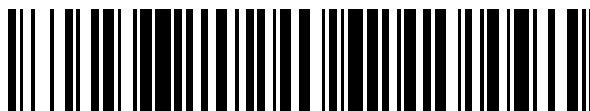


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 619 563**

51 Int. Cl.:

H04M 9/08

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **29.05.2014 PCT/US2014/039871**

87 Fecha y número de publicación internacional: **04.12.2014 WO2014194011**

96 Fecha de presentación y número de la solicitud europea: **29.05.2014 E 14733861 (0)**

97 Fecha y número de publicación de la concesión europea: **22.02.2017 EP 2987313**

54 Título: **Extracción de eco**

30 Prioridad:

31.05.2013 GB 201309773
28.08.2013 US 201314012458

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

26.06.2017

73 Titular/es:

MICROSOFT TECHNOLOGY LICENSING, LLC
(100.0%)
One Microsoft Way
Redmond, WA 98052, US

72 Inventor/es:

AHGREN, PER

74 Agente/Representante:

UNGRÍA LÓPEZ, Javier

ES 2 619 563 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Extracción de eco

5 **Antecedentes**

Un dispositivo puede tener un aparato de entrada audio que puede ser usado para recibir señales audio del entorno circundante. El dispositivo también puede tener un aparato de salida audio que puede ser usado para emitir señales audio al entorno circundante. Por ejemplo, un dispositivo puede tener uno o varios altavoces para enviar señales audio y uno o varios micrófonos para recibir señales audio. Las señales audio que son emitidas por el o los altavoces del dispositivo pueden recibirse como "eco" en la señal audio recibida por el (los) micrófono(s). Puede darse el caso de que este eco no sea deseable en la señal audio recibida. Por ejemplo, el dispositivo puede ser un dispositivo de usuario (tal como un teléfono móvil, tableta, ordenador portátil, PC, etc) que se usa en un evento de comunicación, tal como una llamada audio o vídeo, con otro dispositivo de usuario por una red. Las señales de extremo lejano de la llamada pueden ser emitidas por el altavoz del dispositivo de usuario y pueden ser recibidas como eco en las señales audio recibidas por el micrófono en el dispositivo. Tal eco puede ser perturbador para los usuarios de la llamada, y la calidad percibida de la llamada se puede reducir debido al eco. En particular, el eco puede producir interferencia para señales audio de extremo próximo destinadas a ser recibidas por el micrófono y transmitidas al extremo lejano en la llamada. Por lo tanto, se puede aplicar cancelación de eco y/o supresión de eco a las señales audio recibidas para suprimir por ello el eco en la señal audio recibida. La potencia del eco en la señal audio recibida puede variar dependiendo de la disposición del dispositivo de usuario. Por ejemplo, el dispositivo de usuario puede ser un teléfono móvil y, en ese caso, la potencia del eco en la señal audio recibida sería normalmente más alta cuando el teléfono móvil esté operando en un modo "manos libres" en comparación con cuando el teléfono móvil no está operando en un modo "manos libres".

Las técnicas de cancelación de eco (o "supresión de eco") tienen la finalidad de estimar una señal de eco incluida en la señal audio recibida en el micrófono, en base a conocimiento de la señal audio que es emitida por el altavoz. La estimación de la señal de eco se puede restar entonces de la señal audio recibida quitando por ello al menos parte del eco de la señal audio recibida. La supresión de eco se usa para aplicar supresión dependiente de frecuencia a la señal audio recibida para suprimir por ello el eco en la señal audio recibida.

Luis A. Azpicueta-Ruiz y colaboradores: "Novel schemes for nonlinear acoustic echo cancellation based on filter combinations", Acoustics, Speech and Signal Processing, 2009, describe dos esquemas de cancelación de eco acústico no lineal en base a esquemas de combinación. El primer esquema, 'Combination of filters'-Scheme (CFS), consta de una combinación de un filtro lineal y otro Volterra, mientras que el segundo, 'Combination of kernels'-Scheme (CKS), se basa en la combinación de un núcleo cuadrático y otro todo ceros.

Resumen

Este resumen se ofrece con el fin de presentar una selección de conceptos de forma simplificada que se describen mejor más adelante en la descripción detallada. Este resumen no tiene la finalidad de identificar elementos clave o características esenciales de la materia reivindicada, ni se pretende usarla para limitar el alcance de la materia reivindicada.

Se facilita un método de extraer eco en una señal audio recibida. Como parte de la extracción de eco, se modela un recorrido del eco en la señal audio recibida usando un primer modelo para determinar una primera estimación de modelo del eco en la señal audio recibida. La primera estimación de modelo se usa para determinar un primer valor de rendimiento del recorrido de eco según una métrica de rendimiento. El primer valor de rendimiento se compara con un valor umbral. Entonces se determina si el recorrido de eco puede ser considerado lineal en base a la comparación. Si el recorrido de eco puede ser considerado lineal, la primera estimación de modelo del eco se usa para quitar el eco en la señal audio recibida. De otro modo, el recorrido del eco en la señal audio recibida es modelado usando un segundo modelo en base a la señal audio salida y la señal audio recibida para determinar una segunda estimación de modelo del eco, y la segunda estimación de modelo del eco se usa para quitar el eco en la señal audio recibida.

El método puede ser usado en una llamada (por ejemplo, una llamada que implementa protocolo de voz por Internet (VoIP) para transmitir datos audio entre dispositivos de usuario) en cuyo caso la señal audio emitida puede ser una señal de extremo lejano recibida del extremo lejano de la llamada, y la señal recibida incluye el eco resultante y una señal de extremo próximo para transmisión al extremo lejano de la llamada.

60 **Breve descripción de los dibujos**

Para una mejor comprensión de las realizaciones descritas y para mostrar cómo la misma se puede llevar a cabo, ahora se hará referencia, a modo de ejemplo, a los dibujos siguientes, en los que:

65 La figura 1 representa una ilustración esquemática de un sistema de comunicaciones.

La figura 2 es un diagrama esquemático de bloques de un dispositivo de usuario.

La figura 3 es un diagrama funcional que representa módulos de un dispositivo de usuario para uso en extracción de eco.

Y la figura 4 es un diagrama de flujo para un proceso de quitar eco.

Descripción detallada

Ahora se describirán realizaciones a modo de ejemplo solamente.

La figura 1 representa un sistema de comunicaciones 100 incluyendo un primer usuario 102 ("Usuario A") que está asociado con un primer dispositivo de usuario 104 y un segundo usuario 108 ("Usuario B") que está asociado con un segundo dispositivo de usuario 110. En otras realizaciones, el sistema de comunicaciones 100 puede incluir cualquier número de usuarios y dispositivos de usuario asociados. Los dispositivos de usuario 104 y 110 pueden comunicar por la red 106 en el sistema de comunicaciones 100, permitiendo por ello que los usuarios 102 y 108 comuniquen entre sí por la red 106. El sistema de comunicaciones 100 representado en la figura 1 es un sistema de comunicaciones basado en paquetes, pero se podría usar otros tipos de sistemas de comunicaciones. La red 106 puede ser, por ejemplo, Internet. Cada uno de los dispositivos de usuario 104 y 110 puede ser, por ejemplo, un teléfono móvil, una tableta, un ordenador portátil, un ordenador personal ("PC") (incluyendo, por ejemplo, PCs Windows™, Mac OSTM y Linux™), un dispositivo de juego, una televisión, un asistente digital personal ("PDA") u otro dispositivo embebido capaz de conectar a la red 106. El dispositivo de usuario 104 está dispuesto para recibir y enviar información de/al usuario 102 del dispositivo de usuario 104. El dispositivo de usuario 104 incluye medios de salida tales como una pantalla y altavoces. El dispositivo de usuario 104 también incluye medios de entrada tales como un teclado, una pantalla táctil, un micrófono para recibir señales audio y/o una cámara para capturar imágenes de una señal vídeo. El dispositivo de usuario 104 está conectado a la red 106.

El dispositivo de usuario 104 ejecuta una instancia de un cliente de comunicación, proporcionado por un proveedor de software asociado con el sistema de comunicaciones 100. El cliente de comunicación es un programa de software ejecutado en un procesador local en el dispositivo de usuario 104. El cliente realiza el procesamiento requerido en el dispositivo de usuario 104 para que el dispositivo de usuario 104 transmita y reciba datos por el sistema de comunicaciones 100.

El dispositivo de usuario 110 corresponde al dispositivo de usuario 104 y ejecuta, en un procesador local, un cliente de comunicación que corresponde al cliente de comunicación ejecutado en el dispositivo de usuario 104. El cliente en el dispositivo de usuario 110 realiza el procesamiento requerido para que el usuario 108 pueda comunicar por la red 106 de la misma forma que el cliente en el dispositivo de usuario 104 realiza el procesamiento requerido para que el usuario 102 pueda comunicar por la red 106. Los dispositivos de usuario 104 y 110 son puntos finales en el sistema de comunicaciones 100. La figura 1 representa solamente dos usuarios (102 y 108) y dos dispositivos de usuario (104 y 110) para claridad, pero se puede incluir muchos más usuarios y dispositivos de usuario en el sistema de comunicaciones 100, y pueden comunicar por el sistema de comunicaciones 100 usando respectivos clientes de comunicación ejecutados en los respectivos dispositivos de usuario.

