined by the appended claims instead of the above description, and aims at covering the meanings of the equivalent components falling into the claims and all changes within the scope in this invention. Any reference sign in the claims shall not be deemed as limiting the concerned claims. Besides, apparently the word "comprise/include" does not exclude other components or steps, singular numbers does not exclude complex numbers, the plurality of components or means mentioned in device claims may also be accomplished by one component or means through software or hardware, the wording like first and second are only used to represent names rather than any specific order.

We claim:

1. A method for packet acoustic echo cancellation, wherein the method comprises the following steps:

- a. obtaining one or more source voice packet streams in a multi-party call, wherein acoustic echoes in the one or more source voice packet streams are to be cancelled and each source voice packet stream is from one participant in the multi-party call;
- b. selecting packets from the one or more source voice packet streams so as to obtain a broadcasted stream corresponding to the multi-party call;
- c. cancelling echo in a target packet stream corresponding to the broadcasted stream based on a reference packet stream corresponding to the broadcasted stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream;
- d. broadcasting the echo-cancelled packet stream to participants in the multi-party call.

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

- determining parameter information of the storage space for storing the target packet stream and the reference packet stream based on parameters of the multi-party call, in conjunction with packet voice lengths of the target packet stream and the reference packet stream.

3. The method according to claim 1 or 2, wherein the step c comprises:

- c1. updating the target packet stream and the reference packet stream corresponding to the multi-party call based on the broadcasted stream;
- cancelling echo in the target packet stream based on the reference packet stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream.

4. The method according to claim 3, wherein the step c1 comprises:

- updating the target packet stream corresponding to the multi-party call based on the broadcasted stream;
- determining the reference packet stream corresponding to the target packet stream based on the broadcasted stream and the target packet stream.

5. The method according to any one of claims 1 to 4, wherein the method further comprises:

- updating the reference packet stream based on the echo-cancelled packet stream.

6. The method according to any one of claims 1 to 5, wherein the step c comprises:

- cancelling echo in the target packet stream based on the reference packet stream corresponding to the broadcasted stream, in conjunction with source address information corresponding to each packet in a corresponding target packet stream of the broadcasted stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream.

7. The method according to any one of claims 1 to 6, wherein the step d comprises:

- determining one or more target participants in the multi-party call corresponding to the echo-cancelled packet stream based on source address information corresponding to each packet in the target packet stream;
- broadcasting the echo-cancelled packet stream to the one or more target participants.

8. An echo cancelling device for packet acoustic echo cancellation, wherein the device comprises:

obtaining apparatus for obtaining one or more source voice packet streams in a multi-party call, wherein acoustic echoes in the one or more source voice packet streams are to be cancelled and each source voice packet stream is from one participant in the multi-party call;

selecting apparatus for selecting packets from the one or more source voice packet streams so as to obtain a broadcasted stream corresponding to the multi-party call;

cancelling apparatus for cancelling echo in a target packet stream corresponding to the broadcasted stream based on a reference packet stream corresponding to the broadcasted stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream;

broadcasting apparatus for broadcasting the echo-cancelled packet stream to participants in the multi-party call.

9. The echo cancelling device according to claim 8, wherein the device further comprises:

parameter determining apparatus for determining parameter information of the storage space for storing the target packet stream and the reference packet stream based on parameters of the multi-party call, in conjunction with packet voice lengths of the target packet stream and the reference packet stream.

10. The echo cancelling device according to claim 8 or 9, wherein the cancelling apparatus comprises:

updating unit for updating the target packet stream and the reference packet stream corresponding to the multi-party call based on the broadcasted stream;

cancelling unit for cancelling echo in the target packet stream based on the reference packet stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream.

11. The echo cancelling device according to claim 10, wherein the updating unit is configured to:

- update the target packet stream corresponding to the multi-party call based on the broadcasted stream;
- determine the reference packet stream corresponding to the target packet stream based on the broadcasted stream and the target packet stream.

12. The echo cancelling device according to any one of claims 8 to 11, wherein the device further comprises:

reference updating apparatus for updating the reference packet stream based on the echo-cancelled packet stream.

13. The echo cancelling device according to any one of claims 8 to 12, wherein the cancelling apparatus is configured to:

- cancel echo in the target packet stream based on the reference packet stream corresponding to the broadcasted stream, in conjunction with source address information corresponding to each packet in a corresponding target packet stream of the broadcasted stream, so as to obtain the echo-cancelled packet stream corresponding to the target packet stream.

14. The echo cancelling device according to any one of claims 8 to 13, wherein the broadcasting apparatus is configured to:

- determine one or more target participants in the multi-party call corresponding to the echo-cancelled packet stream based on source address information corresponding to each packet in the target packet stream;
- broadcast the echo-cancelled packet stream to the one or more target participants.

