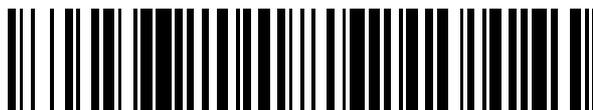


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 528 043**

51 Int. Cl.:

**H04M 9/08** (2006.01)

**H04R 3/00** (2006.01)

**G10L 21/0208** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **06.10.2011 E 11767430 (9)**

97 Fecha y número de publicación de la concesión europea: **10.12.2014 EP 2633522**

54 Título: **Supresión de eco que comprende el modelado de componentes de reverberación tardía**

30 Prioridad:

**10.12.2010 EP 10194586**

**25.10.2010 US 406246 P**

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

**03.02.2015**

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR  
FÖRDERUNG DER ANGEWANDTEN  
FORSCHUNG E.V. (100.0%)  
Hansastraße 27c  
80686 München, DE**

72 Inventor/es:

**KUECH, FABIAN;  
SCHMIDT, MARKUS;  
FALLER, CHRISTOF y  
FAVROT, ALEXIS**

74 Agente/Representante:

**PONTI SALES, Adelaida**

**ES 2 528 043 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 eco que comprende el modelado de componentes de reverberación tardía

## 5 Descripción

[0001] Las formas de realización de la presente invención se refieren a aparatos y procedimientos para calcular coeficientes de filtrado para un filtro adaptativo para filtrar una señal de micrófono a fin de suprimir un eco causado por una señal de altavoz, lo cual puede ser utilizado, por ejemplo, en el ámbito de los sistemas de conferencias, sistemas videoconferencia u otros sistemas de conferencia bidireccional.

[0002] Cuando un micrófono capta tonos, sonidos y ruidos procedentes de un altavoz que se encuentre en la misma sala o en el mismo entorno acústico, se producen ecos acústicos. En los sistemas de telecomunicaciones, esta señal de realimentación acústica se retransmite al abonado que se encuentra al otro lado de la línea, el cual percibe una versión retardada de sus propias palabras. En este contexto, las señales de eco crean una distracción que llega a resultar muy molesta e incluso puede impedir la comunicación interactiva bidireccional. Además, los ecos acústicos pueden generar efectos de acople, entre otras inestabilidades del bucle de realimentación acústica.

[0003] El documento WO2006/111370A1 se refiere a un procedimiento y un aparato para eliminar un eco en una señal de audio multicanal. Los elementos acústicos de control de eco y supresión de ruido constituyen una parte importante de cualquier sistema de telecomunicaciones de manos libres como, por ejemplo, el teléfono o los sistemas de conferencia de audio y vídeo. El procedimiento, descrito en el documento, para procesar señales de audio multicanal de altavoz y al menos una señal de micrófono en este contexto incluye las etapas de transformación de la señal de micrófono de entrada en espectros microfónicos de entrada de tiempo corto; el cálculo de un espectro de tiempo corto de las señales de altavoz combinadas a partir de las señales del altavoz; el cálculo de un espectro de tiempo corto de la señal del micrófono combinada a partir de la señal de micrófono de entrada; la estimación de un espectro de magnitud o un espectro de potencia del eco en el espectro de tiempo corto de la señal de micrófono combinada; el cálculo de un filtro de ganancia para la modificación de la magnitud del espectro de tiempo corto de entrada de micrófono; la aplicación del filtro de ganancia a al menos un espectro de entrada de micrófono; y la conversión del espectro de entrada de micrófono filtrado al dominio temporal.

[0004] Los sistemas de supresión de eco y de cancelación de eco, tal como se denominan hoy en día y que, en general, también pueden englobarse en la denominación de sistemas de eliminación de eco, a menudo presentan el problema de que no responden ante diferentes componentes de sonido, tono y ruido y componentes de eco de manera óptima, a pesar del uso de filtros adaptativos. Por ejemplo, si un componente predomina frente a otro, se puede producir una supresión no óptima del eco de la señal del altavoz en la señal del micrófono de dicho sistema de comunicación. Por otra parte, en el caso de una composición diferente de los componentes de las distintas fuentes, pueden aparecer artefactos tonales al utilizar un sistema de supresión de eco o de cancelación de eco, y dichos artefactos tonales también resultan enormemente molestos.

[0005] El documento EP 1 429 315 A1 se refiere a un procedimiento y un sistema para suprimir ecos y ruido en entornos en condiciones acústicas variables, así como en condiciones de una gran proporción de realimentación. El sistema para cancelar ecos y ruidos en entornos con condiciones acústicas variables y una alta realimentación, como por ejemplo en el interior de un automóvil, que permite la comunicación audible entre los ocupantes, comprende uno o varios micrófonos para una señal de voz y convertidores A/D y D/A, medios de amplificación y filtrado que generan una señal eléctrica reproducida como una señal acústica por un altavoz, y un cancelador de eco que usa la señal y una señal eléctrica procedente del micrófono que integra la señal de realimentación de la señal, otra señal de voz, y ruido captado por el micrófono. El sistema se aplica en la eliminación de la realimentación, mediante la adición de un filtrado de una señal eléctrica de salida del dispositivo de cancelación, tras el tratamiento de la señal, que comprende un filtro variante con el tiempo que suprime el eco acústico residual no cancelado por el sistema y el ruido captado por los micrófonos.

[0006] El documento WO 2009/095161 A1 se refiere a un aparato y un procedimiento para calcular coeficientes de filtrado para la supresión de eco. Los coeficientes de filtrado se deben usar con un filtro adaptativo para filtrar una señal de micrófono con el fin de suprimir un eco causado por una señal de altavoz. El aparato incluye unos medios de extracción para extraer una señal de componente estacionario o una señal de componente no estacionario a partir de la señal del altavoz o a partir de una señal derivada de la señal del altavoz. El aparato también comprende unos medios de cálculo para calcular los coeficientes de filtrado del filtro adaptativo basándose en la señal de componente estacionario y la señal de componente no estacionario extraídas. El aparato y el procedimiento descritos en el documento WO 2006/095161 examinan la señal del extremo remoto en lo que respecta a las propiedades estadísticas de la señal de componente que constituye la señal del altavoz.

[0007] En el documento WO 2006/040734 A1 se describe un dispositivo de cancelación de eco que comprende un primer filtro adaptativo para producir una primera señal de cancelación de eco y1, un segundo filtro adaptativo para producir una segunda señal de cancelación de eco y2, y un postprocesador para suprimir cualquier resto de eco. El primer filtro adaptativo y el segundo filtro adaptativo están diseñados para cancelar,

respectivamente, una primera parte (por ejemplo, directa) de la respuesta al impulso del eco y una segunda parte (por ejemplo, difusa) de la respuesta al impulso del eco. El dispositivo se puede utilizar en un teléfono móvil.

5 **[0008]** Por tanto, partiendo de la citada técnica anterior, el objeto de la presente invención consiste en mejorar la calidad de sonido de un sistema de supresión de eco o sistema de cancelación de eco con un esfuerzo computacional moderado.

10 **[0009]** Este objeto se logra por medio de un aparato de acuerdo con las reivindicaciones 1, 10 y 11, un procedimiento de acuerdo con las reivindicaciones 12, 15 y 16, o un programa de acuerdo con la reivindicación 17.

Resumen

15 **[0010]** Las formas de realización del aparato, el procedimiento y el programa informático de acuerdo con los conceptos descritos en el presente documento se basan en el descubrimiento de que se puede obtener una mejora de la calidad de audio y/o una reducción del esfuerzo computacional mediante el modelado, en especial, de los componentes tardíos del eco mediante un modelo de decaimiento del eco. El modelo de decaimiento del eco puede valerle de cálculos recursivos que requieren, relativamente, poco esfuerzo computacional. El decaimiento del eco del entorno acústico en el que se usa el filtro adaptativo puede presentar propiedades variables que se pueden estimar a partir de muestras de la señal del micrófono y/o la señal del altavoz.

20 **[0011]** A diferencia de los componentes tempranos del eco, que suelen ser bastante similares al sonido original, los componentes tardíos de eco se vuelven cada vez menos nítidos y menos similares al sonido original. El aparato, el procedimiento y el programa informático de acuerdo con lo expuesto en el presente documento distinguen entre la supresión de componentes tempranos del eco, que requieren una mayor precisión, y la supresión de componentes tardíos del eco, que se pueden estimar de forma menos precisa a lo largo de todo el tiempo de reverberación. En lugar de explotar diferentes propiedades estadísticas de diversos elementos constitutivos de la señal, como por ejemplo estacionarios o no estacionarios, que son inherentes a los correspondientes elementos constitutivos de las señales originales del micrófono y del altavoz (es decir, no se han visto influidas por el entorno acústico), los conceptos descritos en el presente documento hacen uso de las propiedades del eco causado por la señal del altavoz. Estas propiedades referentes al eco están relacionadas con el entorno acústico 120 y, por tanto, varían al cambiar o reconfigurar el entorno acústico. No obstante, las propiedades referentes al eco no varían, o al menos no de manera significativa, cuando las señales del altavoz o del micrófono varían, por ejemplo, con respecto a una proporción entre componentes estacionarios y componentes no estacionarios.

35 **[0012]** Más adelante en la presente descripción se describirán otras formas de realización con respecto a sus estructuras y modos de funcionamiento. Dependiendo de la forma de realización de que se trate, el parámetro o los parámetros de decaimiento del eco se pueden estimar a partir de las respectivas señales, por ejemplo la señal del micrófono que se ha visto influida por una respuesta al sistema del entorno acústico. Un aparato de acuerdo con una forma de realización de los conceptos del presente documento puede comprender también unos medios de filtrado de supresión de eco configurados para filtrar la señal del micrófono en función de los coeficientes de filtrado. Un aparato de acuerdo con los planteamientos descritos en el presente documento puede comprender unos medios de estimación de eco para estimar un espectro de frecuencia de eco o un espectro de potencia de eco del eco contenido en la señal del micrófono. La estimación del espectro de frecuencia del eco o del espectro de potencia del eco se puede dividir en una estimación de componentes tempranos del eco y una estimación de componentes tardíos del eco. La estimación de componentes tardíos del eco puede abarcar un intervalo de tiempo relativamente largo tras un sonido original con un esfuerzo computacional relativamente bajo. La estimación de los componentes de eco tempranos puede ser relativamente precisa y puede abarcar un intervalo de tiempo relativamente corto tras el sonido original.

50 Breve descripción de las figuras

**[0013]** A continuación se explicarán de manera más detallada formas de realización de los conceptos descritos en el presente documento, haciendo referencia a los dibujos adjuntos. Por lo tanto, las formas de realización de los conceptos descritos en el presente documento se pueden entender mejor atendiendo a los siguientes dibujos:

- 55 La fig. 1 muestra un esquema general de un problema de eliminación de eco acústico;
- La fig. 2 muestra la estructura general de una respuesta al impulso habitual en una sala;
- 60 La fig. 3 muestra un diagrama de bloques básico del supresor de eco acústico en un dominio espectral de tiempo corto;
- La fig. 4 muestra un decaimiento exponencial de la reverberación;
- 65 La fig. 5 muestra un diagrama de bloques esquemático de un aparato de acuerdo con diversas formas de realización de los conceptos del presente documento;

La fig. 6 muestra un diagrama de bloques esquemático de otra forma de realización de acuerdo con los conceptos del presente documento;

5 La fig. 7 muestra un diagrama de bloques esquemático de otra forma de realización de acuerdo con los conceptos del presente documento;

La fig. 8 muestra un diagrama de bloques esquemático de otra forma de realización de acuerdo con los conceptos del presente documento;

10

La fig. 9 muestra un diagrama de bloques esquemático de otra forma de realización de acuerdo con los conceptos del presente documento;

15 La fig. 10 muestra un diagrama de tiempo que ilustra los impulsos originales y las correspondientes reverberaciones tardías;

La fig. 11 muestra un diagrama de tiempo de estimación de eco y modelado de reverberación; y

20 La fig. 12 muestra un diagrama de bloques de una forma de realización de los conceptos descritos en el presente documento para una aplicación multicanal.

#### Descripción detallada de las formas de realización

25 **[0014]** En la presente descripción, los números de referencia idénticos hacen referencia a medios, unidades y objetos que son idénticos en acción, función, estructura o similares en acción, función o estructura. Esto permite presentar de manera más breve y concisa las formas de realización de la presente invención, ya que los pasajes descriptivos se pueden trasladar de una forma de realización a otra, a menos que se indique lo contrario de manera explícita.

30 **[0015]** Además, en la presente descripción, se utilizarán números de referencia resumidos para medios, estructuras y objetos que aparezcan más de una vez dentro de una figura o forma de realización. Por ejemplo, las dos trayectorias indirectas 180-1, 180-2 que se muestran en la fig. 1 reciben distintos números de referencia en la fig. 1, pero si se hace referencia a las trayectorias indirectas de por sí, o si se describen características generales de los mismos, en la presente descripción solo se usará el número de referencia resumido 180. Este método también  
35 sirve para mejorar facilidad de comprensión y la concisión la presente descripción.

**[0016]** En primer lugar, antes de describir, en referencia a las figs. 2 a 12, diversas formas de realización de los conceptos descritos en el presente documento que permiten la supresión del eco acústico al tiempo que se usa una separación de componentes de señal estacionarios y no estacionarios, se ilustrará un esquema general de un  
40 problema de eliminación de eco acústico haciendo referencia a la fig. 1.

**[0017]** En un sistema de telecomunicaciones de manos libres con capacidad de comunicación bidireccional simultánea, normalmente se requiere un control de eco para suprimir el acople entre el altavoz y el micrófono.

45 **[0018]** La fig. 1 muestra un altavoz 100, que puede situarse, junto con un micrófono 110, en un entorno acústico 120, que puede ser, por ejemplo, una sala. Asimismo, el entorno acústico 120 también puede ser el interior de un vehículo.

**[0019]** En este contexto, una señal de altavoz 130, que en la fig. 1 también se indica como  $x[n]$  con un número  
50 entero que representa un índice de tiempo  $n$  se pone a disposición del altavoz 130. El micrófono 110 capta los ruidos, sonidos y tonos procedentes del entorno acústico 120 y genera una señal de micrófono 140, que también se indica como  $y[n]$  en la fig. 1. Tanto la señal del altavoz 130 como la señal del micrófono 140 se proporcionan como señales de entrada a una unidad de proceso de eliminación de eco 150, que proporciona, en una salida, una señal con eco suprimido 160 de la señal del micrófono 140, que también se indica como  $e[n]$  en la fig. 1.  
55

**[0020]** De este modo, la fig. 1 ilustra básicamente el problema del eco acústico que puede aparecer en sistemas de comunicación bidireccional. La señal del extremo remoto del sistema de telecomunicación emitido por el altavoz alcanza el micrófono siguiendo una trayectoria directa 170 y a través de las trayectorias reflejadas 180-1, 180-2, que también se denominan trayectorias indirectas. Por este motivo, el micrófono 110 no solo capta la voz en el extremo  
60 cercano, sino que registra el eco que después es realimentado hacia el usuario del extremo remoto.

**[0021]** Dicho de otro modo, la señal del altavoz  $x[n]$  se realimenta en la señal del micrófono  $y[n]$ . Idealmente, un proceso de eliminación de eco que se lleve a cabo en la unidad de proceso de eliminación de eco 150 elimina este eco al tiempo que permite el paso de la voz del extremo cercano del sistema de telecomunicación.

65

**[0022]** Un procedimiento convencional para hacer frente a este eco consiste en colocar un cancelador de eco

acústico (AEC) en paralelo con la trayectoria de propagación de la señal de eco, tal como se describe en la referencia [1]. En dicho cancelador de eco acústico, se estima una réplica digital de la señal de eco y posteriormente se resta de la señal del micrófono medida u observada. Las soluciones estándar para cancelar los ecos acústicos se basan en la suposición de que la trayectoria del eco se puede modelar mediante un filtro FIR lineal (FIR: respuesta al impulso finita), y los canceladores de eco se adaptan en consecuencia, como también se describe en la referencia [1]. Debido a que la trayectoria del eco suele ser desconocida y, más aún, a que puede cambiar durante el tiempo de funcionamiento, el filtro lineal de dicho cancelador de eco acústico se suele construir de manera adaptativa. Para poder modelar las típicas trayectorias del eco, se hace uso de filtros FIR de longitudes que suelen ser de centenares de milisegundos, lo cual, con respecto a la respectiva tasa de muestreo, también conlleva un alto grado de complejidad computacional.

**[0023]** En la práctica, las atenuaciones de eco que se pueden lograr para estas soluciones convencionales resultan con frecuencia insuficientes por diversos motivos. Los diversos motivos incluyen, por ejemplo: largos tiempos de reverberación (efecto de cola de eco), que dan lugar a un inframodelado de la trayectoria del eco, componentes de eco no lineales causados, por ejemplo, por efectos de vibración o el comportamiento no lineal de aparatos de audio de coste particularmente bajo, y problemas de convergencia en el caso de un elevado índice de cambio con respecto a las trayectorias de eco, tal como se describe en la referencia [2]. Por lo tanto, los canceladores de eco acústico se combinan con postprocesadores no lineales para eliminar los ecos residuales y componentes del eco que el cancelador de eco no pudo eliminar, tal como se describe en la referencia [3]. La supresión de ecos residuales se lleva a cabo habitualmente de manera selectiva en cuanto a la frecuencia, tal como se describe en la referencia [4]. De hecho, prácticamente todos los canceladores de eco acústico usan dichos postprocesadores porque fallan demasiado a menudo como para reducir el eco lo suficiente como para que resulte inaudible.

**[0024]** Recientemente, se han propuesto varios supresores de eco acústico para el dominio de subbandas, referencias [5, 6], que presentan semejanzas con los postprocesadores no lineales mencionados anteriormente, pero no necesitan un cancelador de eco acústico ni tampoco estimar la respuesta al impulso de la trayectoria del eco. Se afirma que estos sistemas son de baja complejidad computacional y que son robustos, al tiempo que logran un alto grado de bidireccionalidad.

**[0025]** El esquema del supresor de eco propuesto en la referencia [6] aplica una transformada de Fourier de tiempo corto (STFT) para calcular los espectros de las señales del altavoz y del micrófono. A la correspondiente señal del altavoz se le aplica un valor de retardo  $d$  entre los de las transformadas de Fourier de tiempo corto de las dos señales, escogiéndose dicho valor de retardo de manera que se tenga en cuenta la mayor parte del efecto de la respuesta al impulso de la trayectoria del eco.

**[0026]** Posteriormente, se estima una función real de estimación de eco, que imita el efecto de la trayectoria del eco inicial. Para obtener un espectro de magnitud de eco estimado, se aplican a los espectros del altavoz el valor de retardo estimado y la función de estimación de eco. Usando la estimación del espectro de magnitud de eco, se calcula un filtro de supresión de eco de valores reales y se aplica al espectro de la señal del micrófono para suprimir el eco.

**[0027]** Los sistemas de supresión de eco acústico mencionados anteriormente consideran el efecto de la reverberación tardía en la sala a la hora de estimar el espectro de magnitud del eco, es decir, se toma en consideración el efecto de la reverberación tardía del mismo modo que el efecto de las reflexiones tempranas en la trayectoria del eco, aunque las reverberaciones tardías suelen presentar diferentes propiedades en comparación con las reflexiones tempranas. Para hacer frente a los componentes del eco resultantes de la reverberación tardía, en ocasiones se utiliza un suavizado temporal del filtro de eliminación de eco. En la práctica, este tipo de solución puede aumentar la distorsión no deseada en el extremo cercano.

**[0028]** Además, las reverberaciones tardías suelen tender a ser más aleatorias que las reflexiones tempranas, lo cual puede deberse a una superposición de diferentes trayectorias de propagación y reflexión entre el altavoz y el micrófono en el entorno acústico 120.

