

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 987 930**

51 Int. Cl.:

**H04S 7/00**

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **11.04.2017 PCT/EP2017/058611**

87 Fecha y número de publicación internacional: **19.10.2017 WO17178454**

96 Fecha de presentación y número de la solicitud europea: **11.04.2017 E 17715755 (9)**

97 Fecha y número de publicación de la concesión europea: **14.08.2024 EP 3443761**

54 Título: **Aparato y método para proporcionar zonas individuales de sonido**

30 Prioridad:

**12.04.2016 EP 16164984**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**18.11.2024**

73 Titular/es:

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

72 Inventor/es:

**SCHNEIDER, MARTIN;  
WETZEL, STEFAN;  
WALTHER, ANDREAS;  
UHLE, CHRISTIAN;  
HELLMUTH, OLIVER;  
PROKEIN, PETER y  
HABETS, EMANUEL**

74 Agente/Representante:

**ARIZTI ACHA, Monica**

ES 2 987 930 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

Aparato y método para proporcionar zonas individuales de sonido

5 La presente invención se refiere a procesamiento de señales de audio y, en particular, a un aparato y método para proporcionar zonas individuales de sonido.

La reproducción de diferentes escenas acústicas en múltiples zonas de sonido ubicadas cerca sin barreras acústicas entremedio es una tarea bien conocida en procesamiento de señales de audio, que a menudo es llamada como reproducción multizona (ver, [1]). Desde el punto de vista técnico, la reproducción multizona está cercanamente relacionada con formación de puntos o formación de haces de altavoz (ver, [2]) cuando se consideran escenarios de campo cercano, donde la apertura de arreglo de altavoces también puede rodear al oyente.

10 Un problema de un escenario de reproducción multizona, por ejemplo, puede ser proporcionar escenas acústicas sustancialmente diferentes (por ejemplo, diferentes piezas de música o contenido de audio de diferentes películas) a los oyentes que ocupan zonas de sonido individuales.

Un ejemplo ideal simplificado de reproducción multizona se muestra en la figura 2, donde las dos zonas 221, 222 reciben las señales  $u_1(k)$  y  $u_2(k)$  de dos fuentes de señales 211, 212, respectivamente, sin interferencia de la otra fuente y  $k$  que es el instante del tiempo. Se debe señalar que este escenario solamente es un marcador para escenarios más complejos, donde se proporciona audio multicanal a un número arbitrario de zonas. Sin embargo, el ejemplo simple mostrado en la figura 2 es suficiente para las explicaciones en lo siguiente.

15 Cuando se reproducen múltiples señales en un recinto real, es imposible una separación perfecta ya que no se pueden detener ondas acústicas sin una barrera acústica. Por lo tanto, siempre habrá diafonía entre las zonas individuales de sonido, que están ocupadas por oyentes individuales.

La figura 3 ilustra una reproducción de múltiples señales en la realidad. Las señales reproducidas en las zonas individuales de sonido 221, 222, es decir  $y_1(k)$  y  $y_2(k)$ , se obtienen al convolucionar las fuentes de señales  $u_1(k)$  y  $u_2(k)$  de las fuentes de señales 211, 212 con las respuestas de impulso respectivas  $h_{1,1}(k)$ ,  $h_{2,2}(k)$ ,  $h_{1,2}(k)$ , y  $h_{2,1}(k)$  del LEMS (sistema de micrófono-recinto-altavoz) de acuerdo con

$$y_1(k) = y_{1,1}(k) + y_{1,2}(k) = u_1(k) * h_{1,1}(k) + u_2(k) * h_{1,2}(k), \quad (1)$$

$$y_2(k) = y_{2,2}(k) + y_{2,1}(k) = u_2(k) * h_{2,2}(k) + u_1(k) * h_{2,1}(k), \quad (2)$$

en donde \* denota la convolución, como se define por

$$u_1(k) * h_{1,1}(k) = \sum_{n=-\infty}^{\infty} u_1(k)h_{1,1}(k - n). \quad (3)$$

40 Aquí,  $y_{1,2}(k)$  y  $y_{2,1}(k)$  se considera que son componentes de señales de interferencia, indeseados, en contraste con los componentes deseados  $y_{1,1}(k)$  y  $y_{2,2}(k)$ . Cuando  $u_1(k)$  and  $u_2(k)$  describen escenas acústicas completamente diferentes, sólo una contribución muy pequeña de  $u_2(k)$  en  $y_1(k)$  comparación con la convolución de  $u_1(k)$  en  $y_1(k)$  es aceptable. Lo mismo se mantiene para  $y_2(k)$  con índices reservados.

45 Una forma directa para lograr esto es diseñar la configuración de altavoz de tal forma que  $h_{1,1}(k)$  y  $h_{2,2}(k)$  exhiben una mayor energía en comparación con  $h_{1,2}(k)$  and  $h_{2,1}(k)$ , que describen reproducción de zonas cruzadas. Un ejemplo de esto sería utilizar altavoces ubicados cerca de los oyentes (US 2003103636, US 2003142842), donde el uso de auriculares se puede ver como un caso extremo de esta configuración. Sin embargo, la colocación de altavoces demasiado cerca de los oyentes a menudo es inaceptable, debido a que esto puede interferir con el movimiento de los oyentes, de tal forma que este enfoque se limita a aplicaciones prácticas.

50 Un enfoque para lograr esto, es utilizar altavoces direccionales, donde la directividad del altavoz es covalentemente más alta para mayores frecuencias (ver [35]): JP 5345549, y [21]: US 2005/0190935 A1. Desafortunadamente, este enfoque sólo es adecuado para frecuencias más altas (ver, [1]).

Otro enfoque es utilizar un arreglo de altavoces junto con pre-filtros adecuados para una reproducción de audio personalizada.

La figura 4 ilustra un ejemplo mínimo de reproducción multizona con arreglos. En particular, la figura 4 ilustra una configuración rudimentaria con dos fuentes de señales 211, 212, dos altavoces y dos zonas 221, 222. El ejemplo de la figura 4 es un marcador para escenarios más complejos que se presentan en aplicaciones reales.

5 En el ejemplo de la figura 4, la cantidad de reproducción de zonas cruzadas se determina por la cascada de los pre-filtros  $\mathbf{G}(k)$  413, 414 y las respuestas de impulso  $\mathbf{H}(k)$  417 y no sólo por  $\mathbf{H}(k)$  417. Por lo tanto,  $h_{1,2}(k)$  y  $h_{2,1}(k)$  no necesariamente tienen que ser pequeños en magnitud a fin de lograr una atenuación de zonas cruzadas considerable.

10 La figura 6 ilustra un modelo de señal general de reproducción multizona con arreglos. Las fuentes de señales 610, los pre-filtros 615, las respuestas de impulso 417 y las zonas de sonido 221, 222 se representan.

15 Se debe señalar que la reproducción multizona generalmente no se limita a proporcionar dos señales a dos zonas. En realidad, los números de fuentes, altavoces y zonas de escucha pueden ser arbitrarios. Las siguientes explicaciones y definiciones se pueden utilizar para un escenario general con  $N_S$  fuentes de señales,  $N_L$  altavoces, y  $N_M$  posiciones consideradas en las  $N_Z$  zonas de escucha. En este escenario, es posible que se reproduzcan múltiples señales en una zona individual para lograr una reproducción de sonido espacial. El modelo de señal correspondiente se muestra en la figura 6, en donde "zona 1" 221 se suministra con las señales  $y_1(k)$  y  $y_2(k)$ . Los vectores de señal resultantes están dados por:

$$20 \quad \mathbf{u}(k) = (u_1(k), u_2(k), \dots, u_{N_S}(k))^T, \quad (4)$$

$$\mathbf{x}(k) = (x_1(k), x_2(k), \dots, x_{N_L}(k))^T, \quad (5)$$

$$25 \quad \mathbf{y}(k) = (y_1(k), y_2(k), \dots, y_{N_M}(k))^T, \quad (6)$$

$$\mathbf{x}(k) = \mathbf{G}(k) * \mathbf{u}(k), \quad (7)$$

$$\mathbf{y}(k) = \mathbf{H}(k) * \mathbf{x}(k). \quad (8)$$

30 Aquí, se da una reproducción de la Ecuación (3) por

$$\mathbf{G}(k) * \mathbf{u}(k) = \sum_{n=0}^{L_G-1} \mathbf{G}(k) \mathbf{u}(k-n), \quad (9)$$

35 asumiendo que las respuestas de impulso capturadas en se limitan a que son diferentes de cero sólo para  $0 \leq k < L_G$ .

Las matrices  $\mathbf{G}(k)$  y  $\mathbf{H}(k)$  describen las respuestas de impulso del pre-filtro y las respuestas de impulso de la sala de acuerdo con

$$40 \quad \mathbf{G}(k) = \begin{pmatrix} g_{1,1}(k) & g_{1,2}(k) & \dots & g_{1,N_S}(k) \\ g_{2,1}(k) & g_{2,2}(k) & \dots & g_{2,N_S}(k) \\ \vdots & \vdots & \ddots & \vdots \\ g_{N_L,1}(k) & g_{N_L,2}(k) & \dots & g_{N_L,N_S}(k) \end{pmatrix}, \quad (10)$$

$$45 \quad \mathbf{H}(k) = \begin{pmatrix} h_{1,1}(k) & h_{1,2}(k) & \dots & h_{1,N_L}(k) \\ h_{2,1}(k) & h_{2,2}(k) & \dots & h_{2,N_L}(k) \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_M,1}(k) & h_{N_M,2}(k) & \dots & h_{N_M,N_L}(k) \end{pmatrix}. \quad (11)$$

Para cada señal de fuente existen zonas de sonido en las cuales la señal se debe reproducir, las llamadas "zonas brillosas". Al mismo tiempo, existen zonas donde no se debe reproducir la zona individual, las "zonas oscuras".

Por ejemplo, en la figura 3, la fuente de señales 211 se debe reproducir en la zona de sonido 221, pero no en la zona de sonido 222. Además, en la figura 3, la fuente de señales 212 se debe reproducir en la zona de sonido 222, pero no en la zona de sonido 221.

- 5 Para reproducción multizona, los pre-filtros se diseñan covalentemente de tal forma que la relación entre la energía acústica irradiada a las zonas brillosas y la energía acústica radiada a las zonas oscuras se aumenta al máximo. Esta relación a menudo es llamada contraste acústico (ver [3]) y se puede medir al definir  $\mathbf{B}_q(k)$  y  $\mathbf{D}_q(k)$ , que capturan las respuestas de impulso de habitación de cada altavoz, a los puntos de muestreo considerados en las zonas brillosas y oscuras, respectivamente. Ya que esta asignación es diferente para cada señal de fuente, ambas matrices son dependientes del índice de señal de fuente  $q$ . Además, la matriz  $\mathbf{G}(k)$  puede descomponerse en

$$\mathbf{G}(k) = (\mathbf{g}_1(k), \mathbf{g}_2(k), \dots, \mathbf{g}_{N_s}(k)), \quad (12)$$

en donde

15

$$\mathbf{g}_q(k) = (g_{1,q}(k), g_{2,q}(k), \dots, g_{N_{l,q}}(k))^T, \quad (13)$$

captura los coeficientes de filtro individuales  $g_{l,q}(k)$  que están asociados con el altavoz  $l$  y la fuente  $q$ . Eventualmente, el contraste acústico logrado para la fuente  $q$  se puede definir de acuerdo con

20

$$C_q = \frac{(\mathbf{g}_q^T(-k) * \mathbf{B}_q^T(-k)) * (\mathbf{B}_q(k) * \mathbf{g}_q(k))}{(\mathbf{g}_q^T(-k) * \mathbf{D}_q^T(-k)) * (\mathbf{D}_q(k) * \mathbf{g}_q(k))} \Bigg|_{k=0}. \quad (14)$$

Un ejemplo de los niveles de reproducción en zona brillante y oscura con contraste acústico resultante se muestra en la figura 5. En particular, la figura 5 ilustra en (a) un nivel de reproducción de ejemplo en zona brillante y oscura, e ilustra en (b) un contraste acústico resultante.

25

Se debe señalar que si cualquier respuesta de impulso en  $\mathbf{H}(k)$  se asigna ya sea a la zona oscura o a la zona brillante para una fuente, se mantiene lo siguiente:

30

$$\mathbf{H}(k) = \mathbf{B}_q(k) + \mathbf{D}_q(k) \quad \forall q, k. \quad (15)$$

Hay muchos métodos conocidos para determinar  $\mathbf{G}(k)$  de tal manera que  $C_q$  alcanza un alto valor (ver [1], [3], [4], [5] y [6]).

35

Existen dificultades, cuando se lleva a cabo reproducción de sonido direccional.

Algunos de los enfoques mencionados anteriormente intentan lograr reproducción multizona por radiación de sonido direccional. Este enfoque enfrenta la mayoría de los desafíos físicos, que se describen más adelante.

40

Cuando se emite una onda a través de una apertura de tamaño finito, la relación de tamaño de apertura a la longitud de onda determina qué tan bien se puede controlar la dirección de radiación. Se logra un mejor control para longitud de onda más pequeña y tamaños más grandes de apertura. Para la resolución angular de un telescopio éste se describe por la aproximación.

45

$$\Theta \approx 1.22 \frac{\lambda}{D}, \quad (16)$$

donde  $\Theta$  es el ángulo mínimo entre dos puntos que se pueden distinguir,  $\lambda$  es la longitud de onda y  $D$  el diámetro del telescopio, ver:

50

[https://en.wikipedia.org/wiki/Angular\\_resolution](https://en.wikipedia.org/wiki/Angular_resolution) (ver, [63]).

Ya que las ondas acústicas obedecen la misma ecuación de onda, esta regla también es aplicable a ondas acústicas. Eventualmente, las razones técnicas limitan el tamaño de las membranas de altavoz o aperturas de cuerno, que implica un menor límite para las frecuencias para las cuales es efectivamente posible reproducción direccional. Además, lo mismo se mantiene también para arreglos de altavoces, donde el tamaño de los altavoces individuales no es de

55

relevancia, sino las dimensiones del arreglo completo de altavoces. A diferencia de los controles de altavoces individuales, las dimensiones de arreglo se limitan principalmente por razones económicas no por razones técnicas.

5 Cuando se utilizan arreglos de altavoces para reproducción de sonido direccional, la distancia de inter-altavoz mínima implica un límite de frecuencia superior. Esto es debido al teorema de muestreo, ver:

[https://en.wikipedia.org/wiki/Nyquist-Shannon\\_sampling\\_theorem](https://en.wikipedia.org/wiki/Nyquist-Shannon_sampling_theorem) (ver,[64]),

10 También es relevante en el dominio espacial donde se requieren dos puntos de muestreo por longitud de onda a fin de lograr una radiación direccional controlada. La colocación de altavoces suficientemente cerca para controlar la radiación direccional con intervalo de frecuencia audible covalentemente no es un problema. Sin embargo, el tamaño de apertura mínimo resultante (ver, más adelante), y una distancia de inter-altavoz mínima implica un número mínimo de altavoces que depende cuadráticamente del intervalo de frecuencia en el cual se debe controlar la dirección de radiación. Ya que los gastos de un arreglo de altavoces son proporcionales al número de altavoces, hay límites de frecuencia efectivos para soluciones de reproducción de arreglo de altavoces comercialmente viables.

15 Además, el recinto donde se deben crear las múltiples zonas de sonido puede incluir en el propio patrón de radiación logrado. Para mayores frecuencias, recintos más grandes, y paredes rectas, se pueden encontrar modelos para considerar analíticamente la geometría de recinto en el diseño de altavoces direccionales o pre-filtros para reproducción de arreglo de altavoces. Sin embargo, esto ya no es posible cuando el recinto exhibe una curvatura (general), cuando se colocan obstáculos de forma arbitraria en el recinto, o cuando las dimensiones del recinto están en el orden de magnitud de la longitud de onda. Esta configuración existe, por ejemplo, en una cabina de automóvil y será llamada como una configuración compleja en lo siguiente. Bajo estas condiciones, la excitación de un campo de sonido controlado por altavoces direccionales o arreglos dirigidos de forma eléctrica es muy desafiante debido al sonido reflejado del recinto que no se puede modelar de forma exacta. Bajo estas condiciones, incluso altavoces controlados de forma individual, no direccionales pueden exhibir de forma efectiva un patrón direccional no controlado.

Algunos de los documentos de la técnica anterior se refieren a control de ganancia dependiente de señal (cruzada).

30 El documento US 2005/0152562 A1 (ver, [8]) se refiere a reproducción de sonido envolvente en automóvil con diferentes modos de operación relacionados con diferentes patrones de volumen en los asientos individuales y diferentes patrones de ecualización.

35 El documento US 2013/170668 A1 (ver, [9]) describe la mezcla de un sonido de anuncio a una señal de entretenimiento. La mezcla entre ambas señales es individual para cada una de las dos zonas.

40 El documento US 2008/0071400 A1 (ver, [10]) describe procesamiento de señales dependiente de la fuente o información de contenido que considera dos señales diferentes para impedir que el controlador se “sobrecargue acústicamente”.

El documento US 2006/0034470 A1 (ver, [11]) se refiere a ecualización, compresión, y ecualización de “imagen de espejo” para reproducir audio en condiciones de alto ruido con calidad incrementada.

45 El documento US 2011/0222695 A1 (ver, [12]) describe compresión de audio de pistas de audio posteriormente reproducidas, también considerando el ruido ambiente y modelos psicoacústicos.

El documento US 2009/0232320 A1 (ver, [13]) describe compresión para hacer que un sonido de anuncio sea más fuerte que un programa de entretenimiento, con interacción de usuario.

50 El documento US 2015/0256933 A1 (ver, [14]) describe un nivel de balance de contenido de teléfono y entretenimiento para reducir al mínimo fuga acústica de contenido.

El documento US 6,674,865 B1 (ver, [15]) se refiere a control de ganancia automático, para telefonía de manos libres.

55 El documento DE 30 45 722 A1 (ver, [16]) describe compresión paralela a nivel de ruido e incremento de nivel para anuncio.

Otros documentos de la técnica anterior se refieren a reproducción multizona.

60 El documento US 2012/0140945 A1 (ver, [17]) se refiere a la implementación explícita de zonas de sonido. Las altas frecuencias son reproducidas por un altavoz, las bajas frecuencias utilizan interferencias constructivas y destructivas mediante la manipulación de la fase de amplitud y el retardo. Para determinar cómo se tiene que manipular la amplitud, fase y retardo, [17] propone utilizar técnicas especiales, el método “Tan Teta” o resolviendo un problema de valores propios.