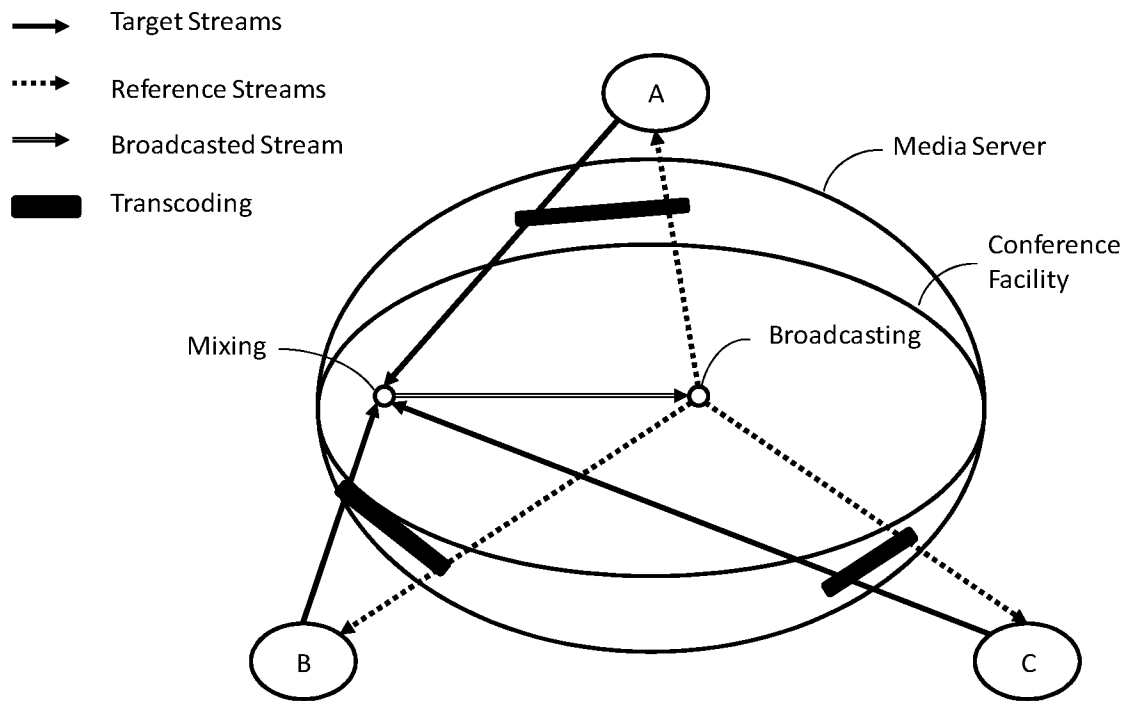


Fig. 1

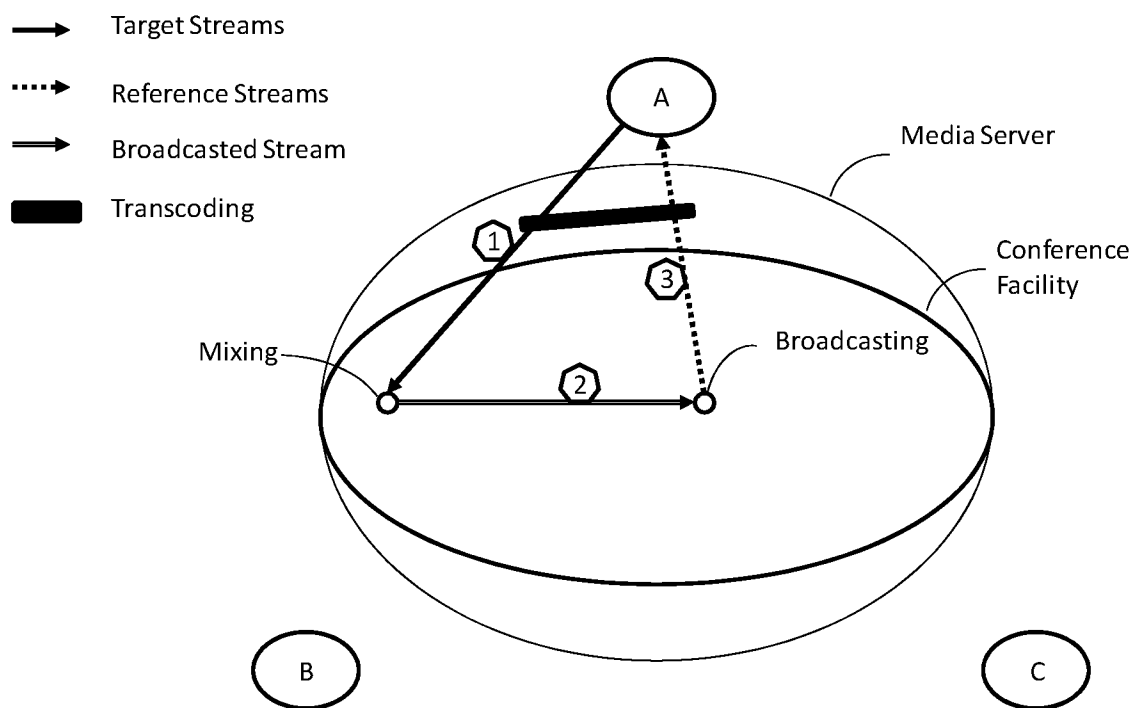


Fig. 2

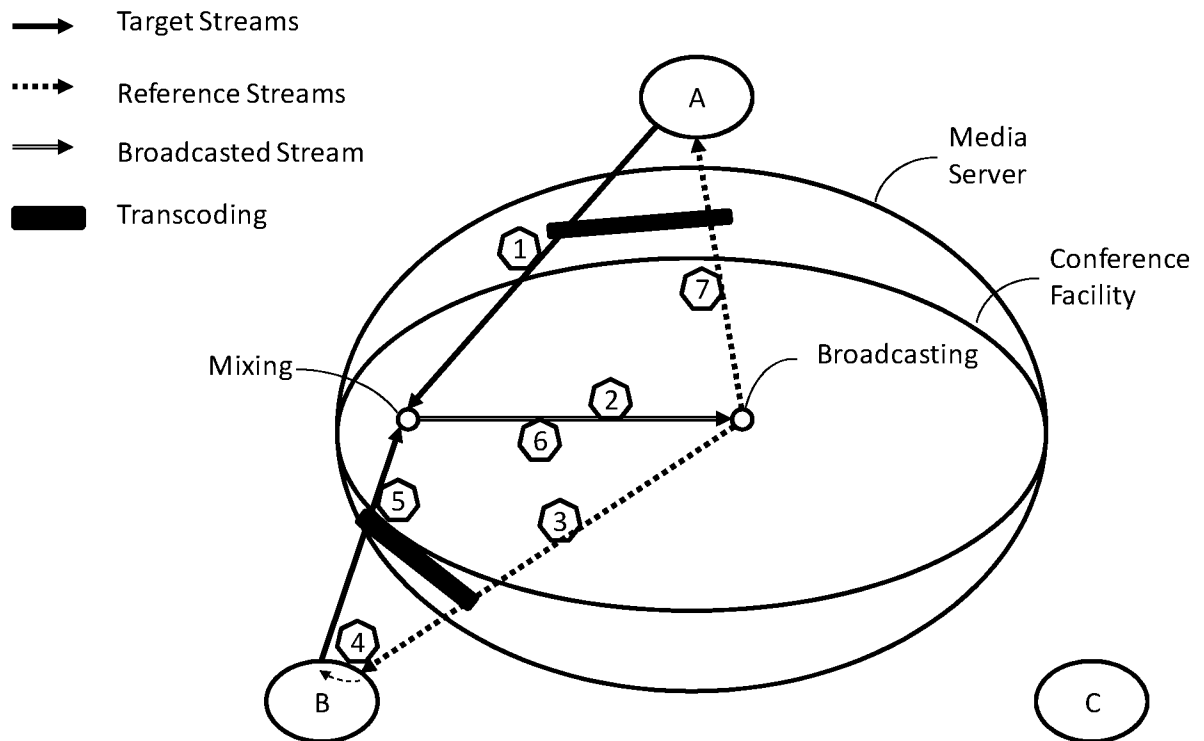


Fig. 3

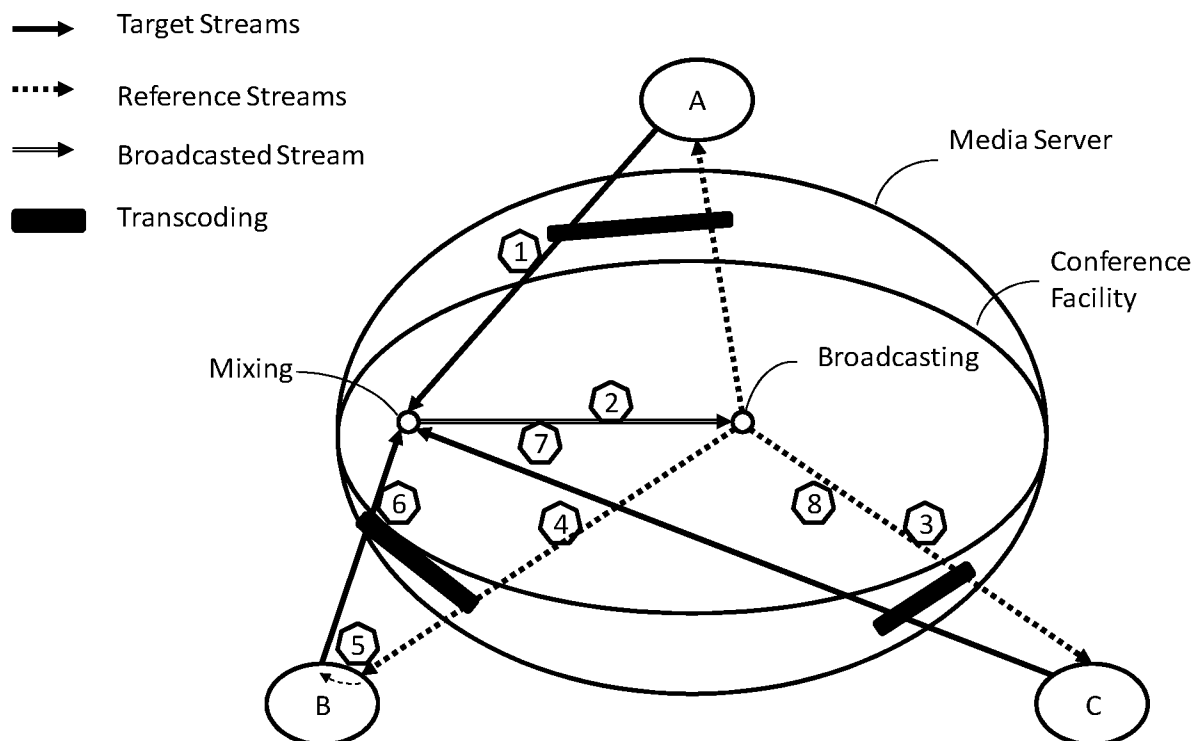


Fig. 4

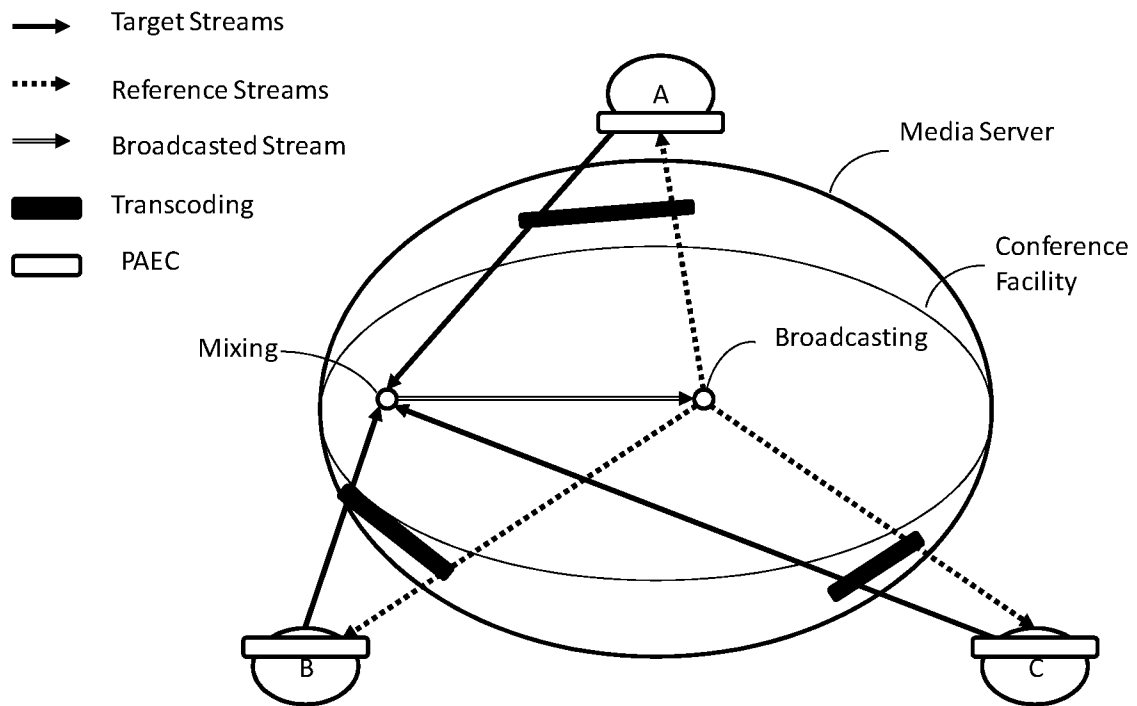


Fig. 5

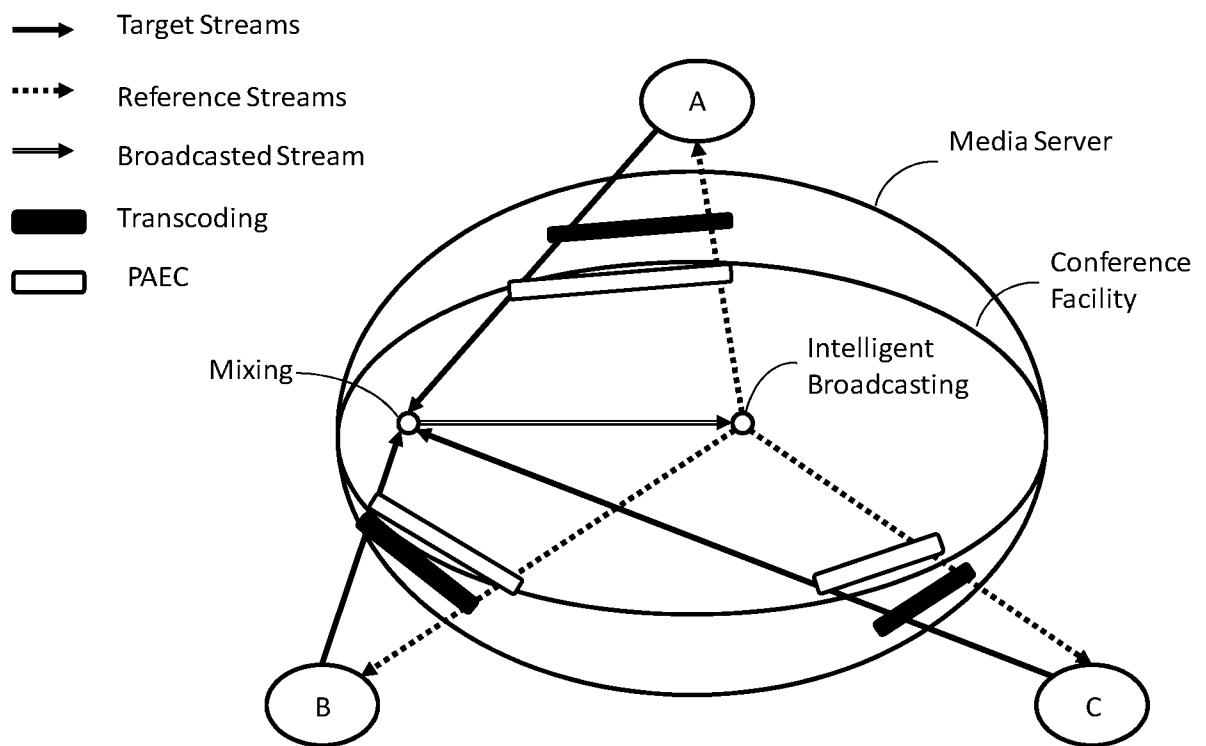


Fig. 6

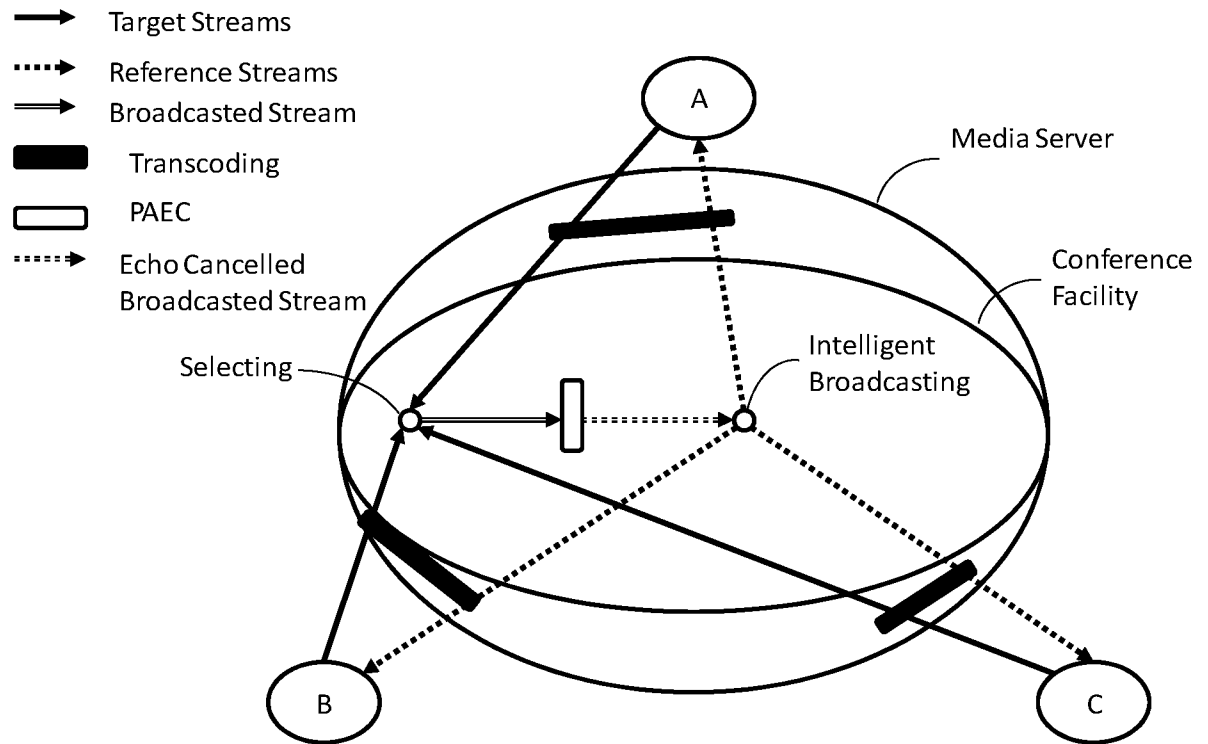