**[0029]** La señal del extremo remoto  $x[n]$  emitida por el altavoz se dirige hacia el micrófono tanto directamente como a través de trayectorias reflejadas. La señal del micrófono  $y[n]$  comprende la voz y el ruido locales del extremo cercano  $w[n]$  y el eco que, de ese modo, se realimenta hacia el usuario del extremo remoto:

$$y[n] = h[n] * x[n] + w[n], \quad (1)$$

en la que  $h$  es la respuesta al impulso de la sala y  $*$  indica una convolución.

**[0030]** Como se ilustra en la fig. 2, la respuesta al impulso de la sala  $h$  se puede descomponer en un sonido directo, una reflexión temprana (o varias reflexiones tempranas) y reverberación tardía. Esta última también se denomina cola de eco. En procedimientos propuestos anteriormente, descritos en la referencia [6], solo se usa un

parámetro de retardo global  $d$  y una función de estimación de eco  $g$  para modelar la trayectoria del eco con el fin de captar el sonido directo y las reflexiones tempranas. Las reflexiones tardías no se modelan en los procedimientos propuestos anteriormente, pero se toman en consideración mediante el suavizado en el tiempo del filtro de supresión de eco. De este modo, la señal del micrófono  $y[n]$  se puede expresar mediante la siguiente fórmula:

5

$$y[n] = g[n] * x[n - d] + w[n]. \quad (2)$$

**[0031]** El parámetro de retardo global  $d$  se puede explicar por el hecho de que la trayectoria de mínima distancia de la señal del altavoz hacia el micrófono viene dada por la distancia desde el altavoz al micrófono, es decir, la señal del altavoz debe recorrer al menos la distancia desde el altavoz hasta el micrófono a través de la trayectoria directa. Las reflexiones tempranas suelen ser relativamente similares a la señal original del altavoz y al sonido directo debido a que las reflexiones tempranas corresponden a varias trayectorias de reflexión primaria 180 (fig. 1), y este número suele ser relativamente pequeño en un entorno acústico normal. La cola de eco se genera a causa de diversos efectos, como la dispersión de sonido, los efectos de resonancia y las reflexiones multipunto, por citar algunos. En el micrófono 110 se produce una superposición de la señal del altavoz  $x[n]$  que ha recorrido una pluralidad de trayectorias diferentes en el entorno acústico 120. Debido al creciente número de trayectorias de propagación posibles entre el altavoz 100 y el micrófono 110 con un creciente retardo de la propagación (es decir, solo hay una única trayectoria directa 170 y tan solo un número limitado de trayectorias de una sola reflexión 180, pero existe una cantidad inmensa de trayectorias de reflexión múltiple), las señales del altavoz superpuestas que alcanzan el micrófono 110 se pueden considerar sustancialmente aleatorias más allá de un intervalo de tiempo dado posterior a la recepción del sonido directo. Los efectos de amortiguamiento del entorno acústico 120 provocan un decaimiento en la reverberación tardía que se puede modelar como un decaimiento exponencial.

**[0032]** Como se ilustra en la fig. 3, se calculan espectros de transformada de Fourier discreta de tiempo corto (STFT), a partir de la señal del altavoz 420 y la señal del micrófono 430. Los intervalos de tiempo considerados para la transformación STFT de la señal del altavoz 420 y la señal del micrófono 430 se indican mediante los signos de referencia 440 y 450, respectivamente, de la fig. 3. Entre la señal del altavoz 420 y la señal del micrófono 430 se introduce un retardo  $d$  (signo de referencia 460), ya que la señal acústica emitida por el altavoz 100 requiere un determinado tiempo mínimo para propagarse hasta el micrófono 110. El retardo  $d$  entre las ventanas STFT aplicadas a la señal del altavoz 420 se escoge de manera que se capte la mayor parte de la energía de la respuesta al impulso de la trayectoria del eco. La representación de la ecuación (2) en el dominio STFT viene dada por:

$$Y[k, m] = G[k, m]X_d[k, m] + W[k, m], \quad (3)$$

en la que  $k$  es el índice del bloque de tiempo y  $m$  representa el índice de frecuencia.  $X_d[k, m]$  es la correspondencia en el dominio de la STFT de la señal del altavoz retardada  $x[n-d]$ . A partir de la ecuación (3), también se pone de manifiesto que la reverberación tardía de la trayectoria del eco acústico no queda reflejada en el modelo, ya que el intervalo temporal que abarca  $G[k, m]$  corresponde a la longitud de bloque de un bloque STFT. Estos bloques suelen tener una longitud de 10 a 30 ms, muy inferior a los tiempos habituales de reverberación, que alcanzan valores de hasta varios centenares de ms.

**[0033]** La supresión del eco acústico propiamente dicha se lleva cabo modificando la magnitud de la STFT de la señal del micrófono  $Y[k, m]$ , al tiempo que se mantiene invariable su fase. Esto se puede expresar mediante la siguiente fórmula:

45

$$E[k, m] = H[k, m]Y[k, m]. \quad (4)$$

en la que el filtro de supresión de eco (ESF) 370 produce como salida un factor de ganancia positivo de valor real  $H[k, m]$ . En la aplicación discreta temporal y frecuencial (como ocurre al tratarse de una transformada de Fourier de tiempo corto), el valor  $H[k, m]$  suele ser un elemento de una matriz  $H$ , que resulta válido para el bloque SFTF que se considere en este momento.

**[0034]** El factor (o factores) de ganancia positiva de valor real  $H[k, m]$  se introduce en el filtro adaptativo 210 (SM: modificación espectral), el cual proporciona la señal del micrófono con eco suprimido  $E[k, m]$  en el dominio de la STFT. La transformación para volver a pasar del dominio de la STFT al dominio temporal se lleva a cabo mediante una transformada de Fourier inversa de tiempo corto (ISTFT) 300 para producir como salida la señal  $e[n]$ .

**[0035]** Volviendo al diagrama de bloques de la fig. 3, el filtro de supresión de eco 370 usa una estimación de la función de transferencia del entorno acústico  $G[k, m]$ , y la estimación se designa con la expresión  $G[k, m]$ . Los valores óptimos para el filtro de supresión de eco 370 y su salida  $H[k, m]$  se pueden obtener minimizando la contribución de los componentes del eco en la señal de salida  $E[k, m]$  en el sentido del error medio cuadrático (MSE). Se obtiene una estimación del espectro de potencia del eco aplicando el retardo estimado y la función de estimación de eco (EEF) 470 al espectro de potencia de la señal del altavoz, por ejemplo:

$$|\hat{P}[k, m]|^2 = |\hat{G}[k, m]|^2 |X_d[k, m]|^2 \quad (5)$$

[0036] Obsérvese que, en la práctica, la función de transferencia de potencia del eco  $|G[k, m]|^2$  es desconocida y debe ser sustituida por su estimación:  $|\hat{G}[k, m]|^2$ .

[0037] La estimación de la función de transferencia del eco  $G[k, m]$  o de la función de transferencia de potencia del mismo dentro de la función de estimación del eco 470 se puede llevar a cabo, por ejemplo, del modo descrito en la referencia [11]. La estimación se puede obtener a partir de la correlación de los espectros de potencia de las señales del micrófono y del altavoz 420, 430, respectivamente.

$$|\hat{G}[k, m]|^2 = \frac{E\{|X_d[k, m]|^2 |Y[k, m]|^2\}}{E\{|X_d[k, m]|^2 |X_d[k, m]|^2\}} \quad (6)$$

[0038] Obsérvese que, en la práctica, la esperanza matemática  $E\{\}$ , que se usa aquí, se puede sustituir por las correspondientes medias de tiempo corto. A modo de ejemplo, consideremos:

$$\Phi_{AB}[k, m] = E\{A[k, m]B[k, m]\}. \quad (7)$$

[0039] La media de tiempo corto  $\hat{\Phi}_{AB}[k, m]$  correspondiente a  $\Phi_{AB}[k, m]$  se puede obtener, por ejemplo, mediante un suavizado recursivo de acuerdo con la siguiente expresión:

$$\hat{\Phi}_{AB}[k, m] = (1 - \alpha_{avg})\hat{\Phi}_{AB}[k - 1, m] + \alpha_{avg}A[k, m]B[k, m] \quad (8)$$

[0040] El factor  $\alpha_{avg}$  determina el grado de suavizado a lo largo del tiempo y se puede ajustar para adaptarse a cualquier necesidad.

[0041] Se ha propuesto una solución práctica para el cálculo del filtro de supresión de eco 370, por ejemplo, en la referencia [6]:

$$H[k, m] = \frac{|Y[k, m]|^2 - \beta|\hat{Y}[k, m]|^2}{|Y[k, m]|^2}, \quad (9)$$

en la que  $\beta$  representa un parámetro de diseño para controlar la cantidad de eco que se ha de suprimir, tal como se describe en la referencia [12].

[0042] El siguiente párrafo resume la estructura de un aparato 200 de acuerdo con una forma de realización relativamente más compleja de los conceptos descritos en el presente documento. Esta forma de realización comprende posiblemente más componentes y/o características de los necesarios para una aplicar de manera básica los planteamientos descritos.

[0043] Un modelo de decaimiento del eco proporciona parámetros de decaimiento del eco a unos medios de determinación de la función de estimación de eco. Los medios de determinación de la función de estimación de eco determinan una función de estimación de eco que tiene en cuenta los parámetros de decaimiento del eco. Después se aplica la función de estimación de eco a la señal del altavoz con el fin de determinar una estimación del eco tal como se produce en el entorno acústico real. Posteriormente, unos medios de determinación del filtro de supresión de eco (también denominados medios de cálculo 270 en la presente descripción) pueden hacer uso de la estimación del eco. Los medios de determinación del filtro de supresión de eco se implementan para determinar un filtro que, al aplicarlo a la señal del micrófono, suprime el componente del eco en la señal del micrófono causado por la señal del altavoz. Para ello, los coeficientes de filtrado que constituyen el filtro de supresión de eco se suministran a una estructura de filtrado correspondiente que recibe la señal del micrófono como entrada y produce una señal con eco suprimido como salida.

[0044] En la fig. 4 se muestra una envolvente  $q[n]$  de la parte tardía de la reverberación de una respuesta al impulso de la sala, si se adopta el modelo de decaimiento exponencial:

$$q[n] = e^{-\frac{n}{\tau}}. \quad (10)$$

[0045] Considerando las dos muestras temporales  $d_\tau$  y  $d_\tau + \Delta d_\tau$ , y los correspondientes valores  $q[d_\tau]$  y  $q[d_\tau + \Delta d_\tau]$ , respectivamente, la constante temporal  $\tau$  se puede estimar basándose en la proporción:

$$\Delta q = \frac{q[d_\tau + \Delta d_\tau]}{q[d_\tau]}, \quad (11)$$

que se puede escribir del siguiente modo:

$$\Delta q = e^{-\frac{\Delta d_\tau}{\tau}}. \quad (12)$$

[0046] Al despejar la constante temporal  $\tau$ , se obtiene:

$$\tau = -\frac{\Delta d_\tau}{\ln \Delta q}. \quad (13)$$

10

[0047] En el caso de que el valor de la envolvente  $q$  se represente en el dominio temporal-frecuencial tras una transformada de Fourier de tiempo corto (STFT), cada banda de frecuencia puede presentar una constante temporal individual  $\tau_m$  denominada constante temporal de bloque. El cálculo de la constante temporal de bloque  $\tau_m$ , tal como se usa en el modelo de dominio de STFT de acuerdo con la ecuación (29) que se explica más adelante, se puede

15

$$Q[k, m] = e^{-\frac{k}{F_s \tau_m}}. \quad (14)$$

[0048] Considerando las dos muestras temporales  $d_\tau$  y  $d_\tau + \Delta d_\tau$ , y los correspondientes valores de envolvente  $Q[d_\tau, m]$  y  $Q[d_\tau + \Delta d_\tau, m]$ , respectivamente, la constante temporal  $\tau_m$  se puede estimar basándose en la proporción

20

$$\Delta Q[m] = \frac{Q[d_\tau + \Delta d_\tau, m]}{Q[d_\tau, m]}, \quad (15)$$

$$\Delta Q[m] = e^{-\frac{\Delta d_\tau}{F_s \tau_m}}, \quad (16)$$

25

$$\tau_m = -\frac{\Delta d_\tau}{F_s \ln(\Delta Q[m])}. \quad (17)$$

[0049] En el siguiente apartado, se presenta un procedimiento para determinar  $\Delta Q[m]$  para un valor dado de  $\Delta d_\tau$ . El procedimiento se basa en un planteamiento similar al usado para determinar la función de estimación de eco en la ecuación (6):

30

$$|\tilde{Q}_{d_\tau}[k, m]|^2 = \frac{\mathbb{E}\{X_{d_\tau}[k, m]Y[k, m]\}}{\mathbb{E}\{X_{d_\tau}[k, m]X_{d_\tau}[k, m]\}}. \quad (18)$$

$$|\tilde{Q}_{d_\tau + \Delta d_\tau}[k, m]|^2 = \frac{\mathbb{E}\{X_{d_\tau + \Delta d_\tau}[k, m]Y[k, m]\}}{\mathbb{E}\{X_{d_\tau + \Delta d_\tau}[k, m]X_{d_\tau + \Delta d_\tau}[k, m]\}}. \quad (19)$$

35

[0050] A partir de lo cual calculamos:

$$\Delta Q[k, m] = \sqrt{\frac{|\tilde{Q}_{d_\tau + \Delta d_\tau}[k, m]|^2}{|\tilde{Q}_{d_\tau}[k, m]|^2}}. \quad (20)$$

[0051] El tiempo de reverberación  $RT_{60}$  para el modelo de decaimiento exponencial se puede calcular como el tiempo  $n$ , en el que el término  $q[n]$  de la ecuación (10) se ha reducido en -60dB:

40

$$e^{-\frac{RT_{60}}{\tau}} = 10^{-6}. \quad (21)$$

[0052] Al resolver para la constante temporal  $\tau$  deseada, se obtiene:

45

$$\tau = \frac{RT_{60}}{\ln(10^6)} \quad (22)$$

**[0053]** Usando esta constante temporal  $\tau$ , se puede determinar un factor de proporcionalidad  $\alpha$  adecuado. El cálculo de un correspondiente factor de proporcionalidad  $\alpha_m$  en el dominio de la STFT a partir del tiempo de reverberación  $RT_{60}$  se puede llevar a cabo de manera análoga.

**[0054]** En lugar de medir el tiempo de reverberación  $RT_{60}$  del modo descrito anteriormente, observando las dos muestras temporales  $d_\tau$  y  $d_\tau + \Delta d_\tau$ , y los correspondientes valores de envolvente  $Q[d_\tau, m]$  y  $Q[d_\tau + \Delta d_\tau, m]$ , respectivamente, también es posible proporcionar valores a priori para el tiempo de reverberación  $RT_{60}$ . Los valores a priori se pueden proporcionar como una función del tipo del entorno acústico 120, por ejemplo, el interior de un coche, una oficina, una sala de conferencias, un auditorio o una sala de conciertos. El tiempo de reverberación de una sala de conciertos (de 2 a 4 segundos) suele ser dos órdenes de magnitud más largo que el tiempo de reverberación del interior de un coche (alrededor de 40 ms). Incluso sin haber medido el tiempo de reverberación real  $RT_{60}$  para un entorno acústico dado, los valores a priori suelen ofrecer una buena aproximación del tipo de entorno acústico correspondiente. El aparato 200 para calcular los coeficientes de filtrado puede comprender un conmutador de selección o similar, mediante el cual el usuario pueda escoger, por ejemplo, un tiempo de reverberación corto, un tiempo de reverberación intermedio o un tiempo de reverberación largo.

**[0055]** Para cada banda de frecuencia considerada en el modelo del dominio de la STFT, se puede determinar una constante temporal individual  $\tau_m$ . La determinación de una pluralidad de constantes temporales  $\tau_m$  individuales para las diversas bandas de frecuencia refleja el hecho de que los tiempos de reverberación observados dentro de las diversas bandas de frecuencia pueden diferir entre sí debido a una respuesta del entorno acústico 120 dependiente de la frecuencia. Por ejemplo, en salas grandes se pueden observar tiempos de reverberación largos para bandas de frecuencia bajas a medias, mientras que las frecuencias más altas tienden a presentar tiempos de reverberación más cortos. La posibilidad alternativa consiste en determinar una única constante temporal  $\tau$  como una media en todas las bandas de frecuencia.

**[0056]** La fig. 5 muestra una primera forma de realización de un aparato 200 para calcular coeficientes de filtrado para un filtro adaptativo 210, en la que dicho aparato utiliza, al igual que el procedimiento asociado, unos medios de modelado del decaimiento del eco 465 con el fin de mejorar la supresión del eco y, de ese modo, mejorar la calidad de audio que se puede lograr. Por consiguiente, las formas de realización de la presente invención permiten aplicar diferentes modos de gestionar la supresión de señales en función de sus propiedades y características de eco, lo que da lugar a una supresión de eco más eficaz y menos propensa a introducir artefactos.

**[0057]** Como introducción, en primer lugar se explicará un diagrama de bloques 200 de acuerdo con una forma de realización de la presente invención, antes de dar más detalles en lo referente a la aplicación, en relación con las figs. 6 a 8, con respecto a diversas formas de realización de la presente invención. En este contexto, resulta útil señalar que aunque en las figuras se muestren y describan diagramas de aparatos de acuerdo con las formas de realización de la presente invención, dichos diagramas de bloques también se pueden entender como diagramas de flujo de los correspondientes procedimientos que ilustran las correspondientes etapas del procedimiento a la vez que indican las direcciones del flujo. Dicho de otro modo, los diagramas de bloques ilustrados en relación con la presente descripción también se pueden entender como diagramas de flujo correspondientes que reflejan cada una de las etapas del procedimiento de cada uno de los medios y unidades.

**[0058]** El aparato 200 se puede implementar, por ejemplo, en la unidad de procesamiento de eliminación de eco 150 que se muestra en la fig. 1.

**[0059]** Antes de explicar de manera más detallada el modo de funcionamiento de la forma de realización, que se muestra en la fig. 5, de un aparato 200, también cabe mencionar que las formas de realización de la presente invención se pueden implementar básicamente tanto en circuitos discretos como en circuitos integrados o en otros circuitos más complejos. Por ejemplo, también se pueden aplicar formas de realización de la presente invención en medios de procesamiento de datos, es decir, procesadores, sistemas integrados (SOC = sistema integrado en un chip), circuitos integrados para aplicaciones específicas (ASIC) u otros circuitos integrados y procesadores de uso especial. En este contexto, es muy posible que se empleen partes de circuito idénticas de los respectivos medios de procesamiento de datos en otros medios diferentes de manera temporalmente consecutiva. Por ejemplo, la misma puerta lógica de una unidad aritmético-lógica (ALU) de un procesador, en primer lugar, dentro del contexto de la funcionalidad de los medios de modelado del decaimiento del eco 465 y, en segundo lugar, dentro del contexto de la funcionalidad de los medios de cálculo 270. No obstante, los dos medios difieren en particular con respecto a otras características como, por ejemplo, en el caso mencionado anteriormente, en referencia a diferentes instrucciones de control, que en conjunto definen, por ejemplo, los diferentes medios. Por lo tanto, resulta bastante posible una superposición parcial o completa de las implementaciones de la ingeniería de circuitos de diferentes medios.

**[0060]** No es esta la única razón por la que, en la presente descripción, los medios componentes y estructuras que se acoplan entre sí se entienden como medios, componentes y estructuras que están directa o indirectamente

interconectados. Si hay una implementación basada en unos medios de procesamiento de datos, por ejemplo, se puede llevar a cabo una conexión mediante una posición de almacenamiento de una memoria cuyo resultado intermedio sea una forma de señal enclavada en la misma.