La figura 2 ilustra una vista detallada del dispositivo de usuario 104 en el que se ejecuta una instancia de cliente de comunicación 206 para comunicar por el sistema de comunicaciones 100. El dispositivo de usuario 104 incluye una unidad central de proceso ("CPU") o "módulo de procesamiento" 202, al que están conectados: dispositivos de salida tales como una pantalla 208, que puede implementarse como una pantalla táctil, y un altavoz (o "parlante") 210 para enviar señales audio; dispositivos de entrada tales como un micrófono 212 para recibir señales audio, una cámara 216 para recibir datos de imagen, y un teclado 218; una memoria 214 para almacenar datos; y una interfaz de red 220 tal como un módem para comunicación con la red 106. El dispositivo de usuario 104 puede incluir otros elementos distintos de los representados en la figura 2. La pantalla 208, el altavoz 210, el micrófono 212, la memoria 214, la cámara 216, el teclado 218 y la interfaz de red 220 pueden estar integrados en el dispositivo de usuario 104 como se representa en la figura 2. En dispositivos de usuario alternativos uno o varios de la pantalla 208, el altavoz 210, el micrófono 212, la memoria 214, la cámara 216, el teclado 218 y la interfaz de red 220 pueden no estar integrados en el dispositivo de usuario 104 y pueden estar conectados a la CPU 202 mediante interfaces respectivas. Un ejemplo de tal interfaz es una interfaz USB. Si la conexión del dispositivo de usuario 104 a la red 106 mediante la interfaz de red 220 es una conexión inalámbrica, entonces la interfaz de red 220 puede incluir una antena para transmitir de forma inalámbrica señales a la red 106 y recibir de forma inalámbrica señales de la red 106.

La figura 2 también ilustra un sistema operativo ("OS") 204 ejecutado en la CPU 202. Sobre el OS 204 se ejecuta el software de la instancia de cliente 206 del sistema de comunicaciones 100. El sistema operativo 204 gestiona los recursos de hardware del ordenador y maneja datos que son transmitidos a y de la red 106 mediante la interfaz de red 220. El cliente 206 comunica con el sistema operativo 204 y gestiona las conexiones por el sistema de comunicaciones. El cliente 206 tiene una interfaz de usuario cliente que se usa para presentar información al usuario

ES 2 619 563 T3

102 y para recibir información del usuario 102. De esta forma, el cliente 206 realiza el procesado requerido para que el usuario 102 pueda comunicar por el sistema de comunicaciones 100.

5 En cancelación de eco acústico la finalidad es quitar la señal de eco $s(t)$ en la señal de micrófono $y(t)$ que se origina a partir de la señal de altavoz $x(t)$. Esto se deberá hacer lo más exactamente posible y de la forma menos obstrusiva posible con el fin de tener el mínimo impacto en la percepción de cualquier señal de extremo próximo $v(t)$. La señal de micrófono puede escribirse como $y(t) = s(t) + v(t)$. La señal de eco es una función de la señal de altavoz como $s(t) = F(x(t))$.

10 Hay dos formas principales de lograr lo anterior, siendo una la sustracción de eco y siendo la otra la supresión de eco. A menudo se combinan estos dos acercamientos.

15 Estos dos métodos de extracción de eco requieren que se estime un modelo del recorrido de eco. Se puede usar múltiples modelos para modelar el recorrido de eco y, dependiendo de qué modelo se use, el modelo puede ser más o menos adecuado para los recorridos de eco disponibles. Un ejemplo es cuando se usa un modelo FIR lineal para modelar el recorrido de eco. Este tipo de modelo es adecuado para cancelación de eco cuando el recorrido de eco es bastante lineal. Sin embargo, cuando el recorrido de eco es más bien no lineal funciona típicamente de forma pobre.

20 Con referencia a las figuras 3 y 4 ahora se describe un método de extraer eco. La figura 3 es un diagrama funcional de una parte del dispositivo de usuario 104 que representa cómo se implementa un proceso de extracción de eco.

25 Como se representa en la figura 3, el dispositivo de usuario 104 incluye el altavoz 210, el micrófono 212, un módulo de modelado 302, y un módulo de extracción de eco 314. El módulo de modelado 302 incluye un primer módulo filtro 304, un segundo módulo filtro 308, y un módulo de determinación de rendimiento 312. El módulo de extracción de eco 314 se describe con referencia a la figura 3 como un módulo de supresión de eco 314.

La figura 4 es un diagrama de flujo para el proceso de suprimir eco.

30 Una señal $x(t)$ a emitir por el altavoz 210 se acopla a una entrada del altavoz 210. Se deberá indicar que en las realizaciones aquí descritas hay solamente un altavoz (indicado con el número de referencia 210 en las figuras), pero en otras realizaciones puede haber más de un altavoz al que se acopla la señal a emitir (para que salga por él). Igualmente, en las realizaciones aquí descritas hay solamente un micrófono (indicado con el número de referencia 212 en las figuras), pero en otras realizaciones puede haber más que un micrófono que reciba señales audio del entorno circundante. La señal a emitir por el altavoz 210 también se acopla al módulo de modelado 302. En particular, la señal a emitir por el altavoz 210 se acopla a una primera entrada del primer módulo filtro 304 y a una primera entrada del segundo módulo filtro 308. Una salida del micrófono 212 se acopla al módulo de modelado 302. En particular, la salida del micrófono 212 se acopla a una segunda entrada del primer módulo filtro 304 y a una segunda entrada del segundo módulo filtro 308. Las salidas del módulo de modelado 302 están acopladas al módulo de supresión de eco 314. En particular una salida del módulo de determinación de rendimiento 312 se acopla a una primera entrada del módulo de supresión de eco 314, y una salida del segundo módulo filtro 308 se acopla a una segunda entrada de módulo de supresión de eco 314. Una salida del primer módulo filtro 304 se acopla a una primera entrada del módulo de determinación de rendimiento 312. La salida del micrófono 212 se acopla a una segunda entrada del módulo de determinación de rendimiento 312. Una salida del módulo de determinación de rendimiento 312 se acopla a una tercera entrada del segundo módulo filtro 308. La salida del micrófono 212 también se acopla a una tercera entrada del módulo de supresión de eco 314. Se usa una salida del módulo de supresión de eco 314 para proporcionar la señal recibida (con supresión de eco que se ha aplicado) para procesado adicional en el dispositivo de usuario 104.

50 En el paso S402 se recibe una señal que ha de ser emitida por el altavoz 210. Por ejemplo, la señal a emitir puede ser una señal de extremo lejano que se ha recibido en el dispositivo de usuario 104 del dispositivo de usuario 110 durante una llamada entre los usuarios 102 y 108 por el sistema de comunicación 100. Cualquier procesado que haya que realizar en la señal recibida (por ejemplo, decodificación usando un codec de voz, despaquetización, etc) se realiza como es conocido en la técnica (por ejemplo, por el cliente 206) para llegar a la señal $x(t)$ que es adecuada para ser emitida por el altavoz 210. La señal $x(t)$ es una señal digital. Al menos parte del procesado de la señal en el dispositivo de usuario 104 antes de la salida de la señal por el altavoz 210 se realiza en el dominio digital. Como es conocido en la técnica, se aplica un convertidor digital a analógico (DAC) a la señal digital $x(t)$ antes de la reproducción por el altavoz 210. Igualmente, se aplica un convertidor analógico a digital (ADC) a la señal capturada por el micrófono 212 para llegar a la señal digital $y(t)$.

60 En otras realizaciones, la señal a emitir puede ser recibida de algún lugar distinto de por el sistema de comunicaciones 100 en una llamada. Por ejemplo, la señal a emitir puede haber sido almacenada en la memoria 214 y el paso S402 puede incluir recuperar la señal de la memoria 214.

65 En el paso S404, la señal audio $x(t)$ es emitida por el altavoz 210. De esta forma, la señal audio $x(t)$ es enviada al usuario 102.

En el paso S406, el micrófono 212 recibe una señal audio. Como se representa en la figura 3, la señal audio recibida puede incluir una señal de extremo próximo que es una señal deseada o "señal primaria". La señal de extremo próximo es la señal que el usuario 102 intenta que reciba el micrófono 212. Sin embargo, la señal audio recibida también incluye una señal de eco resultante de las señales audio emitidas por el altavoz 210 en el paso S404. La señal audio recibida también puede incluir ruido, tal como ruido de fondo. Por lo tanto, la señal audio recibida total $y(t)$ puede darla la suma de la señal de extremo próximo, el eco y el ruido. El eco y el ruido actúan como interferencia para la señal de extremo próximo.

