



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 359 752**

51 Int. Cl.:
H04S 3/02 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **07838488 .0**

96 Fecha de presentación : **19.09.2007**

97 Número de publicación de la solicitud: **2070390**

97 Fecha de publicación de la solicitud: **17.06.2009**

54 Título: **Resolución espacial mejorada del campo sonoro para sistemas de reproducción de audio multicanal mediante derivación de señales con términos angulares de orden superior.**

30 Prioridad: **25.09.2006 US 847322 P**

45 Fecha de publicación de la mención BOPI:
26.05.2011

45 Fecha de la publicación del folleto de la patente:
26.05.2011

73 Titular/es: **DOLBY LABORATORIES LICENSING CORPORATION**
100 Potrero Avenue
San Francisco, California 94103-4813, US

72 Inventor/es: **McGrath, David Stanley**

74 Agente: **Elzaburu Márquez, Alberto**

ES 2 359 752 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

CAMPO TÉCNICO

5 La presente invención incumbe generalmente a audio e incumbe más específicamente a dispositivos y técnicas que pueden ser usados para mejorar la resolución espacial percibida de una reproducción de una señal de audio de resolución espacial baja mediante un sistema de reproducción de audio multicanal.

TÉCNICA ANTECEDENTE

10 Los sistemas de reproducción de audio multicanal ofrecen el potencial para recrear exactamente la sensación auditiva de un suceso acústico tal como una interpretación musical o un acontecimiento deportivo aprovechando las capacidades de altavoces múltiples que rodean a un oyente. Idealmente, el sistema de reproducción genera un campo sonoro multidimensional que recrea la sensación de dirección aparente de sonidos así como reverberación difusa que se espera acompañe a tal suceso acústico.

15 En un acontecimiento deportivo, por ejemplo, un espectador espera normalmente que los sonidos direccionales procedentes de los jugadores en un campo de juego sean acompañados por sonidos envolventes procedentes de otros espectadores. Una recreación exacta de las sensaciones auditivas en el acontecimiento no puede ser conseguida sin este sonido envolvente. De modo similar, las sensaciones auditivas en un concierto dentro de una sala no pueden ser recreadas exactamente sin recrear los efectos de reverberación de la sala de conciertos.

20 El realismo de las sensaciones recreadas por un sistema de reproducción es afectado por la resolución espacial de la señal reproducida. La exactitud de la recreación aumenta generalmente cuando aumenta la resolución espacial. Los sistemas de reproducción de audio comerciales y para consumidores emplean frecuentemente números mayores de altavoces pero, desgraciadamente, las señales de audio que reproducen pueden tener una resolución espacial relativamente baja. Muchas señales de audio difundidas y grabadas tienen una resolución espacial menor que la que puede ser deseada. Como resultado, el realismo que puede ser conseguido por un sistema de reproducción puede ser limitado por la resolución espacial de la señal de audio que ha de ser reproducida. Lo que es necesario es un modo de aumentar la resolución espacial de señales de audio.

25 Los documentos Patente de EE.UU. 5.757.927 y publicación de solicitud de patente internacional nº WO 00/19415 exponen sistemas reproductores Ambisonic que reciben señales de audio de entrada desde micrófonos de orden cero y primer orden. Aunque es conocido que la resolución espacial de un campo sonoro reproducido por estos sistemas puede ser incrementada incluyendo señales que representan el campo sonoro en función de la dirección con términos de orden superior, estos documentos no enseñan cómo derivar términos de orden segundo y superiores de estas señales de audio de entrada.

30

DESCRIPCIÓN DE LA INVENCION

Un objeto de la presente invención es proveer lo necesario para el aumento de resolución espacial de señales de audio que representan un campo sonoro multidimensional.