5 **[0061]** Además, no obstante, las formas de realización de la presente invención no se limitan, básicamente, a aplicaciones digitales, aunque más adelante en la presente descripción se describirán aplicaciones principalmente digitales. Por ejemplo, en principio es factible una aplicación analógica o una aplicación mixta que comprenda componentes analógicos y digitales. En tal caso, se pueden añadir convertidores A/D o D/A (analógico/digital y digital/analógico), por ejemplo, para tener la posibilidad de llevar a cabo una transformación de las señales de un  
10 tipo en otro.

**[0062]** La fig. 5 muestra un diagrama de bloques esquemático de un aparato 200 de acuerdo con diversas formas de realización de los conceptos del presente documento. El aparato 200 comprende unos medios de modelado del decaimiento del eco 465. Los medios de modelado del decaimiento del eco  $\tau$  a los medios de cálculo 270. El parámetro de decaimiento del eco  $\tau$  modela el tiempo de reverberación del entorno acústico 120. Los medios de cálculo 270 determinan una función de filtrado representada por los coeficientes de filtrado  $H[k,m]$ . Los coeficientes de filtrado  $H[k,m]$  se envían al filtro adaptativo 210, que aplica el filtro resultante a la señal del micrófono  $y[n]$ . El filtro adaptativo 210 proporciona como una salida la señal con el eco suprimido  $e[n]$  a los medios de procesamiento de señales  
15 posteriores al filtro adaptativo 210. Aunque la señal del micrófono  $y[n]$ , la señal con el eco suprimido  $e[n]$  y los coeficientes de filtrado  $H[k,m]$  se representan en la fig. 5 como señales digitales de tiempo discreto, también pueden ser señales analógicas de tiempo continuo.

**[0063]** Con respecto al modo de funcionamiento del aparato 200, tal como se ilustra en la fig. 5, existe la  
25 posibilidad de transformar una señal del altavoz en un dominio relacionado con la frecuencia mediante los medios convertidores de tiempo/frecuencia que pueden estar presentes de manera opcional (fig. 6). Los medios convertidores de tiempo/frecuencia garantizan, en el caso de una aplicación que opere con bloques de datos (tramas o *frames*), una conversión a una representación espectral del respectivo bloque de datos, de manera que la representación espectral proporcionada en la salida de los medios convertidores de tiempo/frecuencia 230 (fig. 6)  
30 correspondan al bloque de datos en el dominio temporal. Dependiendo de la implementación concreta, se pueden usar unos medios convertidores basados en la transformada de Fourier, unos medios convertidores basados en subbandas o unos medios convertidores basados en un QMF (QMF: filtro de espejo en cuadratura) dentro del contexto de los medios convertidores de tiempo/frecuencia. Independientemente del modo concreto de funcionamiento de los medios convertidores de tiempo/frecuencia implementados, dichos medios convertidores  
35 convierten la señal (presente en el dominio temporal) recibida en su entrada en una pluralidad de señales de paso de banda. Cada señal de paso de banda lleva asociada una frecuencia característica, que puede ser, por ejemplo, una frecuencia central, una frecuencia de corte inferior de la banda respectiva o una frecuencia de corte superior de la banda respectiva. Dependiendo de la implementación concreta, las señales individuales de paso de banda pueden llevar asociada más de una frecuencia característica u otro parámetro característico.

40 **[0064]** Antes de describir de manera más detallada los componentes individuales del diagrama del bloques, que se muestra en la fig. 5, del aparato 200 en relación con las figs. 6 a 8, en este punto cabe señalar que el procesamiento de la señal del altavoz o de la señal derivada de la señal del altavoz se puede llevar a cabo generalmente en el dominio relacionado con la frecuencia, de manera que una señal individual de paso de banda asociada, una  
45 pluralidad de señales de paso de banda, múltiples señales de paso de banda o todas las señales de paso de banda sean procesadas de acuerdo con ese planteamiento.

**[0065]** También cabe señalar que los medios y filtros individuales pueden funcionar mientras se usan valores relacionados con la energía, por ejemplo, dependiendo de las aplicaciones concretas. Un valor relacionado con la  
50 energía es un valor que se forma como una potencia de un valor de base real con un exponente par, o se forma como una potencia de una magnitud de un valor (valor absoluto) con cualquier exponente. Si, por ejemplo, se procesan espectros de tiempo corto en los filtros individuales o en los medios individuales, dichos espectros pueden operar con valores relacionados con la energía; por ejemplo, con valores de energía que se forman como cuadrados de la magnitud de los coeficientes espectrales asociados. Asimismo, se pueden usar espectros de magnitud, es  
55 decir, valores absolutos de los respectivos coeficientes espectrales, con exponente 1. Dicho de otro modo, se pueden usar valores que sean proporcionales a  $|z|^m$ , siendo  $m$  un número positivo, por ejemplo, un número natural, como valores relacionados con la energía, partiendo de cualquier valor  $z$  que sea un valor real o un valor complejo. En el caso de un valor real  $z$ , los valores que sean proporcionales a  $z^{2m}$  se pueden usar adicionalmente como valores relacionados con la energía.

60 **[0066]** Con respecto al modelado de señales, es preciso señalar que una estimación del espectro de contribución del eco o del espectro de potencia del eco en la que se usa una función de estimación de eco no suele resultar muy precisa en las condiciones existentes en la práctica, ya que solo se puede considerar una parte de la longitud real de la trayectoria del eco. Para evitar que estas inexactitudes den lugar a ecos residuales, se calculan filtros de  
65 eliminación de eco para que supriman los ecos de manera agresiva, de modo que no quede ningún eco residual. Esto se logra sobreestimando el espectro de potencia del eco y llevando a cabo un suavizado temporal, lo que

favorece valores de filtro de ganancia pequeños.

**[0067]** Con frecuencia, los filtros de supresión de eco agresivos mencionados anteriormente dan lugar a una degradación de un ruido del extremo cercano y una voz del extremo cercano.

5

**[0068]** El enfoque propuesto aquí palia este problema mediante el uso de dos modelos diferentes de supresión de eco para señales de eco tempranas y tardías, respectivamente, como también se ilustra en la fig. 6.

**[0069]** La fig. 6 muestra un diagrama de bloques de un aparato 200 de acuerdo con una forma de realización de los conceptos descritos en el presente documento, que comprende un altavoz 100 y un micrófono 110. El altavoz 100 tiene a su disposición una señal del altavoz  $x[n]$ . El bloque DFT (transformada discreta de Fourier) 230 convierte la señal del altavoz  $x[n]$  en una versión DFT:  $X[k,m]$ . Además, los medios de cálculo 270 también tienen a su disposición la señal  $y[n]$  del micrófono 110, de nuevo en forma de la correspondiente representación en el dominio de la DFT,  $Y[k,m]$ , obtenida a partir del bloque DFT 290.

15

**[0070]** Basándose en las señales que tengan a su disposición en cada caso, unos medios de cálculo de filtro de supresión 270 calculan coeficientes de filtrado  $H[k,m]$ . Los coeficientes de filtrado  $H[k,m]$  se proporcionan al filtro adaptativo 210.

**[0071]** Para obtener finalmente una señal con el eco suprimido  $e[n]$  a partir de la señal del micrófono  $y[n]$ , el filtro adaptativo 210 está conectado adicionalmente al micrófono con el fin de recibir la señal del micrófono en una entrada. Una salida del filtro adaptativo 210 está conectada con la transformación discreta inversa de Fourier (IDFT) 300 para proporcionar una representación DFT de la señal con eco suprimido a la IDFT 300. La señal con eco suprimido  $e[n]$  está presente en una salida del bloque de la transformación discreta inversa de Fourier 300. En lugar de usar una transformada discreta de Fourier, se podría usar una transformada de Fourier de tiempo corto (STFT) en los bloques 230 y 290. Asimismo, se puede usar una transformada inversa de Fourier de tiempo corto (ISTFT) en lugar de la transformada discreta inversa de Fourier.

20

25

**[0072]** En lo referente a una estimación de la potencia del eco, se puede obtener una estimación del espectro del eco de la señal de eco aplicando una función de estimación de eco  $\hat{G}[k,m]$  a una versión retardada temporalmente del espectro de potencia del altavoz, de acuerdo con la expresión:

30

$$|\hat{Y}[k,m]|^2 = \hat{G}[k,m]^2 |X[k-d,m]|^2, \quad (23)$$

35 en la que  $|\hat{Y}[k,m]|^2$  indica una estimación del espectro de potencia del eco dentro de la señal del micrófono.

**[0073]** Dependiendo de la implementación concreta de una forma de realización de la presente invención, tal como se ilustra, por ejemplo, en relación con las figs. 6 a 8, la funcionalidad descrita por la ecuación (23) se puede implementar dentro del filtro de supresión 270, por ejemplo.

40

**[0074]** En referencia al filtro de eliminación de eco, el filtro de eliminación de eco asociado  $H[k,m]$  se calcula y se aplica a la señal del micrófono con el fin de suprimir la señal de eco. Esto se realiza de acuerdo con la siguiente expresión:

45

$$E[k,m] = H[k,m]Y[k,m], \quad (24)$$

**[0075]** Los componentes de filtrado del filtro de eliminación de eco se pueden calcular, por ejemplo, de acuerdo con la expresión:

50

$$H[k,m] = \left[ \frac{\max(|Y[k,m]|^{\gamma} - \beta |\hat{Y}[k,m]|^{\gamma}, 10^{-\frac{\gamma}{20}})}{|Y[k,m]|^{\gamma}} \right]^{\frac{1}{\gamma}}. \quad (25)$$

**[0076]** Los parámetros de diseño  $\beta$  y  $\gamma$  se pueden usar para controlar el comportamiento deseado del filtro de eliminación de eco. Dependiendo de la implementación concreta de las formas de realización de la presente invención, dichos parámetros de diseño se pueden seleccionar de manera que sean fijos, o diseñar para que sean adaptables, programables o modificables de cualquier otra manera. Una típica elección del parámetro exponencial es, por ejemplo,  $\gamma = 2$ .

55

**[0077]** El denominado factor de sobreestimación  $\beta$  se usa para controlar la agresividad de la atenuación del eco. Por ejemplo, la agresividad del correspondiente filtro de eliminación de eco se puede incrementar aumentando el factor de sobreestimación. Por lo tanto, una elección típica del filtro de eliminación de eco  $H[k,m]$  con respecto al parámetro  $\beta$  es  $\beta = 2$ , para aplicar solo una atenuación de eco moderada.

60

[0078] Por otra parte, el filtro de eliminación de eco se puede diseñar de manera muy agresiva, para que atenúe de manera eficaz los componentes molestos de la voz en la señal del eco. En este caso se puede escoger un factor de sobreestimación  $\beta$  que sea, por ejemplo,  $\beta = 4$ .

5

[0079] El valor límite L determina la máxima atenuación de eco permitida, en decibelios (dB). Un valor típico de un filtro de eliminación de eco para ruido sustancialmente estacionario es  $L = -10$  dB o  $-15$  dB, lo cual limita oportunamente la atenuación para ecos estacionarios a fin de reducir artefactos accidentales. En el caso de una voz no estacionaria en el extremo remoto, la atenuación debería garantizar la completa atenuación de los correspondientes componentes del eco, que corresponden a un valor límite para L de aproximadamente  $-60$  dB para señales de componentes no estacionarios.

10

[0080] La funcionalidad indicada por la ecuación (25) se puede implementar y materializar, en las formas de realización descritas en las figs. 5 a 8, en el contexto de los medios de cálculo de filtro 270.

15

[0081] En algunas formas de realización de la presente invención, la supresión de eco propiamente dicha no se lleva a cabo directamente mediante la aplicación de filtros de eliminación de eco, tal como se describe en la ecuación (25). En lugar de ello, la correspondiente eliminación del eco se lleva a cabo basándose en una versión correspondiente suavizada en tiempo. El suavizado temporal del filtro o filtros de eliminación de eco puede resultar útil para evitar que se produzcan cambios demasiado repentinos, que pueden molestar al oyente. Al igual que en el caso de los parámetros de diseño descritos anteriormente, los parámetros de suavizado temporal se suelen afinar y optimizar de forma manual y por separado para la supresión de eco de reflexiones tempranas y de reverberación tardía, respectivamente. No obstante, cuando se emplean los conceptos descritos en el presente documento, dicho suavizado temporal deja de ser necesario, o al menos deja de serlo en la medida en que lo es en los sistemas de supresión de eco previos.

20

25

[0082] Dicha funcionalidad de suavizado temporal se puede llevar a cabo, por ejemplo, dentro del contexto de los propios medios de cálculo de filtro 270 o en otros medios cualesquiera posteriores a los mismos. También es posible llevar a cabo dicho suavizado temporal directamente dentro del contexto del filtro adaptativo 210, si fuera preciso.

30

[0083] El siguiente apartado ofrece una presentación y una explicación de los conceptos básicos sobre los que se fundamenta un modelado de reverberación y una estimación del eco de reverberación de ecos tardíos que puede resultar útil para comprender los conceptos descritos en el presente documento. Se propone un procedimiento para tener en cuenta propiedades específicas de la acústica de la sala para ajustar la intensidad de la supresión del eco de manera que no quede ningún eco residual originado por componentes del eco de reverberación tardía. La denominada cola de eco es causada por las reflexiones tardías en la sala. El nivel de las reflexiones se puede estimar a través del tiempo de reverberación  $RT_{60}$ , es decir, el tiempo necesario para que las reflexiones de un sonido directo decaigan 60 dB por debajo del nivel del sonido directo. Ahora se describirá un enfoque adecuado para modelar el efecto de la reverberación sobre el nivel de la señal de eco y la estimación del tiempo de reverberación.

35

40

### Modelado de la reverberación

[0084] Es razonable suponer que la potencia del sonido reverberante en una sala decae de manera aproximadamente exponencial. Para los componentes de reverberación tardía, también es razonable suponer que las reflexiones que llegan al micrófono se consideren aleatorias y estadísticamente independientes. Basándose en este supuesto, las reverberaciones tardías de la respuesta al impulso de la sala  $h$ , que se representa de manera esquemática en la figura 3, se pueden modelar como una secuencia de ruido blanco  $b[n]$  ponderada por una función de decaimiento exponencial:

45

50

$$h_{rev}[n] = b[n]e^{-\frac{n}{\tau}} \quad (26)$$

[0085] La constante temporal  $\tau$  determina el nivel del decaimiento. A continuación nos interesaremos por el decaimiento del nivel del eco tras un sonido de tipo impulso. Sin perder generalidad, suponemos que el impulso ha sido emitido en  $n = 0$ . Entonces, se deduce a partir de (26) que la señal observada correspondiente a la reverberación tardía corresponde a:

55

$$y[n] = y_0 b[n]e^{-\frac{n}{\tau}} \quad (27)$$

[0086] Debido a que los coeficientes  $b[k]$  representan una secuencia de ruido blanco, la potencia de los componentes de reverberación viene dada por la expresión:

60

$$E\{y^2[n]\} = y_0^2 \sigma_b^2 e^{-\frac{2n}{\tau}} \quad (28)$$

en la que  $\sigma_b^2$  indica la varianza del modelo de ruido  $b[n]$ .

**[0087]** La versión STFT de (28) se obtiene aplicando un modelo análogo en el dominio espectral:

$$E\{|Y[k, m]|^2\} = Y_0^2(m) \sigma_b^2 e^{-\frac{2k}{F_s \tau_m}}, \quad (29)$$

**[0088]** De forma análoga a lo que ocurría con  $\tau$  en (27),  $\tau_m$  indica una constante de decaimiento de tiempo de bloque aplicable para la banda de frecuencia  $m$ -ésima.  $F_s$  indica la tasa de muestreo relativa al bloque de la STFT, es decir:

$$F_s = \frac{f_s}{K_{STFT}}, \quad (30)$$

en la que  $f_s$  es la frecuencia de muestreo y  $K_{STFT}$  el tamaño de salto de ventana de la STFT. El tamaño de salto de ventana también se denomina tamaño de la trama (*frame size*), e indica el número de nuevas muestras de tiempo consideradas para el cálculo de un nuevo valor de la STFT.

**[0089]** El modelo para los componentes del eco correspondientes a la reverberación tardía, tal como propone la presente invención, se deduce de (29). Los componentes de reverberación tardía del espectro de potencia del eco  $|\hat{Y}[k, m]|^2$  se obtienen mediante una realización instantánea de (29):

$$|\hat{Y}_{rev}[k, m]|^2 = Y_0^2(m) \sigma_b^2 e^{-\frac{2k}{F_s \tau_m}}. \quad (31)$$

**[0090]** A partir de (31), se deduce inmediatamente que la potencia de los componentes reverberantes del eco en el bloque de tiempo del instante  $k$  se puede estimar a partir de la correspondiente estimación en el instante previo  $k-1$ :

$$|\hat{Y}_{rev}[k, m]|^2 = |\hat{Y}_{rev}[k-1, m]|^2 e^{-\frac{2}{F_s \tau_m}}. \quad (32)$$

#### Estimación del eco de reverberación tardía

**[0091]** La exposición anterior nos conduce al procedimiento para estimar el espectro de potencia del eco considerando la trayectoria directa y las reflexiones tempranas junto con los componentes de reverberación tardía: Supongamos que  $|\hat{Y}_{rev}[k, m]|^2$  indica la estimación del espectro de potencia del eco reverberante tardío y que  $|\hat{Y}[k, m]|^2$  indica la estimación del espectro de potencia de la trayectoria directa y las reflexiones tempranas. La estimación global de la potencia del eco se indica con la expresión:  $|\hat{Y}_{tot}[k, m]|^2$ . La estimación del espectro de potencia de la trayectoria directa y las reflexiones tempranas se lleva a cabo de acuerdo con (5). Por conveniencia, la repetimos aquí:

$$|\hat{Y}[k, m]|^2 = |\hat{G}[k, m]|^2 |X_d[k, m]|^2. \quad (33)$$

**[0092]** Los componentes de reverberación tardía se determinan de manera análoga a la de (32), mediante la ecuación recursiva:

$$|\hat{Y}_{rev}[k, m]|^2 = \alpha_m |\hat{Y}_{tot}[k-1, m]|^2. \quad (34)$$

**[0093]** El factor de proporcionalidad  $\alpha_m$  se puede escoger, por ejemplo, de acuerdo con el modelo de decaimiento exponencial

$$\alpha_m = e^{-\frac{2}{F_s \tau_m}}. \quad (35)$$

**[0094]** De este modo, el factor de proporcionalidad deseado  $\alpha_m$  se puede obtener a partir de las ecuaciones (17) y (35).

**[0095]** Aunque el índice  $m$  en  $\alpha_m$  da a entender que la elección del factor de proporcionalidad es distinta para cada banda de frecuencia, también puede ser igual para todas las bandas de frecuencia.

**[0096]** Obsérvese que, a diferencia de lo que ocurre en (32), el cálculo recursivo del eco reverberante para el procedimiento propuesto (34) se basa en la estimación global de la potencia del eco de la trama anterior  $|\hat{Y}_{tot}[k-$

1,m]]<sup>2</sup>. Esto se ilustra en la fig. 8 y se explica con más detalle en la parte correspondiente de la descripción.

**[0097]** La estimación global del espectro de potencia del eco se determina considerando tanto  $|\hat{Y}_{rev}[k,m]|^2$  (estimación del espectro de potencia del eco reverberante tardío) como  $|Y[k,m]|^2$  (estimación del espectro de potencia de la trayectoria directa y las reflexiones tempranas). Un enfoque adecuado consiste en usar la máxima de las dos como una estimación del espectro de potencia global del eco:

$$|\hat{Y}_{tot}[k,m]|^2 = \max\left\{|\hat{Y}[k,m]|^2, |\hat{Y}_{rev}[k,m]|^2\right\}. \quad (36)$$

**[0098]** Obsérvese que (36) junto con (34) implica que después de que se considere que un componente de eco directo resulta dominante en la estimación global del eco, el modelo recursivo para estimar los componentes de eco reverberante siempre se restablece para los nuevos componentes de eco directo: la estimación recursiva siempre tiene en cuenta la estimación anterior del espectro de potencia global del eco, es decir, no está ligada a la anterior estimación de los componentes de eco reverberante.