El primer módulo filtro 304 toma como entradas la señal audio salida $x(t)$ y la señal audio recibida $y(t)$. En el paso S408, el primer módulo filtro 304 se usa para modelar el eco en la señal audio recibida $y(t)$. En particular, el primer módulo filtro 304 puede funcionar para modelar el recorrido de eco del eco en la señal audio recibida $y(t)$ usando la señal audio salida $x(t)$ y la señal audio recibida $y(t)$ para determinar una estimación del componente de eco en la señal de extremo próximo $y(t)$

El primer módulo filtro 304 puede utilizar cualquier filtro lineal (por ejemplo, un filtro de respuesta finita al impulso (FIR) o un filtro de respuesta infinita al impulso (IIR)) para modelar el recorrido de eco del eco en la señal audio recibida. Así, el primer módulo filtro 304 es adecuado para cancelación de eco cuando el recorrido de eco es bastante lineal.

El recorrido de eco describe los efectos de los recorridos acústicos por los que pasa la señal de extremo lejano desde el altavoz 210 al micrófono 212. La señal de extremo lejano puede ir directamente desde el altavoz 210 al micrófono 212, o puede ser reflejada por varias superficies en el entorno del terminal de extremo próximo. El recorrido de eco atravesado por la señal de extremo lejano emitida por el altavoz 210 puede ser considerado como un sistema que tiene una frecuencia y una respuesta de fase que pueden variar con el tiempo.

Con el fin de quitar el eco acústico $s(t)$ de la señal $y(t)$ registrada en el micrófono de extremo próximo 212 hay que estimar cómo el recorrido de eco cambia la señal de salida de altavoz de extremo lejano deseada $x(t)$ a un componente de eco indeseado en la señal de entrada.

Para un recorrido de eco aproximadamente lineal, el recorrido de eco $h(t)$ describe cómo el eco en la señal audio recibida $y(t)$ se refiere a la señal audio $x(t)$ salida del altavoz 210, por ejemplo, según la ecuación:

$$s(t) = \sum_{n=0}^{N_{true}} h_n(t)x(t - n),$$

donde $s(t)$ es el eco en la señal audio recibida $y(t)$, N_{true} es un número suficientemente grande para cubrir las partes no despreciables de la respuesta de impulso (teóricamente N_{true} es infinito), y $h_n(t)$ son los coeficientes de la respuesta de impulso que describe el recorrido de eco $h(t)$. El recorrido de eco $h(t)$ puede variar tanto en el tiempo como en la frecuencia y se puede denominar aquí $h(t)$ o $h(t,f)$. El recorrido de eco $h(t)$ puede depender de (i) las condiciones medioambientales actuales que rodean el altavoz 210 y el micrófono 212 (por ejemplo, si hay obstrucciones físicas al paso de la señal audio desde el altavoz 210 al micrófono 212, la presión de aire, la temperatura, el viento, etc), y (ii) características del altavoz 210 y/o el micrófono 212 que pueden alterar la señal enviada y/o recibida.

El módulo filtro 304 modela el recorrido de eco $h(t)$ del eco en la señal audio recibida $y(t)$ determinando una suma ponderada de la corriente y un número finito (N) de valores previos de la señal audio emitida $x(t)$. Por lo tanto, el módulo filtro 304 implementa un filtro de orden enésimo que tiene una longitud finita (en el tiempo) sobre la que considera los valores de la señal audio emitida $x(t)$ al determinar la estimación del recorrido de eco $h(t)$. De esta forma, el módulo filtro 304 adapta dinámicamente la estimación de filtro del recorrido de eco $h(t)$. La operación se describe por la ecuación siguiente, que define el eco en la señal audio recibida $y(t)$ en términos de la señal audio salida $x(t)$:

$$\hat{s}_1(t) = \sum_{n=0}^N \hat{h}_n(t)x(t - n).$$

Por lo tanto, se usan $N+1$ muestras de la señal audio emitida $x(t)$, con unos respectivos pesos $N+1$ $\hat{h}_n(t)$. El conjunto de $N+1$ pesos $\hat{h}_n(t)$ se denomina aquí simplemente la estimación del recorrido de eco $h(t)$. En otros términos, la estimación del recorrido de eco $h(t)$ es un vector que tiene $N+1$ valores donde el módulo filtro 304 implementa un filtro de orden enésimo, tomando en cuenta $N+1$ valores (por ejemplo, $N+1$ cuadros) de la señal $x(t)$.

Se puede apreciar que es más fácil adaptar la estimación de filtro del recorrido de eco $h(t)$ cuando el eco es una parte dominante de la señal audio recibida, que es cuando $y(t) \cong s(t)$. Sin embargo, puede ser posible adaptar la estimación de filtro del recorrido de eco $h(t)$ incluso cuando el eco no es una parte dominante de la señal audio recibida $y(t)$ si el eco $s(t)$ es independiente de los otros componentes de señal de $y(t)$.

Los expertos en la técnica apreciarán que la estimación del recorrido de eco $h(t)$ no tiene que ser calculada explícitamente, sino que se podría representar por medio de coeficientes de filtro obtenidos de algoritmos de gradiente estocástico tales como medias de cuadrados mínimos (LMS), medias de cuadrados mínimos normalizados (NLMS), proyección afin rápida (FAP) y cuadrados mínimos recursivos (RLS).

La estimación del recorrido de eco $h(t)$ se usa para proporcionar coeficientes de filtro que filtran la señal de extremo lejano $x(t)$ para generar una estimación del componente de eco en la señal de extremo próximo $y(t)$ según la estimación del recorrido de eco $h(t)$. Independientemente del algoritmo concreto usado, los coeficientes de filtro del primer módulo filtro 304 son actualizados con cada iteración del algoritmo, de modo que los coeficientes del primer módulo filtro 302 son actualizados continuamente con el tiempo independientemente de las condiciones de señal disponibles.

Aunque la descripción anterior se refiere al uso de un modelo FIR de dominio de tiempo del recorrido de eco para estimar el componente de eco en la señal de extremo próximo $y(t)$, los expertos en la técnica apreciarán que esto es solamente un ejemplo y no es limitativo de ninguna forma. Es decir, el primer módulo filtro 304 puede operar para determinar una estimación del recorrido de eco $h(t)$ y por ello una estimación $\hat{s}_1(t)$ del componente de eco en la señal de extremo próximo $y(t)$ en el dominio de tiempo o en el dominio de frecuencia.

En algunas realizaciones, la estimación del componente de eco es enviada desde el primer módulo filtro 304 al módulo de determinación de rendimiento 312 (como se representa en la figura 3).

En otras realizaciones, en el paso S408, la estimación del componente de eco se pasa del primer módulo filtro 304 a un primer módulo de estimación de potencia (no representado en la figura 3). El primer módulo de estimación de potencia estima la potencia de eco en la señal audio recibida en base a la estimación de filtro (determinada por el primer módulo filtro 304) y la señal de extremo lejano $x(t)$. Hay muchas formas de hacerlo, conocidas por expertos en la técnica, y el alcance de esta descripción no se limita a ningún método concreto de determinar una estimación de potencia de eco. En estas realizaciones, el primer módulo de estimación de potencia está dispuesto para enviar su estimación de potencia de eco correspondiente al módulo de determinación de rendimiento 312.

El módulo de determinación de rendimiento 312 toma como una entrada una primera estimación $\hat{s}_1(t)$. La primera estimación $\hat{s}_1(t)$ puede ser una estimación del componente de eco salido del primer módulo filtro 304 o la potencia del error de estimación ($y-s$). El módulo de determinación de rendimiento 312 también toma como una entrada la señal audio recibida $y(t)$.

En el paso S410, el módulo de determinación de rendimiento 312 mide el rendimiento del primer módulo filtro 304 en base a la primera estimación $\hat{s}_1(t)$ y la señal audio recibida $y(t)$ según una métrica de rendimiento. La medición de rendimiento se usa para detectar la linealidad del sistema. El módulo de determinación de rendimiento 312 puede operar para medir el rendimiento del primer módulo filtro 304 en el dominio de tiempo o en el dominio de frecuencia.