35 Este objeto es conseguido por la invención descrita en esta exposición. Según un aspecto de la presente invención, características estadísticas del campo sonoro, expresadas como funciones seno y coseno de primer orden de direcciones angulares de energía acústica en el campo sonoro, son derivadas analizando tres o más señales de audio de entrada que representan el campo sonoro en función de la dirección angular con términos angulares de orden cero y primer orden. Dos o más señales procesadas son derivadas de combinaciones ponderadas de las tres o más señales de audio de entrada. Las tres o más señales de audio son ponderadas en la combinación según las características estadísticas. Las dos o más señales procesadas representan el campo sonoro en función de la dirección angular con términos angulares de uno o más órdenes mayores que uno. Las tres o más señales de audio de entrada y las dos o más señales procesadas representan el campo sonoro en función de la dirección angular con términos angulares de órdenes cero, uno y mayor que uno.

40

45 Las diversas características de la presente invención y sus realizaciones preferidas pueden ser mejor comprendidas refiriéndose a la discusión siguiente y los dibujos adjuntos en los que los números de referencia iguales se refieren a elementos iguales en la diversas figuras. Los contenidos de la discusión siguiente y los dibujos son expuestos como ejemplos solamente y no debería entenderse que representan limitaciones en el alcance de la presente invención.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

50 La Figura 1 es un diagrama esquemático de un suceso acústico captado por un sistema de micrófonos y reproducido subsiguientemente por un sistema de reproducción.

La Figura 2 ilustra un oyente y el acimut aparente de un sonido.

La Figura 3 ilustra una porción de un sistema de reproducción ejemplar que distribuye señales a altavoces para recrear una sensación de dirección.

La Figura 4 es una ilustración gráfica de funciones de ganancia para los canales de dos altavoces adyacente en un sistema de reproducción hipotético.

La Figura 5 es una ilustración gráfica de funciones de ganancia que muestra una degradación en resolución espacial producida por una mezcla de señales de primer orden.

5 La Figura 6 es una ilustración gráfica de funciones de ganancia que incluyen señales de tercer orden.

Las Figuras 7A a 7D son diagramas de bloques esquemáticos de sistemas de reproducción ejemplares hipotéticos.

Las Figuras 8 y 9 son diagramas de bloques esquemáticos de un procedimiento para derivar términos de orden superior a partir de señales en formato B de tres canales (W, X, Y).

10 Las Figuras 10 a 12 son diagramas de bloques esquemáticos de circuitos que pueden ser usados para derivar características estadísticas de señales en formato B de tres canales.

La Figura 13 ilustra diagramas de bloques esquemáticos de circuitos que pueden ser usados para generar señales de segundo y tercer orden a partir de características estadísticas de señales en formato B de tres canales.

La Figura 14 es un diagrama de bloques esquemático de un sistema de micrófonos que incorpora diversos aspectos de la presente invención.

15 Las Figuras 15A y 15B son diagramas esquemáticos de disposiciones alternativas de transductores en un sistema de micrófonos.

La Figura 16 es una ilustración gráfica de funciones de ganancia hipotéticas para canales de altavoces en un sistema de reproducción.

20 La Figura 17 es un diagrama de bloques esquemático de un dispositivo que puede ser usado para implementar diversos aspectos de la presente invención.

MODOS DE REALIZAR LA INVENCION

A. Introducción

25 La Figura 1 proporciona una ilustración esquemática de un suceso acústico 10 y un descodificador 17 que incorpora aspectos de la presente invención, que recibe señales de audio 18 que representan sonidos del suceso acústico captados por el sistema 15 de micrófonos. El descodificador 17 procesa las señales recibidas para generar señales procesadas con resolución espacial aumentada. Las señales procesadas son reproducidas por un sistema que incluye un conjunto de altavoces 19 dispuestos próximos a uno o más oyentes 12 para proporcionar una recreación exacta de las sensaciones auditivas que podrían haber sido experimentadas en el suceso acústico. El sistema 15 de micrófonos capta tanto ondas directas 13 de sonido como ondas indirectas 14 de sonido que llegan después de la reflexión en una o más superficies en algún entorno acústico 16 tal como una habitación o una sala de conciertos.

30 En una implementación, el sistema 15 de micrófonos proporciona señales de audio que se ajustan al formato Ambisonic de señales de cuatro canales (W, X, Y, Z) conocido como formato B. El sistema de micrófonos SPS422B y el sistema de micrófonos MKV obtenible de SoundField Ltd., Wakefield, Inglaterra, son dos ejemplos que pueden ser usados. Detalles de implementación que usa sistemas de micrófonos SoundField son tratados después. Otros sistemas de micrófonos y formatos de señales pueden ser usados si se desea sin apartarse del alcance de la presente invención.