15

**[0099]** La estimación del espectro de potencia del eco se usa entonces para determinar el filtro de supresión de eco de manera análoga a (9):

$$H[k,m] = \frac{|Y[k,m]|^2 - \beta|\hat{Y}_{tot}[k,m]|^2}{|Y[k,m]|^2}, \quad (37)$$

20

**[0100]** Para diferenciar entre las situaciones en las que la señal del altavoz solo proporciona como salida una voz del extremo remoto, se pueden calcular dos parámetros diferentes. Aquí se incluye, inicialmente, la denominada ganancia de predicción, que corresponde al promedio de todas las bandas de funciones de coherencia entre el canal del altavoz y el canal del micrófono. Como segundo parámetro, se hace uso de la actividad vocal dentro del canal del altavoz, que, por ejemplo, se puede obtener a partir de una comparación de los niveles de señal temporales de la señal del altavoz o a partir de parámetros específicos para códecs concretos, como se usan, por ejemplo, específicamente con códecs adecuados para la transmisión de voz. Entre estos códecs se incluyen, por ejemplo, códecs basados en la LPC o códecs basados en la CELP (CELP: predicción lineal con excitación por código, o predicción lineal con excitación por libro de códigos), siendo el vocablo códec un término artificial creado mediante la combinación de las abreviaturas de los equivalentes ingleses a codificador y descodificador.

**[0101]** La ganancia de predicción o ganancia de predicción de eco  $\omega[k]$  describe el nivel de semejanza entre la señal del micrófono y la señal del altavoz retardada. El cálculo de la ganancia de predicción  $\omega[k]$  se lleva a cabo basándose en una función de coherencia cuadrática entre el espectro de potencia retardado de la señal del altavoz  $|X_d[k,m]|^2$  y el espectro de potencia de la señal del micrófono  $|Y[k,m]|^2$ , de acuerdo con la expresión:

35

$$\Gamma_d[k,m] = \frac{(E\{|X_d[k,m]|^2|Y[k,m]|^2\})^2}{E\{|X_d[k,m]|^2|X_d[k,m]|^2\}E\{|Y[k,m]|^2|Y[k,m]|^2\}}, \quad (38)$$

en la que  $E\{\dots\}$  indica el valor de la esperanza matemática. Dicho valor de esperanza matemática se puede obtener dentro del contexto de una estimación de tiempo corto de la función de coherencia  $\Gamma_d[k,m]$  mediante el cálculo o aproximación del valor de la esperanza  $E\{|X_d[k,m]|^2|Y[k,m]|^2\}$  de acuerdo con la siguiente expresión:

40

$$E\{|X_d[k,m]|^2|Y[k,m]|^2\} = \alpha|X_d[k,m]|^2|Y[k,m]|^2 + (1 - \alpha)E\{|X_d[k-1,m]|^2|Y[k-1,m]|^2\} \quad (39)$$

45

**[0102]** En el contexto del cálculo de la ganancia de predicción, el factor  $\alpha$  determina el grado de suavizado de la estimación a lo largo del tiempo. Este factor lleva asociada una constante temporal, ya que la ecuación (39) corresponde aproximadamente a un decaimiento exponencial. La constante temporal  $T_\alpha$  del decaimiento exponencial en segundos es, aproximadamente:

50

$$T_\alpha \propto \frac{1}{\alpha f_s}, \quad (40)$$

en la que  $f_s$  indica la frecuencia de muestreo. Dicho de otro modo, la relación de proporcionalidad (40) ilustra cómo unos factores que en realidad son adimensionales (aquí  $\alpha$ ), relativos a la tasa de muestreo  $f_s$ , se pueden indicar como una constante temporal (aquí  $T_\alpha$ )

55

**[0103]** La ganancia de predicción  $\omega[k]$  se calcula entonces como el valor medio de las funciones de coherencia a

lo largo de las frecuencias que indican los índices  $m = 0, \dots, M - 1$ , de acuerdo con la expresión:

$$\omega[k] = \frac{1}{M} \sum_{m=0}^{M-1} \Gamma_d[k, m], \quad (41)$$

5 en la que M indica el número exacto de bandas de frecuencia.

**[0104]** Un factor de ganancia de eco cercano a 1 significa que la señal del micrófono se puede predecir completamente basándose en la señal del altavoz retardada. Por lo tanto, la probabilidad de que la señal del micrófono contenga únicamente voz del extremo remoto tiende a 1. El parámetro de control  $\beta$  se puede controlar entonces en relación con la ganancia de predicción  $\omega$ . Cuando la ganancia sea elevada, solo habrá presencia de voz del extremo remoto, y la atenuación del eco debería ser lo suficientemente agresiva como para eliminar todas las señales (de eco). De este modo, se elimina el ruido con el valor límite L más bajo en decibelios (dB) escogiendo el parámetro de control  $\beta = 0$ . Cuando la predicción de ganancia sea baja, puede haber presencia de voz tanto del extremo cercano como del remoto, de manera que la supresión de eco debería ser menos agresiva para no introducir artefactos. En este caso, el ruido se procesa y se elimina con el valor límite L en decibelios (dB).

**[0105]** No obstante, en este caso cabe señalar que la ganancia de predicción podría ser elevada si la señal del altavoz contuviera únicamente ruido captado por el micrófono sin presencia de voz. Para evitar que se escoja un valor del parámetro de control  $\beta$  demasiado grande para este caso, lo cual produciría una supresión excesiva, se emplea un segundo parámetro, concretamente la actividad vocal dentro del canal del altavoz. Por lo tanto, las normas descritas para calcular el parámetro de control  $\beta$  como una función de la ganancia de predicción  $\omega$  en realidad solo tienen validez cuando existe actividad vocal dentro del canal del altavoz.

**[0106]** En la forma de realización ilustrada en la fig. 6, esta funcionalidad, que se describe en las ecuaciones (38) a (41), la pueden llevar a cabo los medios de cálculo 270.

**[0107]** Recordemos que el aparato 200 ilustrado en la fig. 6 incluye un altavoz 100 o un terminal para un altavoz 100 o una entrada para una correspondiente señal de altavoz  $x[n]$ . Dicha señal de altavoz  $x[n]$  se transforma en una representación espectral  $X[k, m]$  de la señal de altavoz dentro del contexto de unos medios convertidores de tiempo/frecuencia 230 denominados DFT (transformada discreta de Fourier). Dicha señal del altavoz se envía a unos medios de retardo 480, que generan una versión retardada de la misma  $X[k-d(k, m), m]$ , siendo  $d(k, m)$  el correspondiente valor de retardo.

**[0108]** La señal retardada por los medios de retardo 480 se envía después a una primera función de estimación de eco 240, que genera una señal de estimación del eco  $\hat{Y}[k, m]$  basándose en coeficientes de filtrado  $G[k, m]$ . Dicha señal de estimación de eco  $\hat{Y}[k, m]$  se envía a unos medios de separación 250, que generan, basándose en coeficientes espectrales de esta señal de eco estimada, espectros de potencia tempranos y tardíos de esta señal como señales componentes (derivadas) de la señal del altavoz. De este modo, los medios de extracción 250 generan como salida las señales  $|\hat{Y}_e[k, m]|^2$  (espectro de potencia estimado de componentes de eco temprano) y  $|\hat{Y}_l[k, m]|^2$  (espectro de potencia estimado de componentes de eco tardío) y las envían a unos medios de cálculo 270. Como alternativa a la forma de realización que se muestra en la fig. 6, la función de estimación de eco 240 podría generar como salida los componentes de eco temprano estimados  $|\hat{Y}_e[k, m]|^2$  y los componentes de eco tardío estimados  $|\hat{Y}_l[k, m]|^2$  directamente, en cuyo caso, los medios de separación 250 no serían necesarios.

**[0109]** La señal del micrófono  $y[n]$  de un micrófono 110 también se envía a unos medios convertidores de tiempo/frecuencia 290 configurados como DFT, que generan una representación espectral  $Y[k, m]$  de la señal temporal  $y[n]$  a partir de la misma. Esta señal se hace pasar a través de unos medios de cálculo de valores de energía 490, que, basándose en los componentes espectrales de la señal del micrófono, determina un espectro de potencia de dichos componentes espectrales elevando al cuadrado la magnitud (absoluta) de los valores individuales. El espectro de potencia obtenido también se envía a los medios de cálculo 270, que, junto con los espectros de potencia descritos anteriormente, calculan los dos filtros de eliminación de eco  $H_e[k, m]$  y  $H_l[k, m]$ , es decir los coeficientes de filtrado del filtro adaptativo real  $H[k, m]$  y los envía al filtro adaptativo 210. El filtro adaptativo real  $H[k, m]$  se puede determinar basándose en los dos filtros de eliminación de eco  $H_e[k, m]$  y  $H_l[k, m]$ , por ejemplo, tomando el mínimo de entre  $H_e[k, m]$  y  $H_l[k, m]$ . Los dos filtros de eliminación de eco  $H_e[k, m]$  y  $H_l[k, m]$  se pueden determinar de manera análoga a la descrita anteriormente con respecto a la ecuación (37):

$$H_e[k, m] = \frac{|Y[k, m]|^2 - \beta |\hat{Y}_e[k, m]|^2}{|Y[k, m]|^2}, \quad (42)$$

y  
60

$$H_i[k, m] = \frac{|Y[k, m]|^2 - \beta |\hat{Y}_i[k, m]|^2}{|Y[k, m]|^2}, \quad (43)$$

**[0110]** Como ya se ha mencionado, el filtro adaptativo real  $H[k, m]$  se puede determinar a continuación como el mínimo de los dos filtros de eliminación de eco  $H_e[k, m]$  y  $H_i[k, m]$ :

5

$$H[k, m] = \min(H_e[k, m], H_i[k, m]) \quad (44)$$

**[0111]** El filtro adaptativo 210 también está conectado con la salida de los medios convertidores de tiempo/frecuencia 290 y, por tanto, también recibe los componentes espectrales  $Y[k, m]$  de la señal del micrófono  $y[n]$ , a partir de los cuales genera la señal con eco suprimido en el dominio de frecuencia o el dominio relacionado con la frecuencia  $E[k, m]$ , mientras se tienen en cuenta los coeficientes de filtrado  $H[k, m]$ . Esta señal con eco suprimido se envía después a los medios convertidores de tiempo/frecuencia 300 configurados como una IDTF (DFT inversa), que finalmente vuelven a convertir esta señal en una señal en el dominio temporal, produciendo la señal  $e[n]$ .

15

**[0112]** Para determinar el valor de retardo  $d(k, m)$  para los medios de retardo 480, y para determinar los coeficientes de estimación de eco para la función de estimación de eco 240, tanto las representaciones espectrales de la señal del altavoz  $X[k, m]$  como las representaciones espectrales de la señal del micrófono  $Y[k, m]$  se transmiten a los correspondientes medios de cálculo de energía 500, 510, que están conectados con las salidas de los dos medios convertidores de tiempo/frecuencia 230, 290. Los medios de cálculo de energía 500 están conectados con la salida de los medios convertidores de tiempo/frecuencia 230, y los medios de cálculo de energía 510 están conectados con la salida de los medios convertidores de tiempo/frecuencia 300.

**[0113]** Cada uno de los dos medios de cálculo de valores de energía 500, 510 calcula, por analogía con los medios de cálculo de energía 490, espectros de potencia elevando al cuadrado las magnitudes de los respectivos componentes espectrales y envía estos valores a unos medios de cálculo adicionales 520. Estos medios de cálculo adicionales 520, o una parte de ellos, forman parte de los medios de modelado del decaimiento del eco 465. A continuación los medios de cálculo adicionales 520 determinan, basándose en los valores de los que ha dispuesto, una estimación para el retardo  $d(k, m)$  y valores para los coeficientes  $G[k, m]$  para la función de estimación de eco 240. Los valores correspondientes antes mencionados se transmiten, por una parte, a los medios de retardo 480 y, por otra, a la función de estimación de eco 240, con la que también están conectados los medios de cálculo adicionales 520.

**[0114]** Como se puede observar también en la forma de realización que se muestra en la fig. 6, la separación de las respectivas señales componentes ( $|\hat{Y}_e[k, m]|^2$  y  $|\hat{Y}_i[k, m]|^2$ ) puede llevarse a cabo, por tanto, basándose en una estimación del espectro de la señal de eco  $Y[k, m]$ , que se calcula según la fórmula:

35

$$\hat{Y}[k, m] = \hat{G}[k, m]X[k - d, m]. \quad (45)$$

**[0115]** Este cálculo se lleva a cabo en la función de estimación de eco 240.

**[0116]** La definición de los dos filtros de eliminación de eco  $H_e[k, m]$  y  $H_i[k, m]$  de acuerdo con la ecuación (25) se mantiene sin cambios. Lo mismo ocurre con la determinación del filtro de eliminación de eco combinado  $H[k, m]$ . Obsérvese que, dependiendo de la forma de realización, la combinación de los dos filtros de eliminación de eco  $H_e[k, m]$  y  $H_i[k, m]$  se puede implementar como una selección temporal de uno de los dos filtros de eliminación de eco, basándose en las condiciones de eco observadas en ese momento. En particular, la selección del filtro de eliminación de eco que se encuentre activo en ese momento puede basarse en una evaluación de si lo que predomina en dicho momento son los componentes tempranos del eco o los componentes tardíos del eco. La selección de los filtros de eliminación de eco ( $H_e[k, m]$  o  $H_i[k, m]$ ) se puede implementar de manera indirecta, seleccionando el espectro de potencia estimado de los componentes tempranos del eco  $|\hat{Y}_e[k, m]|^2$  o el espectro de potencia estimado de los componentes tardíos del eco  $|\hat{Y}_i[k, m]|^2$  como la cantidad  $\hat{Y}_{\text{tot}}[k, m]$  de la ecuación (37) mediante la cual se calculan los coeficientes de filtrado  $H[k, m]$ .

**[0117]** Cabe señalar en este momento que la forma de realización que se muestra en la fig. 6 se refiere al caso en el que el espectro estimado de la señal de eco  $\hat{Y}[k, m]$  ya se encuentre disponible. Por supuesto, el procedimiento correspondiente también se puede aplicar cuando solo se conoce el espectro de potencia estimado de la señal de eco estimada  $|\hat{Y}[k, m]|^2$ , de acuerdo con la ecuación (23). Esta situación se describirá más detalladamente en relación con la forma de realización que se muestra en la fig. 7.

**[0118]** Mientras que en la fig. 6 se muestra un diagrama de bloques de un enfoque de la atenuación de eco acústico, el diagrama que se muestra en la fig. 7 ilustra un enfoque similar, pero a diferencia de aquel, está basado en el enfoque de la atenuación de eco acústico que funciona basándose en un espectro de potencia estimado de la

60

señal de eco  $|\hat{Y}[k, m]|^2$ .

**[0119]** Por consiguiente, no solo son muy similares los modos de funcionamiento de las dos formas de realización que se muestran en las figs. 6 y 7, sino también sus estructuras, tal como se mostrará en la siguiente descripción.

5

**[0120]** Más concretamente, la forma de realización que se muestra en la fig. 7 se diferencia de la que se muestra en la fig. 6, básicamente, en que los medios de cálculo de energía 500 no están ya conectados exclusivamente con unidades que preceden a los medios de cálculo adicionales 520 en lo que se refiere a la señal del altavoz  $x[n]$  transferida al dominio de frecuencia, sino que están conectados directamente con la salida de los medios convertidores de tiempo/frecuencia 230 configurados de nuevo como una DFT. De este modo, no solo los medios de cálculo adicionales 520, sino también los medios de retardo 480, la función de estimación de eco 240 y los medios de separación 250 dejan de recibir los componentes espectrales reales y, en cambio, reciben sus espectros de potencia.

10  
15 **[0121]** No obstante, aparte de esto, las dos formas de realización que se muestran en las figs. 6 y 7 solo difieren en que es posible que los respectivos cálculos se desarrollen de manera ligeramente distinta en uno y otro caso, dentro de los componentes o medios individuales. Por ejemplo, el correspondiente cálculo de los valores relacionados con la energía de los componentes espectrales individuales deja de llevarse a cabo en los medios de separación 250, ya que los medios de cálculo de valores de energía 500 ya los han llevado a cabo con anterioridad.

20

**[0122]** En la fig. 8 se muestra otra forma de realización de un aparato o un procedimiento de acuerdo con los conceptos descritos en el presente documento. Los medios de determinación de la función de eco temprano 472 generan como salida una estimación  $|G[k, m]|^2$  de la función de transferencia de potencia del entorno acústico 120. La estimación  $|G[k, m]|^2$  se puede determinar por medio de una correlación entre la señal del micrófono y la señal del altavoz. Los medios de determinación de la función de estimación de eco tardío 475 generan como salida un parámetro de decaimiento del eco  $\alpha_m$ . La estimación  $|G[k, m]|^2$  se envía a un bloque de función 240, es decir, una función de estimación de eco. El bloque de función 240 usa la estimación de la función de transferencia de potencia  $|G[k, m]|^2$  de la respuesta del sistema del entorno acústico 120 y el espectro de potencia de la señal del altavoz  $|X[k, m]|^2$  para determinar un espectro de potencia estimado del micrófono  $|\hat{Y}[k, m]|^2$  como la respuesta actual del sistema del entorno acústico 120 al ser excitado por la señal del altavoz. La estimación  $|\hat{Y}[k, m]|^2$  de la respuesta del sistema a la señal del altavoz que se esté considerando en ese momento suele abarcar, sustancialmente, solo las reflexiones tempranas. El horizonte temporal de las estimaciones  $|G[k, m]|^2$  y  $|\hat{Y}[k, m]|^2$  es limitado. La estimación  $|\hat{Y}[k, m]|^2$  se hace llegar en forma de una de dos entradas a un selector de máximo 550. La otra entrada al selector de máximo 550 la proporciona una salida previa del selector de máximo 550, posiblemente escalada en un factor  $\alpha_m$ , que representa el componente de eco reverberante tardío  $|Y_{rev}[k, m]|^2 = \alpha_m |Y_{tot}[k-1, m]|^2$ . La salida del selector de máximo 550 es una estimación global del espectro de potencia del eco  $|Y_{tot}[k-1, m]|^2$  que se determina teniendo en cuenta tanto  $|\hat{Y}[k, m]|^2$  como  $\alpha_m |Y_{tot}[k-1, m]|^2$ . El segundo término  $\alpha_m |Y_{tot}[k-1, m]|^2$  se calcula a partir de la salida del selector de máximo 550 por medio de un elemento de retardo 560 y un factor escalar 570 para el parámetro  $\alpha_m$ . La disposición del elemento de retardo 560 y el factor escalar 570 representa los medios de estimación de eco 241 y proporciona un cálculo recursivo del eco reverberante basado en un modelo exponencial para la parte reverberante del eco global, como ya se ha explicado. El uso del máximo entre las entradas al selector de máximo 550 representa el descubrimiento de que las reflexiones tempranas suelen predominar durante un primer intervalo de tiempo posterior a una excitación del entorno acústico 120. Estas reflexiones tempranas, que se pueden calcular usando el espectro de potencia de transferencia de respuesta al eco  $|G[k, m]|^2$ , constituyen principalmente el objetivo de la cancelación del eco durante el primer intervalo de tiempo. De hecho, las reflexiones tempranas suelen predominar durante la primera fase posterior a la excitación, de manera que pueden ignorarse otros componentes del eco (reverberaciones tardías). En ausencia de nuevas excitaciones en el entorno acústico 120, la magnitud de las reflexiones tempranas disminuirá y estas acabarán desvaneciéndose completamente con el tiempo. Otra razón para el desvanecimiento de las reflexiones tempranas estimadas, partiendo de  $|\hat{Y}[k, m]|^2$ , es que la función de transferencia  $|G[k, m]|^2$  suele tener solo una longitud limitada. Cuando la magnitud de las reflexiones tempranas estimadas ha descendido hasta un volumen determinado, la otra entrada (es decir, la entrada para  $\hat{Y}_{rev}[k, m]$ ) destinada al selector de máximo 550 puede transportar una señal de una magnitud mayor que el espectro de potencia de las reflexiones tempranas estimadas  $|\hat{Y}[k, m]|^2$ . Como ya se ha mencionado, el modelo recursivo para estimar los componentes de eco reverberantes por medio del elemento de retardo 560 y el factor escalar 570 siempre se reajusta para los nuevos componentes de eco directos  $|\hat{Y}[k, m]|^2$ , a condición de que la magnitud de los nuevos componentes de eco directos sea mayor que la de los componentes de eco reverberantes. Dicho de otro modo, los componentes de eco reverberantes no están ligados a una estimación previa de los componentes de eco reverberantes.