- El documento US 2008/0273713 A1 (ver, [18]) describe zonas de sonido, arreglo de altavoces ubicado cerca de cada asiento, en donde un arreglo de altavoces se asigna explícitamente a cada una de las zonas.
- 5 El documento US 2004/0105550 A1 (ver, [19]) se refiere a zonas de sonido, cerca de la cabeza direccional, no direccional lejos del oyente.
- El documento US 2006/0262935 A1 (ver, [20]) se refiere a zonas de sonido personales explícitamente.
- 10 El documento US 2005/0190935 A1 (ver, [21]) se refiere a altavoces de reposa cabezas o respaldo para reproducción personalizada.
- El documento US 2008/0130922 A1 (ver, [22]) describe implementación de zonas de sonido con altavoces direccionales cerca del asiento frontal, altavoces no direccionales cerca del asiento posterior y procesamiento de señales de tal forma que la parte frontal y la parte posterior cancelan la fuga una de la otra.
- 15 El documento US 2010/0329488 A1 (ver, [23]) describe zonas de sonido en un vehículo con al menos un altavoz y un micrófono asociado con cada zona.
- 20 El documento DE 10 2014 210 105 A1 (ver, [24]) se refiere a zonas de sonido llevadas a cabo con reproducción binaural, utilizando también cancelación de diafonía (entre oídos), y también a una reducción de diafonía en zonas.
- El documento US 2011/0286614 A1 (ver, [25]) describe zonas de sonido con reproducción binaural con base en cancelación de diafonía y rastreo de cabeza.
- 25 El documento US 2007/0053532 A1 (ver, [26]) describe altavoces de reposa cabezas.
- El documento US 2013/0230175 A1 (ver, [27]) se refiere a zonas de sonido, explícitamente que utilizan micrófonos.
- 30 El documento WO 2016/008621 A1 (ver, [28]) describe un simulador de cabeza y torso.
- Documentos de la técnica anterior adicionales se refieren a reproducción direccional.
- El documento US 2008/0273712 A1 (ver, [29]) describe un altavoz direccional montado en un asiento de vehículo.
- 35 El documento US 5,870,484 (ver, [30]) describe reproducción estereofónica con altavoces direccionales.
- El documento US 5,809,153 (ver [31]) se refiere a tres altavoces que apuntan en tres direcciones con circuitos para usarlos como arreglos.
- 40 El documento US 2006/0034467 A1 (ver, [32]) describe zonas de sonido que se refieren a la excitación del techo interior por transductores especiales.
- El documento US 2003/0103636 A1 (ver, [33]) se refiere a reproducción personalizada y silenciado y a arreglos de reposa cabezas para reproducir el campo de sonido en los oídos de oyentes que incluye silenciamiento.
- 45 El documento US 2003/0142842 A1 (ver, [34]) se refiere a altavoces de reposa cabezas.
- El documento JP 5345549 (ver, [35]) describe altavoces paramétricos en asientos frontales que apuntan hacia atrás.
- 50 El documento US2014/0056431 A1 (ver, [36]) se refiere a reproducción direccional.
- El documento US 2014/0064526 A1 (ver, [37]) se refiere a la producción de una señal de audio binaural y localizada a un usuario.
- 55 El documento US 2005/0069148 A1 (véase [38]) describe el uso de altavoces en el techo con un retraso correspondiente.
- 60 Los documentos US 5,081,682 (ver, [39]), DE 90 15 454 (ver, [40]), US 5,550,922 (ver, [41]), US 5,434,922 (ver, [42]), US 6,078,670 (ver, [43]), US 6,674,865 B1 (ver, [44]), DE 100 52 104 A1 (ver, [45]) y US 2005/0135635 A1 (ver, [46]) se refieren a adaptación de ganancia o modificación espectral de señales de acuerdo con ruido ambiente medido o ruido ambiente estimado, por ejemplo, de velocidad.
- El documento DE102 42 558 A1 (ver, [47]) describe un control de volumen anti-paralelo.

El documento US 2010/0046765 A1 (ver, [48]) y DE 10 2010 040 689 (ver, [49]) se refieren a desvanecimiento cruzado optimizado entre escenas acústicas reproducidas de forma posterior.

- 5 El documento US 2008/0103615 A1 (ver, [50]) describe una variación de paneo dependiente de un evento.

El documento US 8,190,438 B1 (ver, [51]) describe un ajuste de reproducción espacial dependiendo de una señal en un flujo de audio.

- 10 El documento WO 2007/098916 A1 (ver, [52]) describe reproducción de un sonido de advertencia.

El documento US 2007/0274546 A1 (ver, [53]) determina qué pieza de música se puede reproducir en combinación con otra.

- 15 El documento US 2007/0286426 A1 (ver, [54]) describe la mezcla de una señal de audio (por ejemplo, de un teléfono) a otro (por ejemplo, música).

Algunos de los documentos de la técnica anterior describen compresión de audio y control de ganancia.

- 20 El documento US 5,018,205 (ver, [55]) se refiere a ajuste selectivo de banda de ganancia en presencia de ruido ambiente.

El documento US 4,944,018 (ver, [56]) describe amplificación controlada por velocidad.

- 25 El documento DE 103 51 145 A1 (ver, [57]) se refiere a amplificación dependiente de la frecuencia para superar un umbral dependiente de la frecuencia.

Algunos documentos de la técnica anterior se refieren a cancelación de ruido.

- 30 El documento JP 2003–255954 (ver, [58]) describe cancelación activa de ruido que utiliza altavoces ubicados cerca de los oyentes.

El documento US 4,977,600 (ver, [59]) describe atenuación de ruido captado para asiento individual.

- 35 El documento US 5,416,846 (ver, [60]) describe cancelación activa de ruido con un filtro adaptativo.

Documentos adicionales de la técnica anterior se refieren a formación de haces de arreglo para audio.

- 40 Los documentos US 2007/0030976 A1 (ver, [61]) y JP 2004–363696 (ver, [62]) describen formación de haces de arreglo para la reproducción de audio, retardo y formador de haces de suma.

Sería altamente deseable si se proporcionaran conceptos mejorados que proporcionen reproducción multizona dentro de un intervalo suficiente del espectro de frecuencia audible.

- 45 El documento US 2010 226499 A1 divulga un dispositivo para procesar datos, que comprende una unidad de detección adaptada para detectar modos de reproducción individuales indicativos de una manera de reproducir los datos por separado para cada uno de una pluralidad de usuarios humanos, y una unidad de procesamiento adaptada para procesar los datos a fin de generar datos reproducibles por separado para cada uno de la pluralidad de usuarios humanos de acuerdo con los modos de reproducción individuales detectados.

- 50 El objeto de la presente invención es proporcionar conceptos mejorados para el procesamiento de señales de audio. El objeto de la presente invención se resuelve el objeto de las reivindicaciones independientes. Se proporcionan realizaciones particulares en las reivindicaciones dependientes. El alcance de la protección estará definido por las reivindicaciones independientes.

- 55 Se proporciona un aparato para generar una pluralidad de señales de altavoz a partir de dos o más señales de fuente de audio, de acuerdo con la reivindicación 1. Cada una de las dos o más señales de fuente de audio se deben reproducir en una o más de dos o más zonas de sonido, y al menos una de las dos señales de fuente de audio no se deben reproducir en al menos una de las dos o más zonas de sonido. El aparato comprende un procesador de audio configurado para modificar cada una de las dos o más señales de audio iniciales para obtener dos o más señales de audio preprocesadas. Además, el aparato comprende un filtro configurado para generar la pluralidad de señales de altavoz dependiendo de las dos o más señales de audio preprocesadas. El preprocesador de audio se configura para utilizar las dos o más señales de fuente de audio como las dos o más señales de audio iniciales, o en donde el preprocesador de audio se configura para generar para cada señal de fuente de audio de las dos o más señales de
- 60

fuerza de audio una señal de audio inicial de las dos o más señales de audio iniciales al modificar la señal de fuente de audio. Además, el preprocesador de audio se configura para modificar cada señal de audio inicial de las dos o más señales de audio iniciales dependiendo de una potencia de señal o un volumen de otra señal de audio inicial de las dos o más señales de audio iniciales. El filtro se configura para generar la pluralidad de señales de altavoz dependiendo de en cuáles de las dos o más zonas de audio se deben reproducir las dos o más señales de fuente de audio y dependiendo de en cuáles de las dos o más zonas de sonido no se deben reproducir las dos o más señales de fuente de audio.

Además, un método para generar una pluralidad de señales de altavoz a partir de dos o más señales de fuente de audio, de acuerdo con la reivindicación 12. Cada una de las dos o más señales de fuente de audio se deben reproducir en una o más de dos o más zonas de sonido, y al menos una de las dos señales de fuente de audio no se deben reproducir en al menos una de las dos o más zonas de sonido. El método comprende:

- Modificar cada una de las dos o más señales de audio iniciales para obtener dos o más señales de audio preprocesadas. Y:

- Generar la pluralidad de señales de altavoz dependiendo de las dos o más señales de audio preprocesadas.

Las dos o más señales de fuente de audio se utilizan como las dos o más señales de audio iniciales, o en donde para cada señal de fuente de audio de las dos o más señales de fuente de audio se genera una señal de audio inicial de las dos o más señales de audio iniciales al modificar la señal de fuente de audio. Cada señal de audio inicial de las dos o más señales de audio iniciales se modifica dependiendo de una potencia de señal o un volumen de otra señal de audio inicial de las dos o más señales de audio iniciales. La pluralidad de señales de altavoz se genera dependiendo de en cuáles de las dos o más zonas de audio se deben reproducir las dos o más señales de fuente de audio y dependiendo de en cuáles de las dos o más zonas de sonido las dos o más señales de fuente de audio no se deben reproducir.

Además, se proporciona un programa informático, de acuerdo con la reivindicación 13.

Algunas realizaciones proporcionan una modificación de nivel dependiente de la señal para reducir la fuga acústica percibida cuando se utilizan medidas para reproducción direccional de señales de entretenimiento independientes.

En ejemplos, de forma opcional, se emplea una combinación de conceptos de reproducción de diferencia para diferentes bandas de frecuencia.

Opcionalmente, algunas realizaciones utilizan filtros FIR optimizados por mínimos cuadrados (FIR = respuesta finita de impulso) con base en respuestas de impulso una vez medidas. Los detalles de algunos ejemplos se describen más adelante, cuando se describe un prefiltro.

Algunos de los ejemplos se emplean opcionalmente en un escenario automotriz, pero no se limitan a este escenario.

Algunos ejemplos se refieren a conceptos que proporcionan contenido de audio individual a oyentes que ocupan el mismo recinto sin el uso de auriculares o similares. Entre otras, estos ejemplos difieren del estado de la técnica por una combinación inteligente de diferentes enfoques de reproducción con un preprocesamiento dependiente de la señal de tal forma que se logra un gran contraste acústico perceptual en tanto que se retiene un alto nivel de calidad de audio.

Algunos de los ejemplos proporcionan un diseño de filtro.

Algunos de los ejemplos emplean procesamiento dependiente de la señal adicional.

En lo siguiente, se describen ejemplos en más detalle con referencia a las figuras, en las cuales:

la figura 1 ilustra un aparato para generar una pluralidad de señales de altavoz de dos o más señales de fuente de audio,

la figura 2 ilustra reproducción ideal multizona,

la figura 3 ilustra una reproducción de múltiples señales en la realidad,

la figura 4 ilustra un ejemplo mínimo de reproducción multizona con arreglos,

la figura 5 ilustra en (a) un nivel de reproducción de ejemplo en zona brillante y oscura, e ilustra en (b) un contraste acústico resultante,

la figura 6 ilustra un modelo de señal general de reproducción multizona con arreglos,

la figura 7 ilustra reproducción multizona con arreglos de acuerdo con una realización,

5 la figura 8 ilustra una implementación de muestra de un preprocesador de audio,

la figura 9 ilustra un diseño de ejemplo de los divisores de banda , en donde (a) ilustra contraste acústico logrado por diferentes métodos de reproducción, y (b) ilustra una repuesta de magnitud elegida del separador de frecuencias de audio,

10

la figura 10 ilustra un diseño de ejemplo de los conformadores espectrales de acuerdo con realizaciones, en donde (a) ilustra contraste acústico logrado por un método de reproducción específico, y en donde (b) ilustra una respuesta de magnitud elegida del filtro de conformación espectral, y

15

la figura 11 ilustra una configuración de altavoz de ejemplo en un recinto.

La figura 1 ilustra un aparato para generar una pluralidad de señales de altavoz de dos o más señales de fuente de audio. Cada una de las dos o más señales de fuente de audio se deben reproducir en una o más de dos o más zonas de sonido, y al menos una de las dos señales de fuente de audio no se deben reproducir en al menos una de las dos o más zonas de sonido.

20

El aparato comprende un preprocesador de audio 110 configurado para modificar cada una de dos o más señales de audio iniciales para obtener dos o más señales de audio preprocesadas. Además, el aparato comprende un filtro 140 configurado para generar la pluralidad de señales de altavoz dependiendo de las dos o más señales de audio preprocesadas. El preprocesador de audio 110 está configurado para usar las dos o más señales de fuente de audio como las dos o más señales de audio iniciales, o en donde el preprocesador de audio 110 está configurado para generar para cada señal de fuente de audio de las dos o más señales de fuente de audio una señal de audio inicial de las dos o más señales de audio iniciales modificando dicha señal de fuente de audio. Además, el preprocesador de audio 110 se configura para modificar cada señal de audio inicial de las dos o más señales de audio iniciales dependiendo de una potencia de señal o un sonido de otra señal de audio inicial de las dos o más señales de audio iniciales.

25

30

El filtro 140 se configura para generar la pluralidad de señales de altavoz dependiendo de en cuales de las dos o más zonas de sonido se deben reproducir las dos o más señales de fuente de audio y dependiendo en cuáles de las dos o más zonas de sonido no se deben reproducir las dos o más señales de fuente de audio.

35

En tanto que los enfoques del estado de la técnica pueden lograr un contraste acústico considerable, el contraste logrado por métodos de la técnica anterior convencionalmente no es suficiente para proporcionar múltiples escenas acústicas no relacionadas a habitantes del mismo recinto, cuando se requiere reproducción de audio de alta calidad.

40

El contraste acústico percibido por los oyentes se debe mejorar, que es dependiente del contraste acústico como se define en la Ecuación (14) anterior, pero no idéntico al mismo. Se debe lograr que el contraste acústico percibido por los oyentes se incremente en lugar de aumentar al máximo el contraste de energía acústica. El contraste acústico percibido será llamado como contraste acústico subjetivo, en tanto que el contraste en energía acústica será llamado como contraste acústico objetivo en lo siguiente. Algunos ejemplos emplean medidas para facilitar reproducción de audio direccional y medidas para conformar la fuga acústica de tal forma que se vuelve menos perceptible.

45

Además de la figura 1, el aparato de la figura 7 comprende además dos divisores de banda (opcionales) 121, 122 y cuatro conformadores espectrales (opcionales) 131, 132, 133, 134.

50

De acuerdo con algunos ejemplos el aparato, por ejemplo puede comprender además dos o más divisores de banda 121, 122 que se configuran para llevar a cabo división de banda en las dos o más señales de audio preprocesadas a una pluralidad de señales de audio de banda dividida. El filtro 140, por ejemplo, se puede configurar para generar la pluralidad de señales de altavoz dependiendo de la pluralidad de señales de audio de banda dividida.

55

En algunos ejemplos, el aparato, por ejemplo, puede comprender además uno o más conformadores espectrales 131, 132, 133, 134 que se configuran para modificar un envolvente espectral de una o más de la pluralidad de señales de audio de banda dividida para obtener una o más señales de audio espectralmente conformadas. El filtro 140 por ejemplo, se puede configurar para generar la pluralidad de señales de altavoz dependiendo de la una o más señales de audio espectralmente conformadas.

60

En la figura 7 se muestra un modelo de señal de una implementación. En particular, la figura 7 ilustra reproducción multizona con arreglos. Este ejemplo se ha elegido por brevedad, se señala que el método en general es aplicable a escenarios con  $N_s$  fuentes de señal,  $N_L$  altavoces, y  $N_z$  zonas de escucha, como se describe anteriormente.

Hay dos fuentes de señal mostradas en la figura 7, que proporcionan dos señales independientes que se alimentan a una etapa de "preprocesamiento". Esta etapa de preprocesamiento, por ejemplo, puede en algunos ejemplos, implementar un procesamiento paralelo para ambas señales (es decir, sin mezcla). A diferencia de los otros pasos de procesamiento, este paso de procesamiento no constituye un sistema LTI (sistema lineal invariante en el tiempo). En su lugar, este bloque de procesamiento determina ganancias variables en el tiempo para todas las señales de fuente procesadas, de tal forma que se reduce su diferencia en nivel de reproducción. La razón detrás de esto es que la fuga acústica en cada zona siempre es linealmente dependiente de las escenas reproducidas en las otras zonas respectivas. Al mismo tiempo, las escenas intencionalmente reproducidas pueden enmascarar la fuga acústica. Por lo tanto, la fuga acústica percibida es proporcional a la diferencia de nivel entre las escenas que se reproducen de forma intencional en las zonas respectivas. Como consecuencia, la reducción de la diferencia del nivel de las escenas reproducidas también reducirá la fuga acústica percibida y, por lo tanto, incrementa el contraste acústico subjetivo. Se puede encontrar una explicación más detallada cuando se describe más adelante el preprocesamiento.

Los divisores de banda (opcionales) 121, 122 llevan a cabo la división de banda de paso de procesamiento (opcional), y dividen la señal en múltiples bandas de frecuencia, al igual que lo haría un separador de frecuencias de audio en un altavoz multidireccional. Sin embargo, a diferencia de separadores de frecuencia de audio en un altavoz, es sólo un segundo objetivo de este divisor de banda aumentar al máximo la potencia acústica radiada. El objetivo principal de este divisor de banda es distribuir las bandas de frecuencias individuales a medidas de reproducción individuales de tal forma que se aumenta al máximo el contraste acústico, dadas ciertas limitaciones de calidad. Por ejemplo, la señal  $w_1(k)$  se alimentará posteriormente a un altavoz individual como señal  $x_1(k)$ . Dado que este altavoz es un altavoz direccional,  $w_1(k)$  se filtrará por paso alto porque la directividad de este altavoz será baja a bajas frecuencias. Por otro lado,  $w_2(k)$  se filtrará posteriormente para obtener  $x_2(k)$  y  $x_3(k)$  de modo que los altavoces correspondientes se utilicen como un arreglo dirigido eléctricamente. En un escenario más complejo, puede haber más salidas del divisor de banda de modo que las señales se distribuyan a múltiples métodos de reproducción de acuerdo con las necesidades de la aplicación (ver también a continuación, donde se describe un sistema de altavoz-caja-micrófono de acuerdo con los ejemplos).