35 Las señales en formato B de cuatro canales (W, X, Y, Z) pueden ser obtenidas desde un conjunto de cuatro transductores acústicos coincidentes. Conceptualmente, un transductor es omnidireccional y tres transductores tienen modelos en forma de dipolo mutuamente ortogonales de sensibilidad direccional. Muchos sistemas de micrófonos en formato B están contruidos a partir de un conjunto tetraédrico de cuatro transductores acústicos direccionales y un procesador de señales que genera las señales en formato B de cuatro canales en respuesta a la salida de los cuatro transductores. La señal de canal W representa una onda sonora omnidireccional y la señales de canales X, Y y Z representan ondas sonoras orientadas a lo largo de tres ejes mutuamente ortogonales que son expresadas típicamente como funciones de la dirección angular con términos angulares θ de primer orden. El eje X está alineado horizontalmente desde atrás hacia delante con respecto a un oyente, el eje Y está alineado horizontalmente desde la derecha la izquierda con respecto al oyente y el eje Z está alineado verticalmente hacia arriba con respecto al oyente. Los ejes X e Y son ilustrados en la Figura 2. La Figura 2 también ilustra el azimut aparente θ de un sonido, que puede ser expresado como un vector (x, y). Obligando a que el vector tenga longitud unitaria, puede verse que:

$$x^2 + y^2 = 1 \tag{1}$$

$$(x, y) = (\cos \theta, \sin \theta) \tag{2}$$

Las señales en formato B de cuatro canales pueden transportar información tridimensional sobre un campo sonoro. Las aplicaciones que requieren solo información bidimensional sobre un campo acústico pueden usar una señal en formato B de tres canales (W, X, Y) que omita el canal Z. Diversos aspectos de la presente invención pueden ser aplicados a sistemas de reproducción bidimensionales y tridimensionales pero la exposición restante hace mención más particular de aplicaciones bidimensionales.

B. Reparto panorámico (panning) de señales

La Figura 3 ilustra una porción de un sistema de reproducción ejemplar con ocho altavoces que rodean al oyente 12. La figura ilustra un estado en el que el sistema está generando un campo sonoro en respuesta a dos señales P y Q de entrada que representan dos sonidos con direcciones aparentes P' y Q', respectivamente. El componente 33 de reparto panorámico (panner) procesa las señales P y Q de entrada para distribuir o repartir panorámicamente (pan) las señales procesadas entre los canales de altavoces para recrear la sensación de dirección. El componente 33 de repartidor panorámico (panner) puede usar un número de procesos. Un proceso que puede ser usado es conocido como el Reparto de Amplitudes de Altavoces Más Próximos (NSAP: Nearest Speaker Amplitude Pan).

El proceso de NSAP distribuye señales a los canales de altavoces adaptando la ganancia para cada canal de altavoz en respuesta a la dirección aparente de un sonido y las ubicaciones de los altavoces con respecto a un oyente o área de audición. En un sistema bidimensional, por ejemplo, la ganancia para la señal P es obtenida de una función del acimut θ_P de la dirección aparente para el sonido que esta señal representa y de los acimuts θ_F y θ_E , de los dos altavoces SF y SE, respectivamente, que están situados en ambos lados de la dirección aparente θ_P . En una implementación, las ganancias para todos los canales de altavoces distintos que los canales para estos dos altavoces más próximos son dispuestas en cero y las ganancias para los canales de los dos altavoces más próximos son calculados según las ecuaciones siguientes:

$$Ganancia_{SE}(\theta_P) = \frac{|\theta_P - \theta_F|}{|\theta_E - \theta_F|} \quad (3a)$$

$$Ganancia_{SF}(\theta_P) = \frac{|\theta_P - \theta_E|}{|\theta_E - \theta_F|} \quad (3b)$$

Cálculos similares son usados para obtener las ganancias para otras señales. La señal Q representa un caso especial donde la dirección aparente θ_Q del sonido que representa está alineada con un altavoz SC. Cualquier altavoz SB o SD puede ser seleccionado como el segundo altavoz más próximo. Como puede verse por las ecuaciones 1a y 1b, la ganancia para el canal del altavoz SC es igual a uno y las ganancias para todos los demás canales de altavoces son cero.

