

(51) International Patent Classification:  
*H04M 9/08* (2006.01)(21) International Application Number:  
PCT/US2013/077353(22) International Filing Date:  
21 December 2013 (21.12.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
1223246.8 21 December 2012 (21.12.2012) GB  
13/889,229 7 May 2013 (07.05.2013) US(71) Applicant: MICROSOFT CORPORATION [US/US];  
One Microsoft Way, Redmond, WA 98052-6399 (US).(72) Inventor: SORENSEN, Karsten, Vandborg; c/o Mi-  
crosoft Corporation, LCA - International Patents, One Mi-  
crosoft Way, Redmond, WA 98052-6399 (US).(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.(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).

## Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))

## Published:

- with international search report (Art. 21(3))

(54) Title: ECHO SUPPRESSION

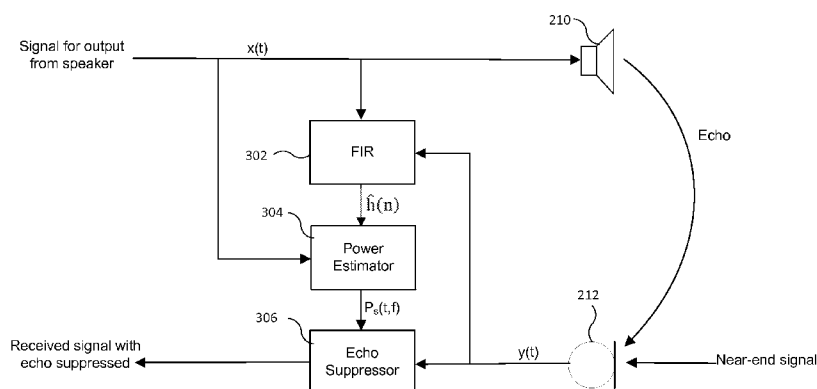


FIG. 3

(57) **Abstract:** Method, user device and computer program product for suppressing echo. An audio signal is output from a speaker. A microphone receives an audio signal, wherein the received audio signal includes an echo resulting from the outputted audio signal. A Finite Impulse Response filter estimate  $\hat{h}(n)$  ( $h$  with  $\wedge$  symbol above  $h$ ) is dynamically adapted in the time domain based on the outputted audio signal and the received audio signal to model an echo path  $h(n)$  related to the echo in the received audio signal. The filter estimate  $\hat{h}(n)$  ( $h$  with  $\wedge$  symbol above  $h$ ) is used in an estimate of the echo power in the received audio signal, and the estimated echo power is used to apply echo suppression to the received audio signal, thereby suppressing the echo in the received audio signal.

## ECHO SUPPRESSION

### BACKGROUND

[0001] A device may have audio input apparatus that can be used to receive audio signals from the surrounding environment. The device may also have audio output apparatus that can be used to output audio signals to the surrounding environment. For example, a device may have one or more speakers for outputting audio signals and one or more microphones for receiving audio signals. Audio signals which are output from the speaker(s) of the device may be received as “echo” in the audio signal received by the microphone(s). It may be the case that this echo is not desired in the received audio signal. For example, the device may be a user device (such as a mobile phone, tablet, laptop, PC, etc.) which is used in a communication event, such as an audio or video call, with another user device over a network. Far-end signals of the call may be output from the speaker at the user device and may be received as echo in the audio signals received by the microphone at the device. Such echo can be disturbing to users of the call, and the perceived quality of the call may be reduced due to the echo. In particular, the echo may cause interference for near-end audio signals which are intended to be received by the microphone and transmitted to the far-end in the call. Therefore echo cancellation and/or echo suppression may be applied to the received audio signals to thereby suppress the echo in the received audio signal. The power of the echo in the received audio signal may vary depending upon the arrangement of the user device. For example, the user device may be a mobile phone and in that case, the power of the echo in the received audio signal would normally be higher when the mobile phone is operating in a “hands-free” mode compared to when the mobile phone is not operating in a “hands-free” mode.

[0002] Echo cancellation (or “echo subtraction”) techniques aim to estimate an echo signal included in the audio signal received at the microphone, based on knowledge of the audio signal which is output from the speaker. The estimate of the echo signal can then be subtracted from the received audio signal thereby removing at least some of the echo from the received audio signal. Echo suppression is used to apply frequency-dependent suppression to the received audio signal to thereby suppress the echo in the received audio signal. In order for echo suppression to be implemented effectively, an echo suppressor needs to have an accurate estimate of the power of the echo in the received audio signal.

## **SUMMARY**

[0003] This Summary is provided to introduce a selection of concepts in a simplified form that are further described below in the Detailed Description. This Summary is not intended to identify key features or essential features of the claimed subject matter, nor is it intended to be used to limit the scope of the claimed subject matter.

[0004] There is provided a method of suppressing echo in a received audio signal. As part of the echo suppression, an estimate of the echo power of the echo is determined using a Finite Impulse Response (FIR) filter that is adapted to approximate the impulse response of the echo path. That is, a Finite Impulse Response filter estimate  $\hat{h}(n)$  is dynamically adapted in the time domain based on the outputted audio signal and the received audio signal to thereby model the impulse response of the echo path  $h(n)$  of the echo in the received audio signal. The filter estimate  $\hat{h}(n)$  is used to estimate the echo power of the echo in the received audio signal, and the estimated echo power is used to apply echo suppression to the received audio signal, thereby suppressing the echo in the received audio signal.