35  
40  
45  
50  
55  
60 **[0123]** A continuación, se usa la salida del selector de máximo 550  $|Y_{tot}[k, m]|^2$  para determinar el filtro de supresión de eco de una manera análoga a la de la ecuación (9). En particular, la estimación del espectro de potencia del eco  $|Y_{tot}[k, m]|^2$  se multiplica por un factor escalar  $\beta_m$  (signo de referencia 580) y después se resta del espectro de potencia de la señal del micrófono  $|Y[k, m]|^2$ . La acción del selector de máximo se puede considerar como una posible implementación para combinar el componente temprano del eco ( $\hat{Y}[k, m]$ ) y el componente tardío del eco ( $\hat{Y}_{rev}[k, m]$ ). Una aplicación alternativa de la combinación de los componentes tempranos y tardíos podría consistir en

65

calcular una media ponderada.

**[0124]** Después, la diferencia resultante se normaliza con respecto al espectro de potencia de la señal del micrófono, tal como se indica en el bloque 590. El factor escalar 580, el punto de suma y el bloque normalizador 590 forman en conjunto los medios de cálculo 270 (o constituyen al menos una parte de los mismos) que proporcionan los coeficientes de filtrado para el filtro adaptativo 210 ( $H[k,m]$ ).

**[0125]** El parámetro  $\alpha_m$  que define la tasa del modelo de decaimiento exponencial usado dentro del factor escalar 570 se puede determinar mediante unos medios de determinación de la función de estimación de eco tardío 475. Seguidamente, se puede ajustar el factor escalar 570 basándose en un valor para el parámetro  $\alpha_m$  determinado por la función de estimación de eco tardío 475 basándose en el espectro de potencia de la señal del altavoz y el espectro de potencia de la señal del micrófono. La estimación del tiempo de reverberación y su relación con el parámetro de decaimiento exponencial ya se han explicado con anterioridad.

**[0126]** Los medios de determinación de la función de estimación de eco tardío 475, en un lado de entrada, pueden recibir, por ejemplo, la constante temporal de decaimiento del eco  $\tau$  (u otras mediciones, parámetros, etc.) procedente de los medios de modelado del decaimiento del eco 465. Los medios de modelado del decaimiento del eco 465 pueden estar conectados con los medios de cálculo de valores de energía 500, 510 y, de este modo, recibir los espectros de potencia de las señales del micrófono y del altavoz. Los medios de modelado del decaimiento del eco 465 pueden determinar, por ejemplo, una correlación entre los espectros de potencia de las señales del altavoz y del micrófono para determinar los parámetros de decaimiento del eco. En la solución alternativa, puede bastar con observar únicamente la señal del micrófono. La función de estimación de eco tardío 475 también puede recibir entradas adicionales para las señales del micrófono y del altavoz (que no se muestran en la fig. 8).

**[0127]** La fig. 8 indica mediante un rectángulo de línea discontinua unos medios de procesamiento del eco 340, que comprenden los medios de modelado del decaimiento del eco 465, los medios de determinación de la función de estimación de eco 472 y 475, la función de estimación de eco 240, el selector de máximo 550, y los medios de estimación de eco tardío 241. Los medios de procesamiento del eco 340 proporcionan una estimación del eco producido por el entorno acústico 120 como respuesta a una señal de altavoz dada. La estimación del eco se puede calcular basándose en los medios de determinación de la función de eco temprano 472 o bien basándose en el modelado recursivo de los componentes reverberantes del eco, implementado en los medios de estimación de eco tardío 241.

**[0128]** En la forma de realización que se muestra en la fig. 8, el selector de máximo 550 usa los espectros de potencia del componente de eco temprano estimado  $|\hat{Y}[k,m]|^2$  y del componente de eco tardío estimado  $|\hat{Y}_{rev}[k,m]|^2$  como valores de entrada y, por consiguiente, también produce un espectro de potencia del eco total  $|\hat{Y}_{tot}[k,m]|^2$  como salida. La otra posibilidad consiste en usar espectros de frecuencia (es decir:  $Y[k,m]$ ,  $\hat{Y}_{rev}[k,m]$  y  $\hat{Y}_{tot}[k,m]$ ) en lugar de los espectros de potencia. En las expresiones «componente temprano del eco» y «componente tardío del eco» usadas en la reivindicación o reivindicaciones dependientes se contemplan tanto espectros de frecuencia como espectros de potencia.

**[0129]** En la fig. 9 se muestra un diagrama de bloques esquemático de una forma de realización que es similar a la que se muestra en la fig. 8. Hasta la función de estimación de eco 240, los diagramas de bloques de las figs. 8 y 9 son sustancialmente idénticos. La función de estimación de eco 240 genera como salida el espectro de potencia del componente de eco temprano estimado  $|\hat{Y}[k,m]|^2$ . A continuación, se distribuye este espectro de potencia del componente de eco temprano estimado  $|\hat{Y}[k,m]|^2$  a unos primeros medios de cálculo 270<sub>e</sub> y a un selector 553. Los medios de cálculo 270<sub>e</sub> corresponden a los medios de cálculo 270 que se muestran en la fig. 8. La diferencia entre la fig. 8 y la fig. 9 consiste en que en la fig. 9 los medios de cálculo 270<sub>e</sub> siempre consideran el espectro de potencia del componente de eco temprano estimado  $|\hat{Y}[k,m]|^2$  como una entrada. De este modo, la salida de los medios de cálculo 270<sub>e</sub> de la fig. 9 se pueden considerar como un filtro de eliminación de eco  $\hat{H}_e[k,m]$  para los componentes tempranos del eco. Además de los medios de cálculo 270<sub>e</sub>, también se distribuye el espectro de potencia del componente de eco temprano estimado  $|\hat{Y}[k,m]|^2$  a una primera entrada del selector 553. Una salida del selector 553 está conectada con unos segundos medios de cálculo 270<sub>i</sub> y también con unos medios de estimación del eco tardío 241 similares a los que se muestran en la fig. 8. Por consiguiente, los medios de estimación del eco tardío 241 comprenden un elemento de retardo 560 y un factor escalar  $\alpha_m$ . Una salida de los medios de estimación del eco tardío 241 está conectada con una segunda entrada del selector 553. Cuando el selector 553 está en una configuración en la que selecciona su segunda entrada, se lleva a cabo un cálculo recursivo de los componentes del eco (tardío) tal como se ya explicó en referencia a la fig. 8. Cuando el selector 553 está en una configuración en la que se selecciona su primera entrada, el cálculo recursivo se reajusta o reinicia con un valor actualizado del espectro de potencia del componente de eco temprano estimado  $|\hat{Y}[k,m]|^2$ .

**[0130]** Los segundos medios de cálculo 270<sub>i</sub> reciben el componente de eco reverberante tardío  $|Y_{rev}[k,m]|^2$  y llevan a cabo un cálculo análogo al de los primeros medios de cálculo 270<sub>e</sub>, lo que produce una estimación del filtro de eliminación de eco  $\hat{H}_i[k,m]$  para los componentes tardíos del eco. El filtro de eliminación del eco temprano  $\hat{H}_e[k,m]$  y el filtro de eliminación del eco tardío  $\hat{H}_i[k,m]$  se envía a unos medios de establecimiento de filtro 551 y a un comparador 552. Los medios de establecimiento de filtro 551 determinan los coeficientes de filtrado  $H[k,m]$  del filtro

adaptativo 210 basándose en el filtro de eliminación del eco temprano  $\hat{H}_e[k,m]$  y en el filtro de eliminación del eco tardío  $\hat{H}_l[k,m]$ . Dicho de otro modo, los medios de cálculo 270 están asimismo configurados para determinar un filtro de eliminación del eco temprano  $\hat{H}_e[k,m]$  y un filtro de eliminación del eco tardío  $\hat{H}_l[k,m]$ . El aparato 200 también comprende unos medios de establecimiento de filtro 551 configurados para determinar los coeficientes de filtrado  $H[k,m]$  del filtro adaptativo 210 basándose en una combinación de coeficientes de filtrado del filtro de eliminación del eco temprano  $\hat{H}_e[k,m]$  y del filtro de eliminación del eco tardío  $\hat{H}_l[k,m]$ . La combinación del filtro de eliminación de eco temprano  $\hat{H}_e[k,m]$  y del filtro de eliminación del eco tardío  $\hat{H}_l[k,m]$  puede ser una combinación lineal, una selección por coeficientes de los coeficientes de filtrado entre el filtro de eliminación de eco temprano  $\hat{H}_e[k,m]$  y el filtro de eliminación de eco tardío  $\hat{H}_l[k,m]$ , u otra combinación adecuada de los dos filtros de eliminación de eco  $\hat{H}_e[k,m]$  y  $\hat{H}_l[k,m]$ . En la forma de realización representada en la fig. 9, los medios de establecimiento de filtro 551 están configurados como un selector de mínimo que selecciona el mínimo entre el filtro de eliminación del eco temprano  $\hat{H}_e[k,m]$  y el filtro de eliminación del eco tardío  $\hat{H}_l[k,m]$ , que después se envía al filtro adaptativo 210. Dicho de otro modo, y tal como se indica anteriormente, los medios de cálculo 270 (incluidos los primeros medios de cálculo 270<sub>e</sub> y los segundos medios de cálculo 270<sub>i</sub>) están asimismo configurados para determinar un filtro de eliminación del eco temprano  $\hat{H}_e[k,m]$  y un filtro de eliminación del eco tardío  $\hat{H}_l[k,m]$ . El aparato 200 también comprende unos medios de establecimiento de filtro 551 en forma del selector de mínimo que está configurado para determinar los coeficientes de filtrado  $H[k,m]$  del filtro adaptativo 210 basándose en una selección del mínimo entre el filtro de eliminación del eco temprano  $\hat{H}_e[k,m]$  y el filtro de eliminación del eco tardío  $\hat{H}_l[k,m]$ .

**[0131]** El comparador 552 genera una señal de control para el selector 553 basándose en una comparación entre el filtro de eliminación de eco temprano  $\hat{H}_e[k,m]$  y el filtro de eliminación de eco tardío  $\hat{H}_l[k,m]$ . En la forma de realización ejemplar representada, la señal de control (línea discontinua indicada como «RESET» en la fig. 9) es alta si el filtro de eliminación del eco temprano  $\hat{H}_e[k,m]$  es más pequeño que el filtro de eliminación del eco tardío  $\hat{H}_l[k,m]$ , lo cual provoca que el selector 553 seleccione su primera entrada, es decir, la entrada que está conectada con la función de estimación de eco 240. De este modo, el cálculo de los componentes de eco reverberantes  $|\hat{Y}_{rev}[k,m]|^2$  se reajusta o reinicia para la estimación actualizada del componente de eco temprano  $|\hat{Y}[k,m]|^2$ .

**[0132]** La continuación del procesamiento es similar al procesamiento ilustrado y descrito con respecto a la forma de realización que se muestra en la fig. 8.

**[0133]** La fig. 10 ilustra el modo en que las reverberaciones tardías causadas por varios impulsos dentro del entorno acústico 120 se pueden solapar. Se han emitido tres impulsos de diferente magnitud al entorno acústico 120 en momentos diferentes. Las reflexiones tempranas no se ilustran en la fig. 10 para mayor claridad. No obstante, se muestran las reverberaciones tardías de los impulsos. En particular, se puede observar que una función con decaimiento exponencial y correspondiente al segundo impulso aún no se ha desvanecido cuando una segunda función con decaimiento exponencial y correspondiente al segundo impulso se superpone a la primera reverberación tardía con decaimiento exponencial. Un efecto similar ocurre cuando aparece la tercera reverberación tardía con decaimiento exponencial correspondiente al tercer impulso. Además, aunque los impulsos originales y los sonidos directos tuvieran una duración relativamente corta, las reverberaciones tardías se prolongarían durante un intervalo de tiempo más largo. Obsérvese que se han exagerado las magnitudes de las reverberaciones tardías con fines ilustrativos. Las reverberaciones tardías suelen ser aproximadamente uno o varios órdenes de magnitud más pequeñas que los impulsos originales.

**[0134]** La fig. 11 ilustra la manera en que el modelado del eco puede cambiar de un modelo de eco temprano a un modelo de decaimiento exponencial. Como ya se ha explicado en el contexto de la fig. 8, se puede determinar un máximo entre las estimaciones de eco temprano y el modelo de decaimiento exponencial. Hasta un instante indicado por  $n_{switch}$  en la fig. 11, las estimaciones de eco temprano son de mayor magnitud que un eco estimado por medio de un modelo de decaimiento exponencial. Por lo tanto, las estimaciones de eco temprano se usan para modelar la respuesta al eco del entorno acústico 120. Posteriormente al instante  $n_{switch}$ , el eco estimado por medio del modelo de decaimiento exponencial es mayor que cualquiera de las estimaciones de eco restantes. La función o funciones de estimación de eco temprano no se suelen implementar para calcular las estimaciones de eco más allá de un horizonte temporal predeterminado, con el fin de mantener el esfuerzo computacional dentro de unos límites razonables; lo cual significa que, una vez superado este horizonte temporal, el modelo de decaimiento exponencial es el único modelo de eco que proporciona una estimación distinta de cero para los componentes tardíos del eco. Antes del horizonte temporal, ambos modelos pueden generar estimaciones distintas de cero y se selecciona el que tenga mayor magnitud.

**[0135]** En la fig. 11 se puede observar que el modelo de decaimiento exponencial siempre se reajusta para la magnitud de las estimaciones de eco temprano, a condición de que las estimaciones de eco temprano tengan una magnitud mayor que la magnitud en dicho momento de un eco estimado por medio del modelo de decaimiento exponencial. En la solución alternativa, sería posible calibrar el modelo de decaimiento exponencial con, por ejemplo, el sonido directo generado por un impulso original.

**[0136]** La fig. 12 muestra otra forma de realización de la presente invención, en la que se proporciona al respectivo aparato 200, por ejemplo, más de una señal de altavoz o más de una señal de micrófono. Dicho de otro modo, la forma de realización representada en la fig. 12 es un aparato multicanal.

**[0137]** Aunque las formas de realización de la presente invención se han explicado y descrito anteriormente únicamente para canales individuales o para el caso de un canal individual, en el que solo estuviera disponible una señal de altavoz y una señal de micrófono, las formas de realización de la presente invención no se limitan al caso de un canal individual, como se explicará más adelante. Dichas formas de realización también se pueden aplicar, por analogía, a sistemas de atenuación de eco acústico con capacidad multicanal.

**[0138]** Debido a que la forma de realización, que se muestra en la fig. 12, del aparato 200 tiene una estructura similar a la que se muestra en la fig. 6, a continuación se hará referencia a la descripción relativa a las figs. 6 y 8, en lo que respecta al modo de funcionamiento, conexión y otros aspectos.

**[0139]** La variante multicanal, que se muestra en la fig. 12, del aparato 200 comprende una pluralidad de entradas 220-1, 220-1..., por las que se pueden conectar varias señales de altavoz al aparato 200. Por consiguiente, el aparato 200 también comprende una correspondiente pluralidad de medios convertidores de tiempo/frecuencia 230-1, 230-2..., que pueden llevar a cabo una traslación o conversión de las correspondientes señales de altavoz desde el dominio temporal a un dominio relativo a la frecuencia.

**[0140]** La pluralidad de medios convertidores de tiempo/frecuencia 230 están conectados con un número correspondiente de entradas de unos medios de agrupamiento 530, que, basándose en las señales de altavoz entrantes, forman una señal de altavoz derivada común que después envían a la primera función de estimación de eco 240 o a los medios de separación 250, dependiendo de la presencia de la primera función de estimación de eco opcional 240. Los medios de separación 250 se pueden conectar a una segunda función de estimación de eco opcional 260 o directamente a los medios de cálculo 270. Por último, dichos medios generan en una de sus salidas los coeficientes de filtrado calculados.

**[0141]** La variante multicanal del aparato 200 de la fig. 12 comprende además unos medios de agrupamiento adicionales 540, que están conectados, en el lado de entrada, con un número correspondiente de entradas 280-1, 280-2..., para las correspondientes señales de micrófono a través de unos medios convertidores de tiempo/frecuencia opcionales 290-1, 290-2... Por analogía con los medios de agrupamiento 530, los medios de agrupamiento adicionales 540 determinan, basándose en las señales de micrófono presentes dentro del dominio temporal o en un dominio relativo a la frecuencia, y que reciben dichos medios, una señal de micrófono derivada, eficaz o común, que se puede poner, opcionalmente, a disposición de los medios de separación 250 o de los medios de cálculo 270.

**[0142]** La variante multicanal del aparato 200 que se muestra en la fig. 12 comprende además un filtro adaptativo 210-1, 210-2..., para cada señal de micrófono o cada una de las entradas de señal de micrófono 280, estando conectado dicho filtro adaptativo 210-1, 210-2..., posiblemente a través de los medios convertidores de tiempo/frecuencia opcionales 290-1, 290-2..., con las respectivas entradas 280-1, 280-2... Por consiguiente, los filtros adaptativos 210-1, 210-2... están conectados, posiblemente a través de una pluralidad de medios convertidores de tiempo/frecuencia opcionales 300-1, 300-2..., con una salida 310-1, 310-2..., respectivamente. Una vez que las señales de salida se han filtrado mediante los filtros adaptativos 210, que se han eliminado sus ecos o se han modificado espectralmente, se hallan a disposición del aparato 200 en dichas salidas 310.

**[0143]** Todos los filtros adaptativos 210-1, 210-2... están conectados en paralelo con la salida de los medios de cálculo 270 en la cual dichos medios de cálculo 270 proporcionan los coeficientes de filtrado para los filtros adaptativos. Dicho de otro modo, todas las señales de micrófono de la pluralidad de señales de micrófono se filtran, en una forma de realización de la presente invención que se ilustra en la fig. 12, con el mismo filtro adaptativo desde un punto de vista funcional, es decir, basado en los mismos coeficientes de filtrado, con el fin de obtener las versiones modificadas espectralmente o desprovistas de eco de las correspondientes señales de micrófono.

**[0144]** De este modo, si  $x_w[k, m]$  representa el altavoz  $w$ -ésimo, donde  $w$  es un número entero entre 0 y  $W-1$ , donde  $W$  indica el número de diferentes altavoces o señales de altavoz, se puede calcular una representación en el dominio STFT  $X_w[k, m]$  de la  $w$ -ésima señal del altavoz.

**[0145]** Después se calcula un espectro de potencia conjunto para todos los canales de altavoz combinando los espectros de los espectros de las señales de altavoz individuales:

$$|X[k, m]|^2 = \sum_{w=0}^{W-1} |X_w[k, m]|^2, \quad (46)$$

donde  $W$  indica el número de canales de altavoz.

