

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 991 393**

51 Int. Cl.:

**H04M 9/08**

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.03.2021** **PCT/US2021/023059**

87 Fecha y número de publicación internacional: **30.09.2021** **WO21194859**

96 Fecha de presentación y número de la solicitud europea: **19.03.2021** **E 21719331 (7)**

97 Fecha y número de publicación de la concesión europea: **04.09.2024** **EP 4128732**

54 Título: **Supresión de residuo de eco**

30 Prioridad:

**23.03.2020 US 202062993149 P**

**25.02.2021 US 202163153523 P**

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

**03.12.2024**

73 Titular/es:

**DOLBY LABORATORIES LICENSING  
CORPORATION (100.0%)**

**1275 Market Street  
San Francisco, CA 94103, US**

72 Inventor/es:

**WANG, NING**

74 Agente/Representante:

**LINAGE GONZÁLEZ, Rafael**

ES 2 991 393 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

Supresión de residuo de eco

5 **Campo**

La presente divulgación se refiere al procesamiento de audio y, en particular, a la supresión de eco.

**Antecedentes**

10

A menos que se indique lo contrario en el presente documento, los enfoques descritos en esta sección no constituyen técnica anterior a las reivindicaciones de esta solicitud y no se admiten como técnica anterior mediante su inclusión en esta sección.

15

Un dispositivo de telecomunicaciones, como un sistema de audioconferencia, generalmente incluye un altavoz y un micrófono. Las dos partes de una comunicación pueden denominarse la parte de extremo cercano y la parte de extremo lejano. La parte de extremo cercano está próxima a un primer dispositivo de telecomunicaciones, y la parte de extremo lejano está en una ubicación diferente a la de la parte de extremo cercano y se comunica usando un segundo dispositivo de telecomunicaciones a través de una red de telecomunicaciones cableada o inalámbrica. El micrófono del dispositivo de extremo cercano captura no solo la locución de la parte de extremo cercano, sino que también puede capturar la locución de la parte de extremo lejano que se ha emitido desde el altavoz de extremo cercano. La salida del altavoz que capta el micrófono generalmente se denomina eco. El dispositivo de telecomunicaciones de extremo cercano generalmente incluye un sistema de gestión de eco para reducir el eco antes de transmitir el audio capturado en el extremo cercano al extremo lejano.

25

El sistema de gestión de eco generalmente incluye un sistema de cancelación de eco seguido de un sistema de supresión de eco. El sistema de cancelación de eco es un sistema lineal y típicamente puede incluir un filtro adaptativo. El eco que queda en la señal de audio después de la cancelación de eco generalmente se denomina "residuo de eco". El sistema de supresión de eco es un sistema no lineal que aplica una atenuación adicional a la señal de audio con el fin de reducir el residuo de eco. La atenuación realizada por el sistema de supresión de eco se denomina no lineal, ya que puede aplicar diferentes ganancias a diferentes frecuencias o bandas de frecuencia.

30

35

Los dispositivos de telecomunicaciones generalmente implementan aproximadamente un agregado de 50 dB reducción de eco basándose en las contribuciones de la pérdida de retorno de eco físico (por ejemplo, debido a la distancia entre el altavoz y el micrófono), la cancelación de eco y la supresión de eco.

El documento EP1978649 divulga un método de cancelación de eco no lineal de dominio espectral en un dispositivo de manos libres.

40

CHEN K et. al "An improved post-filter of acoustic echo canceller based on subband implementation", vol. 70, n.º 6, 1 de junio de 2009, divulga dos versiones modificadas alternativas del postfiltro para reducir los efectos de la distorsión de la locución de extremo cercano, lo que garantiza la calidad de las comunicaciones habladas. Los postfiltros propuestos se combinan perfectamente con el sistema AEC basado en subbandas con una carga computacional bastante pequeña.

45

**Sumario**

50

Una cuestión con los sistemas de gestión de eco existentes es que los ordenadores portátiles se usan cada vez más para telecomunicaciones, como audioconferencias o videoconferencias. En un ordenador portátil, el altavoz y el micrófono están muy cerca, lo que hace difícil alcanzar el objetivo de 50 dB de reducción de eco. Teniendo en cuenta lo anterior, existe la necesidad de mejorar el sistema de supresión de eco con el fin de reducir el residuo de eco.

55

De acuerdo con la invención, se proporciona un método para proporcionar un método implementado por ordenador de procesamiento de audio como se define en la reivindicación 1, un medio no transitorio legible por ordenador que almacena un programa de ordenador como se define en la reivindicación 10, y un aparato para procesamiento de audio como se define en la reivindicación 11.

60

La siguiente descripción detallada y dibujos que se acompañan proporcionan una mejor comprensión de la naturaleza y las ventajas de diversas implementaciones.

**Breve descripción de los dibujos**

65

La figura 1 es un diagrama de bloques de un sistema 100 de gestión de eco.

La figura 2 es un diagrama de bloques del cancelador 201 de eco (véase la figura 1).

La figura 3 es un diagrama de bloques del supresor 301 de eco (véase la figura 1).

5 La figura 4 es un diagrama de bloques de un calculador 400 de potencia de retorno en bucle.

La figura 5 es un histograma de la relación entre la potencia de retorno en bucle  $\tilde{X}(b, t')$  y la función de transferencia de error (residuo de eco)  $\tilde{H}(b, t')$  del filtro adaptativo.

10 La figura 6 es un gráfico que ilustra la función de densidad de probabilidad  $\Pr(x | E^c)$ .

La figura 7 es una arquitectura de dispositivo móvil 700 para implementar las características y procesos descritos en el presente documento, de acuerdo con una realización.

15 La figura 8 es un diagrama de flujo de un método 800 de procesamiento de audio.

### Descripción detallada

20 Se describen en el presente documento técnicas relacionadas con la supresión de eco. En la siguiente descripción, a modo explicativo, se exponen numerosos ejemplos y detalles específicos con el fin de proporcionar una comprensión completa de la presente divulgación. Será evidente, sin embargo, para un experto en la técnica que la presente divulgación tal como se define en las reivindicaciones puede incluir algunas o todas las características de estos ejemplos solos o en combinación con otras características descritas a continuación, y puede incluir además modificaciones y equivalentes de las características y conceptos descritos en el presente documento.

25 En la siguiente descripción se detallan diversos métodos, procesos y procedimientos. Aunque los pasos particulares pueden describirse en un cierto orden, tal orden es principalmente para mayor comodidad y claridad. Un paso particular puede repetirse más de una vez, puede ocurrir antes o después de otros pasos (incluso si esos pasos se describen, por el contrario, en otro orden) y puede ocurrir en paralelo con otros pasos. Se requiere un segundo paso después del primero solo cuando el primer paso debe completarse antes de comenzar el segundo. Tal situación se señalará específicamente cuando no quede clara por el contexto.

30 En este documento se usan los términos “y”, “o” y “y/o”. Estos términos deben interpretarse como si tuvieran un significado inclusivo. Por ejemplo, “A y B” pueden significar al menos lo siguiente: “tanto A como B”, “al menos tanto A como B”. Como otro ejemplo, “A o B” puede significar al menos lo siguiente: “al menos A”, “al menos B”, “tanto A como B”, “al menos tanto A como B”. Como otro ejemplo, “A y/o B” puede significar al menos lo siguiente: “A y B”, “A o B”. Cuando se pretende un “o” exclusivo, se indicará específicamente (por ejemplo, “ya sea A o B”, “como máximo uno de A y B”).

40 Este documento describe diversas funciones de procesamiento que están asociadas con estructuras como bloques, elementos, componentes, circuitos, etc. En general, estas estructuras pueden ser implementadas por un procesador controlado por uno o más programas de ordenador.

45 Una breve visión general que describe diversas realizaciones para la gestión del eco es como viene a continuación. El sistema primero estima la potencia de residuo de eco en cada banda de la señal de entrada modelando la ganancia de potencia del cancelador de eco para cada banda como una variable aleatoria distribuida gaussiana, cuya media y varianza dependen de la potencia de retorno en bucle y de una estimación del perfil de ganancia. A continuación, el sistema estima el perfil de ganancia estadística para cada banda y para cada nivel de entrada diferente. El sistema usa el perfil de ganancia estadística para cuantificar el nivel de entrada en varios segmentos. A continuación, el sistema modela el residuo de eco de banda ancha como un modelo de mezcla gaussiana y calcula la verosimilitud de conversación local. Finalmente, dependiendo de la verosimilitud de conversación local, el sistema aplica diferentes supresiones máximas a las bandas de la señal de entrada. A continuación se ofrecen más detalles.