Las ganancias para los canales de altavoces pueden ser trazadas en función del acimut. El gráfico mostrado en la Figura 4 ilustra las funciones de ganancia para los canales de los altavoces S_E y S_F en el sistema mostrado en la Figura 3 donde los altavoces S_E y S_F están separados entre sí y de sus vecinos inmediatos para un ángulo igual a 45 grados. El acimut es expresado en términos del sistema de coordenadas mostrado en la Figura 2. Cuando un sonido tal como el representado por la señal P tiene una dirección aparente entre 135 grados y 180 grados, las ganancias para los altavoces S_E y S_F estarán entre cero y uno y las ganancias para todos los demás altavoces en el sistema serán dispuestas en cero.

C. Modelos o patrones de ganancias de micrófonos

Los sistemas pueden aplicar el proceso de NSAP a señales que representan sonidos con direcciones discretas para generar campos sonoros que son capaces de recrear exactamente las sensaciones auditivas de un suceso acústico original. Desgraciadamente, los sistemas de micrófonos no proporcionan señales que representan sonidos con direcciones discretas.

Cuando un suceso acústico 10 es captado por el sistema 15 de micrófonos, las ondas sonoras 13, 14 llegan típicamente al sistema de micrófonos desde un gran número de direcciones diferentes. Los sistemas de micrófonos de Sound Field Ltd. mencionado anteriormente generan señales que se ajustan al formato B. Señales en formato B de cuatro canales (W, X, Y, Z) pueden ser generadas para transportar características tridimensionales de un campo sonoro expresadas como funciones de la dirección angular. Ignorando la señal de canal Z, señales en formato B de tres canales (W, X, Y) pueden ser obtenidas para representar características bidimensionales de un campo sonoro que también son expresadas como funciones de la dirección angular. Lo que se precisa es un modo de procesar estas señales de modo que las sensaciones auditivas puedan ser recreadas con una exactitud espacial similar a la que puede ser conseguida por el proceso de NSAP cuando es aplicado a señales que representan sonidos con direcciones discretas. La capacidad de conseguir este grado de exactitud espacial es dificultada por la resolución espacial de las señales que son provistas por el sistema 15 de micrófonos.

La resolución espacial de una señal obtenida de un sistema de micrófonos depende de con cuánta exactitud el modelo direccional real de sensibilidad para el sistema de micrófonos se ajusta a algún modelo ideal, lo que a su vez depende del modelo direccional real de sensibilidad para los transductores acústicos individuales dentro del sistema de micrófonos. El modelo direccional de sensibilidad para transductores reales puede apartarse significativamente de algún modelo ideal pero el procesamiento de señales puede compensar estas desviaciones respecto a los modelos ideales. El procesamiento de señales también puede convertir las señales de salida de transductores en un formato deseado tal como el formato B. El modelo direccional eficaz que incluye el formato de señales del sistema de transductor/procesador es el resultado combinado de sensibilidad direccional de transductor y procesamiento de señales. Los sistemas de micrófonos de SoundField Ltd. mencionados anteriormente son ejemplos de este procedimiento. Este detalle de implementación no es crítico para la presente invención porque no es importante como es conseguido el modelo direccional eficaz. En el resto de esta discusión, términos como "modelo direccional" y "direccionalidad" se refieren a la sensibilidad direccional eficaz del transductor o combinación de transductor/procesador usada para captar un campo sonoro.

Un modelo direccional bidimensional de sensibilidad para un transductor puede ser descrito como un modelo de ganancia que es una función de la dirección angular θ , que puede tener una forma que puede ser expresada por cualquiera de las ecuaciones siguientes:

$$\text{Ganancia}(a, \theta) = (1 - a) + a \cdot \cos \theta \quad (4a)$$

$$\text{Ganancia}(a, \theta) = (1 - a) + a \cdot \text{sen} \theta \quad (4b)$$

donde $a=0$ para un modelo de ganancia omnidireccional;

$a=0,5$ para un modelo de ganancia cardioide; y

$a=1$ para un modelo de ganancia en figura 8.