[0005] The method may be used in a call (e.g. a call implementing voice over internet protocol (VoIP) to transmit audio data between user devices) in which case the outputted audio signal may be a far-end signal received from the far-end of the call, and the received signal includes the resulting echo and a near-end signal for transmission to the far-end of the call.

## **BRIEF DESCRIPTION OF THE DRAWINGS**

[0006] Figure 1 shows a schematic illustration of a communication system in accordance with one or more embodiments;

[0007] Figure 2 is a schematic block diagram of a user device in accordance with one or more embodiments;

[0008] Figure 3 is a functional diagram showing modules of a user device for use in echo suppression in accordance with one or more embodiments; and

[0009] Figure 4 is a flow chart for a process of suppressing echo in accordance with one or more embodiments.

## **DETAILED DESCRIPTION**

[0010] In order for echo suppression to be implemented effectively, an echo suppressor needs to have an accurate estimate of the power of the echo in the received audio signal. As described herein, the echo power can be estimated using the output from a FIR filter that is adapted to approximate the impulse response of the echo path between a loudspeaker

outputting audio signals and a microphone receiving audio signals including the echo resulting from the outputted audio signals.

[0011] The FIR filter might only be used to estimate the echo power, and not to estimate the actual echo signal. This may be advantageous because the requirements for accuracy in the FIR filter are much less when used to estimate the echo power compared to if the FIR filter is used to estimate the actual echo signal. Therefore by estimating the echo power (rather than the echo signal) from the FIR filter, echo suppression is more robust to problems such as clock-drift between the playout (e.g. from a loudspeaker) and recording sides (e.g. at a microphone) in the VoIP client, nonlinearities in the echo path and changes in the echo path. In embodiments described herein, the FIR filter is adapted using time-domain data including phase information.

[0012] Figure 1 shows a communication system 100 comprising a first user 102 (“User A”) who is associated with a first user device 104 and a second user 108 (“User B”) who is associated with a second user device 110. In other embodiments the communication system 100 may comprise any number of users and associated user devices. The user devices 104 and 110 can communicate over the network 106 in the communication system 100, thereby allowing the users 102 and 108 to communicate with each other over the network 106. The communication system 100 shown in Figure 1 is a packet-based communication system, but other types of communication system could be used. The network 106 may, for example, be the Internet. Each of the user devices 104 and 110 may be, for example, a mobile phone, a tablet, a laptop, a personal computer (“PC”) (including, for example, Windows™, Mac OS™ and Linux™ PCs), a gaming device, a television, a personal digital assistant (“PDA”) or other embedded device able to connect to the network 106. The user device 104 is arranged to receive information from and output information to the user 102 of the user device 104. The user device 104 comprises output means such as a display and speakers. The user device 104 also comprises input means such as a keypad, a touch-screen, a microphone for receiving audio signals and/or a camera for capturing images of a video signal. The user device 104 is connected to the network 106.

[0013] The user device 104 executes an instance of a communication client, provided by a software provider associated with the communication system 100. The communication client is a software program executed on a local processor in the user device 104. The client performs the processing required at the user device 104 in order for the user device 104 to transmit and receive data over the communication system 100.

[0014] The user device 110 corresponds to the user device 104 and executes, on a local processor, a communication client which corresponds to the communication client executed at the user device 104. The client at the user device 110 performs the processing required to allow the user 108 to communicate over the network 106 in the same way that the client at the user device 104 performs the processing required to allow the user 102 to communicate over the network 106. The user devices 104 and 110 are endpoints in the communication system 100. Figure 1 shows only two users (102 and 108) and two user devices (104 and 110) for clarity, but many more users and user devices may be included in the communication system 100, and may communicate over the communication system 100 using respective communication clients executed on the respective user devices.

[0015] Figure 2 illustrates a detailed view of the user device 104 on which is executed a communication client instance 206 for communicating over the communication system 100. The user device 104 comprises a central processing unit (“CPU”) or “processing module” 202, to which is connected: output devices such as a display 208, which may be implemented as a touch-screen, and a speaker (or “loudspeaker”) 210 for outputting audio signals; input devices such as a microphone 212 for receiving audio signals, a camera 216 for receiving image data, and a keypad 218; a memory 214 for storing data; and a network interface 220 such as a modem for communication with the network 106. The user device 104 may comprise other elements than those shown in Figure 2. The display 208, speaker 210, microphone 212, memory 214, camera 216, keypad 218 and network interface 220 may be integrated into the user device 104 as shown in Figure 2. In alternative user devices one or more of the display 208, speaker 210, microphone 212, memory 214, camera 216, keypad 218 and network interface 220 may not be integrated into the user device 104 and may be connected to the CPU 202 via respective interfaces. One example of such an interface is a USB interface. If the connection of the user device 104 to the network 106 via the network interface 220 is a wireless connection then the network interface 220 may include an antenna for wirelessly transmitting signals to the network 106 and wirelessly receiving signals from the network 106.

[0016] Figure 2 also illustrates an operating system (“OS”) 204 executed on the CPU 202. Running on top of the OS 204 is the software of the client instance 206 of the communication system 100. The operating system 204 manages the hardware resources of the computer and handles data being transmitted to and from the network 106 via the network interface 220. The client 206 communicates with the operating system 204 and manages the connections over the communication system. The client 206 has a client user interface

which is used to present information to the user 102 and to receive information from the user 104. In this way, the client 206 performs the processing required to allow the user 102 to communicate over the communication system 100.