55 La figura 1 es un diagrama de bloques de un sistema 100 de gestión de eco. El sistema 100 de gestión de eco generalmente elimina el eco del audio capturado por el dispositivo de extremo cercano antes de la transmisión al dispositivo de extremo lejano. Por ejemplo, cuando el audio procedente del extremo lejano es recibido por el dispositivo de extremo cercano y es emitido por el altavoz del dispositivo de extremo cercano, el “eco” corresponde a esta salida del altavoz que es capturada por el micrófono del dispositivo de extremo cercano. El sistema 100 de gestión de eco incluye un banco 101 de filtros de análisis, un cancelador 201 de eco, un supresor 301 de eco y un banco 111 de filtros de síntesis.

60 El banco 101 de filtros de análisis recibe una señal 120 de referencia (también denotada como  $x(t)$ ) y una señal capturada 122 (también denotada como  $y(t)$ ), realiza una transformada y genera una señal 130 de referencia transformada (también denotada como  $X(k, t')$ ) y una señal capturada transformada 132 (también denotada como

$Y(k, t')$ ). La señal 120 de referencia corresponde a la señal recibida desde el dispositivo de extremo lejano que se emite por el altavoz del dispositivo de extremo cercano. La señal capturada 122 corresponde al audio capturado por el micrófono del dispositivo de extremo cercano, que puede incluir conversación local (por ejemplo, la locución de una persona asociada con el dispositivo de extremo cercano) así como el eco de la salida del altavoz (por ejemplo, la locución de una persona asociada con el dispositivo de extremo lejano).

La transformada generalmente corresponde a transformar una señal de entrada de un dominio a otro dominio (por ejemplo, una transformada del dominio tiempo al dominio frecuencia). El banco 101 de filtros de análisis realiza el filtrado y diezmado de la señal 120 de referencia y la señal capturada 122 para generar la señal 130 de referencia transformada y la señal capturada transformada 132. Por ejemplo, el banco 101 de filtros de análisis realiza el filtrado y diezmado de  $x(t)$  para generar  $X(k, t')$ , donde  $k$  indica el índice de filtro de subbanda y  $t'$  indica la nueva variable de tiempo después del diezmado. De manera similar, el banco 101 de filtros de análisis realiza el filtrado y diezmado de  $y(t)$  para generar  $Y(k, t')$ .

El parámetro  $k$  se puede ajustar según se desee para ajustar la finura de los contenedores de transformada para un ancho de banda de señal de entrada determinado. Por ejemplo, los anchos de banda comunes de la señal 120 de referencia (o la señal capturada 122) pueden ser 4 kHz, 8 kHz, 12 kHz, 16 kHz, 24 kHz, etc., con frecuencias de muestreo correspondientes de 8 kHz, 16 kHz, 24 kHz, 32 kHz, 48 kHz, etc. Para un ancho de banda de contenedor de transformada de 25 Hz,  $k$  puede ser 160 contenedores para una señal de 4 kHz, 960 contenedores para una señal de 24 kHz, etc. al realizar la transformada (por ejemplo, una transformada rápida de Fourier).

El cancelador 201 de eco recibe la señal 130 de referencia transformada y la señal capturada transformada 132, realiza la cancelación de eco y genera una señal 140 de residuo de eco (también denotada como  $V(k, t')$ ). En general, la cancelación de eco se refiere a la aplicación de un filtrado lineal a  $Y(k, t')$  basado en  $X(k, t')$  para generar  $V(k, t')$ . El cancelador 201 de eco típicamente implementa un filtro adaptativo. Se proporcionan más detalles del cancelador 201 de eco con referencia a la figura 2. La señal 140 de residuo de eco se denomina "señal de residuo de eco" porque resulta de aplicar la cancelación de eco a la señal capturada transformada 132; el eco restante en la señal después de la cancelación de eco, y antes de la supresión de eco, es "eco residual".

El supresor 301 de eco recibe la señal 130 de referencia transformada y la señal 140 de residuo de eco, realiza la supresión de eco y genera una señal 150 de residuo de eco modificada (también denotada como  $\tilde{V}(k, t')$ ). En general, la supresión de eco se refiere a la aplicación de una atenuación no lineal a  $V(k, t')$  basándose en  $X(k, t')$  para generar la señal de residuo de eco modificada  $\tilde{V}(k, t')$  como una señal suprimida. Se proporcionan más detalles del supresor 301 de eco con referencia a la figura 3.

El banco 111 de filtros de síntesis recibe la señal 150 de residuo de eco modificada, realiza una transformada inversa y genera una señal capturada modificada 160 (también denotada como  $\tilde{v}(t)$ , que es una señal de dominio tiempo). La transformada inversa generalmente corresponde a una inversa de la transformada realizada por el banco 101 de filtros de análisis. La señal capturada modificada 160 puede luego transmitirse desde el dispositivo de extremo cercano al dispositivo de extremo lejano.

Para el sistema 100 de gestión de eco, el eco se puede modelar como la aplicación de un filtro lineal a la señal de reproducción, por ejemplo usando la ecuación (1):

(1)

$$\hat{x}(t) = h(t) * x(t) + n(t)$$

En la ecuación (1),  $\hat{x}(t)$  corresponde al eco que se debe cancelar,  $h(t)$  corresponde a un modelo de la respuesta de impulso de altavoz-sala-micrófono (de la sala en la que se ubica el dispositivo cercano con el altavoz y el micrófono),  $x(t)$  corresponde a la señal 120 de referencia y  $n(t)$  corresponde al ruido.

Cuando no hay voz de enlace ascendente (por ejemplo, no hay conversación local en el extremo cercano y el micrófono captura solo el eco de la salida del altavoz; la voz de enlace ascendente también puede denominarse voz de extremo cercano), se tiene la ecuación (2):

(2)

$$y(t) = h(t) * x(t) + n(t)$$

En las ecuaciones (1) y (2),  $h(t)$  puede estimarse mediante diversas técnicas, incluyendo un proceso de mínimos cuadrados medios (LMS), un proceso de mínimos cuadrados medios normalizados (NLMS), un proceso de mínimos cuadrados recursivos (RLS), etc.

Debido a la naturaleza acústica, una respuesta de impulso de una sala típica tiene más de miles de tomas y, por

lo tanto, el filtro adaptativo requerirá una gran cantidad de recursos computacionales. Esto es altamente indeseable para ordenadores portátiles y de escritorio, y el banco 101 de filtros se usa para implementar un filtro adaptativo de subbanda para reducir el cálculo.

5 La figura 2 es un diagrama de bloques del cancelador 201 de eco (véase la figura 1). El cancelador 201 de eco incluye una pluralidad de filtros adaptativos (por ejemplo, k filtros adaptativos) correspondientes al número de  
10 contenedores de análisis. Se muestran dos filtros adaptativos 251 y 252. El filtro adaptativo 251 recibe  $X(0, t')$  e  $Y(0, t')$ , aplica un filtro adaptativo  $H_0$  y genera  $V(0, t')$ . El filtro adaptativo 252 recibe  $X(k, t')$  e  $Y(k, t')$ , aplica un filtro adaptativo  $H_k$  y genera  $V(k, t')$ . Los otros filtros adaptativos del cancelador 201 de eco que no se muestran  
15 generan los k - 1 componentes restantes. X, Y y V corresponden respectivamente a la señal 130 de referencia transformada, la señal capturada transformada 132 y la señal 140 de residuo de eco (véase la figura 1), con las señales correspondientes a los respectivos contenedores estando denotadas con un sufijo correspondiente al contenedor (por ejemplo, "130-0" para el contenedor 0-ésimo de la señal 130 de referencia, "130-k" para el contenedor k-ésimo, etc.). La respuesta del filtro H corresponde a la respuesta de impulso de la sala  $h(t)$  en el dominio transformada.

Aunque el filtro adaptativo de subbanda implementado por el banco 101 de filtros puede reducir en gran medida la complejidad del algoritmo, el rendimiento del cancelador 201 de eco se degrada debido a la superposición de  
20 espectro entre filtros. Se propone un filtro adaptativo de subbanda de bucle cerrado para aliviar el problema, pero debido al hecho de que la complejidad aumenta en gran medida, el filtro adaptativo de subbanda de bucle abierto de la figura 2 sigue siendo una técnica comúnmente adoptada para la cancelación de eco acústico.

La figura 3 es un diagrama de bloques del supresor 301 de eco (véase la figura 1). Como se mencionó  
25 anteriormente, un filtro adaptativo de subbanda de bucle abierto es una solución práctica (pero no adecuada) para cancelar el eco hasta una cantidad de 50 dB para permitir una buena experiencia de comunicación de voz. Aunque la mayor parte del tiempo durante una llamada de voz, la comunicación dúplex completa (por ejemplo, doble conversación, cuando ambos extremos están hablando) ocurre raramente, y cuando sucede, ninguno de los hablantes activos esperará entender la locución de la otra parte; en cambio, su intención es simplemente interrumpir. Basándose en esto, se puede aplicar una supresión de eco no lineal después de la cancelación de  
30 eco para suprimir aún más el eco pero permitir que la voz de enlace ascendente de la parte de extremo cercano interrumpa a la parte de extremo lejano.