Estos modelos son expresados como funciones de la dirección angular con términos angulares θ de primer orden y son citados en esto como modelos de ganancia de primer orden.

En implementaciones típicas, el sistema 15 de micrófonos usa tres o cuatro transductores con modelos de ganancia de primer orden para proporcionar señales en formato B de tres canales (W, X, Y) cuatro canales (W, X, Y, Z) que transportan información bidimensional o tridimensional sobre un campo sonoro. Refiriéndose a las ecuaciones 4a y 4b, un modelo de ganancia para cada uno de los tres canales (W, X, Y) de señales en formato B puede ser expresado como:

$$\text{Ganancia}_w(\theta) = \text{Ganancia}(a = 0, \theta) = 1 \quad (5a)$$

$$\text{Ganancia}_x(\theta) = \text{Ganancia}(a = 1, \theta) = \cos \theta = x \quad (5b)$$

$$\text{Ganancia}_y(\theta) = \text{Ganancia}(a = 1, \theta) = \text{sen} \theta = y \quad (5c)$$

donde el canal W tiene un modelo de ganancia de orden cero omnidireccional como es indicado por $a=0$ y los canales X e Y tienen un modelo de ganancia de primer orden en figura 8 como es indicado por $a=1$.

D. Resolución de sistema de reproducción

El número y la colocación de altavoces en un conjunto de reproducción puede influir en la resolución espacial percibida de un campo sonoro recreado. Un sistema con ocho altavoces igualmente separados es tratado e ilustrado aquí pero esta disposición es un ejemplo simplemente. Al menos tres altavoces son necesarios para recrear un campo sonoro que rodea a un oyente pero cinco o más altavoces son generalmente preferidos. En implementaciones preferidas de un sistema de reproducción, el descodificador 17 genera una señal de salida para cada altavoz que es descorrelacionada lo más posible respecto a otras señales de salida. Niveles más altos de descorrelación tienden a estabilizar la dirección percibida de un sonido dentro de un área mayor de audición, evitando problemas de localización bien conocidas para oyentes que están situados fuera del denominado lugar ideal ("sweet spot").

En una implementación de un sistema de reproducción según la presente invención, el descodificador 17 procesa señales en formato B de tres canales (W, X, Y) que representan un campo sonoro en función de la dirección solo con términos angulares de orden cero y primer orden para derivar señales procesadas que representan el campo sonoro en función de la dirección con términos angulares de orden superior que son distribuidos a uno o más altavoces. En sistemas convencionales, el descodificador 17 mezcla señales procedentes de cada uno de los tres canales en formato B en una señal procesada respectiva para cada uno de los altavoces usando factores de ganancia que son seleccionados basados en ubicaciones de altavoces. Desgraciadamente, este tipo de proceso de mezcla no proporciona

una resolución espacial tan grande como las funciones de ganancia usadas en el proceso de NSAP para sistemas típicos como se describió anteriormente. Por ejemplo, el gráfico ilustrado en la Figura 5 muestra una degradación en resolución espacial para las funciones de ganancia que resultan de una mezcla lineal de señales en formato B de primer orden.

- 5 La causa de esta degradación en resolución espacial puede ser explicada observando que el acimut preciso θ_P de un sonido P con amplitud R no es medido por el sistema 15 de micrófonos. En cambio, el sistema 15 de micrófonos registra tres señales $W=R$, $S=R \cdot \cos \theta_P$ e $Y=R \cdot \sin \theta_P$ que representan un campo sonoro en función de la dirección con términos angulares de orden cero y primer orden. Por ejemplo, la señal procesada generada para el altavoz SE está compuesta por una combinación lineal de la señales de canales W , X e Y .
- 10 La curva de ganancia para este proceso de mezcla puede ser considerada como una aproximación de Fourier de orden reducido a la función deseada de ganancia de NSAP. La función de ganancia de NSAP para el canal de altavoz SE mostrado en la Figura 4, por ejemplo, puede ser representada por una serie de Fourier

$$\text{Ganancia}_{SE}(\theta) = a_0 + a_1 \cos \theta + b_1 \sin \theta + a_2 \cos 2\theta + b_2 \sin 2\theta + a_3 \cos 3\theta + b_3 \sin 3\theta + \dots \quad (6)$$