**[0146]** Por analogía, se calcula también un espectro de potencia agrupado o común para los canales de micrófono, de acuerdo con la expresión:

$$|Y[k, m]|^2 = \sum_{p=0}^{P-1} |Y_p[k, m]|^2, \quad (47)$$

donde  $Y_p[k, m]$  indica la señal del micrófono p-ésimo 110, y P representa el número de micrófonos. De nuevo, el índice p es un número entero entre 0 y P-1. Este cálculo se puede llevar a cabo, en la forma de realización que se muestra en la fig. 12, mediante los medios de agrupamiento adicionales 540.

**[0147]** Para determinar el filtro de eliminación de eco de acuerdo con la ecuación (25), durante las siguientes etapas algorítmicas se usan los espectros (de potencia) del altavoz  $|X[k, m]|^2$  de acuerdo con la ecuación (46) y el espectro (de potencia) del micrófono  $|Y[k, m]|^2$  de acuerdo con la ecuación (47), tal como se explica en los párrafos precedentes de la descripción. La determinación del parámetro de control  $\beta$  descrito en relación con el control de comportamiento de acuerdo con las ecuaciones (38) a (41) también se puede llevar a cabo basándose en los espectros comunes o agrupados de acuerdo con las ecuaciones (46) a (47).

**[0148]** A continuación se lleva a cabo la supresión de eco propiamente dicha, dentro del contexto de la modificación espectral, de manera individual para cada señal de micrófono, pero usando el mismo filtro de eliminación de eco 210 para cada canal de micrófono de acuerdo con la expresión:

$$E_p[k, m] = H[k, m]Y_p[k, m], \quad (48)$$

para  $p = 0, 1, \dots, P-1$ . Por analogía, también se pueden aplicar los filtros de eliminación de eco 210 de manera diferente, como ya se ha explicado con anterioridad.

**[0149]** En este contexto, conviene señalar que dentro del contexto de la variante multicanal del aparato 200, como se muestra, por ejemplo, en la fig. 12, el número W de señales de altavoz y el número P de señales de micrófono deberían ser idénticos y diferentes el uno del otro. En principio, se puede proporcionar cualquier número de entradas, tanto para las señales de altavoz como para las señales de micrófono. Además, no es absolutamente necesario proporcionar tanto una pluralidad de entradas de señales de altavoz como una pluralidad de entradas de señales de micrófono con los correspondientes medios de agrupamiento 530, 540. En las formas de realización, es muy posible que solo se proporcione una pluralidad de entradas de señales de altavoz con unos correspondientes medios de agrupamiento 530 sin que haya que proporcionar más de una entrada de señal de micrófono con unos correspondientes medios de agrupamiento adicionales 540. Dicho sistema se puede emplear, por ejemplo, cuando existe una única señal de micrófono, pero una pluralidad de altavoces, debido, por ejemplo, a que la señal de comunicación del abonado del extremo remoto se reproduce a través de un sistema de megafonía que comprende varios altavoces, como puede ser el caso de su implementación en automóviles, por ejemplo.

**[0150]** Además, no es necesario proporcionar una posible implementación con más de una entrada para una señal de altavoz con unos correspondientes medios de agrupamiento 530, cuando solo se proporciona un altavoz central, por ejemplo en el contexto de un sistema de conferencias, pero una pluralidad de conferenciantes disponen de un micrófono propio para cada uno. En tal caso, solo podría ser aconsejable la aplicación de los medios de agrupamiento adicionales 540.

**[0151]** Además, en este momento es conveniente señalar que los medios de agrupamiento 530, 540, como es natural, se pueden configurar de manera que estén diseñados para admitir más señales de altavoz o señales de micrófono de las que tengan a su disposición en un momento dado. Por consiguiente, un aparato 200 puede comprender más entradas 220, 280 correspondientes de las que vayan a usarse en un momento dado. En este caso, por ejemplo, un circuito previo, como los medios convertidores de tiempo/frecuencia opcionales 230, 290, o los propios medios de agrupamiento 530, 540, puede determinar el número de canales activos y seleccionar los parámetros W y P en función del mismo. Es evidente que también se puede proporcionar una fuente externa del número de canales y posiblemente del número de señales de micrófono y de altavoz que se deban tener en cuenta.

**[0152]** También cabe señalar en este momento que la forma de realización que se muestra en la fig. 12, como es natural, también funciona con una única señal de altavoz y una única señal de micrófono, si los correspondientes parámetros W y P se transmiten a los medios de agrupamiento 530, 540. En principio, las ecuaciones (46) y (47) también se pueden aplicar a los casos  $P = 1$  y/o  $W = 1$ . De este modo, la forma de realización que se muestra en la fig. 12 representa una extensión con «compatibilidad descendente» de las formas de realización que se muestran en las figs. 6 a 8 de la presente memoria descriptiva.

#### Fluctuaciones temporales de los espectros de potencia

**[0153]** En la referencia [11] se ha propuesto el uso de fluctuaciones temporales de espectros de potencia para mejorar la estimación de parámetros usados en aplicaciones de supresión de eco acústico. Las fluctuaciones temporales de los espectros de potencia se determinan como versiones centradas con respecto a los

correspondientes valores medios:

$$\tilde{Y}[k, m] = |Y[k, m]|^2 - \mathbf{E}\{|Y[k, m]|^2\}, \quad (49)$$

$$\tilde{X}_a[k, m] = |X_a[k, m]|^2 - \mathbf{E}\{|X_a[k, m]|^2\},$$

[0154] En la práctica, los operadores de esperanza se sustituyen por medias temporales, por ejemplo, de acuerdo con la expresión (8). A continuación se usan las fluctuaciones temporales de los espectros de potencia de la señal de altavoz retardada  $\tilde{X}_a[k, m]$  y la señal de micrófono  $\tilde{Y}[k, m]$  en las ecuaciones (18) y (19) con el fin de estimar la constante temporal de decaimiento  $\tau_m$ .

$$|\hat{Q}_{d_\tau}[k, m]|^2 = \frac{\mathbf{E}\{\tilde{X}_{d_\tau}[k, m]\tilde{Y}[k, m]\}}{\mathbf{E}\{\tilde{X}_{d_\tau}[k, m]\tilde{X}_{d_\tau}[k, m]\}},$$

$$|\hat{Q}_{d_\tau+\Delta d_\tau}[k, m]|^2 = \frac{\mathbf{E}\{\tilde{X}_{d_\tau+\Delta d_\tau}[k, m]\tilde{Y}[k, m]\}}{\mathbf{E}\{\tilde{X}_{d_\tau+\Delta d_\tau}[k, m]\tilde{X}_{d_\tau+\Delta d_\tau}[k, m]\}}.$$

**Alternativas a la transformada de Fourier de tiempo corto**

[0155] En lo referente a la resolución de frecuencia, puede que sea aconsejable apartarse de la de una STFT. La resolución espectral uniforme de una STFT resulta en cierta medida diferente a la resolución espectral de la percepción humana. Por lo tanto, podría resultar ventajoso agrupar los coeficientes espaciados uniformemente  $|X[k, m]|^2$  y  $|Y[k, m]|^2$  en un cierto número de particiones o grupos que no se solapen, tal como se muestra también en la referencia [9], y estas particiones o grupos comprenden anchos de banda que imitan la resolución de frecuencia del sistema auditivo humano, tal como se representa, por ejemplo, en la referencia [10].

[0156] Para una tasa de muestreo de 16 kHz, una longitud de bloque DFT para la STFT de 512 muestras y 15 grupos de particiones constituye una elección razonable, en la que cada partición posee un ancho de banda que corresponde aproximadamente al doble del ancho de banda rectangular equivalente (ERB), tal como se describe en la referencia [10].

[0157] Tal como se mostraba en la descripción anterior de formas de realización de la presente invención, en algunos casos las formas de realización de la presente invención comprenden unidades funcionales que incluyen las siguientes etapas a modo de breve resumen. Algunas formas de realización de la presente invención incluyen las operaciones de: recepción de al menos una señal de altavoz, recepción de al menos una señal de micrófono, conversión de la señal del altavoz y la señal del micrófono en espectros de tiempo corto, cálculo de los correspondientes espectros de potencia del altavoz y del micrófono, generación de unos parámetros de decaimiento del eco a partir de un modelo de decaimiento de eco predeterminado o un modelo de decaimiento de eco estimado dinámicamente, cálculo de un filtro de ganancia de eliminación de eco que modela el eco teniendo en cuenta los parámetros de decaimiento del eco, aplicación del filtro de ganancia al espectro del micrófono a fin de suprimir el eco, y conversión al dominio temporal del espectro del micrófono del que se ha suprimido el eco.

[0158] Aunque algunos aspectos se han descrito en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del procedimiento correspondiente, en el que un bloque o dispositivo corresponde a una etapa del procedimiento o a una característica de una etapa del procedimiento.

[0159] De forma análoga, los aspectos descritos en el contexto de una etapa de un procedimiento también representan una descripción de un correspondiente bloque, o elemento o característica de un aparato correspondiente. Algunas o todas las etapas del procedimiento se pueden ejecutar mediante (o usando) un aparato de tipo *hardware*, como, por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas formas de realización, dicho aparato puede ejecutar una o más de las etapas del procedimiento más importantes.

[0160] Dependiendo de ciertos requisitos de su implementación, las formas de realización de la invención se pueden implementar en forma de *hardware* o de *software*. La implementación se puede llevar a cabo usando un medio de almacenamiento digital, por ejemplo un disco flexible, un DVD, un *Blue-Ray*, un CD, una memoria ROM, PROM, EPROM o *Flash*, que tenga almacenadas unas señales de control de lectura electrónica, que cooperen (o puedan cooperar) con un sistema informático programable de manera que se lleve a cabo el respectivo procedimiento. Por lo tanto, el medio de almacenamiento digital puede ser de lectura por ordenador.

[0161] Algunas formas de realización de acuerdo con la invención comprenden un soporte de datos provisto de unas señales de control de lectura electrónica, que son capaces de cooperar con un sistema informático

programable, de manera que se lleve a cabo uno de los procedimientos descritos en el presente documento.

5 **[0162]** En general, las formas de realización de la presente invención se pueden implementar en forma de producto de programa informático con un código de programa, en el que se puede utilizar dicho código de programa para llevar a cabo uno de los procedimientos cuando se ejecuta en un ordenador el producto de programa informático. El código de programa se puede almacenar, por ejemplo, en un soporte de lectura mecánica.

10 **[0163]** Otras formas de realización comprenden el programa informático para llevar a cabo uno de los procedimientos descritos en el presente documento, almacenado en un soporte de lectura mecánica.

**[0164]** Dicho de otro modo, una forma de realización del procedimiento de la invención consiste, por tanto, en un programa informático que posee un código de programa para llevar a cabo uno de los procedimientos descritos en el presente documento, cuando se ejecuta el programa informático en un ordenador.

15 **[0165]** Otra forma de realización de los procedimientos de la invención consiste, por tanto, en un soporte de datos (o un medio de almacenamiento digital, o un medio de lectura por ordenador) que comprende, grabado, el programa informático para llevar a cabo uno de los procedimientos descritos en el presente documento. El soporte de datos, el medio de almacenamiento digital o el medio grabado suelen ser tangibles y/o no transitorios.

20 **[0166]** Otra forma de realización del procedimiento de la invención consiste, por tanto, en un flujo de datos o una secuencia de señales que representan el programa informático para llevar a cabo uno de los procedimientos descritos en el presente documento. El flujo de datos o la secuencia de señales pueden estar configurados para ser transferidos a través de una conexión de comunicación de datos, por ejemplo, a través de Internet.

25 **[0167]** Otra forma de realización comprende unos medios de procesamiento, por ejemplo, un ordenador o un dispositivo lógico programable, configurados o adaptados para llevar a cabo uno de los procedimientos descritos en el presente documento.

30 **[0168]** Otra forma de realización comprende un ordenador que tiene instalado el programa informático para llevar a cabo uno de los procedimientos descritos en el presente documento.

35 **[0169]** Otra forma de realización de acuerdo con la invención comprende un aparato o un sistema configurados para transferir (por ejemplo, por medios electrónicos u ópticos) a un receptor un programa informático para llevar a cabo uno de los procedimientos descritos en el presente documento. El receptor puede ser, por ejemplo, un ordenador, un dispositivo móvil, un dispositivo de memoria o similar. El aparato o sistema puede comprender, por ejemplo un servidor de archivos para transferir el programa informático al receptor.

40 **[0170]** En algunas formas de realización, se puede usar un dispositivo lógico programable (por ejemplo, una matriz de puertas programable in situ) para llevar a cabo algunas o todas las funciones de los procedimientos descritos en el presente documento. En algunas formas de realización, una matriz de puertas programable in situ puede cooperar con un microprocesador para llevar a cabo uno de los procedimientos descritos en el presente documento. En general, los procedimientos se llevan a cabo preferentemente mediante cualquier aparato de tipo *hardware*.

45 **[0171]** Las formas de realización descritas tienen un fin meramente ilustrativo de los principios de la presente invención. Se entiende que a otros expertos en la materia les parecerán evidentes algunas modificaciones y variaciones de las disposiciones y los detalles descritos en el presente documento. Por lo tanto, se considerará que la única limitación es la impuesta por el alcance de las reivindicaciones de patente y no por los detalles específicos presentados a título descriptivo y explicativo de las formas de realización incluidas en el presente documento.

50 **[0172]** El procesador puede estar compuesto por un ordenador, una tarjeta con chip (tarjeta inteligente), un sistema integrado (SOC = sistema integrado en un chip), un circuito integrado para aplicaciones específicas (ASIC) o cualquier otro circuito integrado (IC).

#### Referencias

55

**[0173]**

[1] C. Breining, P. Dreiseitel, E. Hänslér, A. Mader, B. Nitsch, H. Puder, T. Schertler, G. Schmidt, and J. Tilp. Acoustic echo control. IEEE Signal Processing Magazine, 16(4): 42 - 69, July 1999.

60

[2] A. N. Birkett and R. A. Goubran. Limitations of handsfree acoustic echo cancellers due to nonlinear loudspeaker distortion and enclosure vibration effects. In Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, pp. 13 - 16, New Paltz, Oct. 1995.

65

[3] G. Schmidt and E. Hänslér. Acoustic echo and noise control: a practical approach. Hoboken: Wiley, 2004.

- [4] W. L. B. Jeannes, P. Scalart, G. Faucon, and C. Beaugeant. Combined noise and echo reduction in hands-free systems: a survey. *IEEE Transactions on Speech and Audio Processing*, 9(8): 808 - 820, Nov. 2001.
- [5] C. Faller and J. Chen. Suppressing acoustic echo in a sampled auditory envelope space. *IEEE Trans. on Speech and Audio Proc.*, 13(5): 1.048 - 1.062, Sept. 2005.
- [6] C. Faller and C. Tournery. Estimating the delay and coloration effect of the acoustic echo path for low complexity echo suppression. In *Proc. Intl. Works. on Acoust. Echo and Noise Control (IWAENC)*, Sept. 2005.
- 10 [7] W. Etter and G. S. Moschytz. Noise reduction by noise-adaptive spectral magnitude expansion. *J. Audio Eng. Soc.*, 42: 341 - 349, May 1994.
- [8] O. Capp'e. Elimination of the musical noise phenomenon with the ephrain and malah noise suppressor. *IEEE Trans. Speech and Audio Processing*, 2(2): 345 - 349, April 1994.
- 15 [9] C. Faller and F. Baumgarte. Binaural Cue Coding - Part II: Schemes and applications. *IEEE Trans. on Speech and Audio Proc.*, 11(6): 520 - 531, Nov. 2003.
- [10] B. R. Glasberg and B. C. J. Moore. Derivation of auditory filter shapes from notched-noise data. *Hear. Res.*, 47: 20 103 - 138, 1990.
- [11] A. Favrot, C. Faller, M. Kallinger, F. Kuech, and M. Schmidt. Acoustic echo control based on temporal fluctuations of short-time spectra. In *Proc. Intl. Works. on Acoust. Echo and Noise Control (IWAENC)*, Sept. 2008.
- 25 [12] M. Berouti, R. Schwartz, and J. Makhoul. Enhancement of speech corrupted by musical noise. In *Proc. ICASSP*, pages 208-211, 1979.

Lista de números de referencia

- 30 **[0174]**
100. altavoz
110. micrófono
- 35 120. entorno acústico
130. señal del altavoz
- 40 140. señal del micrófono
150. unidad de proceso de eliminación de eco
160. señal con eco suprimido
- 45 170. trayectoria directa
180. trayectoria indirecta
- 50 200. aparato
210. filtro adaptativo
220. entradas
- 55 230. medios convertidores de tiempo/frecuencia
240. filtro de estimación de eco
- 60 241. filtro de estimación de eco tardío
242. filtro de estimación de eco temprano
250. medios de separación
- 65 260. filtro de estimación de eco adicional

- 270. medios de cálculo
- 270e. primeros medios de cálculo (para el filtro de estimación de eco temprano  $H_e[k,m]$ )
- 5 270i. segundos medios de cálculo (para el filtro de estimación de eco tardío  $H_i[k,m]$ )
- 280. entrada
- 10 290. medios convertidores de tiempo/frecuencia
- 300. medios convertidores de frecuencia/tiempo
- 310. salida
- 15 340. medios de procesamiento de eco
- 370. filtro de supresión de eco
- 20 380. medios de combinación / combinador de filtros
- 420. curva (señal del altavoz)
- 430. curva (señal del micrófono)
- 25 440. corchete: intervalo STFT
- 450. corchete: intervalo STFT
- 30 460. retardo entre la señal del micrófono y la señal del altavoz
- 465. medios de modelado del decaimiento del eco
- 470. medios de función de estimación de eco
- 35 472. medios de determinación de filtro de eco temprano
- 475. medios de determinación de filtro de eco tardío
- 40 480. medios de retardo
- 490. medios de cálculo de valores de energía
- 500. medios de cálculo de valores de energía
- 45 510. medios de cálculo de valores de energía
- 520. medios de cálculo adicionales; cálculo de la función o funciones de estimación de eco
- 50 540. medios de agrupamiento adicionales
- 550. selector de máximo
- 551. medios de establecimiento de filtro
- 55 552. comparador
- 553. selector
- 60 560. elemento de retardo
- 570. factor escalar  $\alpha_m$
- 580. factor escalar  $\beta_m$
- 65 590. medios normalizadores

940. filtro de estimación de eco

$d[k,m]$ . valor de retardo para retardo entre señales de micrófono y de altavoz

5

$x[n]$ . señal de altavoz

$y[n]$ . señal de micrófono

10  $E[k,m]$ . espectro de la señal con eco suprimido

$G[k,m]$ . filtro de estimación de eco

$H[k,m]$ . filtro de supresión de eco

15

$H_e[k,m]$ . filtro de supresión de eco temprano

$H_i[k,m]$ . filtro de supresión de eco tardío

20  $X[k,m]$ . espectro de la señal del altavoz

$X_d[k,m]$ . espectro de la señal del altavoz retardada

$Y[k,m]$ . espectro de la señal del micrófono

25

$\hat{Y}_e[k,m]$ . espectro del componente de eco temprano estimado en la señal del micrófono

$\hat{Y}_{rev}[k,m]$ . espectro del componente de eco tardío (reverberante) estimado en la señal del micrófono

30  $\hat{Y}_{tot}[k,m]$ . espectro del eco total estimado en la señal del micrófono