Teniendo esto en mente, durante una llamada de múltiples extremos con tres participantes, la tercera parte (que no son los dos hablantes activos en el escenario dúplex) tendrá una mala experiencia si el supresor deja pasar  
35 demasiado residuo de eco, aunque la persona que está siendo interrumpida no lo haga. Por lo tanto, un buen supresor debe suprimir la mayor cantidad posible de residuo de eco mientras pasa la mayor cantidad posible de voz de enlace ascendente. Este objetivo solo es posible, para un funcionamiento lineal, mediante un filtro adaptativo como el cancelador 201 de eco.

40 Para abordar estas cuestiones y proporcionar una gestión adicional del eco, el supresor 301 de eco usa las estadísticas intrínsecas del residuo de eco para realizar la supresión de eco. El supresor 301 de eco incluye un componente 311 de disposición en bandas, un estimador 312 de perfil de ganancia de potencia, un estimador 313 de residuo de eco, un calculador 314 de ganancia y un combinador 310 de señales.

45 El componente 311 de disposición en bandas recibe la señal 130 de referencia transformada (también denotada como  $X(k, t')$ ) y la señal 140 de residuo de eco (también denotada como  $V(k, t')$ ), realiza disposición en bandas de la frecuencia y genera una señal 320 de referencia en bandas (también denotada como  $\tilde{X}(b, t')$ ) y una señal 322 de residuo de eco en bandas (también denotada como  $\tilde{V}(b, t')$ ). En general, "disposición en bandas" se refiere a agrupar o agregar múltiples subbandas o contenedores de frecuencias de una señal para formar una  
50 banda, donde el número resultante de bandas es menor que el número original de subbandas o contenedores de frecuencias. Por ejemplo, la disposición en bandas se puede realizar sumando múltiples subbandas o contenedores de frecuencias para formar una banda, y puede incluir operaciones adicionales como multiplicación, cálculo de valor absoluto, cálculo de magnitud, etc. El componente 311 de disposición en bandas puede realizar disposición en bandas de la señal 140 de residuo de eco para generar la señal 322 de residuo de  
55 eco en bandas de acuerdo con la ecuación (3):

(3)

$$\tilde{V}(b, t') = \sum_{k \in B(b)} |V(k, t')|^2$$

El componente 311 de disposición en bandas puede realizar disposición en bandas de la señal 130 de referencia transformada para generar la señal 320 de referencia en bandas de acuerdo con la ecuación (4):  
60

(4)

$$X(b, t') = \sum_{k \in B(b)} |X(k, t')|^2$$

En las ecuaciones (3) y (4), B(b) corresponde a un conjunto de contenedores k que están asociados con una banda b, y la salida de la disposición en bandas corresponde a la potencia de la señal para cada banda en particular. La disposición en bandas puede ser la disposición en bandas de ancho de banda rectangular equivalente (ERB), que corresponde a un modelo psicoacústico de la audición humana y asocia anchos de banda mayores con frecuencias más bajas. Por ejemplo, el número de contenedores k puede ser 960 contenedores, el número de bandas b puede ser 61 y cada banda b está asociada con un número de contenedores k, donde más contenedores están asociados con las bandas de frecuencia más bajas que con las bandas de frecuencia más altas.

El estimador 312 de perfil de ganancia de potencia recibe la señal 320 de referencia en bandas y la señal 322 de residuo de eco en bandas, estima un perfil de ganancia de potencia y genera un perfil 324 de ganancia de potencia estimado y una potencia 420 de retorno en bucle (también denotada como  $\tilde{X}(b, t')$ ; véase la figura 4). El perfil 324 de ganancia de potencia estimado corresponde a una estimación del perfil de ganancia de potencia de un cancelador de eco acústico tal como el cancelador 201 de eco (véase la figura 1). El estimador 312 de perfil de ganancia de potencia también puede recibir una señal 326 de control que indica la presencia de doble conversación. La doble conversación se refiere a la situación en la que tanto la parte de extremo lejano como la parte de extremo cercano están hablando al mismo tiempo (por ejemplo, donde la locución de extremo lejano se indica de acuerdo con la señal 120 de referencia y la locución de extremo cercano se indica mediante la señal capturada 122, que también incluye el eco de la locución de extremo lejano emitido por el altavoz y capturado por el micrófono en el extremo cercano).

Cuando el cancelador 201 de eco alcanza un estado estable, y cuando no hay voz de enlace ascendente (por ejemplo, indicado por la ausencia de doble conversación de acuerdo con la señal 326 de control, y la presencia de eco de acuerdo con la señal 120 de referencia), se tendrá la señal 140 de residuo de eco (también denotada como  $V(k, t')$ ) como se representa de acuerdo con la ecuación (5):

(5)

$$V(k, t') = X(k, t) * [\hat{H}(k, t) - H(k, t)]$$

En la ecuación (5),  $\hat{H}(k, t) - H(k, t)$  es la respuesta de error del filtro adaptativo que no coincide con la respuesta real del altavoz-sala-micrófono. Téngase en cuenta que debido a que  $\hat{H}(k, t)$  depende de la entrada  $X(k, t)$ ,  $\hat{H}(b, t')$  también dependerá de la entrada. Y debido a que el filtro adaptativo  $\hat{H}(k, t)$  tiene múltiples tomas, la potencia residual actual  $\mathcal{V}(b, t')$  podría no solo depender de la potencia de entrada actual, sino también de la potencia de entrada anterior. En otras palabras, el residuo de eco  $V(k, t')$  no será una salida del sistema lineal invariante en el tiempo (LTI).

Por sencillez e implementación práctica, se modela la potencia del residuo de acuerdo con la ecuación (6):

(6)

$$\mathcal{V}(b, t') = \tilde{X}(b, t') \cdot \tilde{H}(b, t')$$

Se proporcionan detalles adicionales del estimador 312 de perfil de ganancia de potencia en las figuras 4-5.

La figura 4 es un diagrama de bloques de un calculador 400 de potencia de retorno en bucle. El calculador 400 de potencia de retorno en bucle puede ser un componente del estimador 312 de perfil de ganancia de potencia (véase la figura 3). El calculador 400 de potencia de retorno en bucle generalmente calcula la potencia 420 de retorno en bucle (también denotada como  $\tilde{X}(b, t')$ ; véase también la figura 3) de la señal 320 de referencia en bandas (véase la figura 3; también denotada como  $X(b, t')$ ). En general, la potencia 420 de retorno en bucle corresponde a una combinación ponderada de valores históricos de la señal 320 de referencia en bandas, como se detalla con más detalle a continuación. El calculador 400 de potencia de retorno en bucle incluye un buscador 410 de tomas de filtro adaptativo dominantes, una memoria 412, multiplicadores 414 y 416, y un sumador 418.

El buscador 410 de tomas de filtro adaptativo dominantes recibe coeficientes 402 de filtro (correspondientes a los coeficientes de filtro de los filtros adaptativos  $H_k$  en el cancelador 201 de eco de la figura 2), determina las dos tomas dominantes y proporciona los pesos  $w_0$  y  $w_1$  de las tomas dominantes a los multiplicadores 414 y 416. Una toma dominante es una toma que proporciona una mayor contribución de peso al filtro adaptativo que otra toma.

El buscador 410 de tomas de filtro adaptativo dominantes usa los coeficientes de filtro 402 correspondientes a  $H_k$  en todas las subbandas. Los pesos  $w_0$  y  $w_1$  pueden ser pesos relativos.

La memoria 412 almacena un historial de la señal de referencia en bandas  $X(b, t')$ . La memoria 412 tiene varios  
5 elementos de memoria, para almacenar los valores actuales de la señal de referencia en bandas  $X(b, t')$  y uno o más valores anteriores de la señal de referencia en bandas  $X(b, t' - 1)$ ,  $X(b, t' - 2)$ , ...,  $X(b, t' - n)$ . Como ejemplo específico, la figura 2 muestra la memoria 412 que contiene 4 elementos de memoria.

El multiplicador 414 multiplica uno de los valores históricos de la señal 320 de referencia en bandas  
10 almacenados en la memoria 412 por el peso  $w_0$ , y el multiplicador 416 multiplica otro de los valores históricos de la señal 320 de referencia en bandas almacenados en la memoria 412 por el peso  $w_1$ . El sumador 418 suma los resultados de los multiplicadores 414 y 416. Como resultado, la potencia de retorno en bucle  $\tilde{X}(b, t')$  se puede calcular de acuerdo con la ecuación (7):

(7)

$$\tilde{X}(b, t') = w_0 X(b, t' - n_0) + w_1 X(b, t' - n_1)$$

En la ecuación (7),  $n_0$  y  $n_1$  son los índices de retardo correspondientes respectivamente a las tomas dominantes  
15  $w_0$  y  $w_1$ . En el ejemplo que se muestra en la figura 4,  $n_0$  es 2 y  $n_1$  es 0.