Fig.7

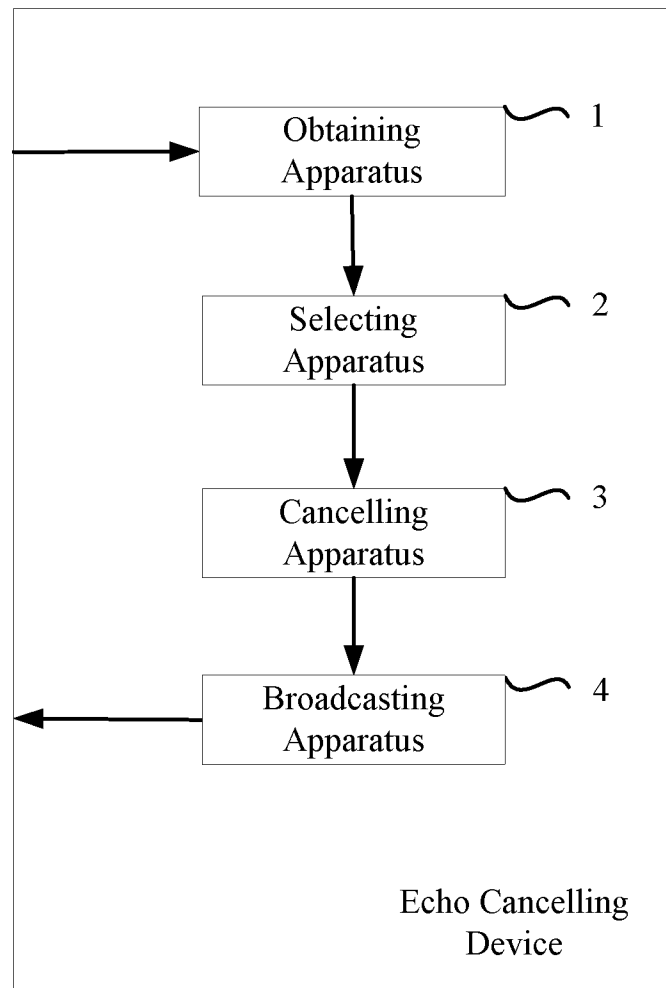


Fig. 8

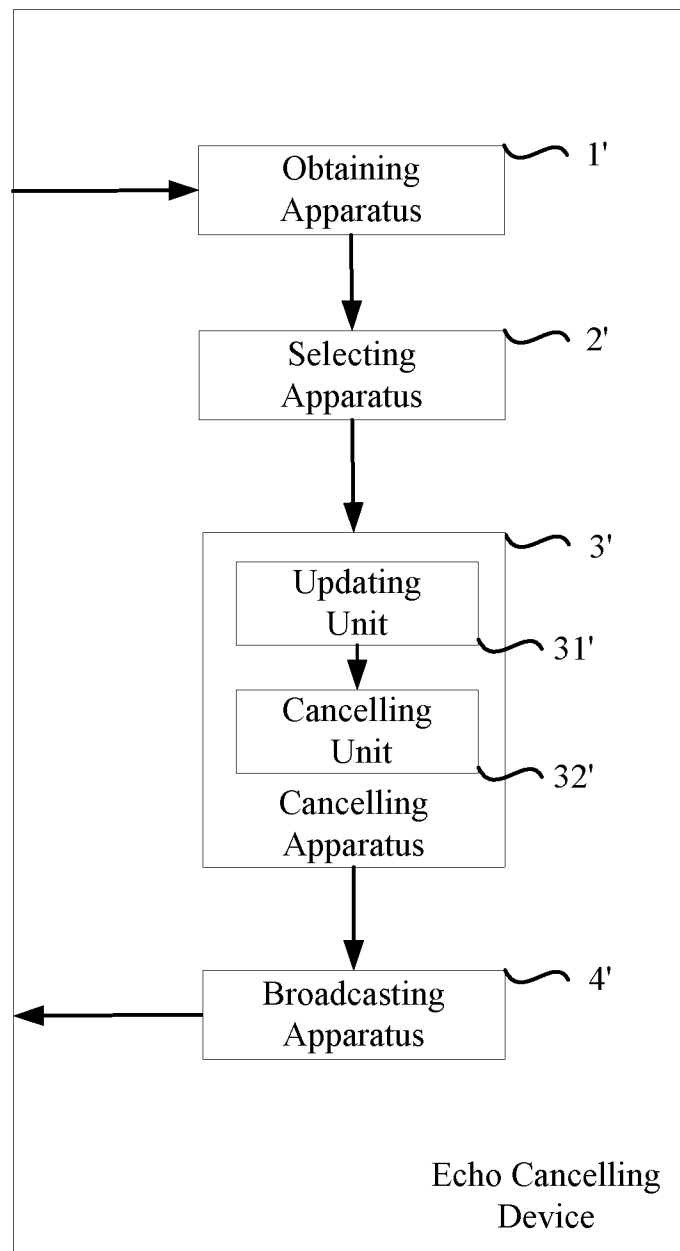


Fig. 9

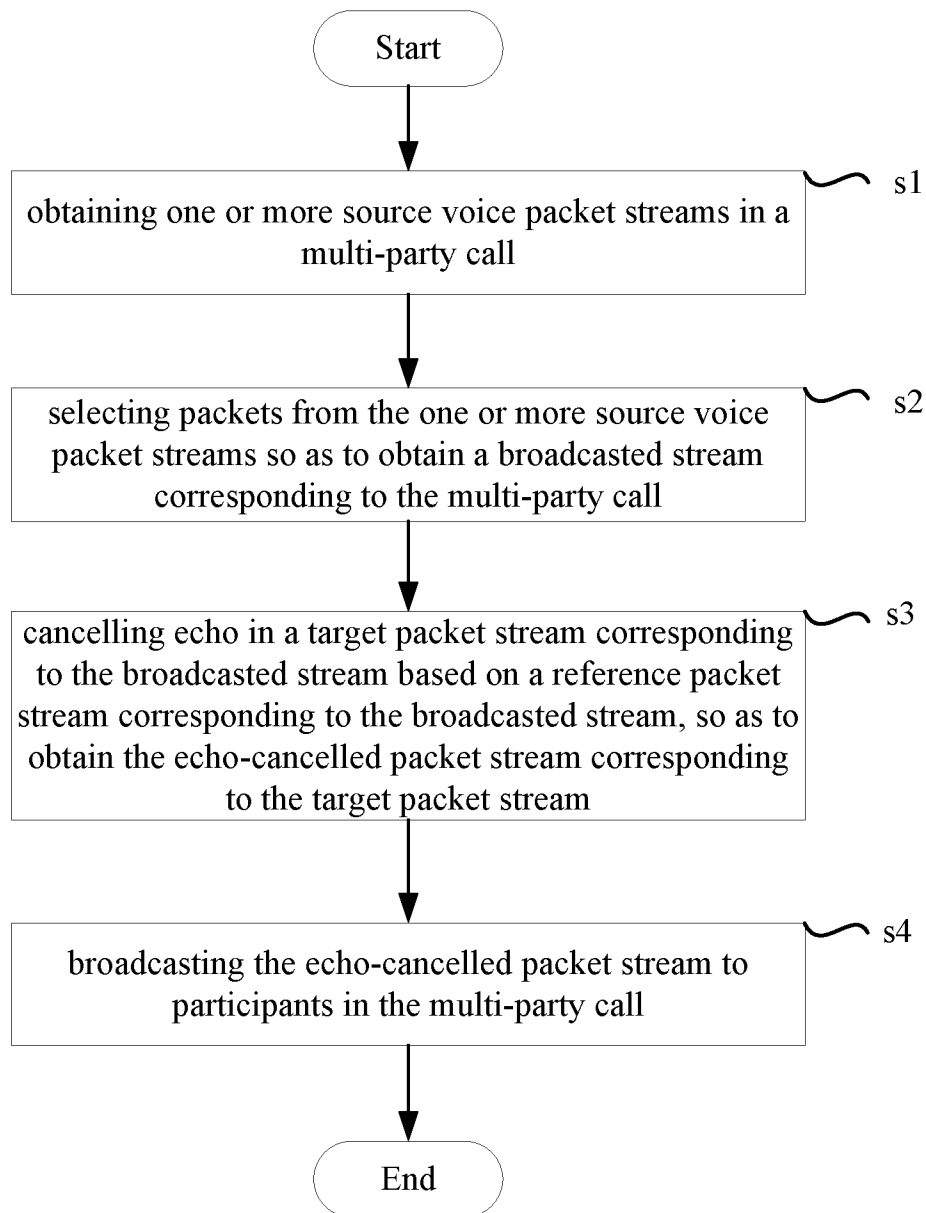


Fig. 10

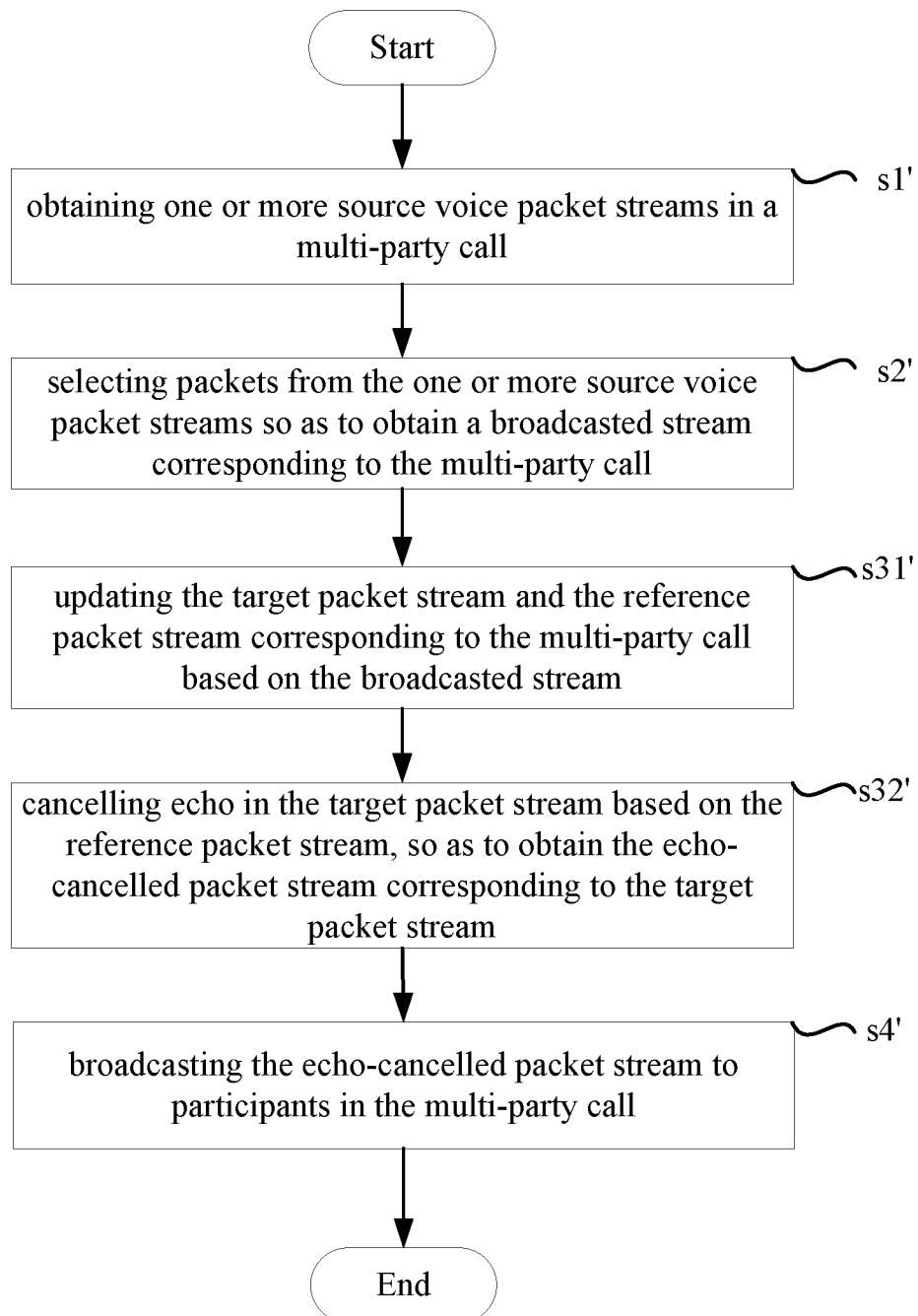


Fig. 11

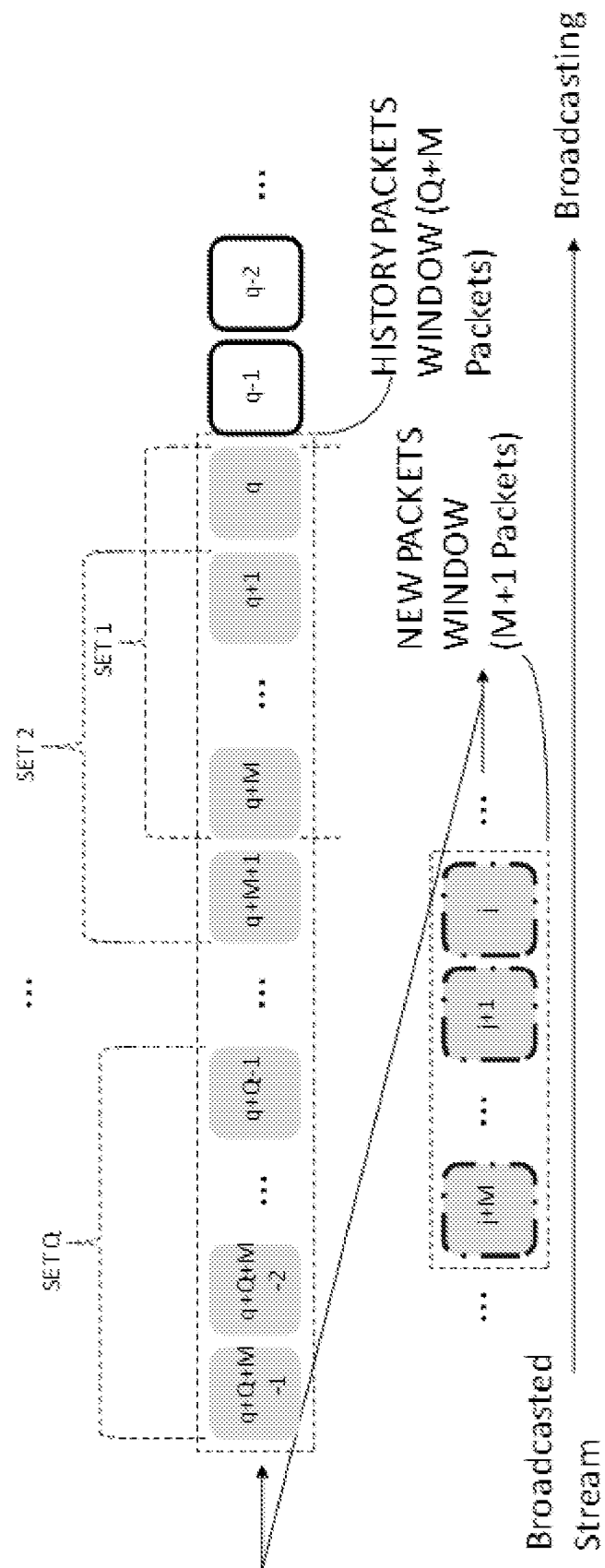


Fig.12

INTERNATIONAL SEARCH REPORT

International application No
PCT/IB2014/003164