[0017] With reference to Figures 3 and 4 there is now described a method of suppressing echo. Figure 3 is a functional diagram of a part of the user device 104 showing how an echo suppression process is implemented, and Figure 4 is a flow chart for the process of suppressing echo.

[0018] As shown in Figure 3, the user device 104 comprises the speaker 210, the microphone 212, a FIR filter module 302, a power estimating module 304 and an echo suppression module 306. A signal  $x(t)$  to be output from the speaker 210 is coupled to an input of the speaker 210. It should be noted that in the embodiments described herein there is just one loudspeaker (indicated by reference numeral 210 in the figures) but in other embodiments there may be more than one loudspeaker to which the signal to be outputted is coupled (for outputting therefrom). Similarly, in the embodiments described herein there is just one microphone (indicated by reference numeral 212 in the figures) but in other embodiments there may be more than one microphone which receive audio signals from the surrounding environment. The signal to be output from the loudspeaker 210 is also coupled to a first input of a FIR filter module 302 and to a first input of the power estimating module 304. An output of the microphone 212 is coupled to a second input of the FIR filter module 302 and to a first input of the echo suppression module 306. An output of the FIR filter module 302 is coupled to a second input of the power estimating module 304. An output of the power estimating module is coupled to a second input of the echo suppression module 306. An output of the echo suppression module 306 is used to provide the received signal (with echo suppression having been applied) for further processing in the user device 104.

[0019] In step S402 a signal is received which is to be outputted from the speaker 210. For example, the signal to be outputted may be a far-end signal that has been received at the user device 104 from the user device 110 during a call between the users 102 and 108 over the communication system 100. Any processing that is required to be performed on the received signal (e.g. decoding using a speech codec, depacketizing, etc.) is performed as is known in the art (e.g. by the client 206) to arrive at the signal  $x(t)$  which is suitable to be outputted from the speaker 210. The signal  $x(t)$  is a digital signal. At least some of the processing of the signal in the user device 104 prior to outputting the signal from the speaker 210 is performed in the digital domain. As is known in the art, a digital to analogue converter (DAC) is applied to the digital signal  $x(t)$  before playout from the loudspeaker

210. Similarly, an analogue to digital converter (ADC) is applied to the signal captured by the microphone 212 to arrive at the digital signal  $y(t)$ .

[0020] In other embodiments, the signal to be outputted may be received from somewhere other than over the communication system 100 in a call. For example, the signal to be  
5 outputted may have been stored in the memory 214 and step S402 may comprise retrieving the signal from the memory 214.

[0021] In step S404 the audio signal  $x(t)$  is outputted from the speaker 210. In this way the audio signal  $x(t)$  is outputted to the user 102.

[0022] In step S406 the microphone 212 receives an audio signal. As shown in Figure 3  
10 the received audio signal may include a near-end signal which is a desired signal or “primary signal”. The near-end signal is the signal that the user 102 intends the microphone 212 to receive. However, the received audio signal also includes an echo signal resulting from the audio signals outputted from the speaker 210 in step S404. The received audio signal may also include noise, such as background noise. Therefore, the total received audio signal  $y(t)$   
15 can be given by the sum of the near-end signal, the echo and the noise. The echo and the noise act as interference for the near-end signal.

[0023] The FIR filter module 302 takes as inputs the outputted audio signal  $x(t)$  and the received audio signal  $y(t)$ . In step S408 the FIR filter module 302 dynamically adapts a FIR filter estimate  $\hat{h}(n)$  in the time domain based on the outputted audio signal  $x(t)$  and the  
20 received audio signal  $y(t)$  to model an echo path  $h(n)$  of the echo in the received audio signal  $y(t)$ . The “impulse response of the echo path  $h(n)$ ” is also referred to herein as the “echo path  $h(n)$ ”.

[0024] For a linear echo path impulse response of length  $N_{true} + 1$  the echo path  $h(n)$  describes how the echo in the received audio signal relates to the audio signal  $x(t)$  output  
25 from the speaker 210, e.g. according to the equation  $y^{echo}(t) = \sum_{n=0}^{N_{true}} h(n)x(t-n)$ , where  $y^{echo}(t)$  is the echo in the received audio signal  $y(t)$ . The echo path  $h(n)$  may vary over time. The echo path  $h(n)$  may depend upon (i) the current environmental conditions surrounding the speaker 210 and the microphone 212 (e.g. whether there are any physical obstructions to the passage of the audio signal from the speaker 210 to the microphone 212,  
30 the air pressure, temperature, wind, etc.), and (ii) characteristics of the speaker 210 and/or the microphone 212 which may alter the signal as it is outputted and/or received.

[0025] The FIR filter module 302 models the echo path  $h(n)$  of the echo in the received audio signal by determining a weighted sum of the current and a finite number (N) of

previous values of the outputted audio signal  $x(t)$ . The FIR filter module 302 therefore implements an Nth order FIR filter which has a finite length (in time) over which it considers the values of the outputted audio signal  $x(t)$  in determining the estimate of the echo path  $\hat{h}(n)$ . In this way, the FIR filter module 302 dynamically adapts the FIR filter estimate  $\hat{h}(n)$ . The operation is described by the following equation, which defines the echo in the received audio signal  $y(t)$  in terms of the outputted audio signal  $x(t)$ :

$$\hat{y}^{echo}(t) = \sum_{n=0}^N \hat{h}(n)x(t-n).$$

[0026] Therefore N+1 samples of the outputted audio signal  $x(t)$  are used, with a respective N+1 weights  $\hat{h}(n)$ . The set of N+1 weights  $\hat{h}(n)$  is referred to herein simply as the estimate of the echo path  $\hat{h}(n)$ . In other words the estimate of the echo path  $\hat{h}(n)$  consists of N+1 values where the FIR filter module 302 implements an Nth order FIR filter, taking N+1 values (e.g. N+1 samples) of the signal  $x(t)$  into account.