Cuando la potencia de entrada es mayor en una banda particular (como corresponde a la potencia de retorno en  
20 bucle  $\tilde{X}(b, t')$ ),  $\hat{H}(b, t')$  también tiende a ser mayor en valor absoluto. Esta relación se detalla más en la figura 5.

La figura 5 es un histograma de la relación entre la potencia de retorno en bucle  $\tilde{X}(b, t')$  (también denominada  
25 como la potencia 420 de retorno en bucle en las figuras 3-4) y la función de transferencia de error (residuo de eco)  $\hat{H}(b, t')$  del filtro adaptativo. Como se puede ver en la figura 5, cuando la potencia de entrada es mayor en una banda particular (como corresponde a la potencia de retorno en bucle  $\tilde{X}(b, t')$  en una banda particular),  $\hat{H}(b, t')$  también tiende a ser mayor en valor absoluto. Por otro lado, cuando la banda contiene principalmente ruido,  $\hat{H}(b, t')$  tiende a ser pequeño ya que el ruido no se cancelará.

Volviendo a la figura 3, el estimador 312 de perfil de ganancia de potencia usa la relación de la figura 5 como un  
30 histograma para estimar un perfil de ganancia para cada banda. El perfil 324 de ganancia de potencia estimado se puede representar como la media  $\mu(b, t')$  y la varianza  $\sigma_b^2(b, t')$  de  $\hat{H}(b, t')$ . (La media y la varianza se denotan como  $\mu_b$  y  $\sigma_b^2$  en los párrafos siguientes, para simplificar). En resumen, el estimador 312 de perfil de ganancia de potencia usa las estadísticas de la señal de residuo de eco (por ejemplo, la señal 140 de residuo de eco) en  
35 combinación con el historial de la señal de entrada (por ejemplo, la señal 130 de referencia) para estimar el perfil de ganancia de potencia.

El estimador 312 de perfil de ganancia de potencia usa la señal 326 de control (por ejemplo, el indicador de doble  
40 conversación) para actualizar el perfil de ganancia. Específicamente, el estimador 312 de perfil de ganancia de potencia actualiza el perfil de ganancia solo cuando tiene confianza en que no hay voz de enlace ascendente (extremo cercano), como se indica mediante la señal 326 de control. La señal 326 de control puede generarse de diversas maneras, incluyendo mediante un detector de doble conversación o usando un método estocástico. Un ejemplo del método estocástico es actualizar la estimación en un momento aleatorio una vez que la señal de  
45 referencia ha superado un primer umbral durante una duración mayor que un segundo umbral.

El estimador 313 de residuo de eco recibe la señal 322 de residuo de eco en bandas (también denotada como  
50  $V(b, t')$ ), la potencia 420 de retorno en bucle (también denotada como  $\tilde{X}(b, t')$ ) y el perfil 324 de ganancia de potencia estimado (la media y la varianza de  $\hat{H}(b, t')$ , denotadas como  $\mu_b$  y  $\sigma_b^2$ ), y calcula una potencia 328 de residuo de eco estimada (también denotada como  $\mathcal{R}(b, t')$ ). El estimador 313 de residuo de eco puede calcular la potencia 328 de residuo de eco estimada basándose en una combinación del perfil 324 de ganancia de potencia estimado y la potencia 420 de retorno en bucle. El estimador 313 de residuo de eco puede calcular la potencia 328 de residuo de eco estimada usando un operador mínimo aplicado a una combinación de la señal 322 de residuo de eco en bandas, la potencia 420 de retorno en bucle y el perfil 324 de ganancia de potencia  
55 estimado. El estimador 313 de residuo de eco puede calcular la potencia 328 de residuo de eco estimada de acuerdo con la ecuación (8):

(8)

$$\mathcal{R}(b, t') = \min \{V(b, t'), (\mu_b + \beta \sigma_b^2) \cdot \tilde{X}(b, t')\}$$

En la ecuación (8),  $\beta$  es un factor determinado por el ajuste. El operador mínimo garantiza que la potencia 328 de residuo de eco estimada nunca superará  $V(b, t')$ , que es una combinación de la potencia de residuo de eco y la potencia de conversación local. La potencia 328 de residuo de eco estimada se puede modelar como una variable aleatoria gaussiana con su centroide en  $\mathcal{R}(b, t')$ . En resumen, el estimador 313 de residuo de eco calcula la potencia 328 de residuo de eco estimada de acuerdo con la potencia 420 de retorno en bucle para cada banda, modelando la ganancia de potencia del cancelador 201 de eco (véanse las figuras 1-2) para cada banda como una variable aleatoria gaussiana parametrizada por el perfil 324 de ganancia de potencia estimado.

El calculador 314 de ganancia recibe la señal 322 de residuo de eco en bandas y la potencia 328 de residuo de eco estimada, y calcula las ganancias 330 en bandas (también denotadas como  $\mathcal{G}(b, t')$ ). Uno de los objetivos del cálculo de las ganancias 330 en bandas es aplicar, a cada banda en particular, diferentes cantidades de supresión para el escenario semidúplex frente al escenario dúplex completo. (El escenario semidúplex describe cuando solo una de las partes está hablando, por ejemplo el extremo cercano o el extremo lejano. El escenario dúplex completo describe cuando ambas partes están hablando, por ejemplo doble conversación). En otras palabras, el cálculo de las ganancias 330 en bandas incluye el cálculo selectivo de la pluralidad de ganancias en bandas en función de la verosimilitud de conversación local en la señal de residuo de eco. Como parte del cálculo de las ganancias 330 en bandas, el calculador 314 de ganancias calcula una característica para orientar el cálculo de ganancia entre los dos escenarios.

Debido a la naturaleza del ruido de la sala, el ruido electrónico y la aleatoriedad de la señal de voz, se supone que la potencia 328 de residuo de eco estimada en dB tiene una distribución gaussiana con media  $\text{dB}\{\mathcal{R}(b, t')\}$  y varianza  $\sigma_b^2$ , que corresponde al perfil 324 de ganancia de potencia estimado calculado por el estimador 312 de perfil de ganancia de potencia. También se supone que la potencia 328 de residuo de eco estimada en dB no tiene correlación cruzada (con el primer supuesto, independiente) entre bandas. Sea  $E^c$  el evento cuando no hay voz en el enlace ascendente; la función de densidad de probabilidad para la potencia de residuo de eco en dB en una banda particular está dada por la ecuación (9):

(9)

$$\Pr(x|E^c) \propto e^{-\frac{(x - \text{dB}\{\mathcal{R}(b, t')\})^2}{\sigma_b^2}}$$

En la ecuación (9),  $x$  es la potencia 328 de residuo de eco estimada en dB, y  $\mathcal{R}(b, t')$  corresponde al centroide de la potencia 328 de residuo de eco estimada.

La figura 6 es un gráfico que ilustra la función de densidad de probabilidad  $\Pr(x|E^c)$  de la Ecuación (9). En la figura 6, el eje  $x$  corresponde a  $x$  y el eje  $y$  corresponde a  $\Pr(x|E^c)$ .

Al modelar el residuo de eco de banda ancha como un modelo de mezcla gaussiana, la verosimilitud logarítmica  $L$  de que no haya voz en el enlace ascendente en todas las bandas  $b$  se da mediante la ecuación (10):

(10)

$$L = \sum_b -\frac{(\text{dB}\{\mathcal{V}(b, t')\} - \text{dB}\{\mathcal{R}(b, t')\})^2}{\sigma_b^2}$$

En la ecuación (10), la verosimilitud logarítmica  $L$  es la suma de todas las bandas  $b$  del componente exponencial

$$-\frac{(\text{dB}\{\mathcal{V}(b, t')\} - \text{dB}\{\mathcal{R}(b, t')\})^2}{\sigma_b^2}$$

de la ecuación (9). En otras palabras, la verosimilitud de conversación local en la señal 140 de residuo de eco es una verosimilitud logarítmica basada en una media y una varianza del perfil de ganancia de potencia estimado sumado sobre una pluralidad de bandas de frecuencia. En resumen, la verosimilitud logarítmica  $L$  se calcula modelando la señal 140 de residuo de eco como un modelo de mezcla gaussiana aplicado sobre todas las bandas  $b$ .

Cuando no hay voz de enlace ascendente (por ejemplo, ninguna conversación local capturada por el dispositivo),  $L(E^c)$  debe ser baja; de lo contrario,  $L(E)$  es alta y la verosimilitud logarítmica  $L$  de la ecuación (10) será la señal para dirigir el cálculo de ganancia.