Una métrica de rendimiento usada para medir el rendimiento del primer módulo filtro 304 es la mejora de pérdida de retorno de eco (ERLE), que es una medida de la cantidad de reducción del eco, la métrica ERLE se define como:

$$ERLE_k(t) = \frac{E[y(t)]^2}{E[\hat{s}_k(t) - y(t)]^2}$$

La métrica ERLE se puede medir en decibelios (dB), según la ecuación siguiente (suponiendo que se use un logaritmo de base 10, es decir, \log_{10}):

$$ERLE_k(t) = 10 \log \frac{E[y(t)]^2}{E[\hat{s}_k(t) - y(t)]^2}$$

En las ecuaciones anteriores, $E[\]$ es el operador de expectación. La medición ERLE se puede aplicar, y se aplica típicamente, a señales no estacionarias. Por lo tanto, en la práctica, los valores de expectación son evaluados usando valores medios de tiempo corto:

$$ERLE_k(t) = \frac{\sum_{k=0}^{K-1} y^2(t-k)}{\sum_{k=0}^{K-1} (\hat{s}(t-k) - y(t-k))^2}$$

La métrica ERLE es una medida de la capacidad de la estimación de modelo de modelar la señal de micrófono $y(t)$. La métrica ERLE es limitada en el sentido de que solamente evalúa la exactitud de modelo cuando la señal de micrófono consta principalmente de eco. Si ése no es el caso, la medición ERLE puede ser baja aunque el modelo sea exacto. Sin embargo, si la medición ERLE es alta, puede deberse solamente a que el modelo es exacto. ERLE siempre es más alto cuando se está quitando más eco, independientemente de si se está midiendo en dB o no.

La métrica de rendimiento, distinta de ERLE, puede ser usada para medir el rendimiento del primer módulo filtro 304. Ejemplos de tal otra métrica de rendimiento son la magnitud del error de estimación:

$$\sum_{k=0}^{K-1} (\hat{s}(t-k) - y(t-k))^2$$

medidas ERLE ponderadas tales como:

$$ERLE_k(t) = \frac{\sum_{k=0}^{K-1} w_k y^2(t-k)}{\sum_{k=0}^{K-1} w_k (\hat{s}(t-k) - y(t-k))^2}$$

y medidas de semejanza de señal tales como la correlación cruzada entre $\hat{s}(t)$ y $y(t)$:

$$\frac{E[y(t)\hat{s}(t)]^2}{E[\hat{s}_k(t)]^2 E[y(t)]^2}$$

Independientemente de la métrica de rendimiento usada, la medición de rendimiento en base a la primera estimación $\hat{s}_1(t)$ tomada en el paso S410 puede ser determinada periódicamente. Por ejemplo, las mediciones de rendimiento en base a la primera estimación $\hat{s}_1(t)$ pueden ser promediadas sobre un número predeterminado de muestras de la señal audio $x(t)$ y la señal audio recibida $y(t)$ en un período de tiempo dado para llegar a una medición de rendimiento final. Es decir, el paso S410 puede incluir determinar la medición de rendimiento en base a la primera estimación $\hat{s}_1(t)$ por cada cuadro de la señal audio recibida $y(t)$; sin embargo, esto es simplemente un ejemplo, y la medición de rendimiento en base a la primera estimación $\hat{s}_1(t)$ tomada en el paso S410 puede determinarse menos o más a menudo que por cada cuadro.

En el paso S412, el módulo de determinación de rendimiento 312 determina si el recorrido de eco puede ser considerado lineal (durante un cierto tiempo y frecuencia) en base a la medición de rendimiento tomada en el paso S410.

Es decir, en el paso S412, el módulo de determinación de rendimiento 312 determina si la primera estimación $\hat{s}_1(t)$ es más exacta que una exactitud umbral comparando la medición de rendimiento (tomada en el paso S410 por el módulo de determinación de rendimiento 312) con un valor umbral.

El valor umbral indica una exactitud umbral y el recorrido de eco puede ser considerado lineal cuando la comparación indica que la primera estimación de modelo $\hat{s}_1(t)$ es más exacta que la exactitud umbral. Este valor umbral puede ser un umbral predeterminado (por ejemplo 10dB) o ser dependiente de la señal.

Con respecto a alguna métrica de rendimiento, la medición de rendimiento tomada en el paso S410 aumenta cuando aumenta la exactitud de estimación del primer módulo filtro 304. ERLE es un ejemplo de este tipo de métrica de rendimiento. Cuando se usa este tipo de métrica de rendimiento, cuando la medición de rendimiento es mayor o igual al valor umbral, el proceso pasa al paso S41., y cuando la medición de rendimiento es menor que el valor umbral, el proceso pasa al paso S416.

Con otra métrica de rendimiento, la medición de rendimiento tomada en el paso S410 disminuye cuando aumenta la exactitud de estimación del primer módulo filtro 304. Por ejemplo, si el error de predicción al cuadrado $(y(t)-s(t))^2$ se usa como la métrica de rendimiento. Cuando se usa este tipo de métrica de rendimiento, cuando la medición de rendimiento es menor que el valor umbral, el proceso pasa al paso S414, y cuando la medición de rendimiento es mayor o igual al valor umbral, el proceso pasa al paso S416.

Independientemente de la métrica de rendimiento concreta usada, el proceso pasa al paso S414, cuando el módulo de determinación de rendimiento 312 determina que la linealidad del sistema está a un nivel suficiente de tal manera que el primer módulo filtro 304 proporcione una estimación exacta del recorrido de eco.

En el paso S414, el módulo de supresión de eco 314 usa la primera estimación $\hat{s}_1(t)$ para aplicar supresión de eco a la señal audio recibida $y(t)$, suprimiendo por ello el eco en la señal audio recibida. La supresión de eco realizada en el paso S414 se describe más adelante.

Independientemente de la métrica de rendimiento concreta usada, el proceso pasa al paso S416, cuando el módulo de determinación de rendimiento 312 determina que la linealidad del sistema no está a un nivel suficiente de tal manera que el primer módulo filtro 304 proporcione una estimación exacta del recorrido de eco. En el paso 416, el módulo de determinación de rendimiento 312 no envía la primera estimación $\hat{s}_1(t)$ al módulo de supresión de eco 314. En cambio, en el paso S416, el módulo de determinación de rendimiento 312 envía una señal de control para

habilitar el segundo módulo filtro 308. En respuesta a la recepción de esta señal de control del módulo de determinación de rendimiento 314, el segundo módulo filtro 308 comienza a modelar el recorrido de eco del eco en la señal audio recibida $y(t)$. Por lo tanto, se apreciará que en cualquier punto dado en el tiempo solamente, uno del primer módulo filtro 304 y el segundo módulo filtro 308 puede ser operativo para modelar el recorrido de eco.

El segundo módulo filtro 308 toma como entradas la señal audio salida $x(t)$ y la señal audio recibida $y(t)$. El segundo módulo filtro 308 puede funcionar para modelar el recorrido de eco del eco en la señal audio recibida $y(t)$ usando la señal audio salida $x(t)$ y la señal audio recibida $y(t)$ para determinar una estimación del componente de eco en la señal de extremo próximo $y(t)$ de la misma forma que el primer módulo filtro 304 como se ha descrito anteriormente.

En comparación con el primer módulo filtro 304, el segundo módulo filtro 308 es un modelo más crudo que es menos sensible a no linealidades. Es decir, el segundo módulo filtro 308 es más adecuado para cancelación de eco cuando el recorrido de eco es más bien no lineal.

En algunas realizaciones, la estimación del componente de eco es enviada desde el segundo módulo filtro 308 al módulo de supresión de eco 314 (como se representa en la figura 3). En estas realizaciones, la estimación del componente de eco es suministrada al módulo de supresión de eco 314.

En otras realizaciones, en el paso S416, la estimación del componente de eco es pasada desde el segundo módulo filtro 308 a un segundo módulo de estimación de potencia (no representado en la figura 3). El segundo módulo de estimación de potencia estima la potencia de eco en la señal audio recibida en base a la estimación de filtro (determinada por el segundo módulo filtro 308) y la señal de extremo lejano $x(t)$. Hay muchas formas de hacerlo, conocidas por los expertos en la técnica, y el alcance de esta descripción no se limita a ningún método particular de determinar una estimación de potencia de eco. En estas realizaciones, el segundo módulo de estimación de potencia 308 está dispuesto para enviar su estimación de potencia de eco correspondiente al módulo de supresión de eco 314.

El módulo de supresión de eco 314 toma como una entrada una segunda estimación $\hat{s}_2(t)$. La segunda estimación $\hat{s}_2(t)$ puede ser una estimación del componente de eco salido del segundo módulo filtro 308 o una estimación de potencia de eco salida del segundo módulo de estimación de potencia. En el paso S418, el módulo de supresión de eco 314 usa la segunda estimación $\hat{s}_2(t)$ para aplicar supresión de eco a la señal audio recibida $y(t)$, suprimiendo por ello el eco en la señal audio recibida. La supresión de eco realizada en el paso S418 se describe más adelante.

Después del paso S414, el proceso supervisa continuamente el rendimiento del primer módulo filtro 304 para determinar si seguir aplicando supresión de eco usando la primera estimación $\hat{s}_2(t)$ o conmutar para iniciar el modelado del recorrido de eco del eco en la señal audio recibida $y(t)$ usando el segundo módulo filtro 308 y usar la estimación $\hat{s}_2(t)$ para aplicar supresión de eco a la señal audio recibida $y(t)$.