- 15 pero el proceso de mezcla de un descodificador típico omite términos por encima del primer orden, lo que puede ser expresado como:

$$\text{Ganancia}_{SR}(\theta) \approx a_0 + a_1 \cos \theta + b_1 \sin \theta \quad (7)$$

La resolución espacial de la función de procesamiento para el descodificador 17 puede ser aumentada incluyendo señales que representan un campo sonoro en función de la dirección con términos de orden superior. Por ejemplo, una función de ganancia para el canal de altavoz SE que incluye términos hasta el tercer orden puede ser expresada como:

20
$$\text{Ganancia}_{SE}(\theta) = a_0 + a_1 \cos \theta + b_1 \sin \theta + a_2 \cos 2\theta + b_2 \sin 2\theta + a_3 \cos 3\theta + b_3 \sin 3\theta \quad (8)$$

Una función de ganancia que incluye términos de tercer orden puede proporcionar una aproximación más exacta a la curva deseada de ganancia de NSAP como se ilustra en la Figura 6.

- 25 Términos angulares de segundo orden y tercer orden podrían ser obtenidos usando un sistema de micrófonos que capta componentes de campo sonoro de segundo orden y tercer orden pero esto exigiría transductores acústicos con modelos direccionales de segundo orden y tercer orden de sensibilidad. Los transductores con sensibilidades direccionales de orden superior son muy difíciles de fabricar. Además, este procedimiento no proporcionaría ninguna solución para la reproducción de señales que fueron registradas usando transductores con modelos direccionales de primer orden de sensibilidad.

- 30 Los diagramas de bloques esquemáticos mostrados en las Figuras 7A a 7D ilustran sistemas de reproducción hipotéticos diferentes que pueden ser usados para generar un campo sonoro multidimensional en respuesta a tipos diferentes de señales de entrada. El sistema de reproducción ilustrado en la Figura 7A excita ocho altavoces en respuesta a ocho señales de entrada discretas. Los sistemas de reproducción ilustrados en las Figuras 7B y 7C excitan ocho altavoces en respuesta a señales de entrada en formato B de primer orden y tercer orden, respectivamente, usando un descodificador 17 que realiza un proceso de descodificación que es apropiado para el formato de las señales de entrada. El sistema de reproducción ilustrado en la Figura 7D incorpora diversas características de la presente invención, en el que el descodificador 17 procesa señales de orden cero y primer orden en formato B de tres canales (W, X, Y) para derivar señales procesadas que se aproximan a la señales que podrían haber sido obtenidas de un sistema de micrófonos que usa transductores con modelos de ganancia de segundo orden y tercer orden. La discusión siguiente describe métodos diferentes que pueden ser usados para derivar estas señales procesadas.

40 **E. Derivar términos de orden superior**

Dos procedimientos básicos para derivar términos angulares de orden superior son descritos a continuación. El primer procedimiento deriva los términos angulares para señales de banda ancha. El segundo procedimiento es una variación del primer procedimiento que deriva los términos angulares para subbandas de frecuencias. Las técnicas pueden ser usadas para generar señales con componentes de orden superior. Además, estas técnicas pueden ser aplicadas a la señales en formato B de cuatro canales para aplicaciones tridimensionales.