La verosimilitud logarítmica  $L$  de la ecuación (10) se usa para calcular una ganancia de supresión máxima global



en todas las bandas  $G_{\max}$  de acuerdo con la ecuación (11):

$$G_{\max}(dB) = \begin{cases} G_0 + \frac{G_1 - G_0}{1 + (L - L_{th})^2}, & \text{cuando } L \geq L_{\text{ésima}} \\ G_1, & \text{de otro modo} \end{cases} \quad (11)$$

- 5 En la ecuación (11),  $G_0$  es la ganancia en dB cuando hay voz de enlace ascendente y  $G_1$  es la ganancia predefinida en dB cuando no hay voz de enlace ascendente.  $L_{\text{ésima}}$  es un umbral predefinido que se puede obtener para cada dispositivo durante la configuración inicial a través del conocimiento previo de si hay conversación local presente o no.
- 10 En otras palabras, el cálculo de las ganancias 330 en bandas incluye el cálculo de una ganancia de supresión máxima global  $G_{\max}$  en la pluralidad de bandas de frecuencia. La ganancia de supresión máxima global se basa en una primera ganancia  $G_0$  correspondiente a una presencia de conversación local, una segunda ganancia  $G_1$  correspondiente a una ausencia de conversación local y la verosimilitud  $L$  de conversación local.
- 15 En cada banda, el calculador 314 de ganancia puede calcular la ganancia para cada banda  $G(b, t')$  usando uno o más procesos de cálculo de ganancia. Los procesos de cálculo de ganancia de ejemplo que son adecuados incluyen los procesos de cálculo de ganancia descritos en la patente de EE.UU. n.º 9.173.025, la patente de EE.UU. n.º 8.712.076 y la patente de EE.UU. n.º 8.804.977, que se incorporan en el presente documento por referencia.
- 20 El combinador 310 de señales recibe la señal 140 de residuo de eco (también denotada como  $V(k, t')$ ) y las ganancias 330 en bandas (también denotadas como  $G(b, t')$ ) y genera una señal 150 de residuo de eco modificada (también denotada como  $\tilde{V}(k, t')$ ). La ganancia calculada para una banda  $b$  dada se aplica a todas las subbandas  $k$  que pertenecen a esa banda ERB dada. El combinador 310 de señales puede generar la señal 150 de residuo de eco modificada de acuerdo con la ecuación (12):
- 25

(12)

$$\tilde{V}(k, t') = \sum_{k \in B(b)} G(b, t') V(k, t')$$

- En otras palabras, la señal 140 de residuo de eco tiene un número de contenedores de frecuencias  $k$ , donde una ganancia en bandas dada  $G(b, t')$  de las ganancias 330 en bandas corresponde a un contenedor de frecuencias dado (por ejemplo, cada banda  $b$  puede estar asociada con un número de contenedores  $k$ ), y generar la señal 150 de residuo de eco modificada incluye, para cada uno de los contenedores  $k$  de la señal 140 de residuo de eco, aplicar una ganancia en bandas correspondiente  $G(b, t')$  para generar la señal 150 de residuo de eco modificada.
- 30
- 35 La figura 7 es una arquitectura 700 de dispositivo móvil para implementar las características y procesos descritos en el presente documento, de acuerdo con una realización. La arquitectura 700 se puede implementar en cualquier dispositivo electrónico, incluidos, entre otros: un ordenador de escritorio, un equipo de audio/video (AV) de consumo, un equipo de transmisión de radio, dispositivos móviles (por ejemplo, un teléfono inteligente, una tableta, un ordenador portátil, un dispositivo portátil), etc. En la realización de ejemplo mostrada, la arquitectura 700 es para un ordenador portátil e incluye procesador(es) 701, interfaz 702 de periféricos, subsistema 703 de audio, altavoces 704, micrófono 705, sensores 706 (por ejemplo, acelerómetros, giroscopios, barómetro, magnetómetro, cámara), procesador 707 de ubicación (por ejemplo, receptor GNSS), subsistemas 708 de comunicaciones inalámbricas (por ejemplo, Wi-Fi, Bluetooth, celular) y subsistema(s) 709 de E/S, que incluye controlador táctil 710 y otros controladores 711 de entrada, superficie táctil 712 y otros dispositivos 713 de entrada/control. También se pueden usar otras arquitecturas con más o menos componentes para implementar las realizaciones divulgadas.
- 40
- 45
- 50 La interfaz 714 de memoria está acoplada a los procesadores 701, a la interfaz 702 de periféricos y a la memoria 715 (por ejemplo, flash, RAM, ROM). La memoria 715 almacena instrucciones y datos del programa de ordenador, incluyendo, entre otros: instrucciones 716 del sistema operativo, instrucciones 717 de comunicación, instrucciones 718 de GUI, instrucciones 719 de procesamiento de sensores, instrucciones 720 de teléfono, instrucciones 721 de mensajería electrónica, instrucciones 722 de navegación web, instrucciones 723 de procesamiento de audio, instrucciones 724 de GNSS/navegación y aplicaciones/datos 725. Las instrucciones 723 de procesamiento de audio incluyen instrucciones para realizar el procesamiento de audio descrito en el presente documento.
- 55

La figura 8 es un diagrama de flujo de un método 800 de procesamiento de audio. El método 800 puede ser realizado por un dispositivo (por ejemplo, un ordenador portátil, un teléfono móvil, etc.) con los componentes de la arquitectura 700 de la figura 7, para implementar la funcionalidad del sistema 100 de gestión de eco (véase la figura 1), el cancelador 201 de eco (véase la figura 2), el supresor 301 de eco (véase la figura 3), el calculador 400 de potencia de retorno en bucle (véase la figura 4), etc., por ejemplo ejecutando uno o más programas de ordenador.

En 802, se reciben una señal de referencia y una señal de residuo de eco. Por ejemplo, el supresor 301 de eco (véase la figura 3) puede recibir la señal 130 de referencia transformada y la señal 140 de residuo de eco. La señal 130 de referencia transformada corresponde a la señal 120 de referencia, como se puede ver en la figura 1.

En 804, se generan una señal de referencia en bandas y una señal de residuo de eco en bandas realizando una disposición en bandas de frecuencia en la señal de referencia y la señal de residuo de eco. Por ejemplo, el componente 311 de disposición en bandas (véase la figura 3) puede realizar disposición en bandas en la señal 130 de referencia transformada y la señal 140 de residuo de eco para generar la señal 320 de referencia en bandas y la señal 322 de residuo de eco en bandas, respectivamente.

En 806, se calcula un perfil de ganancia de potencia estimado y una potencia de retorno en bucle basándose en la señal de referencia en bandas y la señal de residuo de eco en bandas. El perfil 324 de ganancia de potencia estimado corresponde a una estimación del perfil de ganancia de potencia de un cancelador de eco acústico tal como el cancelador 201 de eco (véase la figura 1). Por ejemplo, el estimador 312 de perfil de ganancia de potencia (véase la figura 3) puede calcular el perfil 324 de ganancia de potencia estimado basándose en la señal 320 de referencia en bandas y la señal 322 de residuo de eco en bandas. El calculador 400 de potencia de retorno en bucle (véase la figura 4) puede calcular la potencia 420 de retorno en bucle basándose en la señal 320 de referencia en bandas. El perfil 324 de ganancia de potencia estimado puede indicar una cantidad a largo plazo de cancelación de eco según lo calculado por el calculador de potencia de retorno en bucle 400, por ejemplo usando el historial de la señal 320 de referencia en bandas almacenada en la memoria 412.

En 808, se calcula una potencia de residuo de eco estimada basándose en la señal de residuo de eco en bandas, la potencia de retorno en bucle y el perfil de ganancia de potencia estimado. Por ejemplo, el estimador 313 de residuo de eco (véase la figura 3) puede calcular la potencia 328 de residuo de eco estimada basándose en la señal 322 de residuo de eco en bandas, la potencia 420 de retorno en bucle y el perfil 324 de ganancia de potencia estimado.

En 810, se calcula una pluralidad de ganancias en bandas basándose en la señal de residuo de eco en bandas y la potencia de residuo de eco estimada. Por ejemplo, el calculador 314 de ganancia (véase la figura 3) puede calcular las ganancias 330 en bandas basándose en la señal 322 de residuo de eco en bandas y la potencia 328 de residuo de eco estimada.

En 812, se genera una señal de residuo de eco modificada aplicando la pluralidad de ganancias en bandas a la señal de residuo de eco. Por ejemplo, el combinador 310 de señales (véase la figura 3) puede generar la señal 150 de residuo de eco modificada aplicando las ganancias 330 en bandas a la señal 140 de residuo de eco.