Como se discutió anteriormente, las medidas de reproducción direccional aplicadas posteriormente siempre presentarán una cierta fuga de una zona a otra. Esta fuga se puede medir como degradación en contraste acústico entre las zonas. En una configuración compleja, estas degradaciones pueden presentarse en múltiples puntos en el espectro de frecuencia para cada uno de los métodos de reproducción direccional previstos, que constituyen un mayor obstáculo en la aplicación de aquellos métodos. Es bien conocido que las variaciones de timbre son aceptables a una cierta medida. Estos grados de libertad se pueden utilizar para atenuar bandas de frecuencia de contraste crítico.

Por lo tanto, los conformadores espectrales (opcionales) 131, 132, 133, 134 se diseñan de tal forma que las señales reproducidas posteriormente se atenúan en estas partes del espectro de frecuencia, donde se espera un bajo contraste acústico. A diferencia de los divisores de banda, se propone que los conformadores espectrales modifiquen el timbre del sonido reproducido. Además, esta etapa de procesamiento también puede implicar retardos y ganancias de tal forma que la escena acústica intencionalmente reproducida puede enmascarar espacialmente la fuga acústica.

Los bloques denotados por  $\mathbf{G}_1(k)$  y  $\mathbf{G}_2(k)$  pueden, por ejemplo, describir filtros lineales invariantes en el tiempo que están optimizados para maximizar el contraste acústico objetivo dadas las restricciones de calidad subjetivas. Existen diferentes posibilidades para determinar aquellos filtros, que incluyen (pero no se limitan a) ACC, acoplamiento de presión (ver [4] y [6]), y formación de haces de altavoz. Se ha encontrado, que un enfoque de coincidencia de presión de mínimos cuadrados como se describe a continuación, cuando se describe un prefiltro, es especialmente adecuado, cuando se consideran las respuestas de impulso medidas para la optimización del filtro. Éste puede ser un concepto preferido para la implementación.

Otros ejemplos emplean el enfoque anterior al operar en respuestas de impulso calculadas. En ejemplos particulares, las respuestas de impulso se calculan para representar las respuestas de impulso de campo libre de los altavoces a los micrófonos.

Ejemplos adicionales, emplean el enfoque anterior al operar en respuestas de impulso calculadas que se han obtenido utilizando modelo de fuente de imagen del recinto.

Se debe señalar que las respuestas de impulso se miden una vez de tal forma que no es necesario ningún micrófono durante la operación. La diferencia de ACC, el enfoque de acoplamiento de presión prescribe una magnitud y fase dadas en la zona brillante respectiva. Esto da por resultado una alta calidad de reproducción. Los enfoques de formación de haces tradicionales también son adecuados cuando se deben reproducir altas frecuencias.

El bloque denotado por  $\mathbf{H}(k)$  representa LOS LEMS, donde cada entrada está asociada con un altavoz. Cada una de las salidas está asociada con un oyente individual que recibe la superposición de todas las contribuciones de altavoz en su zona de sonido individual. Los altavoces que se controlan sin utilizar los prefiltros  $\mathbf{G}_1(k)$  y  $\mathbf{G}_2(k)$  son altavoces direccionales que irradian primario en una zona de sonido o altavoces que están dispuestos cerca (o en) una zona de

sonido individual de tal manera que excitan principalmente el sonido en esa zona. Para frecuencias más altas, se pueden construir altavoces direccionales sin esfuerzo significativo. Por lo tanto, estos altavoces se pueden utilizar para proporcionar las frecuencias de alto intervalo a los oyentes, donde los altavoces no tienen que colocarse directamente en los oídos de los oyentes.

5

En lo siguiente, se describen ejemplos en más detalle.

Primero, se describe preprocesamiento de acuerdo con los ejemplos. En particular, se presenta una implementación del bloque denotado por "preprocesamiento" en la figura 7. Para proporcionar un mejor entendimiento, las siguientes explicaciones se concentran sólo en una señal monoaural por zona. Sin embargo, es sencilla una generalización a señales multicanal. Por lo tanto, algunos ejemplos exhiben señales multicanal por zona.

10

La figura 8 ilustra una implementación de muestra de un preprocesador de audio 110 y un modelo de señal correspondiente. Como se describe anteriormente, las dos señales de entrada  $u_1(k)$  y  $u_2(k)$  están destinadas a reproducirse principalmente en la zona 1 y la zona 2, respectivamente. Por otra parte, hay alguna fuga acústica en la reproducción de  $u_1(k)$  a la zona 2 y en la reproducción de  $u_2(k)$  a la zona 1.

15

Las dos señales de entrada  $u_1(k)$  y  $u_2(k)$  también se denominan señales de fuente de audio en lo siguiente.

20

En una primera etapa opcional, la potencia de ambas señales de entrada,  $u_1(k)$  y  $u_2(k)$  (las señales de la fuente de audio) se normaliza para aliviar la elección de parámetros para el siguiente procesamiento.

Por lo tanto, de acuerdo con un ejemplo, el preprocesador de audio (110) puede, por ejemplo, configurarse para generar las dos señales de audio iniciales más  $d_1(k)$  y  $d_2(k)$  normalizando una potencia de cada una de las dos o más señales de fuente de audio  $u_1(k)$  y  $u_2(k)$ .

25

Las estimaciones de potencia obtenidas  $b_1(k)$  y  $b_2(k)$  convencionalmente describen un promedio de largo plazo, en contraste con los estimadores utilizados en una etapa posterior que consideran convencionalmente un lapso de tiempo más pequeño. La actualización de  $b_1(k)$  y  $b_2(k)$  se puede conectar con una detección de actividad para  $u_1(k)$  y  $u_2(k)$ , respectivamente, de modo que se mantenga la actualización de  $b_1(k)$  o  $b_2(k)$ , cuando no hay actividad en  $u_1(k)$  o  $u_2(k)$ . Las señales  $c_1(k)$  y  $c_2(k)$  pueden, por ejemplo, ser inversamente proporcionales a  $b_1(k)$  y  $b_2(k)$ , respectivamente, de modo que una multiplicación de  $c_1(k)$  y  $c_2(k)$  con  $u_1(k)$  y  $u_2(k)$ , respectivamente, produce las señales,  $d_1(k)$  y  $d_2(k)$  que exhibirían una potencia de señal comparable. En tanto que el uso de esta primera etapa no es absolutamente necesario, asegura un punto de trabajo razonable para el procesamiento relativo de las señales  $d_1(k)$  y  $d_2(k)$ , que alivia la búsqueda de parámetros adecuados para los siguientes pasos. Se debe señalar que si se colocan múltiples casos de este bloque de procesamiento después de los bloques de "divisor de banda" o los bloques de "conformador espectral", aún se tiene que aplicar normalización de potencia antes de los bloques de "divisor de banda".

30

35

Por una normalización de las señales, ya está reducida su diferencia de nivel relativa. Sin embargo, esto convencionalmente no es suficiente para el efecto propuesto, debido a que las estimaciones de potencia son a largo plazo, en tanto que las variaciones de nivel de escenas acústicas convencionales son en su lugar procesos de corto plazo. En lo siguiente, se explica cómo se reduce explícitamente la diferencia en potencia relativa de las señales individuales de una forma a corto plazo, que constituye el objetivo principal del bloque de procesamiento.

40

Las dos señales  $d_1(k)$  y  $d_2(k)$  que se supone que se modifican en escala y se reproducen, también son llamadas como señales de audio iniciales en lo siguiente.

45

Como se describe anteriormente, el preprocesador de audio 110, por ejemplo, se puede configurar para generar para cada señal de fuente de audio de las dos o más señales de fuente de audio  $u_1(k)$ ,  $u_2(k)$  una señal de audio inicial de las dos o más señales de audio iniciales  $d_1(k)$ ,  $d_2(k)$  al modificar la señal de fuente de audio, por ejemplo, al llevar a cabo normalización de potencia.

50

En ejemplos alternativos, sin embargo, el preprocesador de audio 110 puede, por ejemplo, configurarse para utilizar las dos o más señales de fuente de audio  $u_1(k)$ ,  $u_2(k)$  como las dos o más señales de audio iniciales  $d_1(k)$ ,  $d_2(k)$ .

55

En la figura 7, las dos señales  $d_1(k)$  y  $d_2(k)$  pueden, por ejemplo, alimentarse a estimadores de sonoridad adicionales, por ejemplo, del preprocesador de audio 110, que proporcionan las señales  $e_1(k)$  y  $e_2(k)$ , respectivamente.

Estas señales, por ejemplo se pueden utilizar para determinar los factores de modificación de escala  $g'_1(k)$  and  $g'_2(k)$  de acuerdo con

60

$$g'_1 = f(e_1, e_2), \quad (17)$$

$$g_2 = f(e_2, e_1), \quad (18)$$

donde, en algunos ejemplos,  $f(x, y)$  es una función que se incrementa de forma monótona con respecto a  $y$  y disminuye de forma monótona con respecto a  $x$ , en tanto que su valor, por ejemplo, se puede limitar a un intervalo absoluto.

5 Como consecuencia, el valor de  $f(x, y)$  también puede, por ejemplo, aumentar monótonamente con la relación  $y/x$ .

Los factores  $g_1(k)$  y  $g_2(k)$  se utilizan para escalar las señales  $d_1(k)$  y  $d_2(k)$ , respectivamente, para obtener las señales de salida  $h_1(k)$  y  $h_2(k)$ . Las señales de salida  $h_1(k)$  y  $h_2(k)$  pueden, por ejemplo, alimentarse a uno o más módulos que están configurados para llevar a cabo la reproducción multizona, por ejemplo, de acuerdo con un método de reproducción multizona arbitrario.

Por lo tanto, en algunos ejemplos, el preprocesador de audio 110, por ejemplo, se puede configurar para modificar cada señal de audio inicial de las dos o más señales de audio iniciales dependiendo de la potencia de señal o el volumen de otra señal de audio inicial de las dos o más señales de audio iniciales al modificar la señal de audio inicial de las dos o más señales de audio iniciales dependiendo de una relación de un primer valor ( $y$ ) a un segundo valor ( $x$ ). El segundo valor ( $x$ ) puede, por ejemplo, depender de la potencia de señal de dicha señal de audio inicial, y el primer valor ( $y$ ) puede, por ejemplo, depender de la potencia de señal de dicha otra señal de audio inicial de las dos o más señales de audio iniciales. O, el segundo valor ( $x$ ), por ejemplo, puede depender del volumen de la señal de audio inicial, y el primer valor ( $y$ ), por ejemplo puede depender del volumen de la otra señal de audio inicial de las dos o más señales de audio iniciales.

De acuerdo con algunos ejemplos, el preprocesador de audio 110, por ejemplo, se puede configurar para modificar cada señal de audio inicial de las dos o más señales de audio iniciales dependiendo de la potencia de señal o el volumen de otra señal de audio inicial de las dos o más señales de audio iniciales al determinar una ganancia para la señal de audio inicial y al aplicar la ganancia en la señal de audio inicial. Además, el preprocesador de audio 110, por ejemplo, se puede configurar para determinar la ganancia dependiendo de la relación entre el primer valor y el segundo valor, la relación que es una relación entre la potencia de señal de la otra señal de audio inicial de las dos o más señales de audio iniciales y la potencia de señal de la señal de audio inicial como el segundo valor, o la relación que es una relación entre el volumen de la otra señal de audio inicial de las dos o más señales de audio iniciales y el volumen de la señal de audio inicial como el segundo valor.

En algunos ejemplos, el preprocesador de audio 110, por ejemplo, se puede configurar para determinar la ganancia dependiendo de una función que incrementa de forma monótona con la relación entre el primer valor y el segundo valor.

De acuerdo con algunos ejemplos, por ejemplo, ninguna de las señales  $u_1(k)$ ,  $d_1(k)$  o  $h_1(k)$  se mezcla con ninguna de las señales  $u_2(k)$ ,  $d_2(k)$  o  $h_2(k)$ .

40 En lo siguiente, se describe en más detalle la implementación del paso de procesamiento. Dado que los pasos de procesamiento para  $u_1(k)$  y  $u_2(k)$  son idénticos, solo se describirán los pasos de procesamiento para  $u_1(k)$ , que también se aplican a  $u_2(k)$  intercambiando los índices 1 y 2.

Una regla para obtener  $b_1(k)$  puede, por ejemplo, ser dada por

$$b_1(k) = \lambda_1 b_1(k-1) + (1-\lambda_1) \sum_{l=1}^L u_1^2(k, l), \quad (19)$$

donde  $\lambda_1$ , puede, por ejemplo, elegirse cerca de pero menor que 1.

50 En la fórmula anterior se asume que  $u_1(k, l)$  comprende uno o más canales de audio.  $L$  indica el número de canales de audio de  $u_1(k)$ .

En un caso simple,  $u_1(k)$  comprende solo un único canal y la fórmula (19) se convierte en:

$$55 \quad b_1(k) = \lambda_1 b_1(k-1) + (1-\lambda_1) u_1^2(k, 1) \quad (19a)$$

$\lambda_1$  puede estar en el intervalo  $0 < \lambda_1 < 1$ . De manera preferente,  $\lambda_1$  puede, por ejemplo, estar cerca de 1. Por ejemplo,  $\lambda_1$  puede, por ejemplo, estar en el intervalo  $0,9 < \lambda_1 < 1$ .

60 En otros casos,  $u_1(k)$ , por ejemplo, comprende dos o más canales.

El factor de escala  $c_1(k)$  se puede determinar de acuerdo con

$$c_1(k) = \frac{1}{\sqrt{b_1(k)}}, \quad (20)$$

de tal forma que

$$d_1(k, l) = c_1(k)u_1(k, l) \quad (21)$$

describe la señal de audio escalada.

Una regla para obtener  $e_1(k)$  puede, por ejemplo, ser dada por

$$e_1(k) = \lambda_2 e_1(k-1) + (1-\lambda_2) \sum_{l=1}^L d_1^2(k, l), \quad (22)$$

$\lambda_2$  puede estar en el intervalo  $0 < \lambda_2 < 1$ .

En ejemplos particulares, para  $\lambda_1$  de la fórmula (19) y  $\lambda_2$  de la fórmula (22):  $\lambda_1 > \lambda_2$ .

En tanto que existe una variedad de otras opciones. Una de éstas, de acuerdo con un ejemplo, es el valor cuadrático medio de  $d_1^2(k, l)$  en una ventana de  $K$  muestras dadas por

$$e_1(k) = \frac{1}{K} \sum_{n=0}^{K-1} \sum_{l=1}^L d_1^2(k-n, l), \quad (23)$$

Otra definición, de acuerdo con otro ejemplo, es el valor cuadrado máximo en dicha ventana

$$e_1(k) = \max_{n=0,1,\dots,K-1, l=1,2,\dots,L} d_1^2(k-n, l). \quad (24)$$

De acuerdo con algunos ejemplos, para determinar  $g'_1(k)$ , el valor  $e_2(k)$  también debe determinarse como se describió anteriormente. Sin embargo, el método real para determinar  $e_2(k)$ , así como los parámetros, pueden diferir de los elegidos para  $e_1(k)$  (por ejemplo, dependiendo de las necesidades de la aplicación). La ganancia real  $g'_1(k)$  puede, por ejemplo, determinarse de manera similar a la regla de ganancia que se usaría para un compresor de audio convencional, ver: [https://en.wikipedia.org/wiki/Dynamic\\_range\\_compression](https://en.wikipedia.org/wiki/Dynamic_range_compression) (ver [65]), pero considerando tanto  $e_1(k)$  y  $e_2(k)$ .

De acuerdo con una realización de la presente invención, una regla de ganancia de un compresor descendente correspondiente para la señal  $d_1(k)$  sería

$$g'_1(k) = \begin{cases} 10^{(\tau_1 - 10 \log_{10}(e_1(k)) + 10 \log_{10}(e_2(k)))(R-1)/(20R)} & \text{para } \tau_1 - 10 \log_{10}(e_1(k)) + 10 \log_{10}(e_2(k)) < 0 \\ 1 & \text{de otra forma,} \end{cases} \quad (25)$$

o

$$g'_1(k) = \begin{cases} 10^{(\tau_1 + v)(R-1)/(20R)} & \text{para } \tau_1 + v < 0 \\ 1 & \text{de otra forma,} \end{cases} \quad (25')$$

40

con  $v = -10\log_{10}(e_1(k)) + 10\log_{10}(e_2(k))$

donde  $T_1$  define el umbral de compresión en dB y  $R$  la relación de compresión, como se usa en un compresor de audio estándar. Por ejemplo,  $1 \leq R \leq 100$ . Por ejemplo,  $1 < R < 100$ . Por ejemplo,  $2 < R < 100$ . Por ejemplo,  $2 < R < 50$ .

5

A diferencia de las fórmulas (25) y (25'), un compresor de audio estándar de acuerdo con el estado de la técnica no consideraría  $e_2(k)$  para determinar una ganancia para  $d_1(k)$ .

10

De acuerdo con otra realización de la presente invención, una implementación de un compresor ascendente se define por

$$g'_1(k) = \begin{cases} 10^{(T_2 - 10\log_{10}(e_1(k)) + 10\log_{10}(e_2(k)))(R-1)/(20R)} & \text{para } T_2 - 10\log_{10}(e_1(k)) + 10\log_{10}(e_2(k)) > 0 \\ 1 & \text{de otra forma,} \end{cases} \quad (25a)$$

o

$$g'_1(k) = \begin{cases} 10^{(T_2 + v)(R-1)/(20R)} & \text{para } T_2 + v > 0 \\ 1 & \text{de otra forma,} \end{cases} \quad (25a')$$

15

con  $v = -10\log_{10}(e_1(k)) + 10\log_{10}(e_2(k))$

20

que es similar excepto por el intervalo operativo (se señala la condición diferente) y parámetros diferentes. Cabe señalar que  $T_2$  define un umbral inferior en contraste con  $T_1$ .

Algunas realizaciones, donde  $T_2 < T_1$ , combinan ambas reglas de ganancia.

25

En realizaciones de la presente invención, la regla resultante para obtener  $g'_1(k)$  y  $g'_2(k)$  puede ser cualquier combinación de compresores ascendentes y descendentes, donde las implementaciones prácticas típicamente requerirán un ajuste limitado a los intervalos considerados de  $e_1(k)$  y  $e_2(k)$ .