Igualmente, después del paso S418, el proceso supervisa continuamente el rendimiento del primer módulo filtro 304 para determinar si seguir aplicando supresión de eco usando la segunda estimación $\hat{s}_2(t)$ o conmutar a usar la primera estimación $\hat{s}_1(t)$ para aplicar supresión de eco a la señal audio recibida $y(t)$.

En las realizaciones descritas anteriormente, la funcionalidad de extracción de eco se basa en la segunda estimación $\hat{s}_2(t)$ para aplicar supresión de eco a la señal audio recibida $y(t)$ a no ser que la comparación de la medición de rendimiento tomada en el paso S410 con el valor umbral indique que la primera estimación $\hat{s}_1(t)$ es más exacta que la exactitud umbral.

En las realizaciones descritas anteriormente, la funcionalidad de extracción de eco conmuta de depender del primer módulo filtro 304 a depender del segundo módulo filtro 308 tan pronto como la comparación de la medición de rendimiento tomada en el paso S410 con el valor umbral indica que la primera estimación $\hat{s}_1(t)$ es menos exacta que la exactitud umbral. En realizaciones alternativas, la conmutación de depender del primer módulo filtro 304 a depender del segundo módulo filtro 308 solamente tiene lugar cuando las comparaciones de la medición de rendimiento tomada en el paso S410 con el valor umbral indican que la primera estimación $\hat{s}_1(t)$ ha sido menos exacta que la exactitud umbral en un período de tiempo predeterminado, es decir, durante todo el período de tiempo predeterminado.

En las realizaciones descritas anteriormente, la funcionalidad de extracción de eco conmuta de depender del segundo módulo filtro 308 a depender del primer módulo filtro 304 tan pronto como la comparación de la medición de rendimiento tomada en el paso S410 con el valor umbral indica que la primera estimación $\hat{s}_1(t)$ es más exacta que la exactitud umbral. En realizaciones alternativas, la conmutación de depender del segundo módulo filtro 308 a depender del primer módulo filtro 304 solamente tiene lugar cuando las comparaciones de la medición de rendimiento tomadas en el paso S410 con el valor umbral indican que la primera estimación $\hat{s}_1(t)$ ha sido más exacta que la exactitud umbral en un período de tiempo predeterminado, es decir, durante todo el período de tiempo predeterminado.

El esquema descrito anteriormente puede ampliarse a implementarse para bandas secundarias de frecuencia

separadas dentro de un rango de frecuencia. Es decir, la señal audio salida $x(t)$ y la señal audio recibida $y(t)$ procesadas por el módulo de modelado 302 se dividen en una pluralidad de bandas secundarias de frecuencia dentro de un rango de frecuencia, y la medición de rendimiento descrita anteriormente se implementa en base a banda secundaria.

5 Por ejemplo, durante un período de tiempo dado, por cada banda secundaria de frecuencia, el primer módulo filtro 304 modela el recorrido de eco del eco en la señal audio recibida $y(t)$ usando la señal audio salida $x(t)$ y la señal audio recibida $y(t)$ para determinar una estimación del componente de eco en la señal de extremo próximo $y(t)$. El módulo de determinación de rendimiento 312 hace una medición del rendimiento del primer módulo filtro 304 en base a la primera estimación $\hat{s}_1(t)$ y la señal audio recibida $y(t)$ según una métrica de rendimiento concreta para cada banda secundaria de frecuencia. Cada una de estas mediciones de rendimiento se compara con el valor umbral para determinar si el recorrido de eco en la banda secundaria de frecuencia respectiva puede ser considerado lineal (para un cierto tiempo y frecuencia). El módulo de determinación de rendimiento 312 puede determinar entonces si el recorrido de eco en el período de tiempo dado puede ser considerado lineal en base al número de bandas secundarias de frecuencia (que están dentro de un cierto rango de frecuencia) en el que el recorrido de eco es considerado lineal (y así inherentemente en el número de bandas secundarias de frecuencia en el que el recorrido de eco se considera no lineal). Por ejemplo, el módulo de determinación de rendimiento 312 puede determinar que el recorrido de eco en el período de tiempo dado se considera lineal si el número de bandas secundarias de frecuencia de menos de 4 kHz en el que el recorrido de eco se considera lineal es mayor que el número de bandas secundarias de frecuencia en el que el recorrido de eco se considera no lineal, es decir, la mayoría de las bandas de frecuencia por debajo de 4 kHz se considera lineal. En otras implementaciones, el módulo de determinación de rendimiento 312 puede determinar solamente que el recorrido de eco en el período de tiempo dado se considera lineal si una cierta proporción de las bandas de bandas secundarias de frecuencia (están son dentro de un cierto rango de frecuencia) se considera lineal. Por ejemplo, el módulo de determinación de rendimiento 312 puede determinar que el recorrido de eco en el período de tiempo dado se considera lineal si 75% de las bandas de bandas secundarias de frecuencia (que están dentro de un cierto rango de frecuencia) se considera lineal. Se apreciará que estos valores ejemplares se usan simplemente para ilustrar los conceptos y no se ha previsto que sean limitativos de ninguna forma.

30 En otras realizaciones en las que el esquema descrito anteriormente se amplía a implementarse para bandas de frecuencia separadas dentro de un rango de frecuencia, una vez que el módulo de determinación de rendimiento 312 ha determinado si el recorrido de eco en la respectiva banda secundaria de frecuencia puede ser considerado lineal (para un cierto tiempo y frecuencia), el módulo de determinación de rendimiento 312 toma la decisión sobre si aplicar supresión de eco a la señal audio recibida $y(t)$ usando la primera estimación de modelo $\hat{s}_1(t)$, o controlar el segundo módulo filtro 308 para modelar el recorrido de eco del eco en la señal audio recibida $y(t)$ y usar la segunda estimación $\hat{s}_1(t)$ para aplicar supresión de eco a la señal audio recibida $y(t)$, en base por banda secundaria de frecuencia. Así, puede surgir un posible escenario en el que, durante un período de tiempo dado, el primer módulo filtro 304 se usa para modelar el recorrido de eco del eco en la señal audio recibida $y(t)$ para bandas de frecuencia más bajas dentro del rango de frecuencia y el segundo módulo filtro 308 se usa para modelar el recorrido de eco del eco en la señal audio recibida $y(t)$ para bandas de frecuencia más altas dentro del rango de frecuencia.

Ahora se describe la supresión de eco realizada en los pasos S414 y S418.

45 La finalidad del supresor de eco es suprimir el eco de altavoz presente en la señal de micrófono, por ejemplo, en un cliente VoIP, a un nivel suficientemente bajo para no ser apreciable/molesto en presencia de los sonidos de extremo próximo (sonidos no eco) captados por el micrófono 212. Para poder elegir la cantidad apropiada de supresión de eco, se necesita un modelo exacto del recorrido de eco, y como se ha descrito anteriormente, éste se obtiene modelando el recorrido de eco usando uno de dos modelos, por lo que se usa una medición de rendimiento de un modelo respectivo como un detector para cuándo conmutar entre los dos modelos para asegurar que se use el modelo más adecuado para modelar el recorrido de eco. El módulo de supresión de eco 314 está diseñado para aplicar supresión dependiente de señal que varía en el tiempo y la frecuencia con respecto a la señal audio recibida $y(t)$. Los métodos de supresión de eco son conocidos en la técnica. Además, el método de supresión de eco aplicado por el módulo de supresión de eco 314 puede implementarse en formas diferentes. Por lo tanto, como tales, los detalles exactos del método de supresión de eco no se describen aquí en detalle.

55 El módulo de supresión de eco 314 envía la señal recibida, con el eco que ha sido suprimido, para procesamiento adicional en el dispositivo de usuario 104. Por ejemplo, la señal salida del módulo de supresión de eco 314 puede ser procesada por el cliente 206 (por ejemplo, codificada y paquetizada) y luego transmitida por la red 106 al dispositivo de usuario 110 en una llamada entre los usuarios 102 y 108. Además o alternativamente, la señal salida del módulo de supresión de eco 314 puede ser usada para otros fines por el dispositivo de usuario 104, por ejemplo, la señal puede ser almacenada en la memoria 214 o usada como una entrada a una aplicación que se ejecute en el dispositivo de usuario 104.

65 Como se ha descrito anteriormente, el primer módulo filtro 304 es actualizado continuamente independientemente de las condiciones de señal disponibles. Se puede usar opcionalmente un esquema de ajuste de tamaño de paso en relación al primer módulo filtro 304 en las realizaciones descritas anteriormente.

Como se ha descrito anteriormente, los coeficientes de filtro para el primer módulo filtro 304 pueden obtenerse ejecutando un algoritmo de gradiente estocástico. En particular el primer módulo filtro 304 ejecuta un algoritmo de gradiente estocástico para identificar los coeficientes del módulo filtro 304 que minimiza una señal de error $e(t)$.