El método 800 puede incluir pasos adicionales correspondientes a las otras funcionalidades del sistema 100 de gestión de eco, el cancelador 201 de eco, el supresor 301 de eco, el calculador 400 de potencia de retorno en bucle, etc. como se describe en el presente documento. Por ejemplo, la señal de referencia puede emitirse mediante un altavoz, como los altavoces 704 (véase la figura 7). Como otro ejemplo, el audio de extremo cercano puede ser capturado por un micrófono (por ejemplo, el micrófono 705 de la figura 5), se puede generar una señal de audio de extremo cercano y la señal de residuo de eco puede generarse realizando una cancelación de eco en la señal de audio de extremo cercano (por ejemplo, mediante el cancelador 201 de eco de la figura 1). El audio de extremo cercano puede incluir conversación local (por ejemplo, locución generada por la persona de extremo cercano), eco (por ejemplo, la locución de extremo lejano emitida por el altavoz de extremo cercano), etc. Como otro ejemplo, el residuo de eco modificado puede ser transmitido por el dispositivo de extremo cercano a un dispositivo de extremo lejano, por ejemplo como parte de una teleconferencia, una videoconferencia, etc.

#### Detalles de implementación

Una realización puede implementarse en hardware, módulos ejecutables almacenados en un medio legible por ordenador o una combinación de ambos (por ejemplo, matrices lógicas programables). A menos que se especifique lo contrario, los pasos ejecutados por las realizaciones no necesitan estar inherentemente relacionados con ningún ordenador u otro aparato en particular, aunque pueden estarlo en ciertas realizaciones. En particular, se pueden usar diversas máquinas de propósito general con programas escritos de acuerdo con las enseñanzas del presente documento, o puede ser más conveniente construir aparatos más especializados (por ejemplo, circuitos integrados) para realizar los pasos del método requeridos. De este modo, las realizaciones

pueden implementarse en uno o más programas de ordenador que se ejecutan en uno o más sistemas de ordenador programables, cada uno de los cuales comprende al menos un procesador, al menos un sistema de almacenamiento de datos (incluyendo elementos de memoria y/o almacenamiento volátiles y no volátiles), al menos un dispositivo o puerto de entrada, y al menos un dispositivo o puerto de salida. El código de programa se aplica a los datos de entrada para realizar las funciones descritas en el presente documento y generar información de salida. La información de salida se aplica a uno o más dispositivos de salida, de manera conocida.

Cada uno de estos programas de ordenador se almacena o se descarga preferiblemente en un medio o dispositivo de almacenamiento (por ejemplo, una memoria o medio de estado sólido, o un medio magnético u óptico) legible por un ordenador programable de propósito general o especial, para configurar y operar el ordenador cuando el medio o dispositivo de almacenamiento es leído por el sistema de ordenador para realizar los procedimientos descritos en el presente documento. El sistema inventivo también puede considerarse implementado como un medio de almacenamiento legible por ordenador, configurado con un programa de ordenador, donde el medio de almacenamiento así configurado hace que un sistema de ordenador funcione de una manera específica y predefinida para realizar las funciones descritas en el presente documento. (El software per se y las señales intangibles o transitorias quedan excluidos en la medida en que sean materia no patentable).

Los aspectos de los sistemas descritos en el presente documento pueden implementarse en un entorno de red de procesamiento de sonido basado en ordenador apropiado para procesar archivos de audio digitales o digitalizados. Porciones del sistema de audio adaptativo pueden incluir una o más redes que comprenden cualquier número deseado de máquinas individuales, incluyendo uno o más enrutadores (no mostrados) que sirven para almacenar en búfer y enrutar los datos transmitidos entre los ordenadores. Tal red puede construirse sobre diversos protocolos de red diferentes, y puede ser Internet, una red de área amplia (WAN), una red de área local (LAN) o cualquier combinación de ellas.

Uno o más de los componentes, bloques, procesos u otros componentes funcionales pueden implementarse a través de un programa de ordenador que controla la ejecución de un dispositivo informático basado en procesador del sistema. También debe tenerse en cuenta que las diversas funciones divulgadas en el presente documento pueden describirse usando cualquier número de combinaciones de hardware, firmware y/o como datos y/o instrucciones materializados en diversos medios legibles por máquina o legibles por ordenador, en términos de su comportamiento, transferencia de registro, componente lógico y/u otras características. Los medios legibles por ordenador en los que se pueden materializar tales datos y/o instrucciones formateados incluyen, pero no están limitados a, medios de almacenamiento físicos (no transitorios) no volátiles en diversas formas, tales como medios de almacenamiento ópticos, magnéticos o semiconductores.

La descripción anterior ilustra diversas realizaciones de la presente divulgación junto con ejemplos de cómo se pueden implementar aspectos de la presente divulgación. Los ejemplos y realizaciones anteriores no deben considerarse las únicas realizaciones y se presentan para ilustrar la flexibilidad y las ventajas de la presente divulgación tal como se define en las siguientes reivindicaciones. Basándose en la divulgación anterior y las reivindicaciones siguientes, otras disposiciones, realizaciones, implementaciones y equivalentes serán evidentes para aquellos expertos en la técnica y se pueden emplear sin alejarse del alcance de la divulgación tal como se define en las reivindicaciones.

## Referencias

Patentes de EE.UU. n.º 10.115.411, 8.111.833, 9.338.551, 9.363.600, 10.382.092, 10.313.789, 10.477.031, 9.503.815, 9.185.506, 8.462.958, 6.163.608, 9.420.390, 9.173.025, 8.712.076, 8.804.977.

Publicaciones de solicitud de patente de EE.UU. n.º 2016/0241955, 2019/0349471.

C. Yemdji, M. M. Idrissa, N. W. D. Evans y C. Beaugeant, "Efficient low delay filtering for residual echo suppression", en la 18.ª Conferencia Europea de Procesamiento de Señales de 2010, Aalborg, 2010, págs. 16-20.

## REIVINDICACIONES

1.- Un método implementado por ordenador, de procesamiento de audio, comprendiendo el método:

5 recibir (802) una señal (130) de referencia y una señal (140) de residuo de eco;

generar (804) una señal (320) de referencia en bandas y una señal (322) de residuo de eco en bandas mediante la realización de una disposición en bandas de la frecuencia en la señal de referencia y en la señal de residuo de eco, respectivamente;

10 calcular (806) un perfil (324) de ganancia de potencia estimado de un cancelador (201) de eco acústico y una potencia (420) de retorno en bucle, en el que la potencia de retorno en bucle corresponde a una combinación ponderada de valores de potencia históricos de la señal de referencia en bandas, el perfil de ganancia de potencia estimado se calcula mediante el uso de estadísticas de la señal de residuo de eco en bandas en combinación con los valores de potencia históricos de la señal de referencia en bandas;

15 calcular (808) una potencia (328) de residuo de eco estimada para cada banda basándose en la señal de residuo de eco en bandas, la potencia de retorno en bucle y el perfil de ganancia de potencia estimado;

20 calcular (810) una pluralidad de ganancias (330) en bandas basándose en la señal de residuo de eco en bandas y la potencia de residuo de eco estimada; y

generar (812) una señal (150) de residuo de eco modificada mediante la aplicación de la pluralidad de ganancias en bandas a la señal de residuo de eco.

25 2.- El método de la reivindicación 1, que comprende además:

emitir, mediante un altavoz, la señal de referencia.

30 3.- El método de una cualquiera de las reivindicaciones 1-2, que comprende además:

capturar, mediante un micrófono, audio de extremo cercano;

35 generar una señal de audio de extremo cercano correspondiente al audio de extremo cercano; y

generar la señal de residuo de eco mediante la realización de cancelación de eco en la señal de audio de extremo cercano.

40 4.- El método de una cualquiera de las reivindicaciones 1 a 3, que comprende además:

transmitir, a un dispositivo de extremo lejano, la señal de residuo de eco modificada.

45 5.- El método de una cualquiera de las reivindicaciones 1 a 4, en el que la señal de residuo de eco tiene una pluralidad de contenedores de frecuencias, en el que la señal de residuo de eco en bandas tiene una pluralidad de bandas de frecuencia, y en el que la generación de la señal de residuo de eco en bandas comprende:

generar la señal de residuo de eco en bandas mediante la realización de una disposición en bandas de ancho de banda rectangular equivalente en la señal de residuo de eco, en el que la disposición en bandas rectangular equivalente agrega la pluralidad de contenedores de frecuencias en la pluralidad de bandas de frecuencia de acuerdo con un modelo psicoacústico de audición humana.