## REIVINDICACIONES

1. Un aparato (200) para calcular coeficientes de filtrado ( $H[k,m]$ ) para un filtro adaptativo (210) para filtrar una señal de micrófono (140, 430) captada por un micrófono (110) a fin de suprimir un eco debido a una señal de altavoz (130, 420) generada por un altavoz (100), que comprende:
- unos medios de modelado del decaimiento del eco (465) para modelar un comportamiento de retardo de un entorno acústico (120) y para proporcionar un parámetro de decaimiento del eco ( $\tau$ ;  $\alpha_m$ ) correspondiente;
- 10 unos medios de cálculo (270; 370) para calcular los coeficientes de filtrado ( $H[k,m]$ ) del filtro adaptativo (210) basándose en el parámetro de decaimiento del eco ( $\tau$ ;  $\alpha_m$ ), en el que los medios de cálculo (270) están dispuestos para determinar un filtro de eliminación de eco temprano  $H_e[k,m]$  y un filtro de eliminación de eco tardío  $H_l[k,m]$ ; y
- unos medios de establecimiento de filtro (551) dispuestos para determinar los coeficientes de filtrado  $H[k,m]$  del filtro adaptativo (210) basándose en una combinación lineal de coeficientes de filtrado del filtro de eliminación de eco temprano  $H_e[k,m]$  y del filtro de eliminación de eco tardío  $H_l[k,m]$ .
- 15
2. El aparato (200) de acuerdo con la reivindicación 1, en el que el parámetro de decaimiento del eco es un coeficiente de decaimiento exponencial ( $\tau$ ) y en el que el aparato (200) también comprende unos medios de estimación de eco (241) para estimar un componente de eco reverberante por medio de un cálculo recursivo en el que se usa el coeficiente de decaimiento exponencial ( $\tau$ ).
- 20
3. El aparato (200) de acuerdo con la reivindicación 1, que también comprende unos medios de procesamiento de eco (340) para estimar un espectro de frecuencia de eco ( $Y_{tot}[k,m]$ ) o un espectro de potencia de eco ( $|Y_{tot}[k,m]|^2$ ) del eco contenido en la señal del micrófono (140, 430);
- 25
- en el que los medios de procesamiento de eco (340) están dispuestos para estimar un componente temprano del eco ( $Y[k,m]$ ) y un componente tardío del eco ( $Y_{rev}[k,m]$ ), y al menos la estimación del componente tardío del eco usa el parámetro de decaimiento del eco ( $\tau$ ;  $\alpha_m$ ); estando también dispuestos para estimar el espectro de frecuencia del eco ( $Y_{tot}[k,m]$ ) o el espectro de potencia del eco ( $|Y_{tot}[k,m]|^2$ ) basándose en el componente temprano del eco ( $Y[k,m]$ ) y el componente tardío del eco ( $Y_{rev}[k,m]$ );
- 30
- en el que los medios de cálculo (270; 370) están dispuestos para calcular los coeficientes de filtrado ( $H[k,m]$ ) basándose en el espectro de frecuencia del eco ( $Y_{tot}[k,m]$ ) o espectro de potencia del eco ( $|Y_{tot}[k,m]|^2$ ).
- 35
4. El aparato (200) de acuerdo con la reivindicación 3, en el que los medios de procesamiento de eco (340) comprenden un selector de máximo (550) para seleccionar una estimación de eco máxima ( $|Y_{tot}[k,m]|^2$ ;  $Y_{tot}[k,m]$ ) entre la estimación de eco temprano  $Y[k,m]$  y la estimación de eco tardío ( $Y_{rev}[k,m]$ ) como espectro de frecuencia del eco ( $Y_{tot}[k,m]$ ) o el espectro de potencia del eco ( $|Y_{tot}[k,m]|^2$ ).
- 40
5. El aparato (200) de acuerdo con una de las reivindicaciones 3 o 4, en el que los medios de procesamiento de eco están dispuestos para determinar el componente de eco tardío ( $Y_{rev}[k,m]$ ) basándose en un valor previo del espectro de frecuencia del eco estimado ( $Y_{tot}[k-1,m]$ ) o un valor previo del espectro de potencia del eco ( $|Y_{tot}[k-1,m]|^2$ ), determinados en un instante previo, y basándose en el parámetro de decaimiento del eco ( $\tau$ ;  $\alpha_m$ ) aplicado a la estimación de eco total previa ( $Y_{tot}[k-1,m]$ ;  $|Y_{tot}[k-1,m]|^2$ ) para modelar un decaimiento de la estimación de eco tardío entre el instante previo y un instante posterior.
- 45
6. El aparato (200) de acuerdo con una cualquiera de las reivindicaciones precedentes, que además comprende:
- 50
- unos medios para transformar (230; 290) al menos una de las dos señales, la señal del micrófono (140, 430) y la señal del altavoz (130; 420), para que pasen de una representación en el dominio temporal a una representación de tiempo-frecuencia en un dominio de tiempo-frecuencia; y unos medios para transformar (300) una señal con eco suprimido ( $E[k,m]$ ,  $e[n]$ ) para que vuelva a pasar del dominio de tiempo-frecuencia al dominio temporal.
- 55
7. El aparato (200) de acuerdo con una cualquiera de las reivindicaciones precedentes, en el que los medios de modelado del decaimiento del eco (465) comprenden un estimador del parámetro de decaimiento del eco (520) para evaluar la señal del micrófono (140, 430) con respecto a un decaimiento del eco observable en la señal del micrófono (140, 430) y para proporcionar el parámetro de decaimiento del eco resultante ( $\tau$ ;  $\alpha_m$ ) para que continúe el procesamiento.
- 60
8. El aparato (200) de acuerdo con una cualquiera de las reivindicaciones precedentes, en el que los medios de cálculo (270; 370) están dispuestos para determinar los coeficientes de filtrado  $H[k,m]$  basándose en una diferencia entre un espectro de potencia ( $|Y[k,m]|^2$ ) de la señal del micrófono (140, 430) y un espectro de potencia de una estimación de eco total ( $|Y_{tot}[k,m]|^2$ ).
- 65

9. El aparato (200) de acuerdo con una cualquiera de las reivindicaciones precedentes, en el que el modelo de decaimiento del eco (465) está dispuesto para determinar una media temporal de un espectro de potencia de la señal del micrófono y una media temporal de un espectro de potencia de la señal del altavoz, y también están dispuestos para determinar fluctuaciones temporales de los espectros de potencia de la señal del micrófono (140, 430) restando la media temporal de los correspondientes espectros de potencia instantánea ( $|Y[k,m]|^2$ ,  $|X[k,m]|^2$ ), y también está dispuesto para usar las fluctuaciones temporales para la determinación del parámetro de decaimiento del eco ( $\tau$ ;  $\alpha_m$ ).
10. Un aparato (200) para calcular coeficientes de filtrado ( $H[k,m]$ ) para un filtro adaptativo (210) para filtrar una señal de micrófono (140, 430) captada por un micrófono (110) a fin de suprimir un eco debido a una señal de altavoz (130, 420) generada como salida por un altavoz (100), que comprende:
- unos medios de modelado del decaimiento del eco (465) para modelar un comportamiento de retardo de un entorno acústico (120) y para proporcionar un parámetro de decaimiento del eco ( $\tau$ ;  $\alpha_m$ ) correspondiente;
- unos medios de procesamiento de eco (340) para estimar un espectro de frecuencia del eco ( $Y_{tot}[k,m]$ ) o un espectro de potencia del eco ( $|Y_{tot}[k,m]|^2$ ) del eco contenido en la señal del micrófono (140, 430), estando dichos medios de procesamiento de eco (340) dispuestos para estimar un componente temprano del eco ( $Y[k,m]$ ) y un componente tardío del eco ( $Y_{rev}[k,m]$ ), y al menos la estimación del componente tardío del eco usa el parámetro de decaimiento del eco ( $\tau$ ;  $\alpha_m$ ), y estando también dispuestos para estimar el espectro de frecuencia del eco ( $Y_{tot}[k,m]$ ) o el espectro de potencia del eco ( $|Y_{tot}[k,m]|^2$ ) basándose en el componente temprano del eco ( $Y[k,m]$ ) y el componente tardío del eco ( $Y_{rev}[k,m]$ ); y
- unos medios de cálculo (270; 370) para calcular los coeficientes de filtrado ( $H[k,m]$ ) del filtro adaptativo (210) basándose en el parámetro de decaimiento del eco ( $\tau$ ;  $\alpha_m$ ) y basándose en el espectro de frecuencia del eco ( $Y_{tot}[k,m]$ ) o el espectro de potencia del eco ( $|Y_{tot}[k,m]|^2$ ) del eco contenido en la señal del micrófono (140; 340);
- caracterizado porque** los medios de procesamiento de eco (340) comprenden un selector de máximo (550) para seleccionar una estimación de eco máxima ( $|Y_{tot}[k,m]|^2$ ;  $Y_{tot}[k,m]$ ) entre la estimación de eco temprano  $Y[k,m]$  y la estimación de eco tardío ( $Y_{rev}[k,m]$ ) como espectro de frecuencia del eco ( $Y_{tot}[k,m]$ ) o espectro de potencia del eco ( $|Y_{tot}[k,m]|^2$ ).
11. Un aparato (200) para calcular coeficientes de filtrado ( $H[k,m]$ ) para un filtro adaptativo (210) para filtrar una señal de micrófono (140, 430) captada por un micrófono (110) a fin de suprimir un eco debido a una señal de altavoz (130, 420) generada como salida por un altavoz (100), que comprende:
- unos medios de modelado del decaimiento del eco (465) para modelar un comportamiento de retardo de un entorno acústico (120) y para proporcionar un parámetro de decaimiento del eco ( $\tau$ ;  $\alpha_m$ ) correspondiente;
- unos medios de cálculo (270; 370) para calcular los coeficientes de filtrado ( $H[k,m]$ ) del filtro adaptativo (210) basándose en el parámetro de decaimiento del eco ( $\tau$ ;  $\alpha_m$ ), y dichos medios de cálculo (270) también están dispuestos para determinar un filtro de eliminación de eco temprano  $H_e[k,m]$  y un filtro de eliminación de eco tardío  $H_i[k,m]$ ;
- unos medios de establecimiento de filtro (551) dispuestos para determinar los coeficientes de filtrado  $H[k,m]$  del filtro adaptativo (210) basándose en una selección de mínimo entre el filtro de eliminación de eco temprano  $H_e[k,m]$  y el filtro de eliminación de eco tardío  $H_i[k,m]$ .
12. Un procedimiento para calcular coeficientes de filtrado ( $H[k,m]$ ) para un filtro adaptativo (210) para filtrar una señal de micrófono (140, 430) a fin de suprimir un eco debido a una señal de altavoz (130, 420), que comprende las operaciones de:
- generación de un parámetro de decaimiento del eco ( $\tau$ ;  $G[k,m]$ ;  $\alpha_m$ ) determinado mediante unos medios de modelado del decaimiento del eco (465);
- cálculo de los coeficientes de filtrado ( $H[k,m]$ ) del filtro adaptativo (210) basándose en el parámetro de decaimiento del eco ( $\tau$ ;  $\alpha_m$ ), en los que el cálculo comprende la determinación de un filtro de eliminación de eco temprano  $H_e[k,m]$  y un filtro de eliminación de eco tardío  $H_i[k,m]$ ;
- determinación de los coeficientes de filtrado ( $H[k,m]$ ) del filtro adaptativo (210) basándose en una combinación lineal de los coeficientes de filtrado del filtro de eliminación de eco temprano  $H_e[k,m]$  y del filtro de eliminación de eco tardío  $H_i[k,m]$ .
13. El procedimiento de acuerdo con la reivindicación 12, en el que el parámetro de decaimiento del eco

es un coeficiente de decaimiento exponencial ( $\tau$ ) y en el que el procedimiento también comprende:

La estimación de un componente de eco reverberante por medio de un cálculo recursivo en el que se usa el coeficiente de decaimiento exponencial ( $\tau$ ).

5

14. El procedimiento de acuerdo con la reivindicación 12, que además comprende:

la estimación de un componente temprano del eco ( $Y[k,m]$ ) y un componente tardío del eco ( $Y_{rev}[k,m]$ ), y al menos en la estimación del componente tardío del eco se usa el parámetro de decaimiento del eco ( $\tau$ ;  $\alpha_m$ ); y

10

la determinación de un espectro de frecuencia del eco ( $Y[k,m]$ ) o un espectro de potencia del eco ( $|Y[k,m]|^2$ ) del eco contenido en la señal del micrófono (140, 430) como una combinación del componente temprano del eco ( $Y[k,m]$ ) y el componente tardío del eco ( $Y_{rev}[k,m]$ );

15 el cálculo de los coeficientes de filtrado ( $H[k,m]$ ) basándose en el espectro de frecuencia del eco ( $Y_{tot}[k,m]$ ) o el espectro de potencia del eco ( $|Y_{tot}[k,m]|^2$ ).

15. Un procedimiento para calcular coeficientes de filtrado ( $H[k,m]$ ) para un filtro adaptativo (210) para filtrar una señal de micrófono (140, 430) a fin de suprimir un eco debido a una señal de altavoz (130, 420), que comprende las operaciones de:

20

generación de parámetros de decaimiento del eco ( $\tau$ ;  $G[k,m]$ ;  $\alpha_m$ ) determinados mediante unos medios de modelado del decaimiento del eco (465);

25 estimación de un espectro de frecuencia de eco ( $Y_{tot}[k,m]$ ) o un espectro de potencia de eco ( $|Y_{tot}[k,m]|^2$ ) del eco contenido en la señal del micrófono (140, 430), mediante la estimación de un componente temprano del eco ( $Y[k,m]$ ) y un componente tardío del eco ( $Y_{rev}[k,m]$ ), y en al menos una estimación del componente tardío del eco se usa el parámetro de decaimiento del eco ( $\tau$ ;  $\alpha_m$ );

30 estimación del espectro de frecuencia del eco ( $Y_{tot}[k,m]$ ) o del espectro de potencia del eco ( $|Y_{tot}[k,m]|^2$ ) basándose en el componente temprano del eco ( $Y[k,m]$ ) y en el componente tardío del eco ( $Y_{rev}[k,m]$ );

selección de una máxima estimación de eco ( $|Y_{tot}[k,m]|^2$ ;  $Y_{tot}[k,m]$ ) entre la estimación del eco temprano ( $Y[k,m]$ ) y la estimación del eco tardío ( $Y_{rev}[k,m]$ ), como espectro de frecuencia del eco ( $Y_{tot}[k,m]$ ) o espectro de potencia del eco ( $|Y_{tot}[k,m]|^2$ ); y

35

cálculo de los coeficientes de filtrado ( $H[k,m]$ ) del filtro adaptativo (210) basándose en los parámetros de decaimiento del eco ( $\tau$ ;  $\alpha_m$ ) y basándose en un espectro de frecuencia de eco ( $Y_{tot}[k,m]$ ) o un espectro de potencia de eco ( $|Y_{tot}[k,m]|^2$ ) del eco contenido en la señal del micrófono (140; 340).

40

16. Un procedimiento para calcular coeficientes de filtrado ( $H[k,m]$ ) para un filtro adaptativo (210) para filtrar una señal de micrófono (140, 430) a fin de suprimir un eco debido a una señal de altavoz (130, 420), que comprende las operaciones de:

45 generación de parámetros de decaimiento del eco ( $\tau$ ;  $G[k,m]$ ;  $\alpha_m$ ) determinados mediante unos medios de modelado del decaimiento del eco (465); y

cálculo de los coeficientes de filtrado ( $H[k,m]$ ) del filtro adaptativo (210) basándose en parámetros de decaimiento del eco ( $\tau$ ;  $\alpha_m$ ), mediante la determinación de un filtro de eliminación de eco temprano  $H_e[k,m]$  y un filtro de eliminación de eco tardío  $H_t[k,m]$ ; y

50

determinación de los coeficientes de filtrado  $H[k,m]$  del filtro adaptativo (210) basándose en una selección de un mínimo entre el filtro de eliminación de eco temprano  $H_e[k,m]$  y el filtro de eliminación de eco tardío  $H_t[k,m]$ .

55 17. Un programa que comprende un código de programa adaptado para llevar a cabo el procedimiento de acuerdo con una cualquiera de las reivindicaciones 12 a 16, cuando el programa se ejecuta en un procesador.