[0027] It can be appreciated that it is easier to adapt the FIR filter estimate  $\hat{h}(n)$  when the echo is a dominant part of the received audio signal, that is when  $y(t) \cong \hat{y}^{echo}(t)$ . For example, in some embodiments it may be possible to detect when the power of the near-end signal is greater than the power of the echo (e.g. when the user 102 is speaking), and whilst that is the case the FIR estimate  $\hat{h}(n)$  is not adapted, but when the power of the near-end signal is less than the power of the echo in the received audio signal  $y(t)$  (e.g. when the user 102 is not speaking) the FIR estimate  $\hat{h}(n)$  is adapted.

[0028] However, it may be possible to adapt the FIR filter estimate  $\hat{h}(n)$  even when the echo is not a dominant part of the received audio signal. For example, if the adaptation is slower, there is a gain in robustness to disturbances.

[0029] The FIR filter estimate  $\hat{h}(n)$  is passed from the FIR filter module 302 to the power estimating module 304. In step S410 the power estimating module 304 estimates the echo power of the echo in the received audio signal based on the filter estimate  $\hat{h}(n)$  determined in step S408. Step S410 might not comprise estimating the echo signal  $\hat{y}^{echo}(t)$  in the received audio signal  $y(t)$ . The echo power is estimated as a function of time and frequency. Using the estimated echo path  $\hat{h}(n)$  and the input signal  $x(t)$  to estimate the echo power for an echo suppressor is much less sensitive to inaccuracies in the FIR filter approximation of the echo path, than using an estimate of the echo signal  $\hat{y}^{echo}(t)$  for an echo subtractor/canceller. For example, the usage of the estimate of the echo path  $\hat{h}(n)$  provided by the FIR filter module 302 and the input signal  $x(t)$  for echo power estimation is much



less sensitive to problems such as clock-drift between the playout and recording sides on the local endpoint, suboptimal FIR filter updates, nonlinearities in the echo path, and time-variations in the echo path. In echo suppression a rather low accuracy of the echo power estimate is sufficient to achieve good echo suppression. In this way, a linear filter is used for estimating a power spectrum used for echo suppression. This is more robust than using a linear filter for estimating a time domain echo signal for an echo subtractor.

[0030] Step S410 may comprise estimating the echo power  $\hat{P}_s(t, f)$ , which is a scalar power bin for time  $t$  and frequency  $f$ . This can be calculated using the Fourier transform, as known by people skilled in the art.

[0031] In this way, the FIR filter estimate  $\hat{h}(n)$ , that has been adapted using the speaker and microphone signals  $x(t)$  and  $y(t)$  to approximate the time-varying echo path  $h(n)$  of the VoIP client, is used, with the outputted audio signal samples  $x(t)$ , to estimate the power  $\hat{P}_s(t, f)$  of the echo signal at time  $t$  and frequency  $f$ .

[0032] The estimate of the echo power  $\hat{P}_s(t, f)$  is output from the power estimating module 304 and received by the echo suppression module 306. The echo suppression module 306 also receives the audio signal  $y(t)$  from the microphone 212. In step S412 the echo suppression module 306 uses the estimate of the echo power  $\hat{P}_s(t, f)$  to apply echo suppression to the received audio signal  $y(t)$ , thereby suppressing the echo in the received audio signal. The estimate of the echo power  $\hat{P}_s(t, f)$  is frequency dependent and the suppression applied by the echo suppression module 306 is also frequency dependent.

[0033] The purpose of the echo suppressor is to suppress the loudspeaker echo present in the microphone signal, e.g. in a VoIP client, to a level sufficiently low for it not to be noticeable/disturbing in the presence of the near-end sounds (non-echo sounds) picked up by the microphone 212. In order to be able to choose the proper amount of echo suppression a good estimate of the echo power (e.g. as a function of frequency and time) is needed, and as described above this is provided to the echo suppression module 306 by the power estimating module 304. The echo suppression module 306 is designed to apply signal dependent suppression that varies both over time and frequency to the received audio signal  $y(t)$ . Echo suppression methods are known in the art. Furthermore, the echo suppression method applied by the echo suppression module 306 may be implemented in different ways. As such, the exact details of the echo suppression method are therefore not described in detail herein.

[0034] The echo suppression module 306 outputs the received signal, with the echo having been suppressed, for further processing at the user device 104. For example, the signal output from the echo suppression module 306 may be processed by the client 206 (e.g. encoded and packetized) and then transmitted over the network 106 to the user device 110 in a call between the users 102 and 108. Additionally or alternatively, the signal output from the echo suppression module 306 may be used for other purposes by the user device 104, e.g. the signal may be stored in the memory 214 or used as an input to an application which is executing at the user device 104.

[0035] There is therefore described herein the use of an FIR filter module 302 to model the echo path to estimate the power of the loudspeaker echo signal picked up by the microphone 212, for the purpose of computing and applying an echo suppression effect/filter (e.g. for use by the VoIP client 206).