50 6.- El método de una cualquiera de las reivindicaciones 1 a 5, en el que el cálculo del perfil de ganancia de potencia estimado incluye:

55 para cada uno de una pluralidad de filtros adaptativos, calcular, mediante un buscador de tomas de filtro adaptativo dominantes, dos tomas dominantes, en el que la pluralidad de filtros adaptativos corresponde al número de contenedores de análisis de la señal (130) de referencia y cada uno de la pluralidad de filtros adaptativos recibe la señal (130) de referencia y una señal capturada transformada (132) que corresponden a los respectivos contenedores y genera la señal (140) de residuo de eco correspondiente a los respectivos contenedores, en el que las tomas dominantes proporcionan una mayor contribución de peso a un filtro adaptativo actual de la pluralidad de filtros adaptativos en comparación con otras tomas del filtro adaptativo actual de la pluralidad de filtros adaptativos;

60 calcular la potencia de retorno en bucle de la señal de referencia en bandas usando las dos tomas dominantes y los valores de potencia históricos de la señal de referencia en bandas; y

calcular el perfil de ganancia de potencia estimado usando un histograma basándose en la señal de referencia en bandas.

5 7.- El método de la reivindicación 6, en el que el perfil de ganancia de potencia estimado incluye una media y una varianza.

10 8.- El método de una cualquiera de las reivindicaciones 1 a 7, en el que la potencia de retorno en bucle se calcula sobre una pluralidad de bandas formadas por la disposición en bandas de frecuencia en la señal de referencia y la señal de residuo de eco, en el que el cálculo de la potencia de residuo de eco estimada incluye:  
calcular la potencia de residuo de eco estimada de acuerdo con la potencia de retorno en bucle para cada banda de la pluralidad de bandas modelando la ganancia de potencia de un cancelador de eco para cada banda como una variable aleatoria gaussiana parametrizada por el perfil de ganancia de potencia estimado.

15 9.- El método de una cualquiera de las reivindicaciones 1 a 8, en el que el cálculo de la pluralidad de ganancias en bandas incluye calcular selectivamente la pluralidad de ganancias en bandas basándose en una verosimilitud de conversación local en la señal de residuo de eco, en el que la conversación local corresponde a audio capturado por un dispositivo que implementa el método de una cualquiera de las reivindicaciones 1 a 5.

20 10.- El método de la reivindicación 9, en el que la verosimilitud de conversación local en la señal de residuo de eco es:

25 calculada modelando la señal de residuo de eco como un modelo de mezcla gaussiana aplicado sobre una pluralidad de bandas formadas por la disposición en bandas de frecuencia en la señal de referencia y la señal de residuo de eco; o

30 una verosimilitud logarítmica basada en una media y una varianza del perfil de ganancia de potencia estimado sumado a lo largo de una pluralidad de bandas de frecuencia formadas por la disposición en bandas de frecuencia en la señal de referencia y la señal de residuo de eco.

11.- El método de la reivindicación 9, en el que el cálculo de la pluralidad de ganancias en bandas incluye calcular una ganancia de supresión máxima global a través de la pluralidad de bandas de frecuencia.

35 12.- El método de la reivindicación 11, en el que el cálculo de la ganancia de supresión máxima global se basa en una primera ganancia correspondiente a una presencia de la conversación local, una segunda ganancia correspondiente a una ausencia de la conversación local y la verosimilitud de la conversación local.

40 13.- El método de una cualquiera de las reivindicaciones 1 a 12, en el que la señal de residuo de eco tiene una pluralidad de contenedores de frecuencias, en el que una ganancia en bandas dada de la pluralidad de ganancias en bandas corresponde a un contenedor de frecuencias dado de la pluralidad de contenedores de frecuencias, y en el que la generación de la señal de residuo de eco modificada incluye:

45 para cada uno de la pluralidad de contenedores de frecuencias de la señal de residuo de eco, aplicar una correspondiente ganancia en bandas de la pluralidad de ganancias en bandas para generar la señal de residuo de eco modificada.

50 14.- Un medio no transitorio legible por ordenador que almacena un programa de ordenador que, cuando es ejecutado por un procesador, controla un aparato para ejecutar un procesamiento que incluye el método de una cualquiera de las reivindicaciones 1 a 13.

15.- Un aparato para procesamiento de audio, comprendiendo el aparato:

un procesador,

55 en el que el procesador está configurado para controlar el aparato para realizar el método de una cualquiera de las reivindicaciones 1 a 13.