30

Cuando se consideran más de dos señales  $e_1(k)$ ,  $e_2(k)$ ,  $e_3(k)$ , ...,  $e_N(k)$ , por ejemplo, N señales, la fórmula (25) puede, por ejemplo, convertirse en:

$$g'_1(k) = \begin{cases} 10^{(T_1 + v_1)(R-1)/(20R)} & \text{para } T_1 + v_1 < 0 \\ 1 & \text{de otra forma,} \end{cases}$$

con

$$v_1 = -10\log_{10}(e_1(k)) + 10\log_{10}\left(\sum_{i=2}^N e_i(k)\right) \quad (25b)$$

35

Para otras ganancias  $g'_2(k)$ ,  $g'_3(k)$ , ...,  $g'_N(k)$ , la fórmula (25) puede, por ejemplo, convertirse en:

$$g'_r(k) = \begin{cases} 10^{(T_1 + v_r)(R-1)/(20R)} & \text{para } T_1 + v_r < 0 \\ 1 & \text{de otra forma,} \end{cases}$$

con

$$v_2 = -10\log_{10}(e_r(k)) + 10\log_{10}\left(-e_r(k) + \sum_{i=1}^N e_i(k)\right) \quad (25c)$$

5

La Fórmula (25a) puede, por ejemplo, convertirse en:

$$g'_1(k) = \begin{cases} 10^{(T_2+v_1)(R-1)/(20R)} & \text{para } T_2 + v_1 > 0 \\ 1 & \text{de otra forma,} \end{cases}$$

con

10

$$v_1 = -10\log_{10}(e_1(k)) + 10\log_{10}\left(\sum_{i=2}^N e_i(k)\right) \quad (25b)$$

Para otras ganancias  $g'_2(k)$ ,  $g'_3(k)$ , ...,  $g'_N(k)$ , la fórmula (25a) puede, por ejemplo, convertirse en:

$$g'_i(R) = \begin{cases} 10^{(T_2+v_i)(R-1)/(20R)} & \text{para } T_2 + v_i > 0 \\ 1 & \text{de otra forma,} \end{cases}$$

15

con

$$v_2 = -10\log_{10}(e_r(k)) + 10\log_{10}\left(-e_r(k) + \sum_{i=1}^N e_i(k)\right) \quad (25c)$$

20

Se pueden definir reglas alternativas adicionales para reducir la diferencia de energía entre ambas escenas según lo dado por

$$g'_1(k) = (1 - \alpha) + \alpha \sqrt{\frac{e_2(k)}{e_1(k)}} \quad (25d)$$

25

donde  $\alpha = 1$  haría que la señal  $h_1(k)$  tuviera la misma energía que la señal  $d_2(k)$ . Por otro lado,  $\alpha = 0$  no tendría ningún efecto, un parámetro elegido  $0 < \alpha < 1$  se puede utilizar para variar la influencia prevista de ese paso.

Otra oportunidad es el uso de una función sigmoide para limitar el exceso de energía de  $h_2(k)$  en comparación con  $d_1(k)$

$$g'_1(k) = \sqrt{\frac{e_2(k)}{e_1(k)}} f\left(\frac{e_1(k)}{e_2(k)}\right) \quad (25e)$$

30

en donde  $f(x)$  puede ser una de

$$f(x) = \frac{x}{\sqrt{1+x^2}}$$

$$f(x) = \frac{x}{1 + |x|},$$

$$f(x) = \tanh(x),$$

$$f(x) = \frac{2}{\pi} \arctan\left(\frac{\pi}{2}x\right),$$

que todas se limitan por  $-1 < f(x) < 1$  en tanto que se mantiene  $f(0) = 1$ .

- 10 En algunos ejemplos, el preprocesador de audio 110, por ejemplo, se puede configurar para modificar una señal de audio inicial de las dos o más señales de audio iniciales dependiendo de la potencia de señal o el volumen de otra señal de audio inicial de dos o más señales de audio iniciales al determinar una ganancia  $g_1'(k)$  para la señal de audio inicial y al aplicar la ganancia  $g_1'(k)$  en la señal de audio inicial, y el preprocesador de audio 110, por ejemplo, se puede configurar para determinar la ganancia  $g_1'(k)$  de acuerdo con una o más de las fórmulas anteriores.

15 A continuación, se describen otras características del preprocesamiento de acuerdo con los ejemplos.

De acuerdo con un ejemplo, la ramificación de las señales  $e_1(k)$  y  $e_2(k)$  que se alimentan al lado respectivamente opuesto, por ejemplo, se puede filtrar a través de un filtro que describe el acoplamiento acústico real de las dos zonas.

20 Además, de acuerdo con un ejemplo, los estimadores de potencia, por ejemplo, pueden operar en señales que se han procesado por un filtro de ponderación, por ejemplo, que se han procesado por un filtro de ponderación descrito en:

[https://en.wikipedia.org/wiki/Weighting\\_filter](https://en.wikipedia.org/wiki/Weighting_filter) (ver [66]).

25 De acuerdo con un ejemplo, los estimadores de potencia, por ejemplo, se pueden reemplazar por estimadores de volumen, como, por ejemplo, se describe por la Recomendación ITU-R BS.1770-4. Esto permitirá una calidad de reproducción mejorada debido a que el volumen percibido se hace coincidir mejor por este modelo.

- 30 Además, de acuerdo con un ejemplo, se puede utilizar un umbral de nivel para excluir que se tome en cuenta silencio para las estimaciones  $b_1(k)$  y  $b_2(k)$  en la normalización de potencia absoluta.

35 Además, en un ejemplo, una derivada en el positiva de la potencia estimada de forma separada se puede utilizar como un indicador para actividad de las señales de entrada  $u_1(k)$  y  $u_2(k)$ . Las estimaciones  $b_1(k)$  y  $b_2(k)$  se actualizan sólo entonces cuando se detecta actividad.

40 En lo siguiente, se describe un divisor de banda de acuerdo con los ejemplos. En particular, se presenta una implementación del bloque denotado por "divisor de banda" mostrado en la figura 7. En un ejemplo, este bloque puede, por ejemplo, realizarse como un separador de audio digital, por ejemplo, como un separador de audio digital como se describe en:

[https://en.wikipedia.org/wiki/Audio\\_crossover#Digital](https://en.wikipedia.org/wiki/Audio_crossover#Digital) (ver, [67]).

- 45 La respuesta de frecuencia deseada de las rutas de entrada a salida, por ejemplo, puede ser un paso de banda con una respuesta plana de frecuencia en la banda de paso y una alta atenuación en la banda de parada. Los bordes de las bandas de paso y bandas de parada se eligen dependiendo del intervalo de frecuencia en el cual las medidas de reproducción conectadas a salidas individuales pueden lograr un contraste acústico suficiente entre las zonas de sonido respectivas.

- 50 La figura 9 ilustra un diseño de ejemplo del uno o más divisores de banda de acuerdo con los ejemplos, en donde (a) ilustra contraste acústico logrado por diferentes métodos de reproducción, y en donde (b) ilustra una respuesta de magnitud elegida del separador de frecuencias de audio. En particular, la figura 9 ilustra un diseño de ejemplo de la respuesta de magnitud de filtro con respecto al contraste acústico logrado.

- 55 Como se puede ver de la figura 9, el conformador espectral, por ejemplo, se puede configura para modificar un envolvente espectral de una señal de audio dependiendo del contraste acústico.

60 Se pueden emplear diferentes conceptos para llevar a cabo la implementación real del uno o más divisores de banda. Por ejemplo, algunos ejemplos emplean filtros FIR, otras realizaciones emplean un filtro IIR, y realizaciones adicionales emplean filtros analógicos. Se puede emplear cualquier concepto posible para llevar a cabo divisores de banda, por

ejemplo, cualquier concepto que se presente en literatura general en ese tema.

5 Algunos de los ejemplos, por ejemplo, pueden comprender un conformador espectral para llevar a cabo conformación espectral. Cuando se lleva a cabo conformación espectral en una señal de audio, el envolvente espectral de la señal de audio, por ejemplo, se puede modificar y se puede obtener, por ejemplo, una señal de audio espectralmente conformada.

10 En lo siguiente, se describe un conformador espectral de acuerdo con los ejemplos, en particular, un conformador espectral como se ilustra en la figura 7. Los conformadores espectrales constituyen filtros que exhiben respuestas de frecuencias similares a aquellas conocidas para ecualizadores, tal como combinaciones de filtro de primer orden o segundo orden, ver:

[https://en.wikipedia.org/wiki/Equalization\\_\(audio\)#\\_Funciones\\_de\\_filtro\\_\(ver\\_\[68\]\)](https://en.wikipedia.org/wiki/Equalization_(audio)#_Funciones_de_filtro_(ver_[68])).

15 Sin embargo, las respuestas de frecuencia eventuales del filtro espectral se diseñan de una forma completamente diferente en comparación con los ecualizadores: Los filtros espectrales consideran la distorsión espectral máxima que se aceptará por el oyente, y los filtros espectrales se diseñan de tal forma que atenúan aquellas frecuencias que se conoce que produce fuga acústica.

20 La razón detrás de esto es que la percepción humana es sensible de forma diferente a distorsiones espectrales de escenas acústicas a ciertas frecuencias, dependiendo de la excitación de las frecuencias circundantes y dependiendo de si la distorsión es una atenuación o una amplificación.

25 Por ejemplo, si se aplica un filtro de ranura con un ancho de banda pequeño a una señal de audio de banda ancha, los oyentes sólo percibirán una pequeña diferencia, si la hay. Sin embargo, si un filtro de pico con el mismo ancho de banda se aplica a la misma señal, los oyentes percibirán más probablemente una diferencia considerable.

30 Los ejemplos se basan en el hallazgo de que este hecho se puede aprovechar debido a que una degradación de banda limitada en contraste acústico da por resultado un pico en fuga acústica (ver, figura 5). Si la escena acústica reproducida en la zona brillante se filtra por un filtro de ranura conforme, será más probable que no se perciba por los oyentes en esta zona. Por otra parte, el pico de fuga acústica que se percibe en la zona oscura se compensará por esta medida.

35 Un ejemplo de la respuesta de filtro correspondiente se muestra en la figura 10. En particular, la Fig. 10 ilustra un diseño a modo de ejemplo de los conformadores espectrales, en donde (a) ilustra el contraste acústico logrado por un método de reproducción específico, y en donde (b) ilustra una respuesta de magnitud elegida del filtro de conformación espectral.

40 Como se señala anteriormente, el filtro 140 se configura para generar la pluralidad de señales de altavoz dependiendo de en cuáles de las dos o más zonas de sonido se deben reproducir las dos o más señales de fuente de audio y dependiendo de en cuáles de las dos o más zonas de sonido no se deben reproducir las dos o más señales de fuente de audio.

45 En lo siguiente, se describe un filtro 140, por ejemplo un prefiltro.

En un ejemplo, se deben reproducir una o más señales de fuente de audio en una primera zona de sonido, pero no en una segunda zona de sonido y al menos una señal de fuente de audio adicional se debe reproducir en la segunda zona de sonido, pero no en la primera zona de sonido.

50 Ver, por ejemplo, figuras 2 y 3, donde se debe reproducir una primera señal de fuente de audio  $u_1(k)$  en la zona de sonido 1, pero no en la zona de sonido 2, y donde se debe reproducir una segunda señal de fuente de audio  $u_2(k)$ , en la zona de sonido 2, pero no en la zona de sonido 1.

55 Como cada una de las dos o más señales de audio preprocesadas  $h_1(k)$ ,  $h_2(k)$  se ha generado con base en una de las dos o más señales de fuente de audio  $u_1(k)$ ,  $u_2(k)$ , resulta que en esta realización, se deben reproducir una o más señales de audio preprocesadas  $h_1(k)$  en la zona de sonido 1, pero no en la zona de sonido 2 (es decir, estas una o más señales de audio preprocesadas  $h_1(k)$  que se han generado al modificar las una o más señales de fuente de audio  $u_1(k)$  que se deben reproducir en la zona de sonido 1, pero no en la zona de sonido 2). Además, resulta que al menos una señal de audio preprocesada adicional  $h_2(k)$  se debe reproducir en la zona de sonido 2, pero no en la zona de sonido 1 (es decir, aquellas de una o más señales de audio preprocesadas  $h_2(k)$  que se han generado al modificar la una o más fuentes de sonido  $u_2(k)$  que se deben reproducir en la zona de sonido 2, pero no en la zona de sonido 1).

Se pueden emplear medios adecuados que logren que se reproduzca una señal de fuente de audio en una primera zona pero no en una segunda zona, o que logren al menos que la señal de fuente de audio se reproduzca en la primera

zona de sonido son un mayor volumen que en la segunda zona de sonido (y/o que logren al menos que la señal de fuente de audio se reproduzca en la primera zona de sonido con una mayor energía de señal que en la segunda zona de sonido).

5 Por ejemplo, un filtro 140 se puede emplear, y los coeficientes de filtro, por ejemplo, se pueden elegir de tal forma que una primera señal de fuente de audio que se debe reproducir en la primera zona de sonido, pero no en la zona de sonido se reproduce en la primera zona de sonido con un mayor volumen (y/o con una mayor energía de señal) que en la segunda zona de sonido. Además los coeficientes de filtro, por ejemplo, se pueden elegir de tal forma que una  
10 segunda señal de fuente de audio que se debe reproducir en la segunda zona de sonido, pero no en la primera zona de sonido, se reproduce en la segunda zona de sonido con un mayor volumen (y/o con una mayor energía de señal) que en la primera zona de sonido.

Por ejemplo, se puede emplear un filtro FIR (filtro de respuesta de impulso finita), por ejemplo, y los coeficientes de filtro se pueden elegir adecuadamente, por ejemplo, como se describe a continuación.

15 O, se puede emplear por ejemplo, síntesis de campo de onda (WFS), bien conocida en la técnica de procesamiento de audio (para información general acerca de la síntesis de campo de onda, ver, por ejemplo, como uno de muchos ejemplos [69]).

20 O, se puede emplear, por ejemplo, Ambisonics de mayor orden, bien conocido en la técnica del procesamiento de audio (para información general acerca de Ambisonics de mayor orden, por ejemplo, como uno de muchos ejemplos [70]).

Ahora, se describe en más detalle un filtro 140 de acuerdo con algunos ejemplos particulares.

25 En particular, se presenta una implementación del bloque denotado por  $\mathbf{G}_1(\mathbf{k})$  y  $\mathbf{G}_2(\mathbf{k})$  mostrado en la figura 7. Un prefiltro se puede, por ejemplo, asociar con un arreglo de altavoces. Un conjunto de múltiples altavoces se considera como un arreglo de altavoces, siempre que un prefiltro alimente al menos una señal de entrada a múltiples altavoces que se excitan principalmente en el mismo intervalo de frecuencia. Es posible que un altavoz individual sea parte de  
30 múltiples arreglos y que múltiples señales de entrada se alimenten a un arreglo, que entonces se irradian hacia diferentes direcciones.

Hay diferentes métodos bien conocidos para determinar prefiltros lineales de tal forma que un arreglo de altavoces no direccionales exhibirá un patrón de radiación direccional, ver, por ejemplo, [1], [3], [4], [5] y [6].

35 Algunos ejemplos llevan a cabo un enfoque de acoplamiento de presión con base en respuestas de impulso medidas. Algunos de esos ejemplos, que emplean este enfoque, se describen en lo siguiente, donde sólo se considera un solo arreglo de altavoces. Otros ejemplos utilizan múltiples arreglos de altavoces. Es sencilla la aplicación a múltiples arreglos de altavoces.

40 Para la descripción de estos ejemplos, se utiliza una notación que es más adecuada para obtener filtros FIR en comparación con la notación anterior, que también cubriría filtros IIR. Para este fin, se capturan los coeficientes de filtro  $g_{l,q}(k)$  en los vectores.

$$\mathbf{g}_q = (g_{q,1}(0), \dots, g_{q,1}(L_G - 1), g_{q,2}(0), \dots, g_{q,2}(L_G - 1), \dots, g_{q,N_L}(0), \dots, g_{q,N_L}(L_G - 1))^T$$

(26)

45

Para la optimización, la respuesta de impulso convolucionada de los prefiltros y la respuesta de impulso de habitación (RIR) se pueden considerar, que están dadas por

$$z_m(k) = \sum_{l=1}^{N_L} \sum_{n=0}^{L_G-1} h_{m,l}(n)g_l(k-n),$$

(27)

50

en donde  $g_l(k)$  y  $h_{m,l}(k)$  se asumen que con cero para  $k < 0$  y  $k \geq L_G$  o  $k \geq L_H$ , , respectivamente.

Como resultado, las respuestas de impulso general  $z_m(k)$  tienen una longitud de  $L_G + L_H - 1$  muestras y se pueden capturar por el vector

55

$$\mathbf{z} = (z_1(0), z_1(1), \dots, z_1(L_G + L_H - 2), \\ z_2(0), z_2(1), \dots, z_2(L_G + L_H - 2), \dots, \\ z_{N_M}(0), z_{N_M}(1), \dots, z_{N_M}(L_G + L_H - 2))^T. \quad (28)$$

Ahora, es posible definir la matriz de convolución  $\mathbf{H}$ , de tal forma que

$$\hat{\mathbf{z}} = \mathbf{H}\mathbf{g} \quad (29)$$

describe la misma convolución como lo hace la Ecuación (27). Para la optimización, se puede definir el impulso deseado  $d_{m,q}(k)$  de acuerdo con las necesidades de la aplicación.

Una forma de definir  $d_{m,q}(k)$  es considerar cada altavoz como fuente potencial que se va a reproducir con su campo de sonido original en la parte brillante pero sin radiación a la zona oscura. Esto se describe por

$$d_{m,q}(k) = \begin{cases} h_{m,q}(k - \Delta k) & \text{si } h_{m,q}(k) \text{ pertenece a } \mathbf{B}_q(k), \\ 0 & \text{si } h_{m,q}(k) \text{ pertenece a } \mathbf{D}_q(k), \end{cases} \quad (30)$$

donde el retardo  $\Delta k$  se utiliza para asegurar causalidad. Se describe una reproducción perfecta por

$$\mathbf{d}_q = \mathbf{H}\mathbf{g}_q \quad (31)$$

pero convencionalmente no será posible debido a las limitaciones físicas. Se debe señalar que esta definición es sólo una entre muchas, que tiene algún mérito práctico debido a su simplicidad, en tanto que otras definiciones pueden ser más adecuadas, dependiendo del escenario de aplicación.