[0036] In the embodiments described above, the echo suppression is implemented in a VoIP system (e.g. the received audio signal may include speech of the user 102 for transmission to the user device 110 during a call between the users 102 and 108 over the communication system 100). However, the echo suppression methods described herein can be applied in any suitable system in which echo suppression is to be applied.

[0037] In the embodiments described above, and shown in the Figures, echo cancellation (or “echo subtraction”) is not applied to the received audio signal  $y(t)$ . That is, there is no echo cancellation module in the user device 104 and the echo suppression is applied to the received audio signal  $y(t)$  without a prior step of applying echo cancellation to the received audio signal  $y(t)$ .

[0038] However, in other embodiments, echo cancellation may be applied, by an echo cancellation module, to the received audio signal  $y(t)$ . In particular, the echo suppression applied by the echo suppression module 306 may be applied downstream (i.e. after) of the echo cancellation in the processing of the received audio signal  $y(t)$ . The echo cancellation module would subtract an estimate of the echo signal from the received audio signal, but due to inaccuracies in the estimate of the echo signal, a residual echo would most-likely remain in the received audio signal. It is the residual echo that would then be suppressed by the echo suppression module 306. This echo suppression could be applied in the same way as described herein in the embodiments in which no echo cancellation is applied. If echo subtraction is used, the effect of it, would be taken into account in the echo suppression.

[0039] The methods described herein may be implemented by executing a computer program product (e.g. the client 206) at the user device 104. That is, a computer program

product may be configured to suppress echo in the received audio signal  $y(t)$ , wherein the computer program product is embodied on a computer-readable storage medium (e.g. stored in the memory 214) and configured so as when executed on the CPU 202 to perform the operations of any of the methods described herein. The term “computer-readable storage  
5 medium” is intended to cover all statutory forms of media and thus excludes non-statutory forms of media.

[0040] Generally, any of the functions described herein (e.g. the functional modules shown in Figure 3 and the functional steps shown in Figure 4) can be implemented using software, firmware, hardware (e.g., fixed logic circuitry), or a combination of these  
10 implementations. The modules and steps shown separately in Figures 3 and 4 may or may not be implemented as separate modules or steps. For example, the echo suppression module 306 may perform the function of the power estimating module 304, such that a separate power estimating module 304 is not required in addition to the echo suppression module 306. The terms “module,” “functionality,” “component” and “logic” as used  
15 herein generally represent software, firmware, hardware, or a combination thereof. In the case of a software implementation, the module, functionality, or logic represents program code that performs specified tasks when executed on a processor (e.g. CPU or CPUs). The program code can be stored in one or more computer readable memory devices. The features of the techniques described herein are platform-independent, meaning that the  
20 techniques may be implemented on a variety of commercial computing platforms having a variety of processors.

[0041] For example, the user devices may also include an entity (e.g. software) that causes hardware of the user devices to perform operations, e.g., processors functional blocks, and so on. For example, the user devices may include a computer-readable medium that may  
25 be configured to maintain instructions that cause the user devices, and more particularly the operating system and associated hardware of the user devices to perform operations. Thus, the instructions function to configure the operating system and associated hardware to perform the operations and in this way result in transformation of the operating system and associated hardware to perform functions. The instructions may be provided by the  
30 computer-readable medium to the user devices through a variety of different configurations.

[0042] One such configuration of a computer-readable medium is signal bearing medium and thus is configured to transmit the instructions (e.g. as a carrier wave) to the computing device, such as via a network. The computer-readable medium may also be configured as a computer-readable storage medium and thus is not a signal bearing medium. Examples

of a computer-readable storage medium include a random-access memory (RAM), read-only memory (ROM), an optical disc, flash memory, hard disk memory, and other memory devices that may use magnetic, optical, and other techniques to store instructions and other data.

- 5   **[0043]** Although the subject matter has been described in language specific to structural features and/or methodological acts, it is to be understood that the subject matter defined in the appended claims is not necessarily limited to the specific features or acts described above. Rather, the specific features and acts described above are disclosed as example forms of implementing the claims.

**CLAIMS**

1. A method of suppressing echo, the method comprising:  
outputting an audio signal;  
receiving an audio signal, wherein the received audio signal includes an echo resulting from said outputted audio signal;  
dynamically adapting a Finite Impulse Response filter estimate  $\hat{h}(n)$  in the time domain based on the outputted audio signal and the received audio signal to model an echo path  $h(n)$  of the echo in the received audio signal;  
using the filter estimate  $\hat{h}(n)$  in an estimate of the echo power in the received audio signal; and  
using the estimated echo power to apply echo suppression to the received audio signal, thereby suppressing the echo in the received audio signal.
2. The method of claim 1 wherein said echo suppression is applied to the received audio signal without a prior step of applying echo cancellation to the received audio signal.
3. The method of claim 1 or 2 wherein the outputted audio signal is used in the estimate of the echo power in the received audio signal.
4. The method of claim 1 further comprising applying echo cancellation to the received audio signal, wherein said echo suppression is applied after the echo cancellation in the processing of the received audio signal.
5. The method of any preceding claim wherein said echo suppression which is applied to the received audio signal is signal dependent suppression that varies over time and frequency.
6. The method of any preceding claim wherein the method is performed at a user device for use in a communication event, and wherein the received audio signal comprises speech of a user for transmission from the user device in the communication event.