5 Los coeficientes de filtro actualizados para el módulo filtro 304 son generados en respuesta a la señal de error $e(t)$, la señal de entrada $x(t)$ y los coeficientes de filtro previos.

10 El algoritmo de gradiente estocástico opera de manera recursiva en el tiempo. Esto significa que no se adapta de forma instantánea a cambios en el sistema, en cambio el algoritmo converge iterativamente a una aproximación del sistema sobre un intervalo de tiempo finito.

15 Los coeficientes de filtro del primer módulo filtro 304 filtran la señal de extremo lejano $x(t)$ para generar una estimación del componente de eco en la señal de extremo próximo $y(t)$. La señal de error $e(t)$ es obtenida por un sustractor (no representado en la figura 3) que resta la primera estimación de módulo filtro del componente de eco en la señal de extremo próximo $y(t)$ de la señal de extremo próximo $y(t)$ y suministra la señal de error $e(t)$ al primer módulo filtro 304. Se apreciará que es deseable que la señal de error $e(t)$ sea pequeña. Por ejemplo, cuando no hay señal de extremo próximo $v(t)$ en la señal de micrófono, idealmente la señal de error es igual a cero.

20 Los algoritmos de gradiente estocástico tienen un parámetro de convergencia en forma de un tamaño de paso para la actualización de los parámetros de modelo. Éste se puede elegir como fijo en algunas aplicaciones, pero en muchos casos se logra un mejor rendimiento si se elige en dependencia de la señal. El tamaño de paso controla la sensibilidad de la actualización al ruido en la señal de micrófono $y(t)$. Si se elige de modo que sea pequeño, la velocidad de actualización es lenta, pero es menos insensible al ruido, pero si se elige de modo que sea grande la velocidad de actualización es en cambio rápida, pero más sensible al ruido. La referencia a "velocidad de actualización" o "velocidad de adaptación" se usa aquí para hacer referencia a la rapidez con la que el modelo es capaz de adaptarse a las condiciones de señal disponibles en el sistema. Es decir, la utilización de un tamaño de paso más pequeño dará lugar a una señal de error eventual más pequeña $e(t)$, sin embargo, la convergencia a una aproximación del sistema será más lenta debido al mayor número de pasos de iteración requeridos (menor tasa de convergencia), y el uso de un tamaño de paso más grande dará lugar a una señal de error eventual más grande $e(t)$, sin embargo, la convergencia a una aproximación del sistema será más rápida debido al menor número de pasos de iteración requeridos (tasa de convergencia más rápida).

35 Con el fin de lograr estimaciones de exactitud muy alta, el tamaño de paso tiene que ser pequeño con el fin de evitar el rebasamiento de las verdaderas estimaciones debido a un tamaño de paso demasiado alto.

40 En el esquema de ajuste de tamaño de paso, la exactitud de la estimación $\hat{s}_1(t)$ se determina según una medición de mejora de pérdida de retorno de eco. Esta medición de mejora de pérdida de retorno de eco puede ser la misma medición de exactitud hecha por el módulo de determinación de rendimiento 312 en el paso S410. Alternativamente esta medición de mejora de pérdida de retorno de eco puede ser una medición separada de la medición de exactitud hecha por el módulo de determinación de rendimiento 312 cuando se usan en el paso S410 medidas distintas de ERLE.

45 Un módulo de selección de parámetro de convergencia (no representado en la figura 3) determina un valor para un parámetro de convergencia (tamaño de paso) usado en el algoritmo ejecutado en el módulo filtro 304 en base a la exactitud de la estimación $\hat{s}_1(t)$. En particular, el módulo de selección de parámetro de convergencia selecciona el parámetro de convergencia para controlar la velocidad de adaptación del módulo filtro 304 como una función no creciente de la medición de mejora de pérdida de retorno de eco.

50 La medición de mejora de pérdida de retorno de eco puede compararse con un valor umbral, y el módulo de selección de parámetro de convergencia regula el parámetro de convergencia en base a esta comparación.

55 Dado que la medición de mejora de pérdida de retorno de eco tiene la propiedad de que la exactitud de modelo siempre es alta cuando la medición de mejora de pérdida de retorno de eco es alta, se puede usar para ralentizar la velocidad de adaptación cuando la medición de mejora de pérdida de retorno de eco es alta (es decir, más alta que el valor umbral predeterminado) con el fin de lograr estimaciones cada vez más exactas, y aumentar la velocidad de adaptación cuando la medición de mejora de pérdida de retorno de eco es baja (es decir, inferior al valor umbral predeterminado) con el fin de rastrear rápidamente cambios en los parámetros de modelo.

60 El esquema de ajuste de tamaño de paso asegura que se logre una rápida adaptación cuando la exactitud del modelo sea desconocida (mediante la alta velocidad de actualización cuando la medición de la mejora de pérdida de retorno de eco sea baja), y que se logren estimaciones cada vez más exactas cuando se sepa que el modelo es exacto (mediante la disminución de la velocidad de actualización cuando la medición de la mejora de pérdida de retorno de eco sea alta).

65 En las realizaciones descritas anteriormente, la extracción de eco se implementa en un sistema VoIP (por ejemplo,

la señal audio recibida puede incluir voz del usuario 102 para transmisión al dispositivo de usuario 110 durante una llamada entre los usuarios 102 y 108 por el sistema de comunicaciones 100). Sin embargo, los métodos de extracción de eco aquí descritos pueden aplicarse en cualquier sistema adecuado en el que se haya de aplicar extracción de eco.

5 En las realizaciones descritas anteriormente, y representadas en las figuras, el módulo de extracción de eco 314 implementa supresión de eco.

10 En las realizaciones descritas anteriormente, y representadas en las figuras, no se aplica cancelación de eco (o "sustracción de eco") a la señal audio recibida $y(t)$. Es decir, no hay módulo de cancelación de eco en el dispositivo de usuario 104 y la supresión de eco se aplica a la señal audio recibida $y(t)$ sin un paso anterior de aplicar cancelación de eco a la señal audio recibida $y(t)$.

15 Sin embargo, en otras realizaciones, la cancelación de eco puede ser aplicada, por un módulo de cancelación de eco, a la señal audio recibida $y(t)$. En particular, la supresión de eco aplicada por el módulo de supresión de eco 314 puede ser aplicada hacia abajo (es decir, después) de la cancelación de eco en el procesado de la señal audio recibida $y(t)$. El módulo de cancelación de eco restaría una estimación de la señal de eco de la señal audio recibida, pero debido a inexactitudes en la estimación de la señal de eco, quedaría muy probablemente un eco residual en la señal audio recibida. El eco residual es el que sería suprimido entonces por el módulo de supresión de eco 314.
20 Esta supresión de eco podría aplicarse de la misma forma que la aquí descrita en las realizaciones en las que no se aplica cancelación de eco. Si se usa sustracción de eco, su efecto puede tomarse en cuenta en la supresión de eco.

25 En otras realizaciones, el módulo de extracción de eco 314 implementa cancelación de eco. Es decir, el módulo de extracción de eco 314 está dispuesto para restar una estimación de la señal de eco $\hat{s}_1(t)$ o $\hat{s}_2(t)$ de la señal audio recibida $y(t)$.

30 Los métodos aquí descritos pueden implementarse ejecutando un producto de programa de ordenador (por ejemplo, el cliente 206) en el dispositivo de usuario 104. Es decir, un producto de programa de ordenador puede estar configurado para quitar eco en la señal audio recibida $y(t)$, donde el producto de programa de ordenador se realiza en un medio de almacenamiento legible por ordenador (por ejemplo, almacenado en la memoria 214) y configurado de modo que, cuando sea ejecutado en la CPU 202, realice las operaciones de alguno de los métodos descritos aquí.

35 En general, cualquiera de las funciones aquí descritas (por ejemplo, los módulos funcionales representados en la figura 3 y los pasos funcionales representados en la figura 4) pueden implementarse usando software, microprogramas, hardware (por ejemplo, circuitería lógica fija), o una combinación de estas implementaciones. Los módulos y pasos representados por separado en las figuras 3 y 4 y referidos anteriormente pueden implementarse o no como módulos o pasos separados. Por ejemplo, el módulo de supresión de eco 314 puede realizar la función del módulo de determinación de rendimiento 312. Los términos "módulo", "funcionalidad", "componente" y "lógica" en el
40 sentido en que se usan aquí indican en general software, microprogramas, hardware, o su combinación. En el caso de una implementación de software, el módulo, la funcionalidad o la lógica representa un código de programa que realiza tareas especificadas cuando se ejecuta en un procesador (por ejemplo, CPU o CPUs). El código de programa puede almacenarse en uno o varios dispositivos de memoria legibles por ordenador. Las características de las técnicas aquí descritas son independientes de plataforma, lo que quiere decir que las técnicas pueden
45 implementarse en una variedad de plataformas informáticas comerciales que tengan una variedad de procesadores.