Ahora, el error de reproducción de mínimos cuadrados se puede definir como:

$$E_q = (\hat{\mathbf{z}} - \mathbf{d}_q)^H \mathbf{W}_q^H \mathbf{W}_q (\hat{\mathbf{z}} - \mathbf{d}_q), \quad (32)$$

$$= (\mathbf{g}^H \mathbf{H}^H - \mathbf{d}_q^H) \mathbf{W}_q^H \mathbf{W}_q (\mathbf{H}\mathbf{g} - \mathbf{d}_q), \quad (33)$$

donde  $\mathbf{W}_q$  es una matriz que se puede elegir de tal forma que se logra una ponderación dependiente de la frecuencia y/o ponderación dependiente de la posición.

Cuando se obtienen  $\mathbf{B}_q$  y  $\mathbf{D}_q$  de  $\mathbf{B}_q(k)$  y  $\mathbf{D}_q(k)$ , respectivamente, de la misma forma como  $\mathbf{H}$  se obtuvo de  $\mathbf{H}(k)$ , la Ecuación (14) se puede representar por

$$C_q = \frac{\mathbf{g}_q^H \mathbf{B}_q^H \mathbf{B}_q \mathbf{g}_q}{\mathbf{g}_q^H \mathbf{D}_q^H \mathbf{D}_q \mathbf{g}_q}. \quad (34)$$

Se debe señalar que la ecuación de maximización (34) se puede resolver como un problema generalizado de problemas propios [3].

El error  $E_q$  se puede reducir al mínimo al determinar el gradiente complejo de la Ecuación (33) y ajustándolo a cero [7]. El gradiente complejo de la Ecuación (33) está dado por

$$\frac{\partial E_q}{\partial \mathbf{g}_q^H} = \mathbf{H}^H \mathbf{W}_q^H \mathbf{W}_q \mathbf{H} \mathbf{g}_q - \mathbf{H}^H \mathbf{W}_q^H \mathbf{W}_q \mathbf{d}_q. \quad (35)$$

Que da por resultado

$$\mathbf{g}_q = (\mathbf{H}^H \mathbf{W}_q^H \mathbf{W}_q \mathbf{H})^{-1} \mathbf{H}^H \mathbf{W}_q^H \mathbf{W}_q \mathbf{d}_q \quad (36)$$

como la solución óptima de mínimos cuadrados.

- 5 Aunque, se formulan muchos algoritmos para mínimos cuadrados no ponderados, se pueden utilizar para implementar mínimos cuadrados ponderados al reemplazar simplemente  $\mathbf{H}$  y  $\mathbf{d}_q$  con  $\mathbf{W}_q \mathbf{H}$  y  $\mathbf{W}_q \mathbf{d}_q$ , respectivamente.

La matriz de ponderación  $\mathbf{W}_q$  es en general una matriz de convolución similar a  $\mathbf{H}$  definida por (26) a (29).

- 10 La matriz  $\mathbf{H}$  consiste de varias submatrices  $\mathbf{H}_{m,l}$ :

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_{1,1} & \mathbf{H}_{1,2} & \dots & \mathbf{H}_{1,N_L} \\ \mathbf{H}_{2,1} & \mathbf{H}_{2,2} & \dots & \mathbf{H}_{2,N_L} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_{L,1} & \mathbf{H}_{L,2} & \dots & \mathbf{H}_{L,N_L} \end{pmatrix} \quad (36a)$$

- 15 Un ejemplo para  $\mathbf{H}_{m,l}$  puede estar dado asumiendo

$$\begin{aligned} h_{1,1}(0) &= 5 \\ h_{1,1}(1) &= 4 \\ h_{1,1}(2) &= 3 \\ h_{1,1}(3) &= 2 \\ h_{1,1}(4) &= 1 \end{aligned} \quad (36b)$$

donde

$$\mathbf{H}_{1,1} = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 4 & 5 & 0 & 0 \\ 3 & 4 & 5 & 0 \\ 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (36c)$$

20

A partir de ese esquema es claro para el experto como (27) y (29) definen la estructura de  $\mathbf{H}$ .

- 25 Para facilitar una ponderación dependiente del micrófono y dependiente de la frecuencia mediante  $\mathbf{W}_q$ , las respuestas de impulso  $w_{m,q}(k)$  de acuerdo con los métodos de diseño de filtro bien conocidos. Aquí,  $w_{m,q}(k)$  define la ponderación para la fuente  $q$  y el micrófono  $m$ . A diferencia de  $\mathbf{H}$ ,  $\mathbf{W}_q$  es una matriz diagonal de bloques:

$$\mathbf{W}_q = \begin{pmatrix} \mathbf{W}_{1,q} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_{2,q} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{W}_{N_M,q} \end{pmatrix} \quad (36d)$$

en donde  $\mathbf{W}_{m,q}$  está estructurado como  $\mathbf{H}_{m,l}$ .

5 Con respecto al cálculo de los coeficientes de filtro, se señala que (36) da los coeficientes de filtro necesarios de forma explícita, su cálculo es muy demandante en la práctica. Debido a la similitud de este problema con el problema resuelto para ecualización de sala de escucha, también se pueden aplicar los métodos utilizados aquí.

10 Por lo tanto, se describe un algoritmo muy eficiente para calcular (36) en [71]: SCHNEIDER, Martin; KELLERMANN, Walter. Iterative DFT-domain inverse filter determination for adaptive listening room equalization. In: Acoustic Signal Enhancement; Proceedings of IWAENC 2012; International Workshop on. VDE, 2012, S. 1-4.

15 En lo siguiente, se describe un sistema de altavoz-recinto-micrófono (LEMS) de acuerdo con ejemplos. En particular, se analiza el diseño de un LEMS de acuerdo con ejemplos. En algunos ejemplos, las medidas descritas anteriormente, por ejemplo, pueden basarse en las distintas propiedades del LEMS.

20 La figura 11 ilustra una configuración de altavoz a modo de ejemplo en un recinto. En particular, la figura 11 ilustra un LEMS de ejemplo con cuatro zonas de sonido. Se debe reproducir una escena acústica individual en cada una de aquellas zonas de sonido. Para este fin, los altavoces mostrados en la figura 11 se utilizan de formas específicas, dependiendo de su posición con respecto el uno al otro y con respecto a las zonas de sonido.

25 Los dos arreglos de altavoces denotados por "arreglo 1" y "arreglo 2" se utilizan junto con prefiltros determinados en consecuencia (ver más arriba). De esta forma, es posible dirigir eléctricamente la radiación de aquellos arreglos hacia "zona 1" y "zona 2". Asumiendo que ambos arreglos exhiben una distancia de inter-altavoz de unos cuantos centímetros en tanto que los arreglos exhiben un tamaño de apertura de unos cuantos decímetros, es posible direccionamiento efectivo para frecuencias de intervalo medio.

30 Aunque no es evidente, los altavoces omni-direccionales "LS 1", "LS 2", "LS 3", y "LS 4", que por ejemplo, pueden estar ubicados 1 a 3 metros distantes el uno del otro también se pueden accionar como un arreglo de altavoces cuando se consideran frecuencias por debajo de, por ejemplo, 300 Hz. Se pueden determinar prefiltros conformes se utiliza el método descrito anteriormente.

35 Los altavoces "LS 5" y "LS 6" son altavoces direccionales que proporcionan audio de alta frecuencia a las zonas 3 y 4, respectivamente.

40 Como se describe anteriormente, las medidas para reproducción direccional algunas veces pueden no conducir a resultados suficientes para el intervalo de frecuencia audible completo. Para compensar este problema, puede haber, por ejemplo, altavoces ubicados en la vecindad cercana o dentro de las zonas de sonido respectivas. Aunque este posicionamiento es subóptimo con respecto a la calidad de sonido percibida, la diferencia en distancia de los altavoces a la zona asignada en comparación con la distancia a las otras zonas permite una reproducción espacialmente enfocada, independiente de la frecuencia. Por lo tanto, estos altavoces, por ejemplo, se pueden utilizar en intervalos de frecuencia donde los otros métodos no conducen a resultados satisfactorios.

A continuación, se describen aspectos adicionales de acuerdo con algunos de los ejemplos:

45 En algunos de los ejemplos, el bloque de "preprocesamiento" se coloca después de los bloques de "divisor de banda" o después de los bloques de "conformador espectral". En ese caso, un bloque de preprocesamiento puede, por ejemplo, implementarse para cada una de las bandas de frecuencia "divididas". En el ejemplo que se muestra en la Fig. 7, un bloque de "preprocesamiento" consideraría  $w_1(k)$  y  $w_4(k)$  y otro  $w_2(k)$  y  $w_3(k)$ . Aun así, se tiene que colocar un aspecto del preprocesamiento en la posición antigua, como se describe anteriormente, donde se describe el pre procesamiento.

50 Ya que la fuga acústica depende del método de reproducción que a menudo se elige de forma diferente para cada banda de frecuencia, esta implementación tiene la ventaja de que los parámetros de preprocesamiento se pueden hacer coincidir con las demandas del método de reproducción. Además, cuando se elige esta implementación, la compensación de la fuga en una banda de frecuencia no afectará otra banda de frecuencia. Dado que el bloque de "preprocesamiento" no es un sistema LTI este intercambio implica un cambio en la funcionalidad del sistema total, aunque el sistema resultante aún resolverá de forma confiable el mismo problema.

55 Además, cabe señalar que algunos de los ejemplos pueden utilizar una medición de las respuestas de impulso de todos los altavoces a múltiples micrófonos antes de la operación. Por lo tanto, no es necesario ningún micrófono durante la operación.

60 El método propuesto es generalmente adecuado para cualquier escenario de reproducción multizona, por ejemplo, escenarios en automóvil.

Si bien se han descrito algunos aspectos en el contexto de un aparato, es obvio que estos aspectos también representan una descripción del método correspondiente, en el cual un bloque o dispositivo corresponde a una etapa del método o a una característica de una etapa del método. De manera análoga, los aspectos descritos en el contexto de una etapa del método también representan una descripción de un bloque o elemento correspondiente o de una característica de un aparato correspondiente. Algunos o todos las etapas del método pueden ser ejecutados por medio de (o utilizando) un aparato de hardware, como por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, uno o más de las etapas más importantes del método pueden ser ejecutadas por dicho aparato.

Dependiendo de ciertos requisitos de implementación, las formas pueden ser implementadas en hardware o en software o por lo menos parcialmente en software, o por lo menos parcialmente en hardware. La implementación se puede realizar empleando un medio de almacenamiento digital, por ejemplo un disco blando, un DVD, un Blu-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene almacenadas en la misma señales control legibles electrónicamente, que cooperan (o tienen capacidad para cooperar) con un sistema de informático programable de tal manera que se ejecute el método respectivo. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.

Algunos ejemplos de acuerdo a la invención comprenden un portador de datos que tiene señales de control electrónicamente leibles, que son capaces de cooperar con un sistema de computadora programable, de tal forma que se lleva a cabo uno de los métodos descritos en la presente.

En general, los ejemplos de la presente invención se pueden implementar como un producto de programa de computadora con un código de programa, el código de programa que es operativo para llevar a cabo uno de los métodos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede ser almacenado, por ejemplo, en un portador legible por una máquina.

Otras realizaciones comprenden el programa informático para ejecutar uno de los métodos aquí descritos, almacenado en un portador legible por una máquina.

En otras palabras, una realización del método inventivo es, por lo tanto, un programa informático que tiene un código de programa para realizar uno de los métodos descritos en la presente, cuando el programa informático se ejecuta en un ordenador.

Un ejemplo adicional de los métodos inventivos, por lo tanto, es un portador de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, grabado en el mismo, el programa informático para llevar a cabo uno de los métodos descritos en la presente. El portador de datos, el medio de almacenamiento digital o el medio grabado son por lo general tangibles y/o no transitorios.

Un ejemplo adicional del método inventivo, por lo tanto, es un flujo de datos o una secuencia de señales que representan el programa informático para llevar a cabo uno de los métodos descritos en la presente. El flujo de datos o la secuencia de señales pueden estar configurados, por ejemplo, para ser transferida a través de una conexión de comunicación de datos, por ejemplo a través de Internet.

Un ejemplo adicional comprende un medio de procesamiento, por ejemplo un ordenador o un dispositivo lógico programable, configurado para o adaptado para llevar a cabo uno de los métodos descritos en la presente.

Una realización adicional comprende un ordenador que tiene instalada en el mismo el programa informático para llevar a cabo uno de los métodos descritos en la presente.

En algunos ejemplos, un dispositivo lógico programable (por ejemplo, un arreglo de compuertas programable en el campo) se puede utilizar para llevar a cabo algunas o todas las funcionalidades de los métodos descritos en la presente. En algunos ejemplos, un arreglo de compuertas programable en el campo puede cooperar con un microprocesador a fin de llevar a cabo uno de los métodos descritos en la presente. Por lo general, los métodos son ejecutados preferentemente por cualquier aparato de hardware.

El aparato aquí descrito puede implementarse utilizando un aparato de hardware, o utilizando un ordenador, o utilizando una combinación de un aparato de hardware y un ordenador.

Los métodos aquí descritos pueden llevarse a cabo utilizando un aparato de hardware, o utilizando un ordenador, o utilizando una combinación de un aparato de hardware y un ordenador.

Los ejemplos descritos anteriormente son meramente ilustrativos de los principios de la presente invención. Se entiende que modificaciones y variaciones de las disposiciones y los detalles aquí descritos resultarán evidentes para los expertos en la técnica. Es la intención, por consiguiente, limitarse sólo por el alcance de las reivindicaciones de

patente que siguen y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones en el presente documento.

## Referencias

- 5
- [1] W. Druyvesteyn y J. Garas, "Personal sound," *Journal of the Audio Engineering Society*, vol. 45, no. 9, pp. 685–701, 1997.
- 10
- [2] F. Dowlá y A. Spiridon, "Spotforming with an array of ultra-wideband radio transmitters", en *Ultra Wideband Systems and Technologies*, 2003 IEEE Conference on, noviembre de 2003, págs. 172-175.
- [3] J.-W. Choi y Y.-H. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources", *Journal of the Acoustical Society of America*, vol. 111, no. 4, pp. 1695-1700, 2002.
- 15
- [4] M. Poletti, "An investigation of 2-d multizone surround sound systems", en *Audio Engineering Society Convention 125*, octubre de 2008. [En línea]. Disponible: <http://www.aes.org/e-lib/browse.cfm?elib=14703> .
- [5] Y. Wu y T. Abhayapala, "Spatial multizone soundfield reproduction," in *Acoustics, Speech and Signal Processing*, 2009. ICASSP 2009. IEEE International Conference on, abril de 2009, pp. 93-96.
- 20
- [6] Y. J. Wu y T. D. Abhayapala, "Spatial multizone soundfield reproduction: Theory and design," *Audio, Speech, and Language Processing*, IEEE Transactions on, vol. 19, no. 6, pp. 1711–1720, 2011.
- [7] D. Brandwood, "A complex gradient operator and its application in adaptive array theory", *Microwaves, Optics and Antennas*, IEE Proceedings H, vol. 130, no. 1, pp. 11 - 16, febrero de 1983.
- 25
- [8] US 2005/0152562 A1.
- [9] US 2013/170668 A1.
- 30
- [10] US 2008/0071400 A1.
- [11] US 2006/0034470 A1.
- 35
- [12] US 2011/0222695 A1.
- [13] US 2009/0232320 A1.
- [14] US 2015/0256933 A1.
- 40
- [15] US 6,674,865 B1.
- [16] DE 30 45 722 A1.
- 45
- [17] US 2012/0140945 A1.
- [18] US 2008/0273713 A1.
- [19] US 2004/0105550 A1.
- 50
- [20] US 2006/0262935 A1.
- [21] US 2005/0190935 A1.
- 55
- [22] US 2008/0130922 A1.
- [23] US 2010/0329488 A1.
- [24] DE 10 2014 210 105 A1.
- 60
- [25] US 2011/0286614 A1.
- [26] US 2007/0053532 A1.

[27] US 2013/0230175 A1.  
[28] WO 2016/008621 A1.  
5 [29] US 2008/0273712 A1.  
[30] US 5,870,484.  
[31] US 5,809,153.  
10 [32] US 2006/0034467 A1.  
[33] US 2003/0103636 A1.  
15 [34] US 2003/0142842 A1.  
[35] JP 5345549.  
[36] US2014/0056431 A1.  
20 [37] US 2014/0064526 A1.  
[38] US 2005/0069148 A1.  
25 [39] US 5,081,682.  
[40] DE 90 15 454.  
[41] US 5,550,922.  
30 [42] US 5,434,922.  
[43] US 6,078,670.  
35 [44] US 6,674,865 B1.  
[45] DE 100 52104A1.  
[46] US 2005/0135635 A1.  
40 [47] DE102 42 558 A1.  
[48] US 2010/0046765 A1.  
45 [49] DE 10 2010 040 689.  
[50] US 2008/0103615 A1.  
[51] US 8,190,438 B1.  
50 [52] WO 2007/098916 A1.  
[53] US 2007/0274546 A1.  
55 [54] US 2007/0286426 A1.  
[55] US 5,018,205.  
[56] US 4,944,018.  
60 [57] DE 103 51 145 A1.  
58] JP 2003-255954.

[59] US 4,977,600.

[60] US 5,416,846.

5 [61] US 2007/0030976 A1.

62] JP 2004-363696.

10 [63] Wikipedia: "Angular resolution", [https://en.wikipedia.org/wiki/Angular\\_resolution](https://en.wikipedia.org/wiki/Angular_resolution) , recuperado de Internet el 8 de abril de 2016.

[64] Wikipedia: "Nyquist-Shannon sampling theorem", [https://en.wikipedia.org/wiki/Nyquist-Shannon\\_sampling\\_theorem](https://en.wikipedia.org/wiki/Nyquist-Shannon_sampling_theorem) , recuperado de Internet el 8 de abril de 2016.

15 [65] Wikipedia: "Dynamic range compression", [https://en.wikipedia.org/wiki/Dynamic\\_range\\_compression](https://en.wikipedia.org/wiki/Dynamic_range_compression) , recuperado de Internet el 8 de abril de 2016.

[66] Wikipedia: "Filtro de ponderación", [https://en.wikipedia.org/wiki/Weighting\\_filter](https://en.wikipedia.org/wiki/Weighting_filter) , recuperado de Internet el 8 de abril de 2016.

20 [67] Wikipedia: "Audio crossover - Digital" , [https://en.wikipedia.org/wiki/Audio\\_crossover#Digital](https://en.wikipedia.org/wiki/Audio_crossover#Digital) , recuperado de Internet el 8 de abril de 2016.