7. The method of claim 6 wherein the communication event is a voice-over-internet-protocol (VoIP) call, and wherein the outputted audio signal comprises far-end speech signals of the VoIP call which are included in the echo in the received audio signal.
8. A device configured to implement echo suppression, the device comprising:  
audio output apparatus configured to output an audio signal;  
audio input apparatus configured to receive an audio signal, wherein the received audio signal includes an echo resulting from said outputted audio signal;  
a filter estimating module configured to dynamically adapt a Finite Impulse Response filter estimate  $\hat{h}(n)$  in the time domain based on the outputted audio signal and the received audio signal to model an echo path  $h(n)$  related to the echo in the received audio signal;  
a power estimating module configured to use the filter estimate  $\hat{h}(n)$  in an estimate of the echo power of the echo in the received audio signal; and  
an echo suppression module configured to use the estimated echo power to apply echo suppression to the received audio signal, thereby suppressing the echo in the received audio signal.
9. The device of claim 8 wherein the audio output apparatus comprises a speaker configured to output the outputted audio signal, and wherein the audio input apparatus comprises a microphone configured to receive the received audio signal.
10. A computer program product configured to suppress echo in a received audio signal, the computer program product being embodied on a computer-readable storage medium and configured so as when executed on a processor to perform the operations of any of claims 1 to 7.