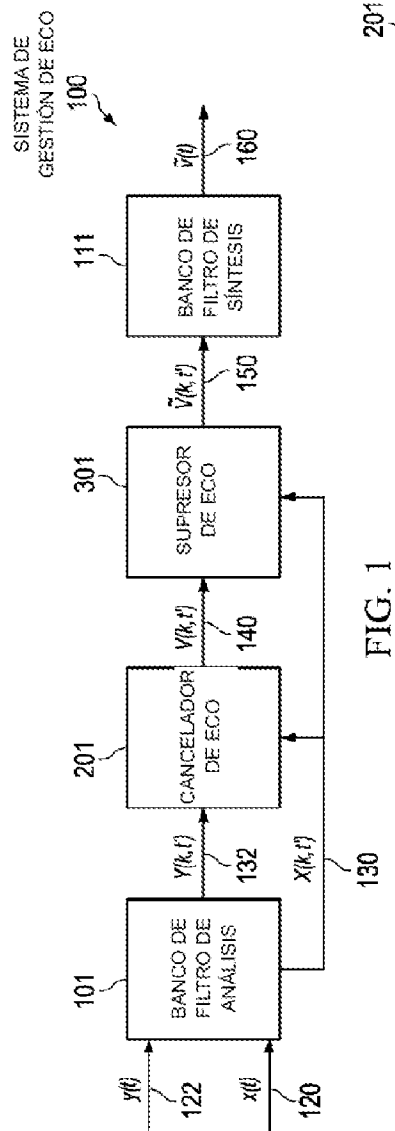


FIG. 1

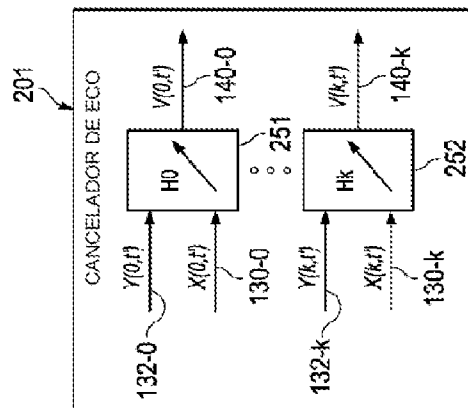


FIG. 2

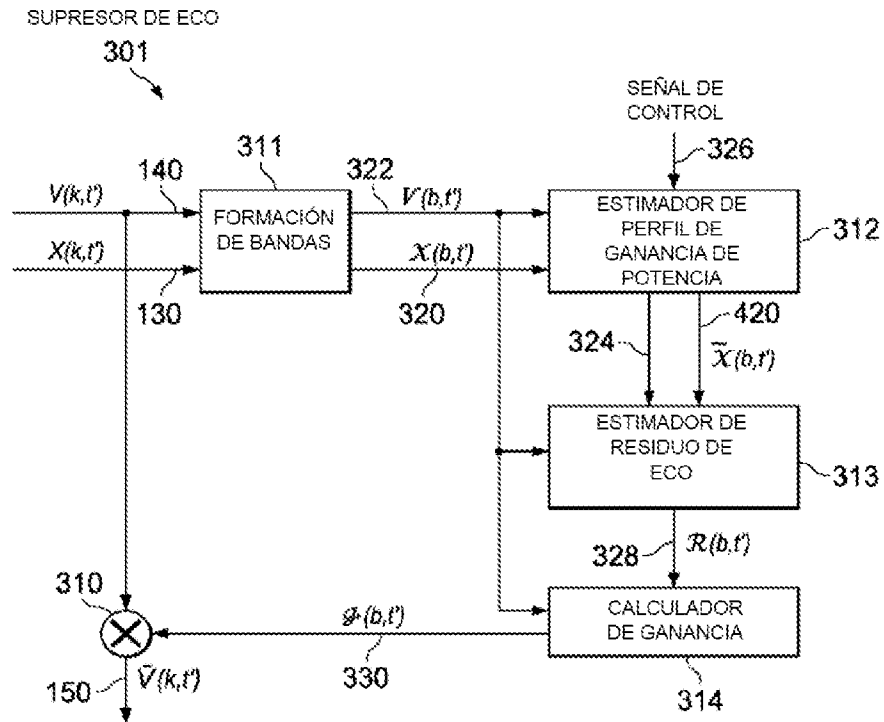


FIG. 3

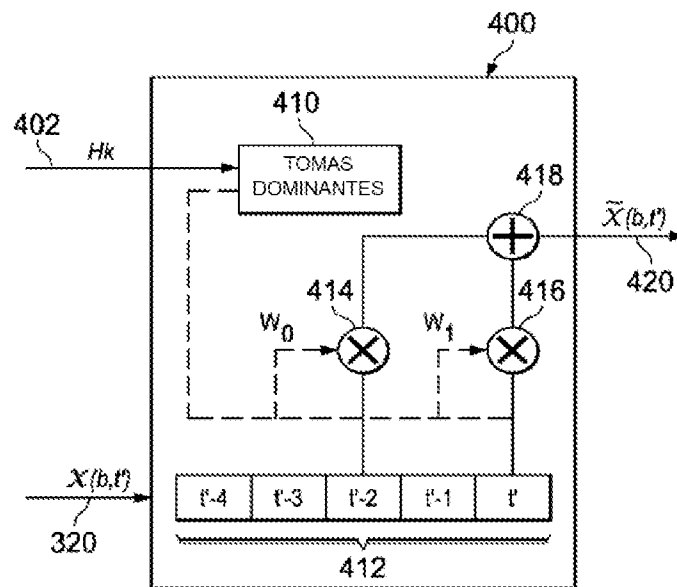


FIG. 4

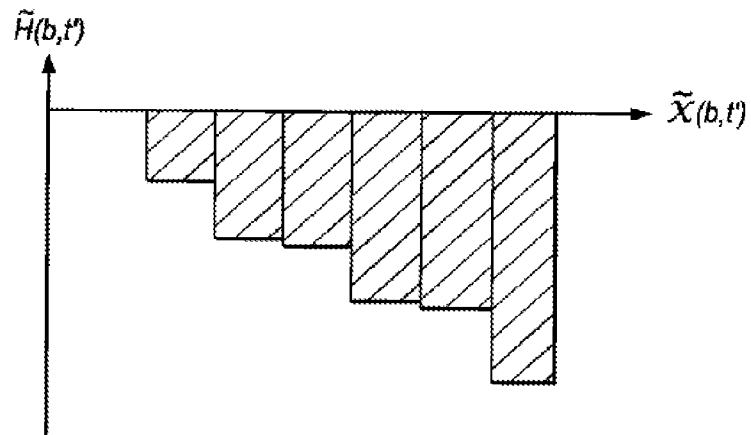


FIG. 5

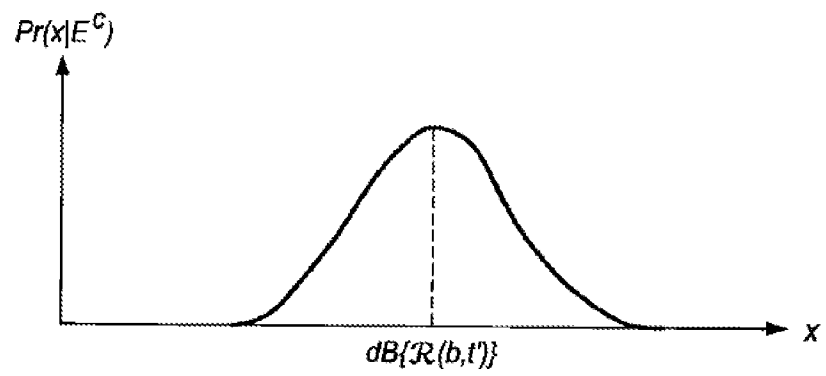
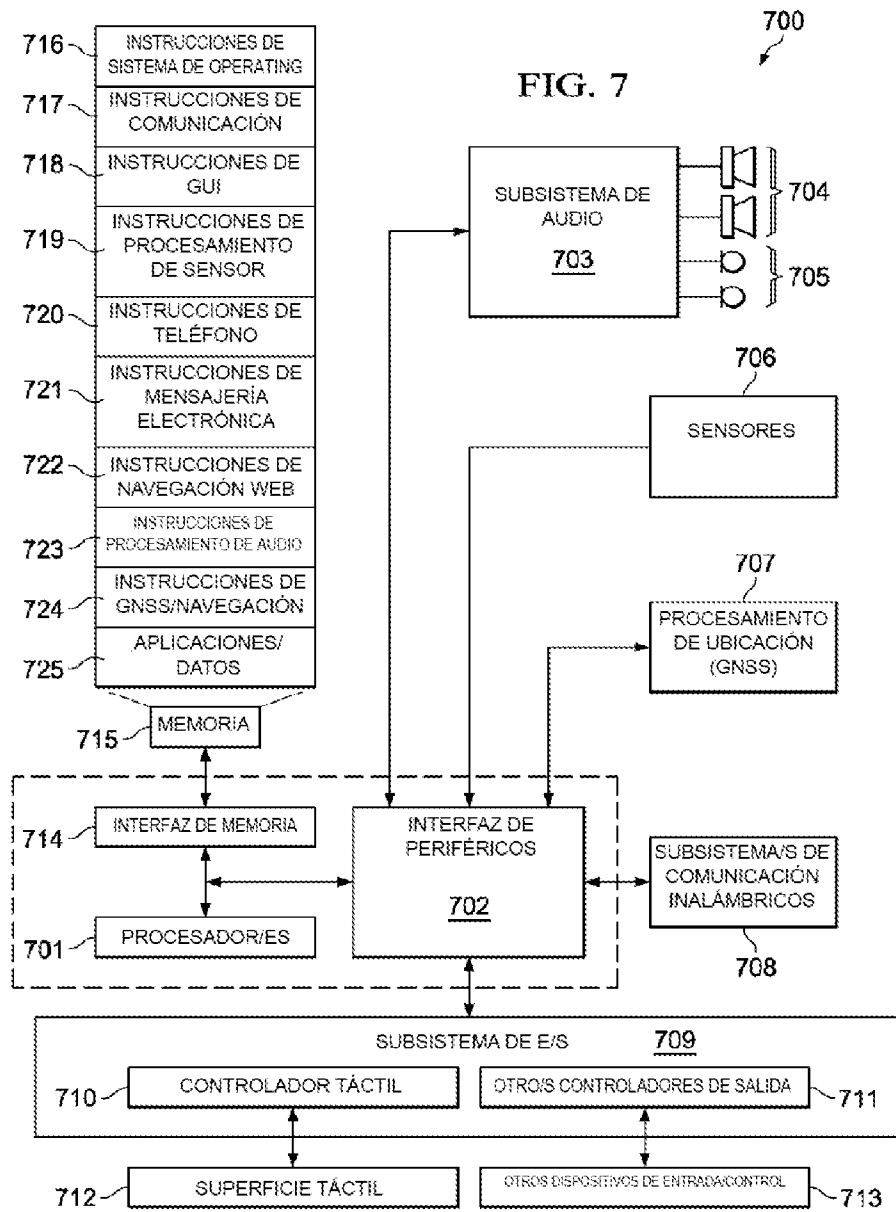


FIG. 6





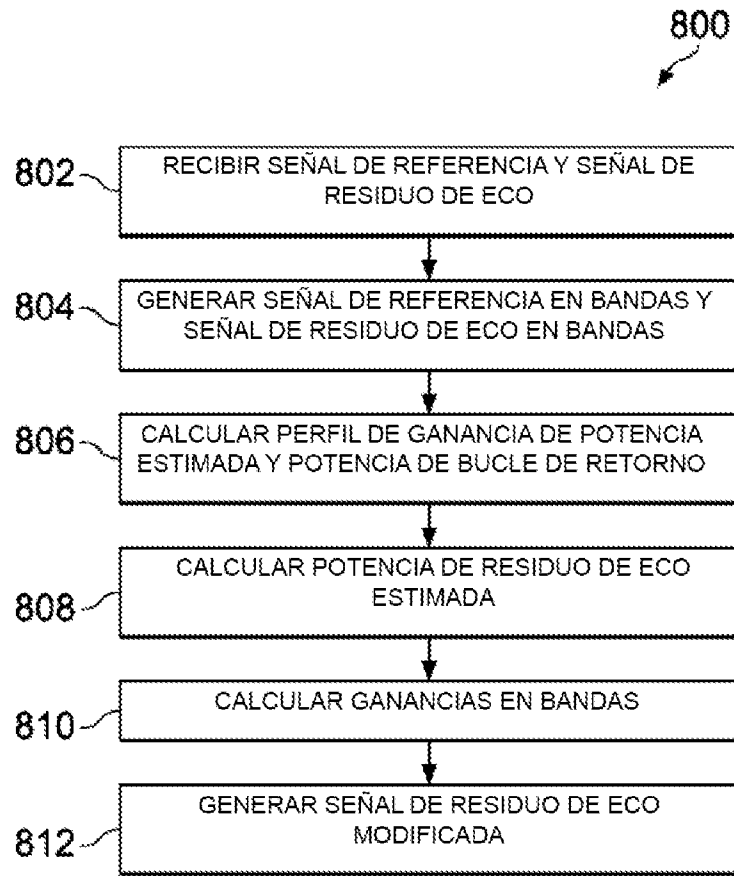


FIG. 8