25 [68] Wikipedia: "Equalization (audio) - Filter functions", [https://en.wikipedia.org/wiki/Equalization\\_\(audio\)#\\_Funciones\\_de\\_filtro](https://en.wikipedia.org/wiki/Equalization_(audio)#_Funciones_de_filtro), recuperadas de Internet el 8 de abril de 2016.

[69] WO 2004/114725 A1.

[70] EP 2 450 880 A1.

30 [71] SCHNEIDER, Martin; KELLERMANN, Walter: "Iterative DFT-domain inverse filter determination for adaptive listening room equalization." In: Acoustic Signal Enhancement; Proceedings of IWAENC 2012; International Workshop on. VDE, 2012, S. 1-4.

35

## REIVINDICACIONES

1. Un aparato para generar una pluralidad de señales de altavoz de dos o más señales de fuente de audio, en donde cada una de las dos o más señales de fuente de audio se deben reproducir en una o más de dos zonas de sonido, y en donde al menos una de las dos o más señales de fuente de audio no se deben reproducir en al menos una de las dos o más zonas de sonido, en donde el aparato comprende:
- 5 un preprocesador de audio (110) configurado para modificar cada una de las dos o más señales de audio iniciales para obtener dos o más señales de audio preprocesadas, y
- 10 un filtro (140) configurado para generar la pluralidad de señales de altavoz dependiendo de las dos o más señales de audio preprocesadas,
- 15 en donde el preprocesador de audio (110) está configurado para utilizar las dos o más señales de fuente de audio como las dos o más señales de audio iniciales, o en donde el preprocesador de audio (110) está configurado para generar para cada señal de fuente de audio de las dos o más señales de fuente de audio una señal de audio inicial de las dos o más señales de audio iniciales al modificar la señal de fuente de audio,
- 20 en donde el preprocesador de audio (110) está configurado para modificar cada señal de audio inicial de las dos o más señales de audio iniciales dependiendo de una potencia de señal o un volumen de otra señal de audio inicial de las dos o más señales de audio iniciales, y
- 25 en donde el filtro (140) está configurado para generar la pluralidad de señales de altavoz dependiendo de en cuáles de las dos o más zonas de sonido se deben reproducir las dos o más señales de fuente de audio y dependiendo de en cuáles de las dos o más zonas de sonido no se deben reproducir las dos o más señales de fuente de audio.
- 30 en donde el preprocesador de audio (110) está configurado para modificar cada señal de audio inicial de las dos o más señales de audio iniciales dependiendo de la potencia de señal o el volumen de otra señal de audio inicial de las dos o más señales de audio iniciales al determinar una ganancia para la señal de audio inicial y al aplicar la ganancia en la señal de audio inicial,
- 35 en donde el preprocesador de audio (110) está configurado para determinar la ganancia dependiendo de la relación entre un primer valor y un segundo valor, siendo dicha relación una relación entre la potencia de señal de dicha otra señal de audio inicial de las dos o más señales de audio iniciales como el primer valor y la potencia de señal de dicha señal de audio inicial como el segundo valor, o siendo dicha relación una relación entre la sonoridad de dicha otra señal de audio inicial de las dos o más señales de audio iniciales como el primer valor y la sonoridad de dicha señal de audio inicial como el segundo valor,
- 40 caracterizado porque:
- 45 el preprocesador de audio (110) está configurado para modificar una señal de audio inicial de las dos o más señales de audio iniciales determinando la ganancia  $g'_1(k)$  para dicha señal de audio inicial y aplicando la ganancia  $g'_1(k)$  en dicha señal de audio inicial, en donde el preprocesador de audio (110) está configurado para determinar la ganancia  $g'_1(k)$  de acuerdo con

$$g'_1(k) = \begin{cases} 10^{(T_1 + v_1)(R-1)/(20R)} & \text{para } T_1 + v_1 < 0 \\ 1 & \text{de otra forma,} \end{cases}$$

50 o de acuerdo con

$$g'_1(k) = \begin{cases} 10^{(T_2 + v_1)(R-1)/(20R)} & \text{para } T_2 + v_1 > 0 \\ 1 & \text{de otra forma,} \end{cases}$$

con

$$v_1 = -10\log_{10}(e_1(k)) + 10\log_{10}\left(\sum_{i=2}^N e_i(k)\right)$$

5 en donde  $k$  es un índice de tiempo, en donde  $T_1$  indica un primer valor umbral y  $T_2$  indica un segundo valor umbral, en donde  $e_1(k)$  indica una potencia de señal o un volumen de dicha señal de audio inicial, en donde  $N$  indica un número de las dos o más señales de audio iniciales, en donde  $e_i(k)$  indica una potencia de señal o un volumen de una señal de audio inicial adicional de las dos o más señales de audio iniciales, y en donde  $R$  indica un número, con  $1 \leq R \leq 100$ ,

o

10 el preprocesador de audio (110) está configurado para modificar cada señal de audio inicial de las dos o más señales de audio iniciales dependiendo de la potencia de la señal o el volumen de otra señal de audio inicial de las dos o más señales de audio iniciales determinando la ganancia  $g'_1(k)$  para dicha señal de audio inicial y aplicando la ganancia  $g'_1(k)$  a dicha señal de audio inicial, en donde

15 el preprocesador de audio (110) está configurado para determinar la ganancia  $g'_1(k)$  de acuerdo con

$$g'_1(k) = \begin{cases} 10^{(T_1+v)(R-1)/(20R)} & \text{para } T_1 + v < 0 \\ 1 & \text{de otra forma,} \end{cases}$$

20 o de acuerdo con

$$g'_1(k) = \begin{cases} 10^{(T_2+v)(R-1)/(20R)} & \text{para } T_2 + v > 0 \\ 1 & \text{de otra forma,} \end{cases}$$

25 con

$$v = -10\log_{10}(e_1(k)) + 10\log_{10}(e_2(k))$$

30 en donde  $k$  es un índice de tiempo, en donde  $T_1$  indica un primer valor umbral y  $T_2$  indica un segundo valor umbral, en donde  $e_1(k)$  indica una potencia de señal o una sonoridad de dicha señal de audio inicial, en donde  $e_2(k)$  indica una potencia de señal o volumen de dicha otra señal de audio inicial de las dos o más señales de audio iniciales, y en donde  $R$  indica un número, con  $1 \leq R \leq 100$ .

35 2. Un aparato de acuerdo con la reivindicación 1, en donde el preprocesador de audio (110) está configurado para determinar la ganancia dependiendo de una función que incrementa de forma monótona con la relación entre el primer valor y el segundo valor.

3. Un aparato de acuerdo con una de las reivindicaciones anteriores,

40 en donde el preprocesador de audio (110) está configurado para modificar cada señal de audio inicial de las dos o más señales de audio iniciales de acuerdo con

$$e_1(k) = \lambda_2 e_1(k-1) + (1-\lambda_2) \sum_{l=1}^L d_1^2(k, l),$$

45 o de acuerdo con

$$e_1(k) = \frac{1}{K} \sum_{n=0}^{K-1} \sum_{l=1}^L d_1^2(k-n, l),$$

o de acuerdo con

$$e_1(k) = \max_{n=0,1,\dots,K-1, l=1,2,\dots,L} d_1^2(k-n, l).$$

en donde  $e_1(k)$  indica una potencia de señal de dicha señal de audio inicial,

5 en donde  $k$  indica un índice de tiempo,

en donde  $\lambda_2$  es un valor en el intervalo  $0 < \lambda_2 < 1$ ,

10 en donde  $L$  es un número de canales de audio de la señal de audio inicial, en donde  $L \geq 1$ ,

en donde  $d_1$  indica dicha señal de audio inicial,

en donde  $K$  indica un número de muestras de una ventana.

15 4. Un aparato de acuerdo con una de las reivindicaciones anteriores, en donde el preprocesador de audio (110) está configurado para generar las dos o más señales de audio iniciales normalizando una potencia de cada una de las dos o más señales de fuente de audio.

20 5. Un aparato de acuerdo con la reivindicación 4,

en donde el preprocesador de audio (110) está configurado para generar cada señal de audio inicial de las dos o más señales de audio iniciales normalizando una potencia de cada señal de fuente de audio de las dos o más señales de fuente de audio de acuerdo con

25 
$$d_1(k, l) = c_1(k) u_1(k, l)$$
 ,

y de acuerdo con

$$c_1(k) = \frac{1}{\sqrt{b_1(k)}}.$$

30 en donde  $k$  es un índice de tiempo,

en donde  $l$  indica uno de uno o más canales de audio de dicha señal de fuente de audio,

35 en donde  $d_1$  indica dicha señal de audio inicial,

en donde  $u_1$  indica dicha señal de fuente de audio,

40 en donde  $b_1$  indica una media de una potencia de dicha señal de fuente de audio  $u_1$ .

6. Un aparato de acuerdo con la reivindicación 5,

en donde el preprocesador de audio (110) está configurado para determinar la media  $b_1$  de la potencia de dicha señal de fuente de audio  $u_1$  de acuerdo con

45 
$$b_1(k) = \lambda_1 b_1(k-1) + (1-\lambda_1) \sum_{l=1}^L u_1^2(k, l),$$

en donde  $0 < \lambda_1 < 1$ .

50 7. Un aparato de acuerdo con una de las reivindicaciones anteriores, en donde el filtro (140) está configurado para generar la pluralidad de señales de altavoz dependiendo de en cuáles de las dos o más zonas de sonido se deben reproducir las dos o más señales de fuente de audio y dependiendo de en cuáles de las dos o más zonas de sonido no se deben reproducir las dos o más señales de fuente de audio, mediante la determinación de los coeficientes de filtrado de un filtro FIR.

55 8. Un aparato de acuerdo con la reivindicación 7,

en donde el filtro (140) está configurado para generar la pluralidad de señales de altavoz dependiendo de en cuál de las dos o más zonas de sonido se reproducirán las dos o más señales de fuente de audio y

dependiendo de en cuál de las dos o más zonas de sonido no se reproducirán las dos o más señales de fuente de audio mediante la determinación de los coeficientes de filtrado del filtro FIR de acuerdo con la fórmula

$$\mathbf{g}_q = (\mathbf{H}^H \mathbf{W}_q^H \mathbf{W}_q \mathbf{H})^{-1} \mathbf{H}^H \mathbf{W}_q^H \mathbf{W}_q \mathbf{d}_q$$

en donde  $\mathbf{g}_q$  es un vector que comprende los coeficientes de filtro del filtro FIR de acuerdo con

$$\mathbf{g}_q = (g_{q,1}(0), \dots, g_{q,1}(L_G - 1), g_{q,2}(0), \dots, g_{q,2}(L_G - 1), \dots, g_{q,N_L}(0), \dots, g_{q,N_L}(L_G - 1))^T$$

en donde H es una matriz de convolución que depende de una respuesta al impulso de la habitación,

en donde W es una matriz de ponderación,

en donde  $\mathbf{d}_q$  indica respuestas de impulso deseadas,

en donde  $g_{q,i}$  indica uno de los coeficientes de filtro con  $1 < i < N_L$ ,

en donde  $N_L$  indica una cantidad de altavoces, y

en donde  $L_G$  indica una longitud del filtro FIR.

9. Un aparato de acuerdo con una de las reivindicaciones 1 a 6, en donde el filtro (140) está configurado para generar la pluralidad de señales de altavoz dependiendo de en cuáles de las dos o más zonas de sonido se deben reproducir las dos o más señales de fuente de audio y dependiendo de en cuáles de las dos o más zonas de sonido no se deben reproducir las dos o más señales de fuente de audio, realizando la síntesis del campo de onda.

10. Un aparato de acuerdo con una de las reivindicaciones anteriores, en donde el aparato comprende además dos o más divisores de banda (121, 122 que se configuran para llevar a cabo división de banda en las dos o más señales de audio preprocesadas a una pluralidad de señales de audio de banda dividida,

en donde el filtro (140) está configurado para generar la pluralidad de señales de altavoz dependiendo de la pluralidad de señales de audio de banda dividida.

11. Un aparato de acuerdo con la reivindicación 10, en donde el aparato comprende además uno o más conformadores espectrales (131, 132, 133, 134 que se configuran para modificar un envolvente espectral de una o más de la pluralidad de señales de audio de banda dividida para obtener una o más señales de audio espectralmente conformadas,

en donde el filtro (140) está configurado para generar la pluralidad de señales de altavoz en función de una o más señales de audio espectralmente conformadas.

12. Un método generar una pluralidad de señales de altavoz de dos o más señales de fuente de audio, en donde cada una de las dos o más señales de fuente de audio se deben reproducir en una o más de dos zonas de sonido, y en donde al menos una de las dos o más señales de fuente de audio no se deben reproducir en al menos una de las dos o más zonas de sonido, en donde el método comprende:

modificar cada una de las dos o más señales de audio iniciales para obtener dos o más señales de audio preprocesadas, y

generar la pluralidad de señales de altavoz dependiendo de las dos o más señales de audio preprocesadas,

en donde las dos o más señales de fuente de audio se utilizan como las dos o más señales de audio iniciales, o en donde para cada señal de fuente de audio de las dos o más señales de fuente de audio se genera una señal de audio inicial de las dos o más señales de audio iniciales al modificar la señal de fuente de audio,

en donde cada audio inicial de las dos o más señales de audio iniciales se modifica dependiendo de una potencia de señal o un volumen de otra señal de audio inicial de las dos o más señales de audio iniciales, y

en donde la pluralidad de señales de altavoz se genera dependiendo de en cuáles de las dos o más zonas de audio se deben reproducir las dos o más señales de fuente de audio y dependiendo de en cuáles de las dos o más zonas de sonido las dos o más señales de fuente de audio no se deben reproducir,

5 en donde la modificación de cada señal de audio inicial de las dos o más señales de audio iniciales dependiendo de la potencia de señal o el volumen de otra señal de audio inicial de las dos o más señales de audio iniciales al determinar una ganancia para la señal de audio inicial y al aplicar la ganancia en la señal de audio inicial,

10 en donde la determinación de la ganancia se lleva a cabo dependiendo de la relación entre un primer valor y un segundo valor, siendo dicha relación una relación entre la potencia de señal de dicha otra señal de audio inicial de las dos o más señales de audio iniciales como el primer valor y la potencia de señal de dicha señal de audio inicial como el segundo valor, o siendo dicha relación una relación entre la sonoridad de dicha otra señal de audio inicial de las dos o más señales de audio iniciales como el primer valor y el volumen de dicha señal de audio inicial como el segundo valor,

caracterizado porque:

20 el método comprende modificar una señal de audio inicial de las dos o más señales de audio iniciales determinando la ganancia  $g'_1(k)$  para dicha señal de audio inicial y aplicando la ganancia  $g'_1(k)$  en dicha señal de audio inicial, en donde el preprocesador de audio (110) está configurado para determinar la ganancia  $g'_1(k)$  de acuerdo con

$$g'_1(k) = \begin{cases} 10^{(T_1+v_1)(R-1)/(20R)} & \text{para } T_1 + v_1 < 0 \\ 1 & \text{de otra forma,} \end{cases}$$

25 o de acuerdo con

$$g'_1(k) = \begin{cases} 10^{(T_2+v)(R-1)/(20R)} & \text{para } T_2 + v > 0 \\ 1 & \text{de otra forma,} \end{cases}$$

30 con

$$v_1 = -10\log_{10}(e_1(k)) + 10\log_{10}\left(\sum_{i=2}^N e_i(k)\right)$$

en donde  $k$  es un índice de tiempo,

35 en donde  $T_1$  indica un primer valor umbral y  $T_2$  indica un segundo valor umbral, en donde  $e_i(k)$  indica una potencia de señal o un volumen de dicha señal de audio inicial, en donde  $N$  indica un número de las dos o más señales de audio iniciales, en donde  $e_i(k)$  indica una potencia de señal o un volumen de una señal de audio inicial adicional de las dos o más señales de audio iniciales, y en donde  $R$  indica un número, con  $1 \leq R \leq 100$ ;

o

45 el método comprende modificar cada señal de audio inicial de las dos o más señales de audio iniciales dependiendo de la potencia de la señal o el volumen de otra señal de audio inicial de las dos o más señales de audio iniciales determinando la ganancia  $g'_1(k)$  para dicha señal de audio inicial y aplicando la ganancia  $g'_1(k)$  a dicha señal de audio inicial, en donde el método comprende determinar la ganancia  $g'_1(k)$ , de acuerdo con

$$g'_1(k) = \begin{cases} 10^{(T_1+v_1)(R-1)/(20R)} & \text{para } T_1 + v_1 < 0 \\ 1 & \text{de otra forma,} \end{cases}$$

50 o de acuerdo con

$$g'_1(k) = \begin{cases} 10^{(T_2+v)(R-1)/(20R)} & \text{para } T_2 + v > 0 \\ 1 & \text{de otra forma,} \end{cases}$$

con

5

$$v = -10\log_{10}(e_1(k)) + 10\log_{10}(e_2(k))$$

10

en donde  $k$  es un índice de tiempo, en donde  $T_1$  indica un primer valor umbral y  $T_2$  indica un segundo valor umbral, en donde  $e_1(k)$  indica una potencia de señal o una sonoridad de dicha señal de audio inicial, en donde  $e_2(k)$  indica una potencia de señal o volumen de dicha otra señal de audio inicial de las dos o más señales de audio iniciales, y en donde  $R$  indica un número, con  $1 \leq R \leq 100$ .

13. Un programa informático para implementar el método de la reivindicación 12 cuando se ejecuta en un ordenador o procesador de señales.