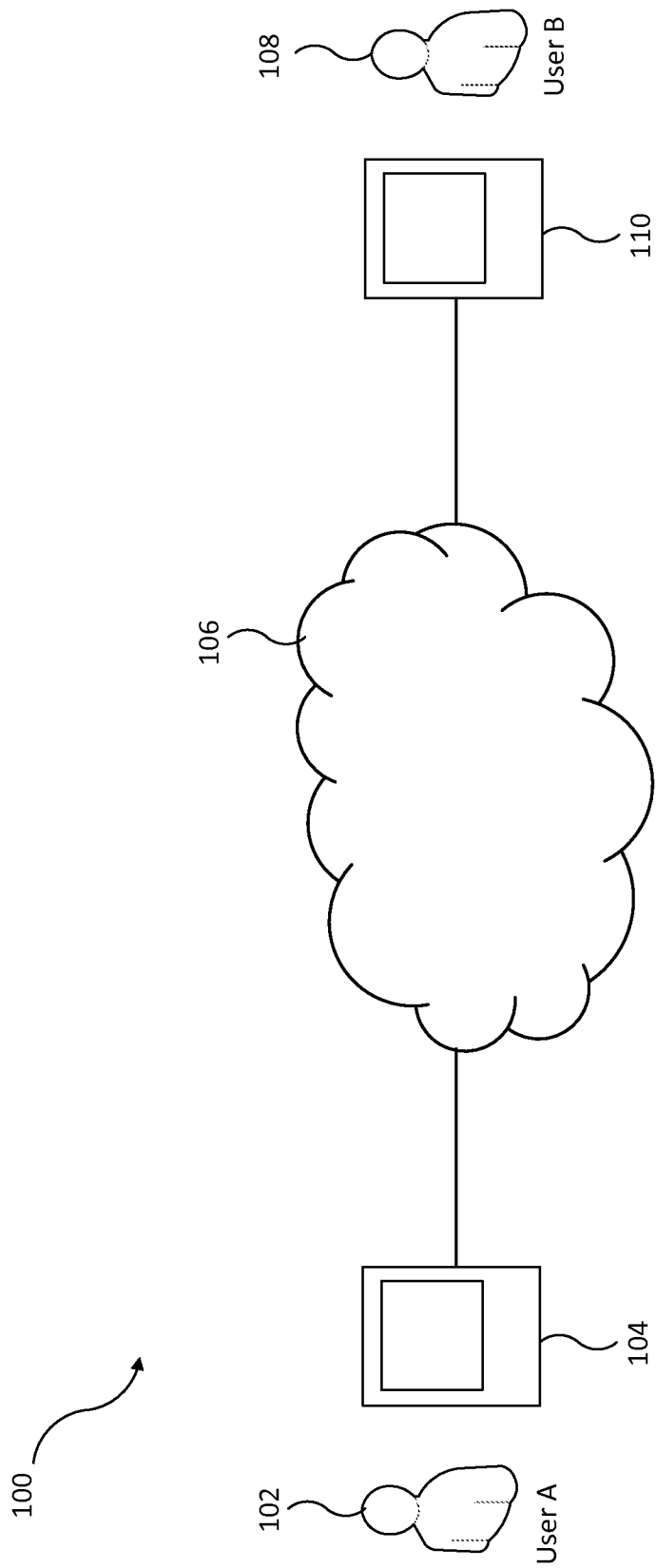


FIG. 1

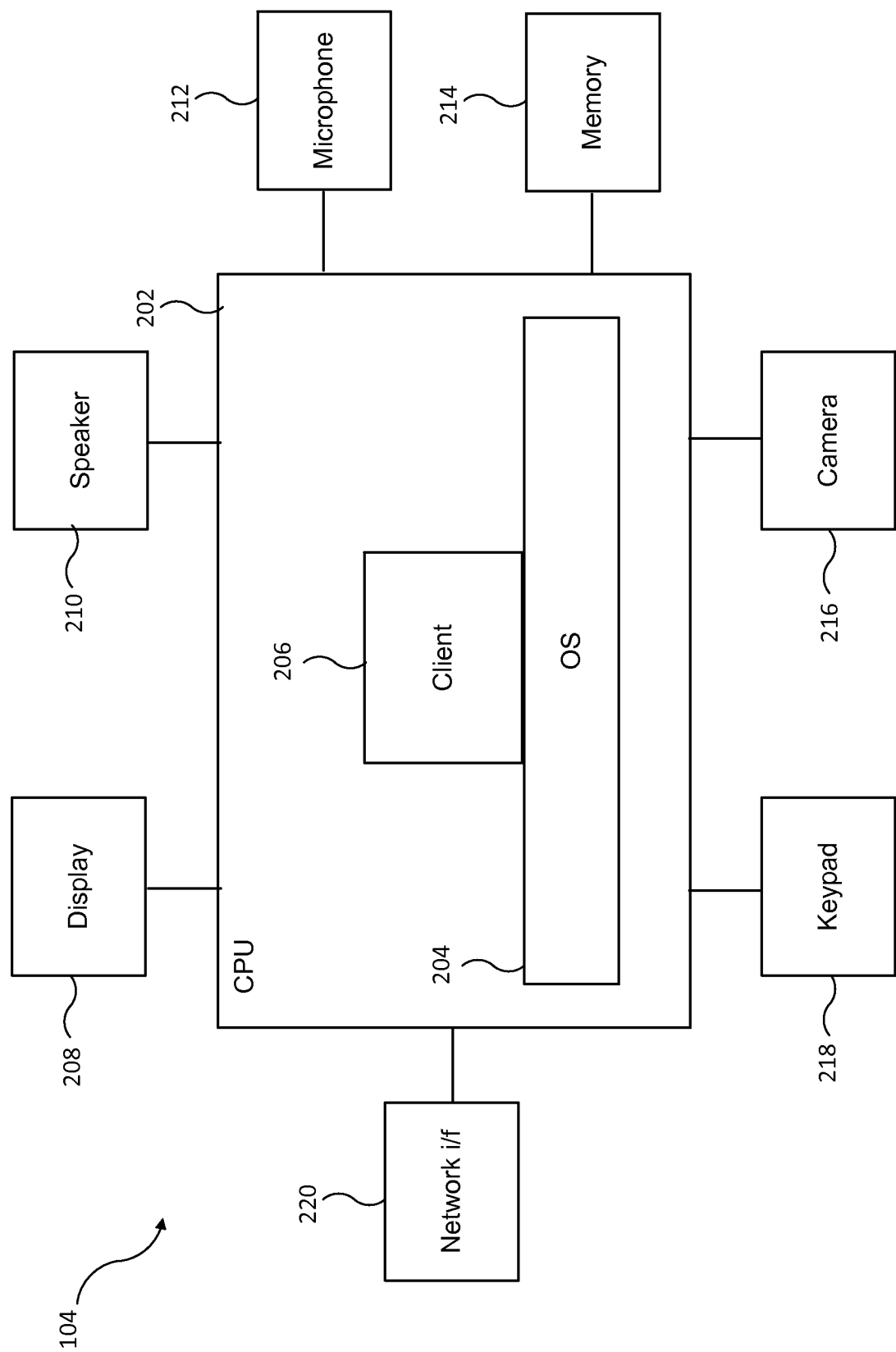


FIG. 2



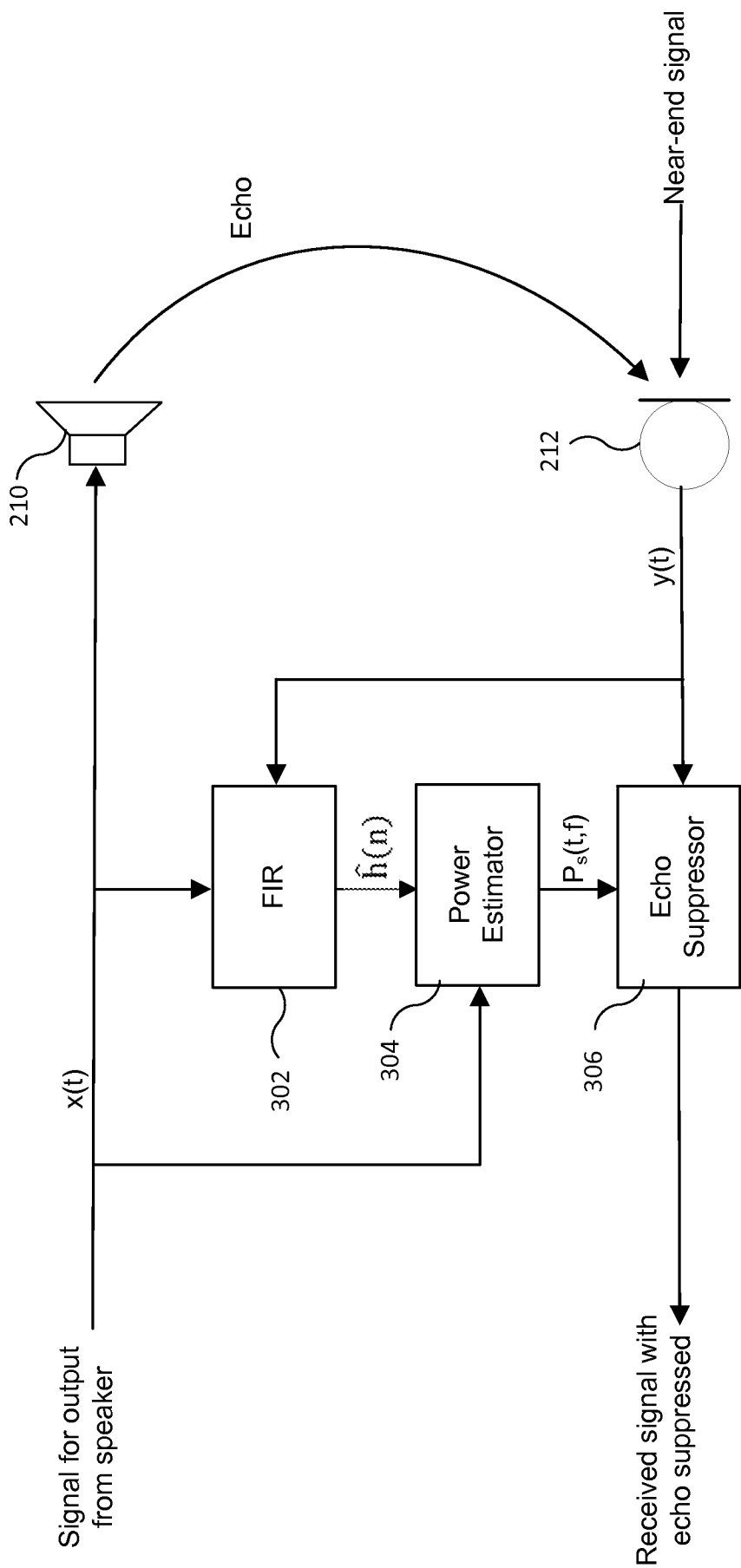
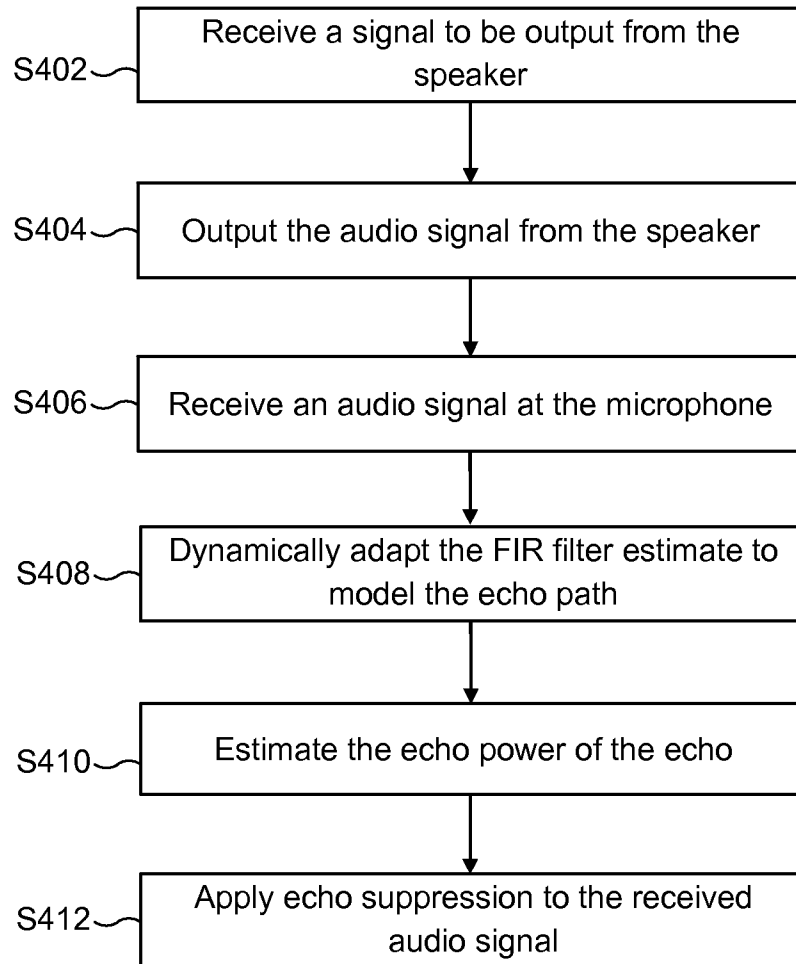


FIG. 3

4/4

**FIG. 4**

## INTERNATIONAL SEARCH REPORT

International application No

PCT/US2013/077353

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 H04B

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)

EPO-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	JP 2009 021741 A (YAMAHA CORP) 29 January 2009 (2009-01-29) abstract; figure 1	1-10
X	----- WO 2004/064365 A1 (PHILIPS INTELLECTUAL PROPERTY [DE]; KONINKL PHILIPS ELECTRONICS NV [NL] 29 July 2004 (2004-07-29) page 3, line 26 - page 4, line 6 page 7, line 12 - page 8, line 30; figure 1	1-10
A	----- US 2004/057574 A1 (FALLER CHRISTOF [CH]) 25 March 2004 (2004-03-25) paragraphs [0023] - [0035], [0057] - [0067] -----	1-10



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

25 February 2014

Date of mailing of the international search report

06/03/2014

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

Information on patent family members

International application No

PCT/US2013/077353

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
JP 2009021741 A	29-01-2009	JP 4978352 B2 JP 2009021741 A	18-07-2012 29-01-2009
WO 2004064365 A1	29-07-2004	AT 459195 T AU 2003285714 A1 CN 1736091 A EP 1584178 A1 JP 4376793 B2 JP 2006513621 A US 2006153359 A1 WO 2004064365 A1	15-03-2010 10-08-2004 15-02-2006 12-10-2005 02-12-2009 20-04-2006 13-07-2006 29-07-2004
US 2004057574 A1	25-03-2004	NONE	