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04M9/08
ADD.

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)
H04M

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)

EP0-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2011/235499 A1 (RAJASEKAR BADRI N [US] ET AL) 29 September 2011 (2011-09-29)	1,8
Y	paragraphs [0013], [0015] - [0016], [0024], [0028] - [0038], [0070] - [0074]; figure 3	2-7,9-14
Y	----- US 2009/168673 A1 (KALAMPOUKAS LAMPROS [US] ET AL) 2 July 2009 (2009-07-02)	2-7,9-14
A	paragraphs [0018], [0030], [0031], [0042] - [0046]	1,8
A	----- US 2013/155924 A1 (SUKKAR RAFID A [US]) 20 June 2013 (2013-06-20) paragraphs [0034] - [0050]	1-14
	----- -/--	



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

1 July 2015

Date of mailing of the international search report

08/07/2015

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Sorrentino, Andrea

INTERNATIONAL SEARCH REPORT

International application No
PCT/IB2014/003164

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 6 320 958 B1 (SEKINE AKIHIRO [JP]) 20 November 2001 (2001-11-20) column 1, line 17 - line 20 column 3, line 26 - line 30 column 4, line 41 - line 65 -----	1-14

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/IB2014/003164

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2011235499 A1	29-09-2011	NONE	
US 2009168673 A1	02-07-2009	CN 101933306 A	29-12-2010
		EP 2245826 A1	03-11-2010
		JP 4922455 B2	25-04-2012
		JP 2011515881 A	19-05-2011
		KR 20100096218 A	01-09-2010
		KR 20120102820 A	18-09-2012
		US 2009168673 A1	02-07-2009
		WO 2009088431 A1	16-07-2009
US 2013155924 A1	20-06-2013	NONE	
US 6320958 B1	20-11-2001	JP 2907279 B2	21-06-1999
		JP H10164239 A	19-06-1998
		US 6320958 B1	20-11-2001