15

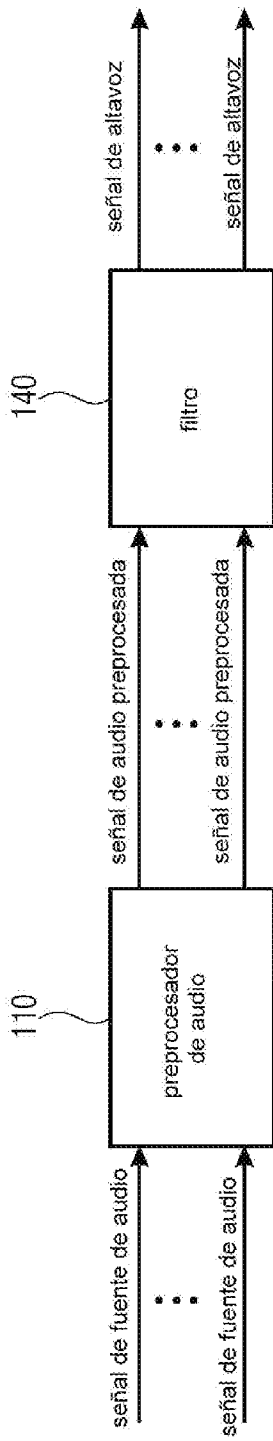


Fig. 1

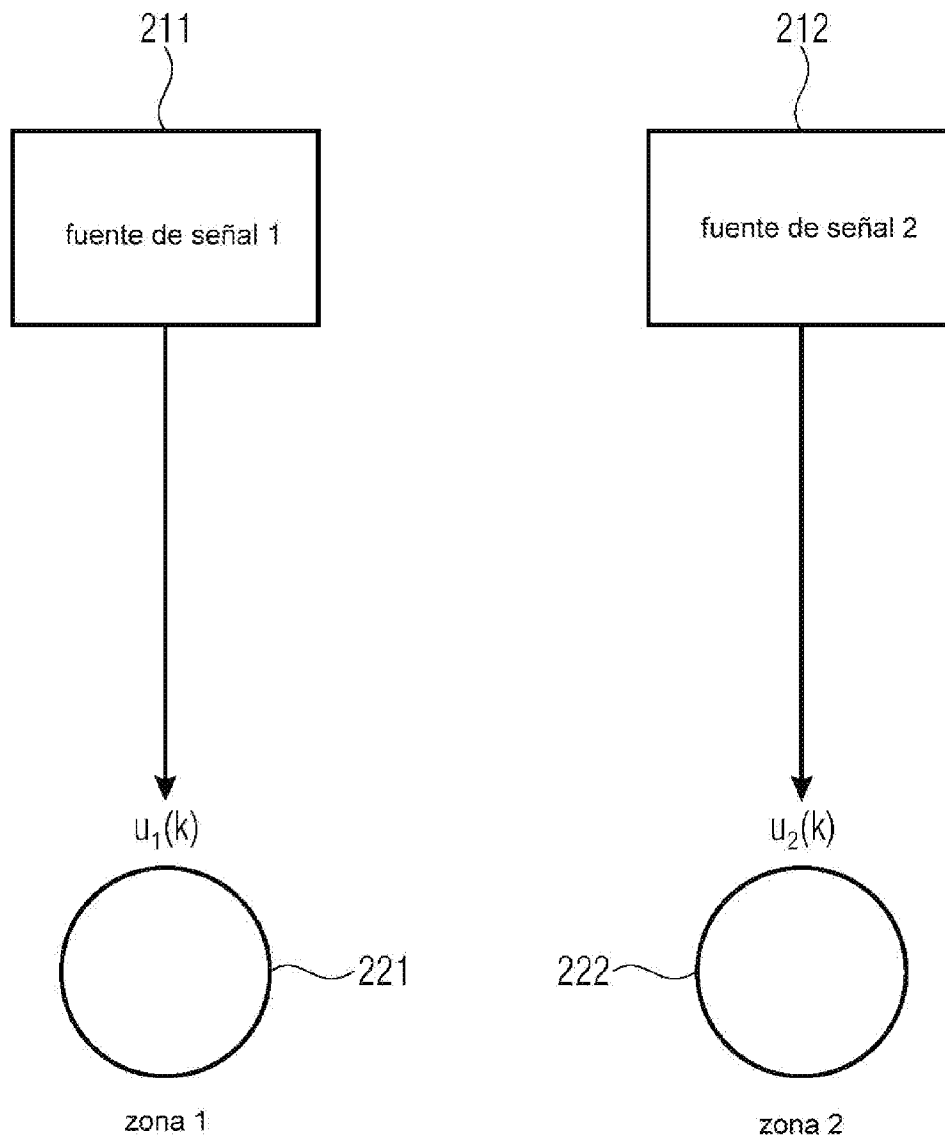


Fig. 2

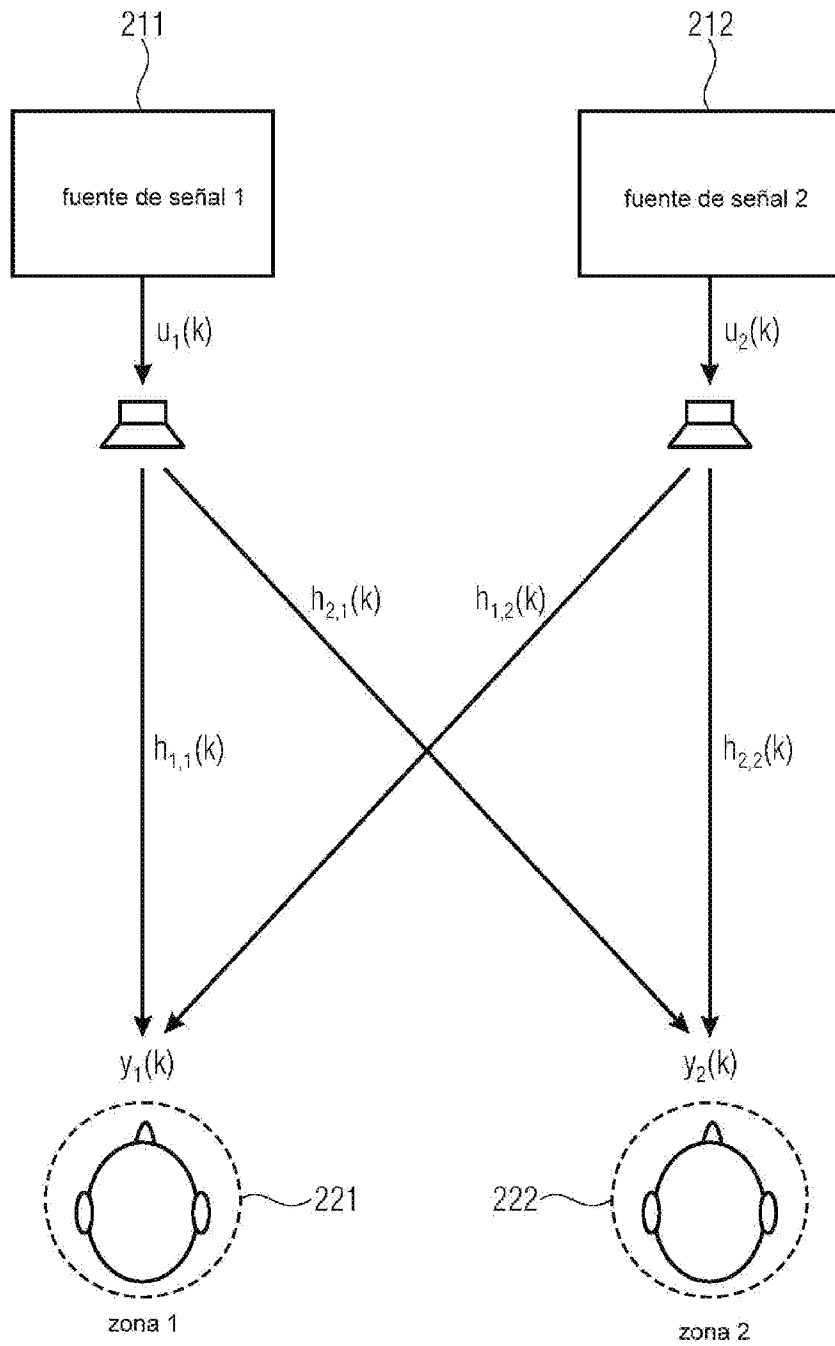


Fig. 3

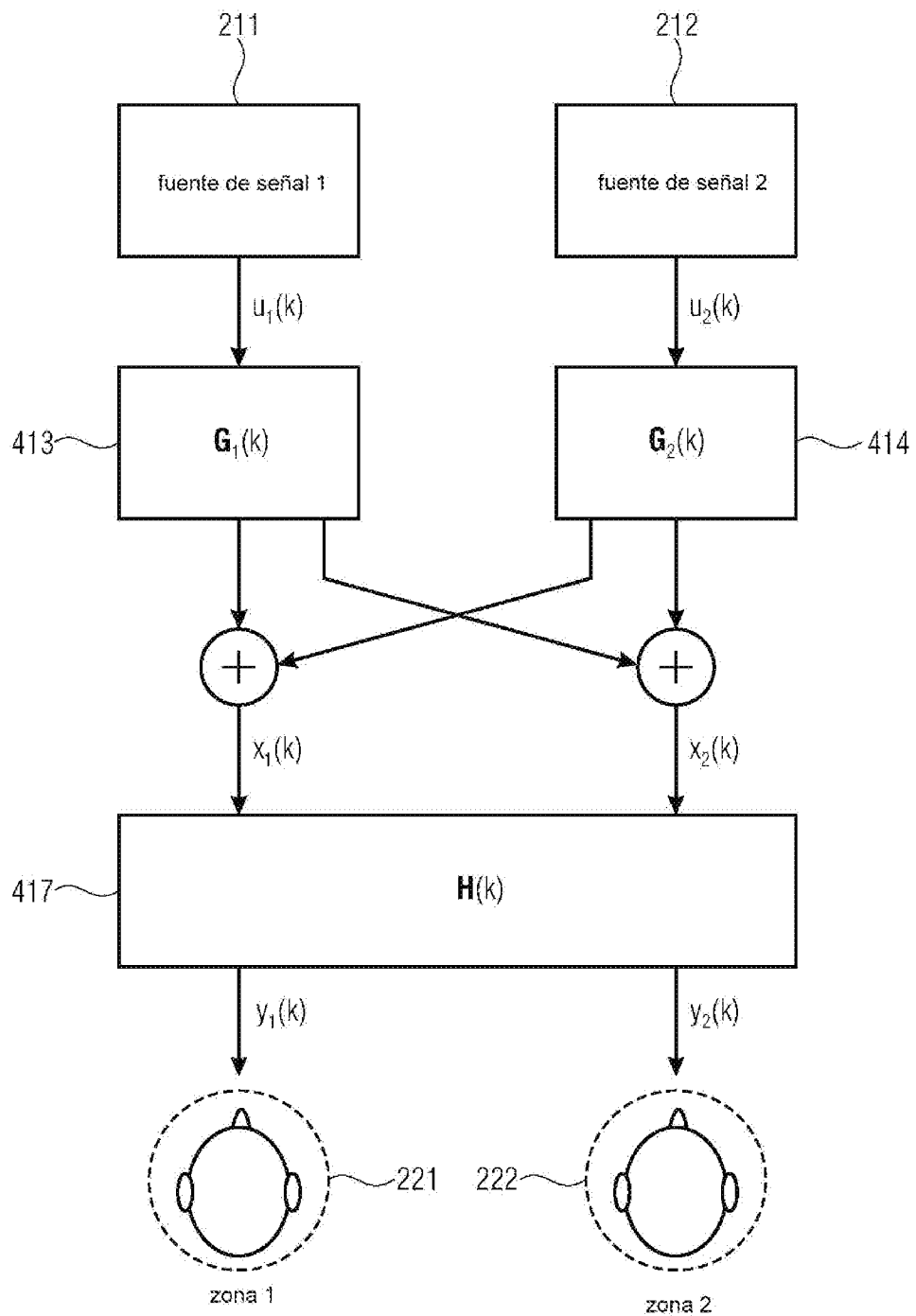


Fig. 4

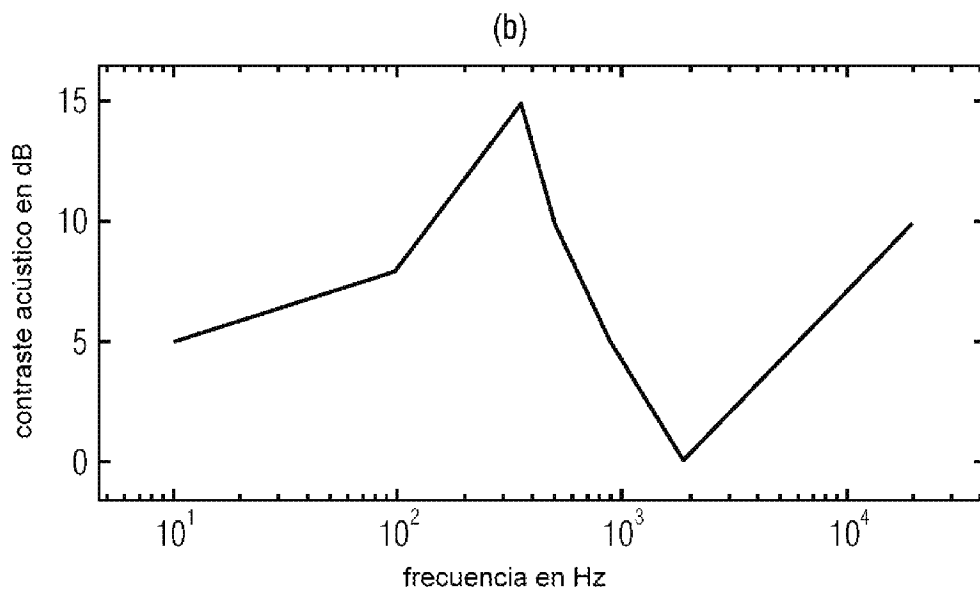
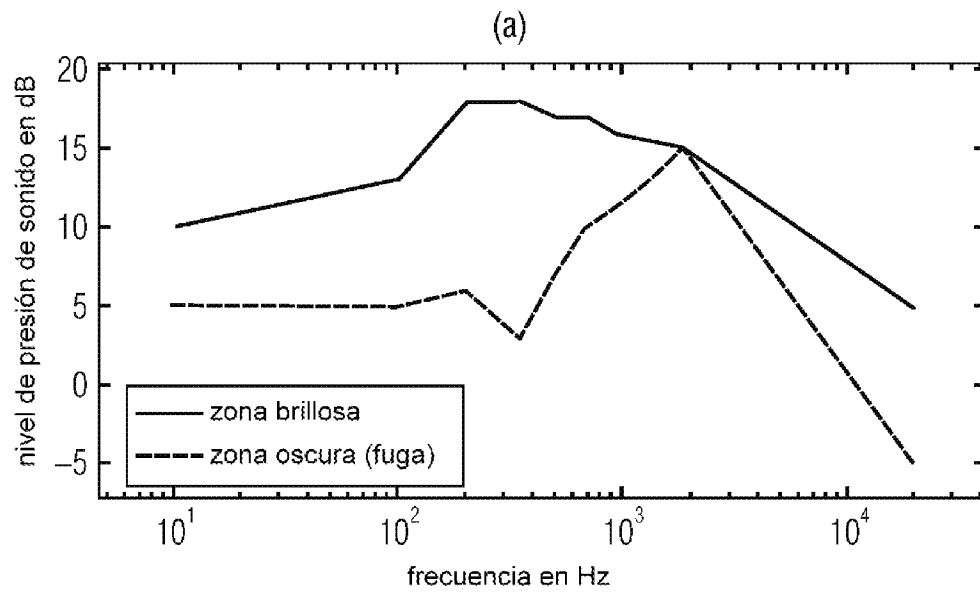