50 Por ejemplo, los dispositivos de usuario también pueden incluir una entidad (por ejemplo, software) que haga que el hardware de los dispositivos de usuario realice operaciones, por ejemplo, bloques funcionales de procesador, etc. Por ejemplo, los dispositivos de usuario pueden incluir un medio legible por ordenador que puede estar configurado para mantener instrucciones que hagan que los dispositivos de usuario, y más en concreto el sistema operativo y el hardware asociado de los dispositivos de usuario, realice operaciones. Así, las instrucciones sirven para configurar el sistema operativo y el hardware asociado para realizar las operaciones y de esta forma dan lugar a transformación del sistema operativo y el hardware asociado para realizar funciones. Las instrucciones pueden ser
55 suministradas por el medio legible por ordenador a los dispositivos de usuario a través de varias configuraciones diferentes.

Una configuración de un medio legible por ordenador es un medio de soporte de señal y por ello está configurado para transmitir las instrucciones (por ejemplo, como una onda portadora) al dispositivo informático, tal como mediante una red. El medio legible por ordenador también puede estar configurado como un medio de
60 almacenamiento legible por ordenador y por ello no es un medio de soporte de señal. Los ejemplos de un medio de almacenamiento legible por ordenador incluyen una memoria de acceso aleatorio (RAM), una memoria de lectura solamente (ROM), un disco óptico, memoria flash, memoria de disco duro, y otros dispositivos de memoria que pueden usar técnicas magnéticas, ópticas y otras para almacenar instrucciones y otros datos.

65 Aunque la materia se ha descrito en lenguaje específico de características estructurales y/o hechos metodológicos, se ha de entender que la materia definida en las reivindicaciones anexas no se limita necesariamente a las

características específicas o hechos descritos anteriormente. Más bien, las características específicas y los hechos descritos anteriormente se describen como formas ejemplares de implementar las reivindicaciones.

REIVINDICACIONES

1. Un método de quitar eco, incluyendo el método:
- 5 enviar una señal audio (S404);
- recibir una señal audio (S406), donde la señal audio recibida incluye eco resultante de dicha señal audio emitida;
- 10 modelar un recorrido de eco del eco en la señal audio recibida usando un primer modelo (304) en base a la señal audio salida y la señal audio recibida para determinar (S408) una primera estimación de modelo del eco en la señal audio recibida;
- usar la primera estimación de modelo para determinar (S410) un primer valor de rendimiento según una métrica de rendimiento;
- 15 comparar (S412) el primer valor de rendimiento con un valor umbral;
- caracterizado porque** el método incluye;
- 20 determinar si el recorrido de eco puede ser considerado lineal en base a dicha comparación y en base a dicha comparación selectivamente:
- usar (S414) la primera estimación de modelo del eco para quitar el eco en la señal audio recibida, o
- 25 modelar el recorrido de eco del eco en la señal audio recibida usando un segundo modelo (308) en base a la señal audio salida y la señal audio recibida para determinar (S416) una segunda estimación de modelo del eco,
- y usar (S418) la segunda estimación de modelo del eco para quitar el eco en la señal audio recibida.
- 30 2. El método de la reivindicación 1, donde el método incluye:
- usar la primera estimación de modelo del eco para quitar el eco en la señal audio recibida si se determina que el recorrido de eco puede ser considerado lineal; y
- 35 modelar el recorrido de eco del eco en la señal audio recibida usando un segundo modelo en base a la señal audio salida y la señal audio recibida para determinar una segunda estimación de modelo del eco, y usando la segunda estimación de modelo del eco para quitar la supresión de eco en la señal audio recibida, si se determina que el recorrido de eco no puede ser considerado lineal.
- 40 3. El método de la reivindicación 1 o 2, donde el valor umbral indica una exactitud umbral y el recorrido de eco puede ser considerado lineal cuando dicha comparación indica que la primera estimación de modelo es más exacta que la exactitud umbral.
- 45 4. El método de cualquier reivindicación precedente, donde la métrica de rendimiento es una métrica de mejora de pérdida de retorno de eco.
5. El método de cualquier reivindicación precedente, donde el método incluye:
- 50 dividir la señal audio salida y la señal audio recibida en un período de tiempo predeterminado en una pluralidad de bandas secundarias de frecuencia dentro de un rango de frecuencia e implementar el método en base por banda secundaria de frecuencia para determinar si el recorrido de eco en cada banda secundaria de frecuencia respectiva dentro de dicho rango de frecuencia puede ser considerado lineal;
- determinar si el recorrido de eco puede ser considerado lineal en dicho período de tiempo predeterminado en base al número de bandas secundarias de frecuencia consideradas lineales dentro de dicho rango de frecuencia; y
- 55 usar la primera estimación de modelo del eco para quitar el eco en la señal audio recibida si el recorrido de eco puede ser considerado lineal en dicho período de tiempo predeterminado, o
- 60 modelar el recorrido de eco del eco en la señal audio recibida usando un segundo modelo si el recorrido de eco no puede ser considerado lineal en dicho período de tiempo predeterminado, y usar la segunda estimación de modelo del eco para quitar el eco en la señal audio recibida.
- 65 6. El método de cualquiera de las reivindicaciones 1 a 4, donde el método incluye:
- dividir la señal audio salida y la señal audio recibida en un período de tiempo predeterminado en una pluralidad de

bandas secundarias de frecuencia dentro de un rango de frecuencia e implementar el método en base por banda secundaria de frecuencia para determinar si el recorrido de eco en cada banda secundaria de frecuencia respectiva dentro de dicho rango de frecuencia puede ser considerado lineal; y para cada banda secundaria de frecuencia:

5 usar la primera estimación de modelo del eco para quitar el eco en la señal audio recibida si el recorrido de eco puede ser considerado lineal; o

10 modelar el recorrido de eco del eco en la señal audio recibida usando el segundo modelo si el recorrido de eco no puede ser considerado lineal, y usar la segunda estimación de modelo del eco para quitar el eco en la señal audio recibida.

7. El método de cualquier reivindicación precedente, donde:

15 el primer modelo ejecuta un primer algoritmo para determinar coeficientes de filtro y usa dichos coeficientes de filtro para filtrar la señal audio salida para determinar la primera estimación de modelo del eco; y el segundo modelo ejecuta un segundo algoritmo para determinar coeficientes de filtro y usa dichos coeficientes de filtro para filtrar la señal audio salida para determinar la segunda estimación de modelo del eco.

20 8. El método de la reivindicación 7, donde el primer algoritmo incluye un parámetro de convergencia y la métrica de rendimiento es una medición de mejora de pérdida de retorno de eco, incluyendo el método además:

25 actualizar dicho parámetro de convergencia en base a dicho primer valor de rendimiento, donde el parámetro de convergencia se selecciona para controlar una velocidad de adaptación del primer modelo como una función no creciente del primer valor de rendimiento.

9. Un dispositivo (104) incluyendo:

un aparato de salida audio (210) configurado para la salida de una señal audio;

30 un aparato de entrada audio (212) configurado para recibir una señal audio, donde la señal audio recibida incluye un eco resultante de dicha señal audio salida;

35 un módulo de modelado (302) configurado para modelar un recorrido de eco del eco en la señal audio recibida usando un primer modelo en base a la señal audio salida y la señal audio recibida para determinar una primera estimación de modelo del eco en la señal audio recibida, usar la primera estimación de modelo para determinar un primer valor de rendimiento según una métrica de rendimiento, comparar el primer valor de rendimiento con un valor umbral y determinar si el recorrido de eco puede ser considerado lineal en base a dicha comparación,

40 donde el módulo de modelado está configurado además para modelar selectivamente el recorrido de eco del eco en la señal audio recibida usando un segundo modelo en base a la señal audio salida y la señal audio recibida, en base a dicha comparación, para determinar una segunda estimación de modelo del eco; y

un módulo de extracción de eco (314);

45 **caracterizado porque** el módulo de extracción de eco está configurado para usar selectivamente la primera estimación de modelo o la segunda estimación de modelo del eco para quitar el eco en la señal audio recibida en base a dicha comparación.

50 10. Un producto de programa de ordenador configurado para quitar eco en una señal audio recibida, dicho eco resultante de una señal audio salida, realizándose el producto de programa de ordenador en un medio de almacenamiento legible por ordenador y configurado, cuando sea ejecutado en un procesador (202), para:

55 modelar un recorrido de eco del eco en la señal audio recibida usando un primer modelo (304) en base a la señal audio salida y la señal audio recibida para determinar (S408) una primera estimación de modelo del eco en la señal audio recibida;

usar la primera estimación de modelo para determinar (S410) un primer valor de rendimiento según una métrica de rendimiento;

60 comparar (S412) el primer valor de rendimiento con un valor umbral;

caracterizado por determinar si el recorrido de eco puede ser considerado lineal en base a dicha comparación, y en base a dicha comparación selectivamente:

65 usar (S414) la primera estimación de modelo del eco para quitar eco en la señal audio recibida; o

modelar el recorrido de eco del eco en la señal audio recibida usando un segundo modelo (308) en base a la señal audio salida y la señal audio recibida para determinar (S416) una segunda estimación de modelo del eco, y usar (S418) la segunda estimación de modelo del eco para quitar el eco en la señal audio recibida.