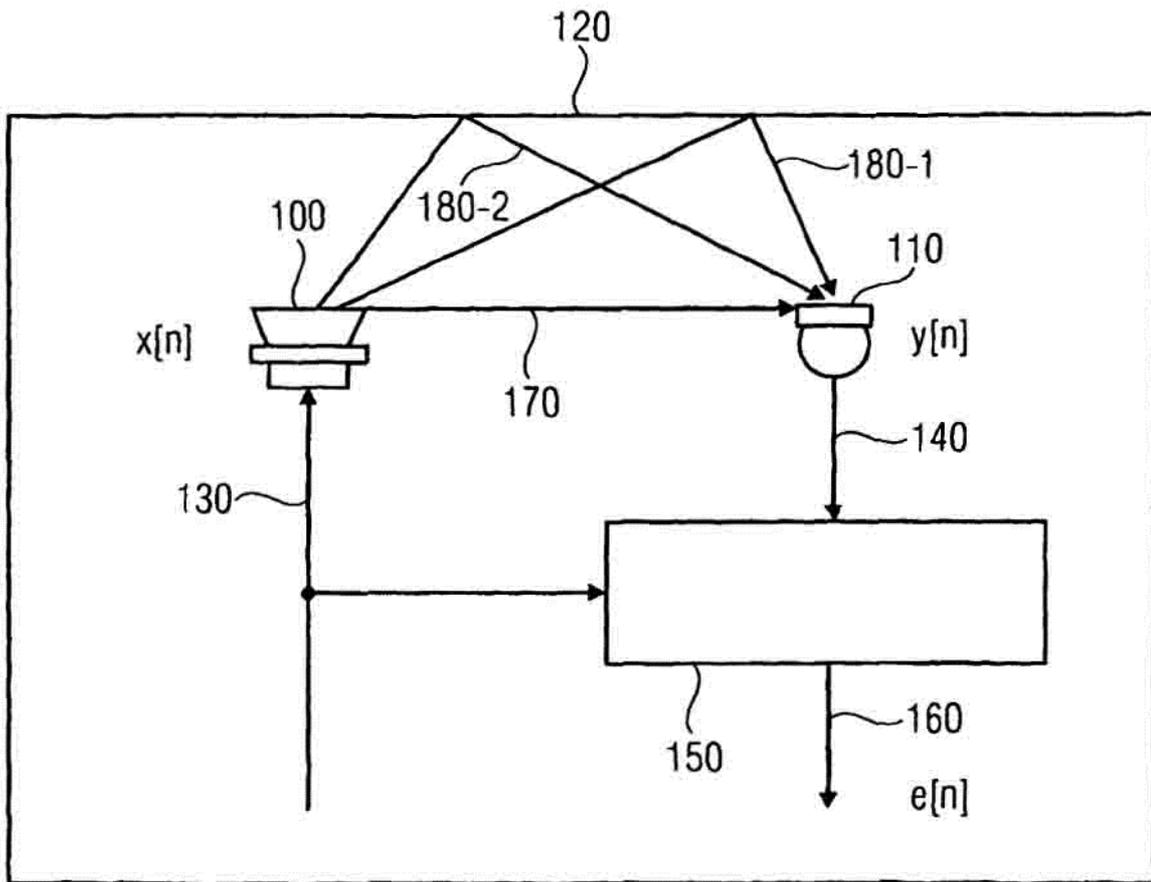


FIG 1

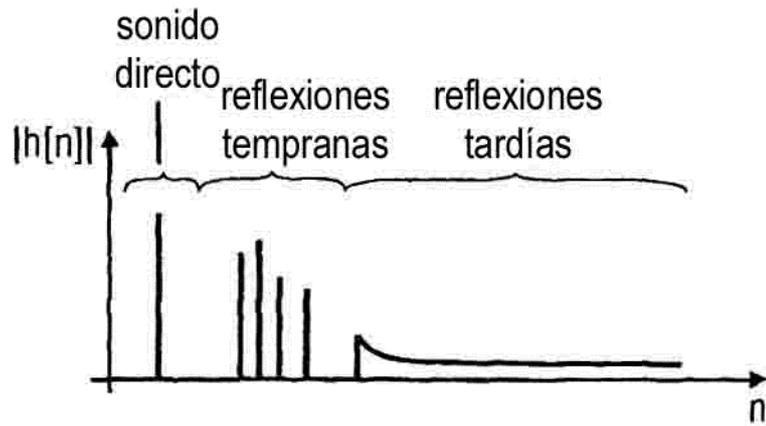


FIG 2

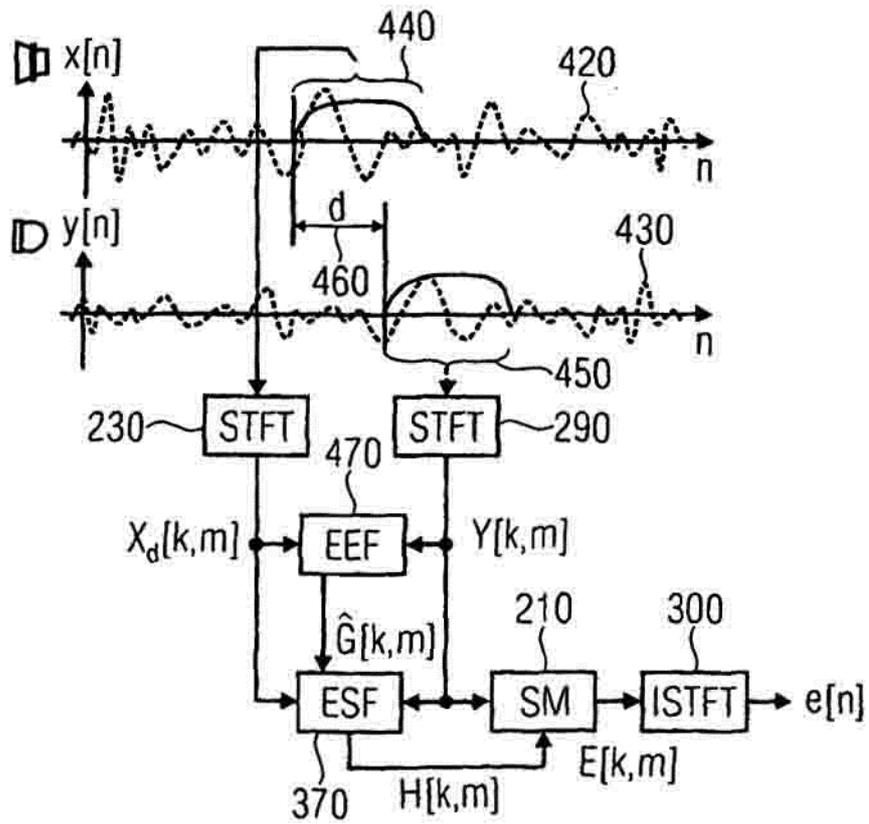


FIG 3

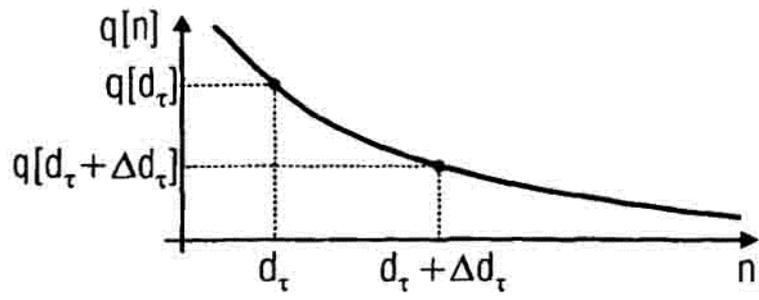


FIG 4

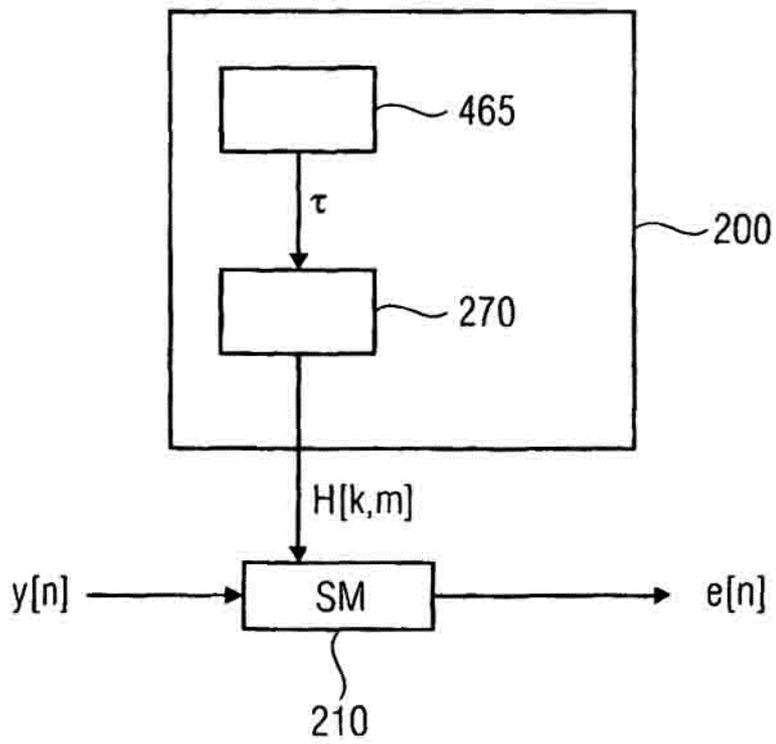


FIG 5

200

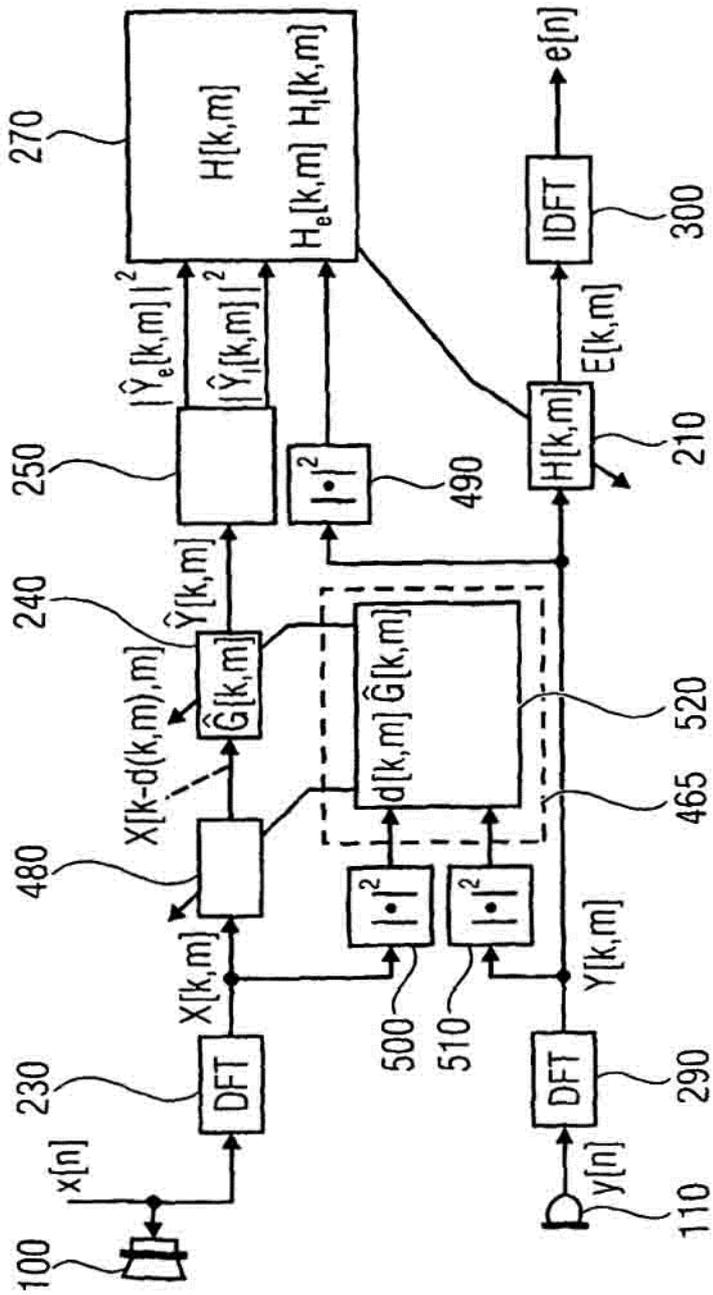


FIG 6



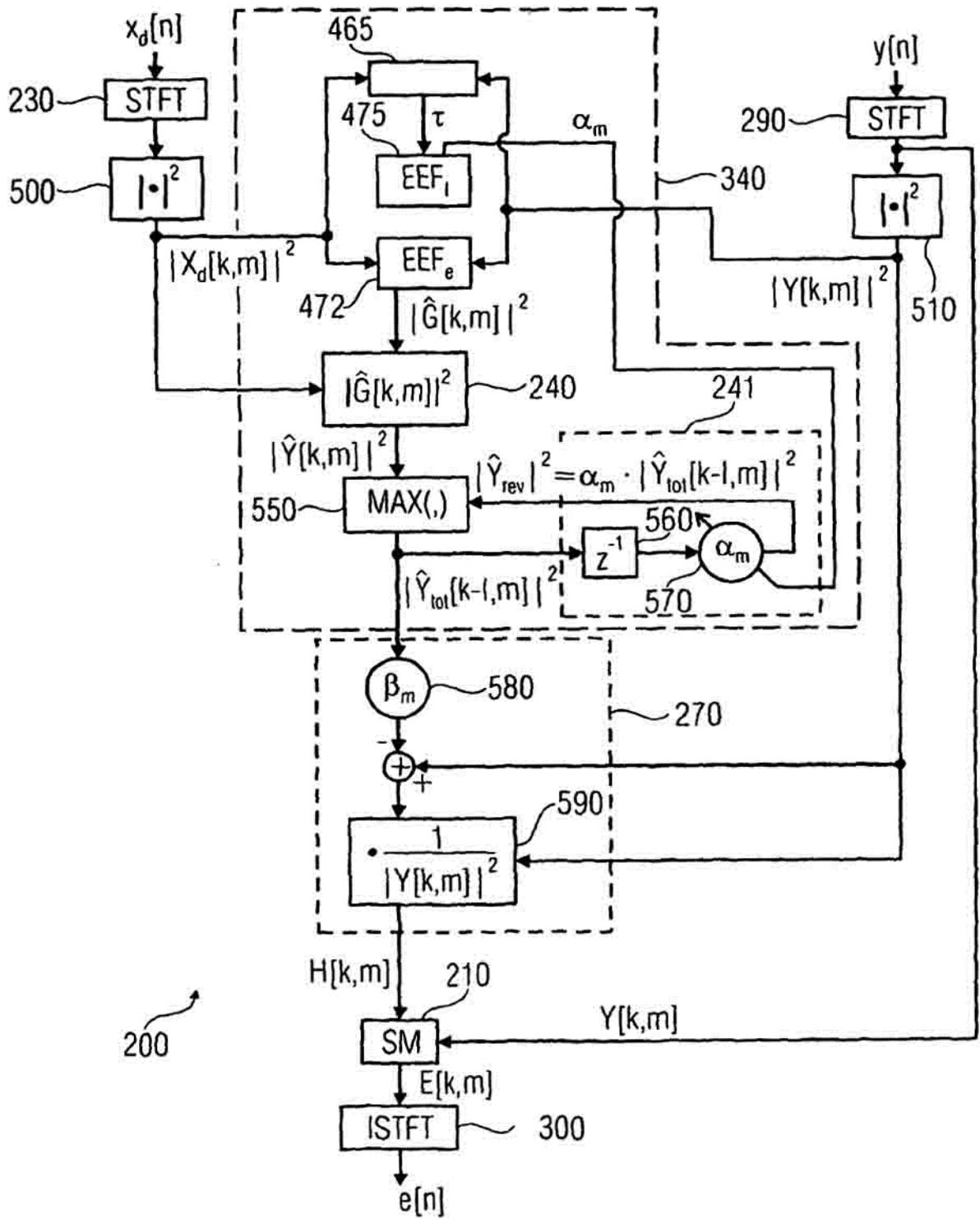


FIG 8



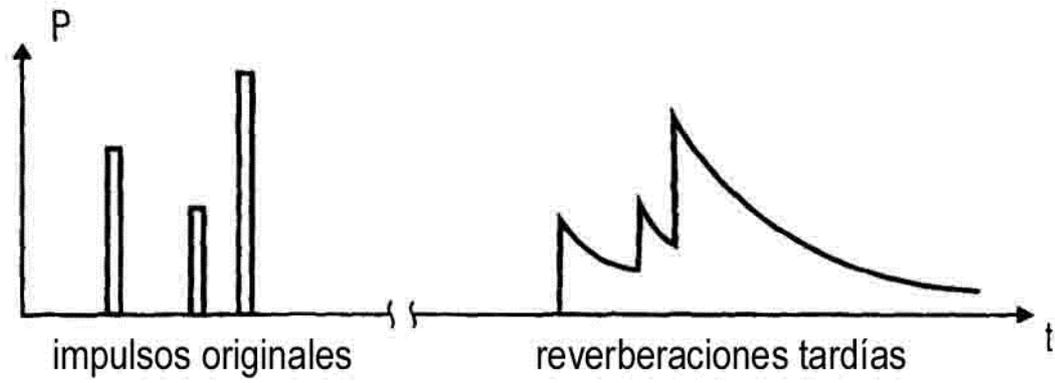


FIG 10

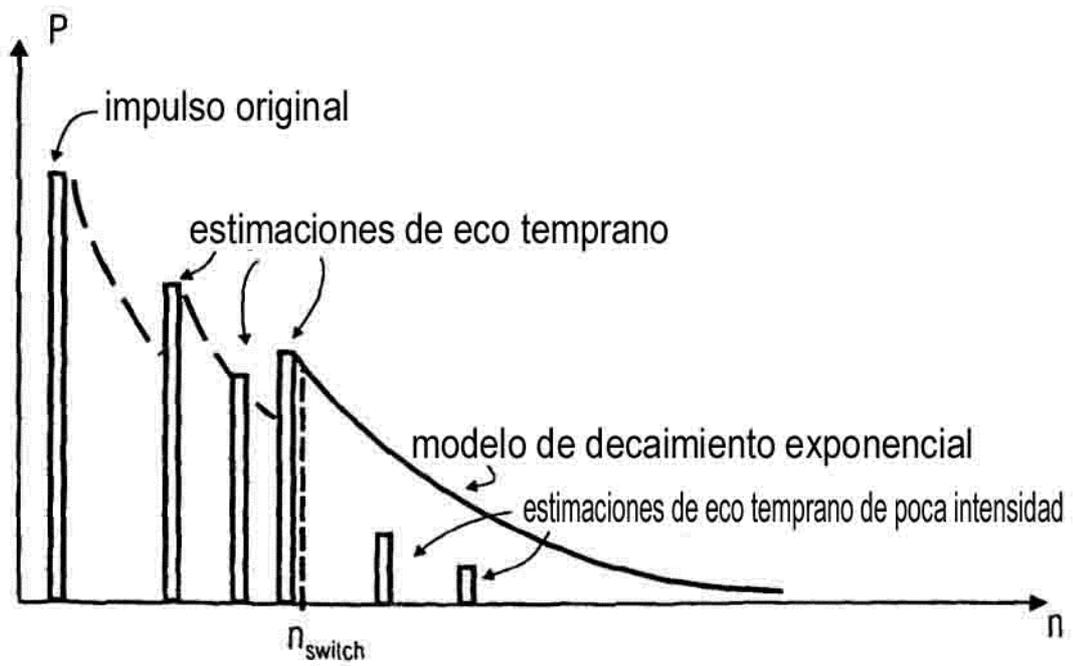


FIG 11

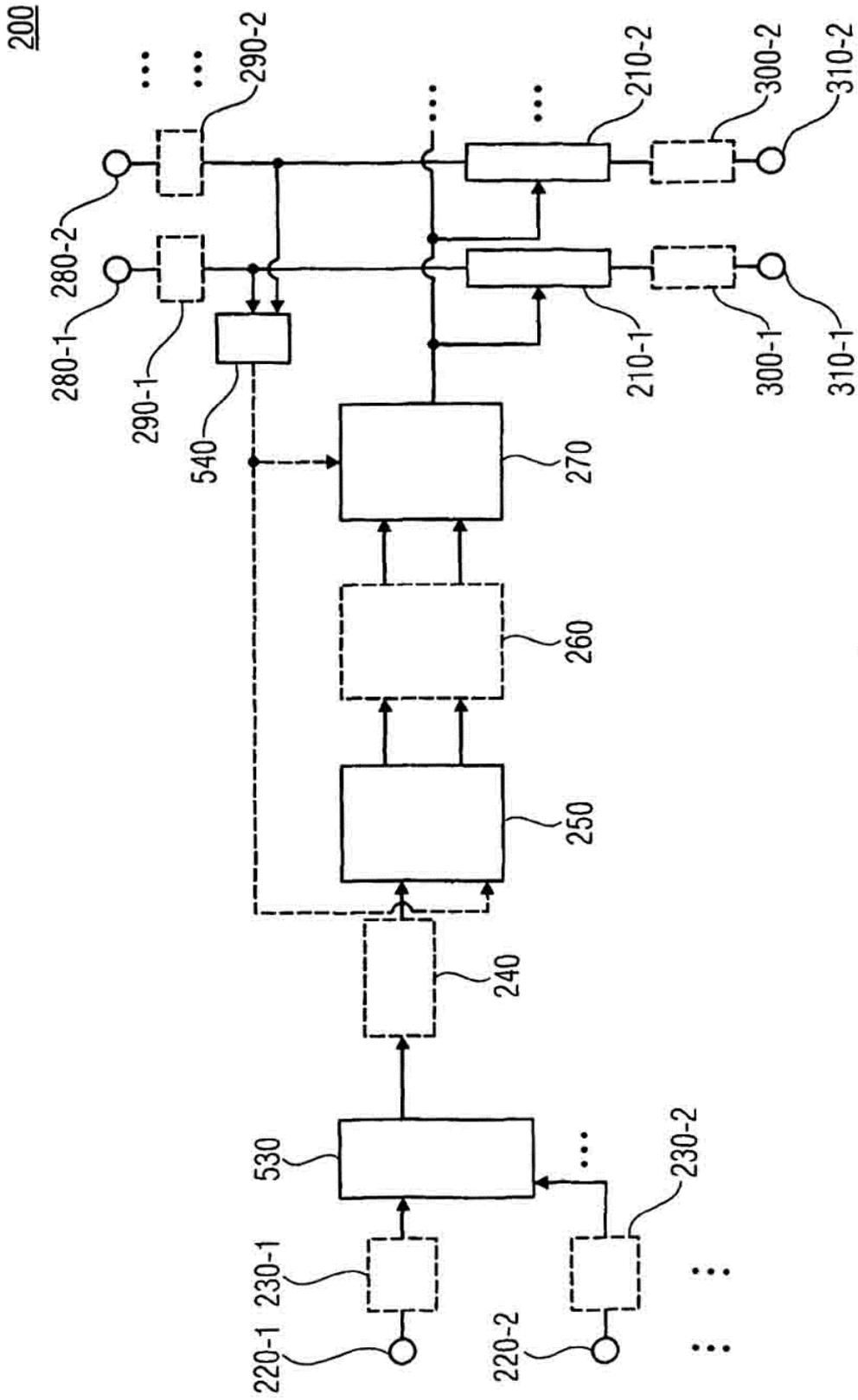


FIG 12