Fig. 5

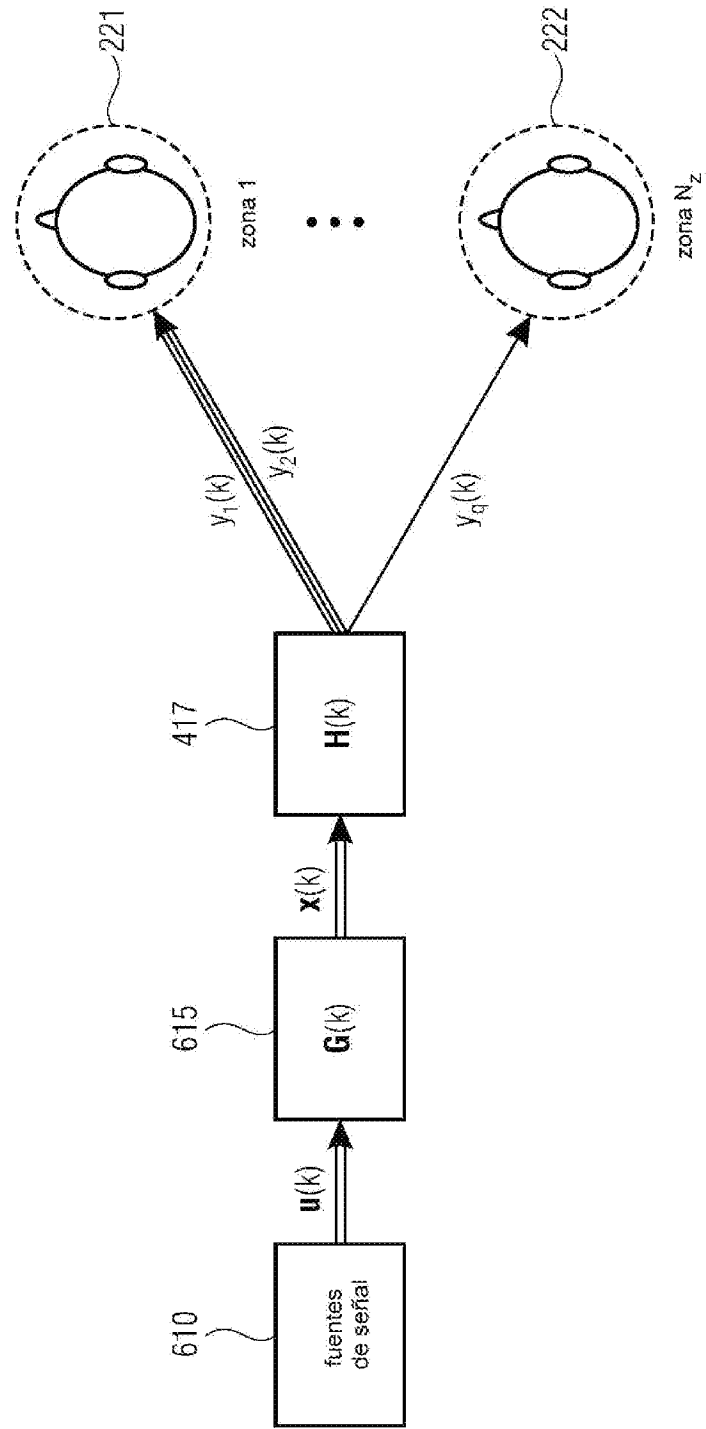


Fig. 6

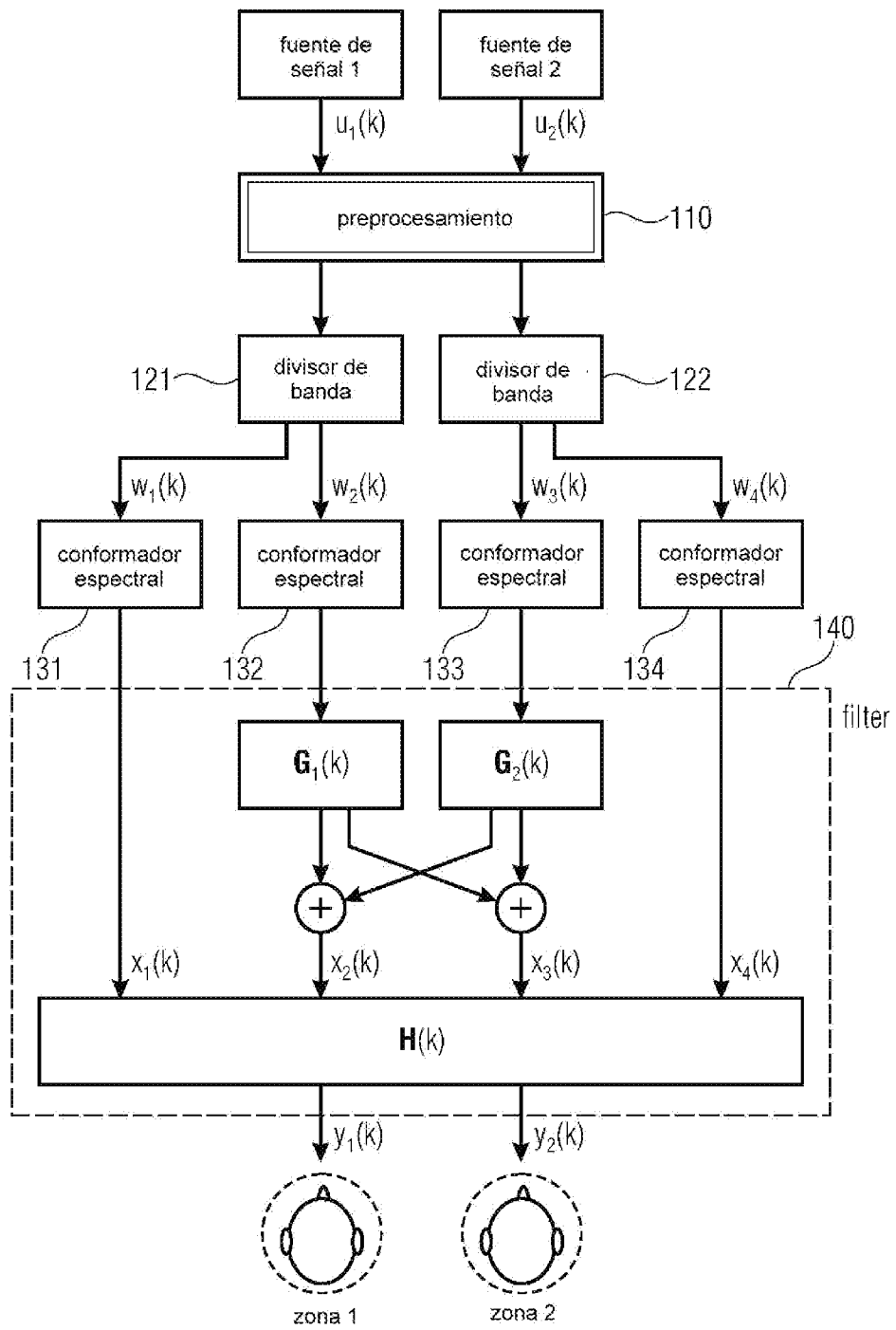


Fig. 7

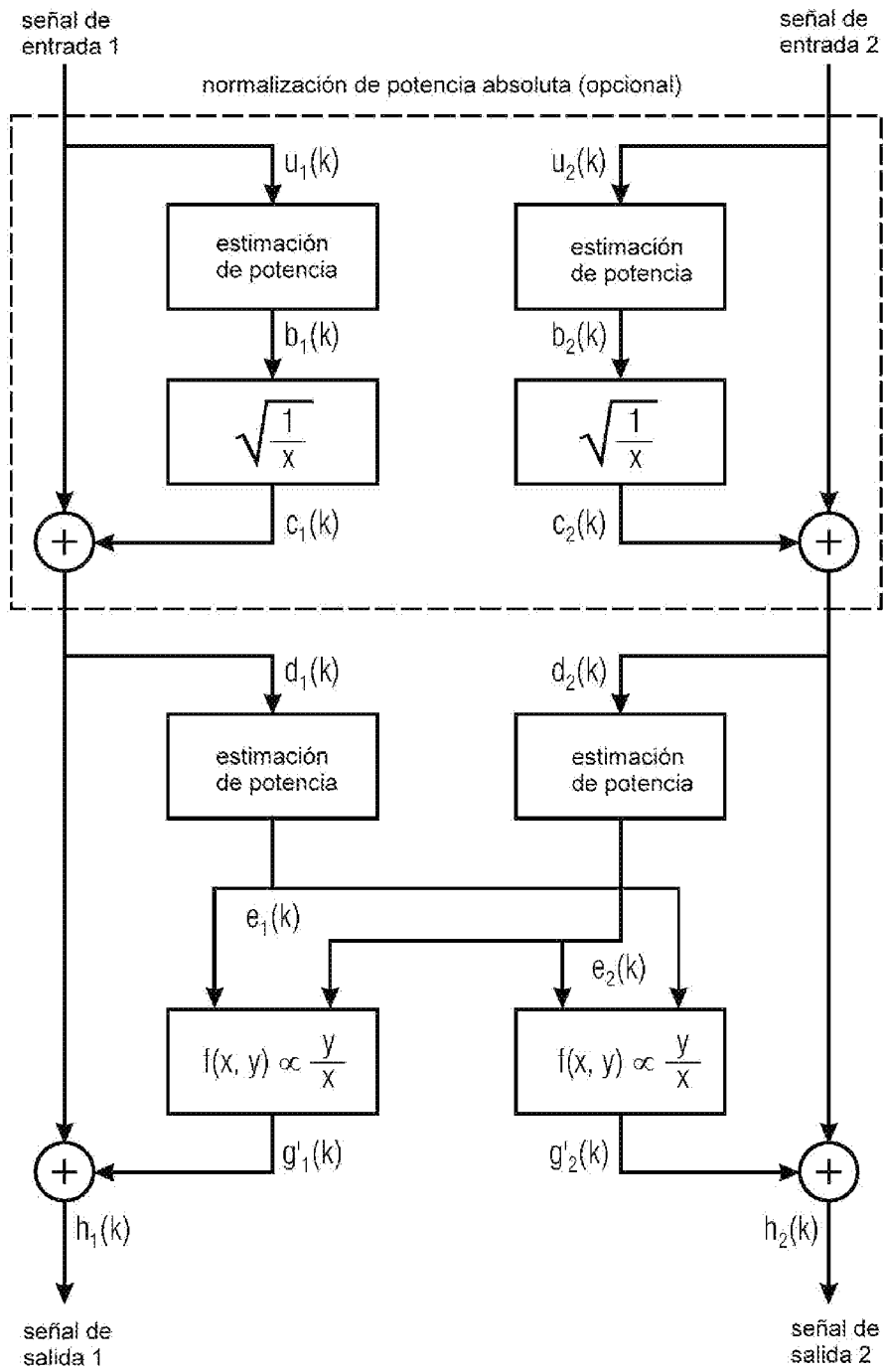


Fig. 8

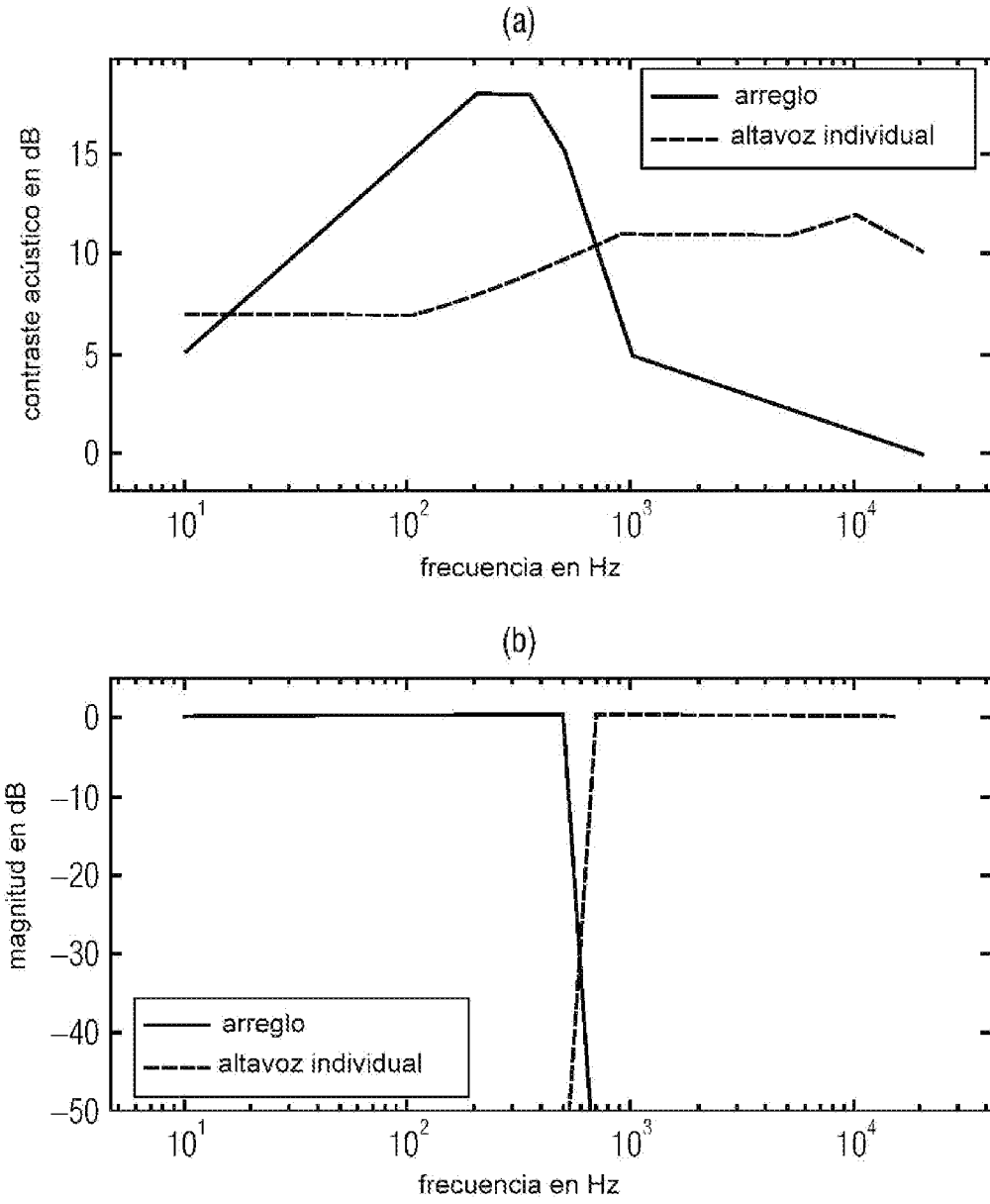


Fig. 9

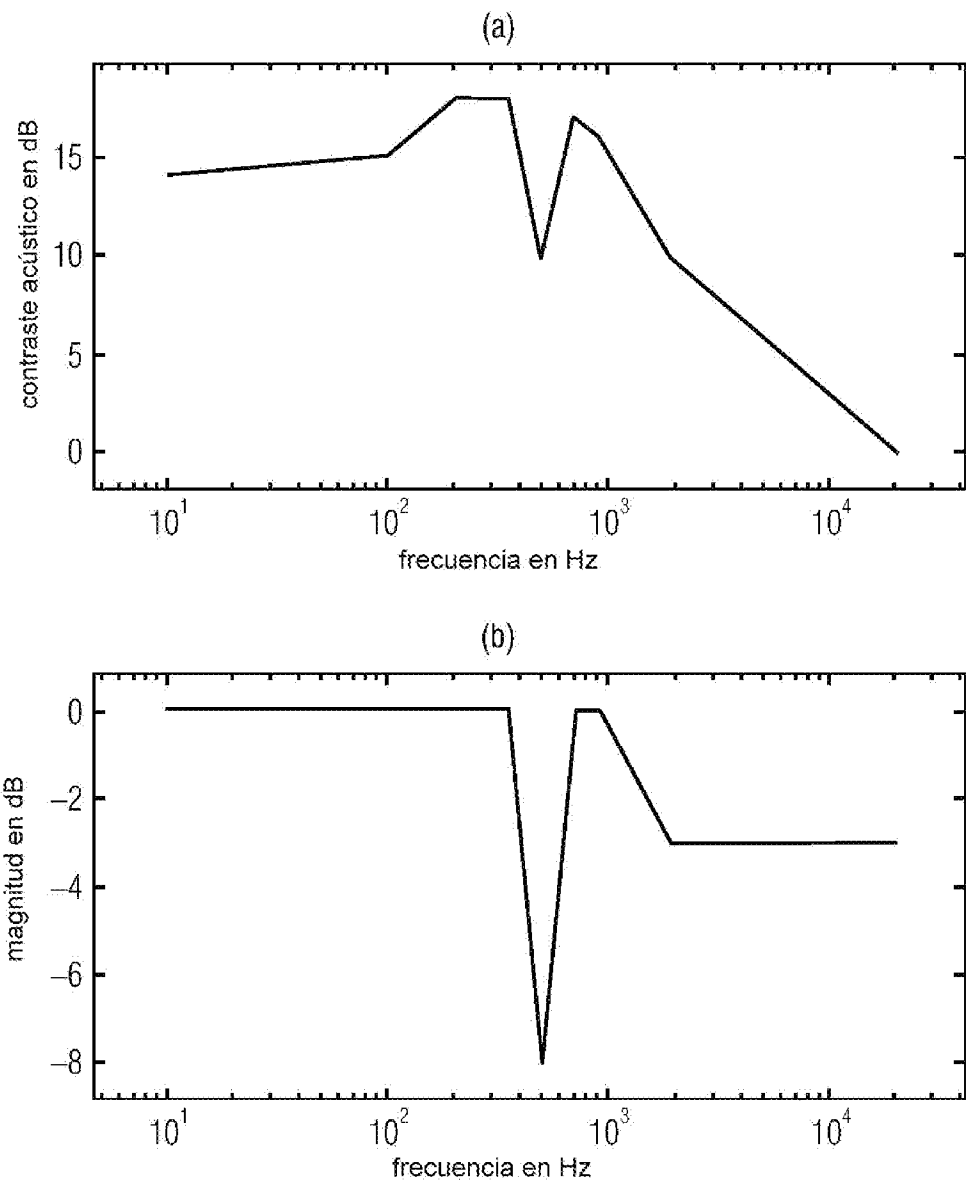


Fig. 10

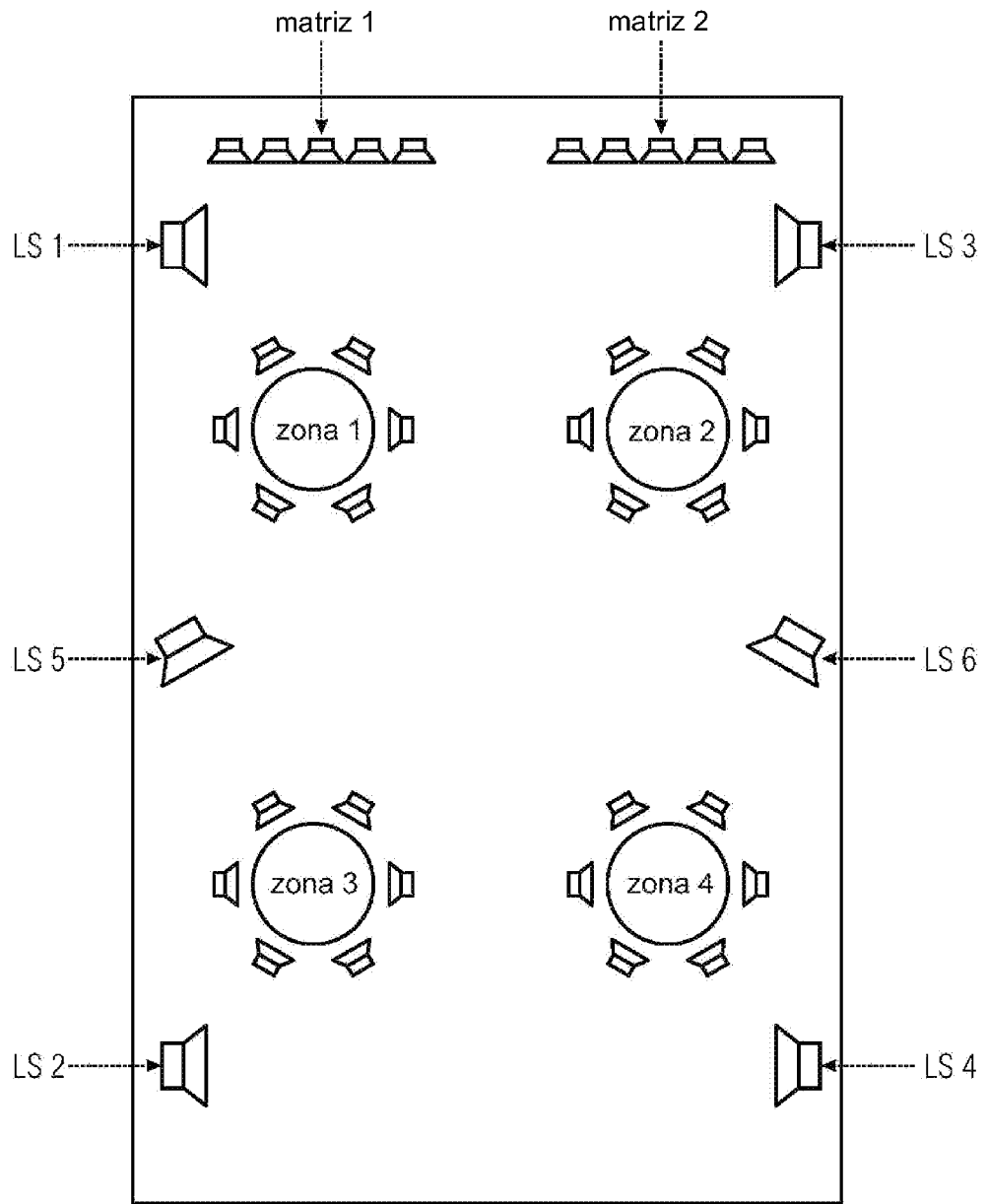


Fig. 11