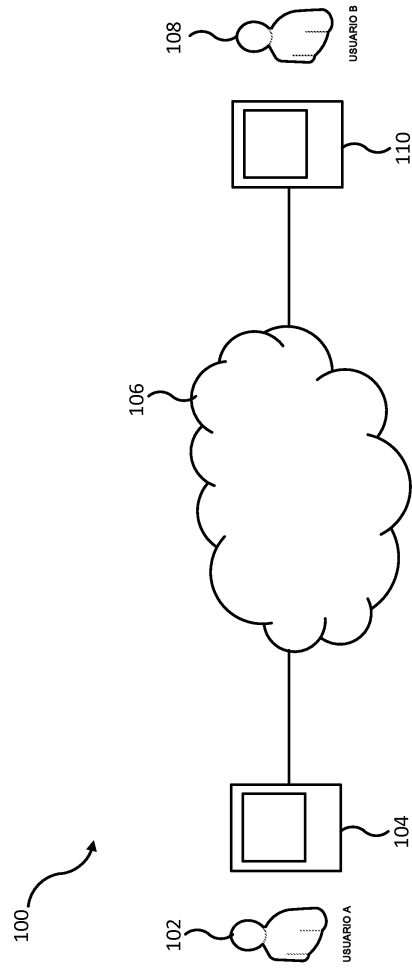


FIG. 1

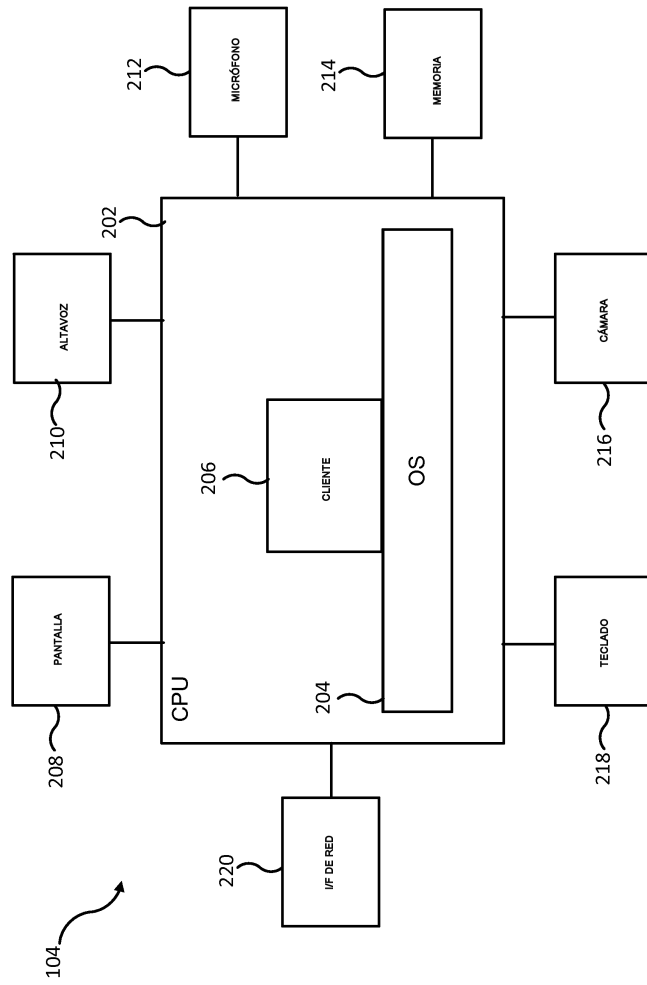


FIG. 2

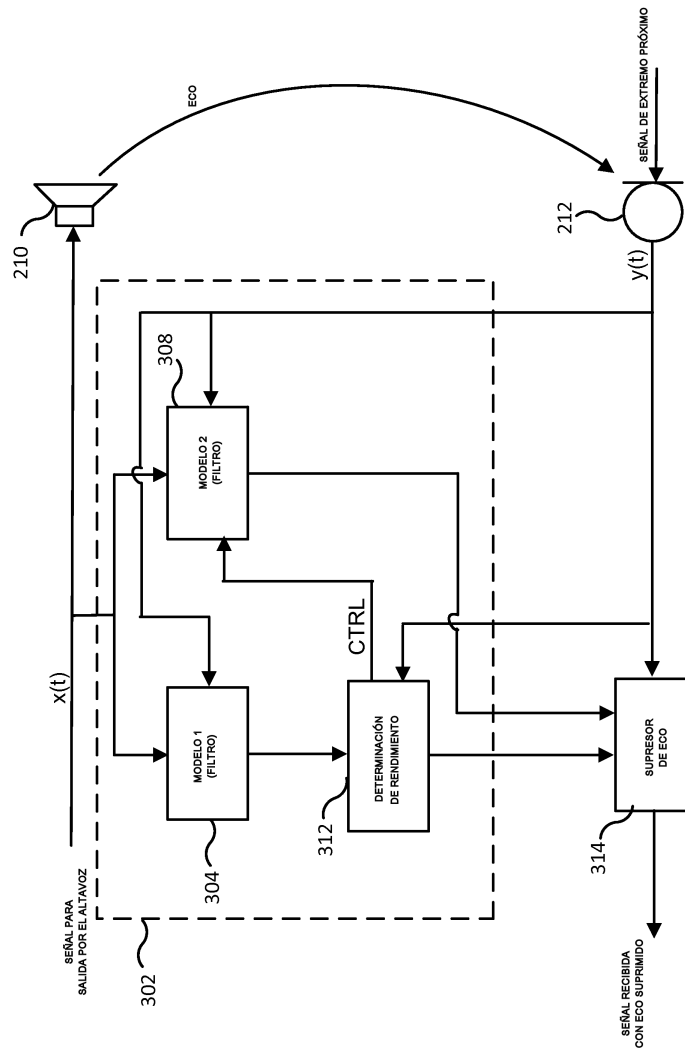


FIG. 3

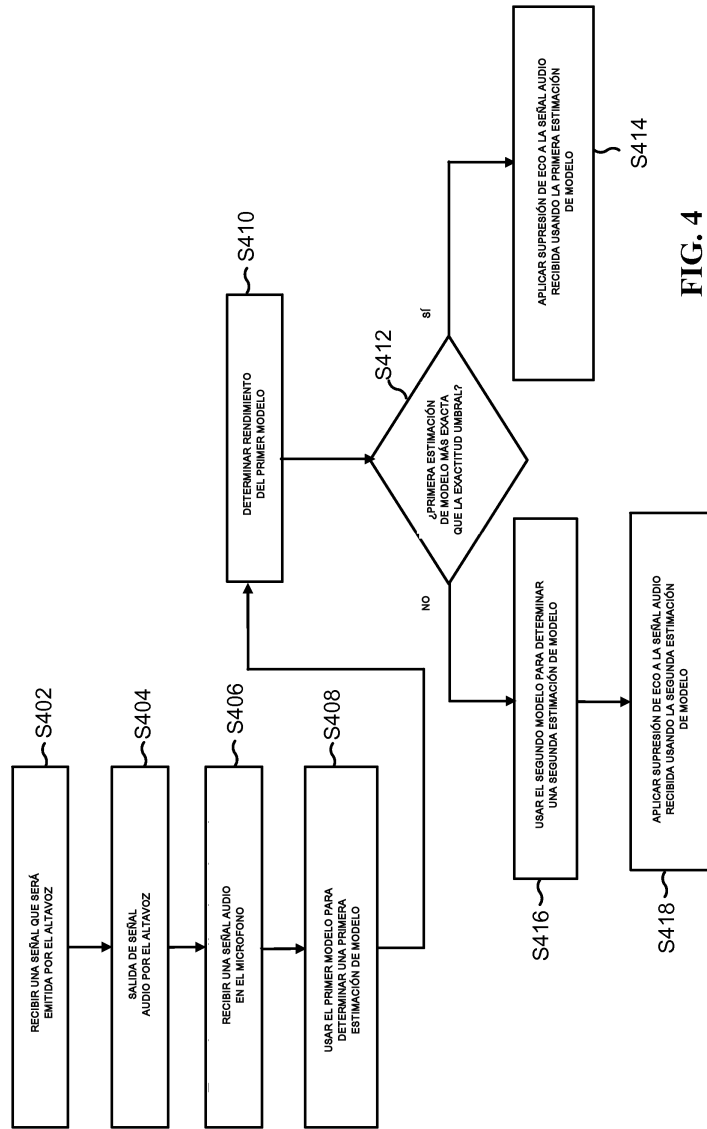


FIG. 4