45 **1. Procedimiento de banda ancha**

La Figura 8 es un diagrama de bloques esquemático de un procedimiento de banda ancha para derivar términos de orden superior a partir de señales en formato B de tres canales (W, X, Y). Cuatro características estadísticas indicadas como

- 50 C_1 = una estimación de $\cos \theta(t)$;

S_1 = una estimación de $\sin \theta(t)$;

C_2 = una estimación de $\cos 2\theta(t)$; y

S_2 = una estimación de $\sin 2\theta(t)$;

5 son derivadas de un análisis de las señales en formato B y estas características son usadas para generar estimaciones de los términos de segundo orden y tercer orden, que son indicados como:

$$X_2 = \text{Señal} \cdot \cos 2\theta(t)$$

$$Y_2 = \text{Señal} \cdot \sin 2\theta(t)$$

$$X_3 = \text{Señal} \cdot \cos 3\theta(t)$$

$$Y_3 = \text{Señal} \cdot \sin 3\theta(t)$$

10 Una técnica para obtener las cuatro características estadísticas supone que en cualquier instante particular t , la mayor parte de la energía acústica incidente en el sistema 15 de micrófonos llega desde una sola dirección angular, lo que hace el acimut una función de tiempo que puede ser designada como $\theta(t)$. Como resultado, se supone que las señales de canales W, X e Y son esencialmente de la forma:

$$W = \text{Señal}$$

$$X = \text{Señal} \cdot \cos \theta(t)$$

$$Y = \text{Señal} \cdot \sin \theta(t)$$

20 Estimaciones de las cuatro características estadísticas de direcciones angulares de la energía acústica pueden ser derivadas de las ecuaciones 9a a 9d mostradas a continuación, en las que la notación $Av(x)$ representa un valor medio de la señal x . Este valor medio puede ser calculado en un período de tiempo que es relativamente breve en comparación con el intervalo en el que las características de señal cambian significativamente.

$$C_1 = \frac{2Av(W \times X)}{Av(W^2) + Av(X^2) + Av(Y^2)} \quad (9a)$$

$$= \frac{2Av(\text{Señal} \cdot \text{Señal} \cdot \cos \theta)}{Av(\text{Señal}^2 + \text{Señal}^2 \cdot \cos^2 \theta + \text{Señal}^2 \cdot \sin^2 \theta)} = \cos \theta$$

$$S_1 = \frac{2Av(W \times Y)}{Av(W^2) + Av(X^2) + Av(Y^2)} \quad (9b)$$

$$= \frac{2Av(\text{Señal} \cdot \text{Señal} \cdot \sin \theta)}{Av(\text{Señal}^2 + \text{Señal}^2 \cdot \cos^2 \theta + \text{Señal}^2 \cdot \sin^2 \theta)} = \sin \theta$$

$$C_2 = \frac{2Av(X^2) - 2Av(Y^2)}{Av(W^2) + Av(X^2) + Av(Y^2)} \quad (9c)$$

$$= \frac{2Av(\text{Señal}^2 \cdot \cos^2 \theta - \text{Señal}^2 \cdot \sin^2 \theta)}{Av(\text{Señal}^2 + \text{Señal}^2 \cdot \cos^2 \theta + \text{Señal}^2 \cdot \sin^2 \theta)} = \cos^2 \theta - \sin^2 \theta = \cos 2\theta$$

$$S_2 = \frac{4Av(X \times Y)}{Av(W^2) + Av(X^2) + Av(Y^2)} \quad (9d)$$

$$= \frac{4Av(\text{Señal}^2 \cdot \cos\theta \cdot \text{sen}\theta)}{Av(\text{Señal}^2 + \text{Señal}^2 \cdot \cos^2\theta + \text{Señal}^2 \cdot \text{sen}^2\theta)} = 2 \cos\theta \cdot \text{sen}\theta = \text{sen}2\theta$$

Otras técnicas pueden ser usadas para obtener estimaciones de las cuatro características estadísticas S_1 , C_1 , S_2 , C_2 como se trata después.

5 Las cuatro señales X_2 , Y_2 , X_3 , Y_3 mencionados anteriormente pueden ser generadas a partir de combinaciones ponderadas de las señales de canales W, X e Y que usan las cuatro características estadísticas como ponderaciones en cualquiera de varios modos usando las identidades trigonométricas siguientes:

$$\cos 2\theta \equiv \cos^2 \theta - \text{sen}^2 \theta$$

$$\text{sen } 2\theta \equiv 2 \cos \theta \cdot \text{sen} \theta$$

$$\cos 3\theta \equiv \cos \theta \cdot \cos 2\theta - \text{sen} \theta \cdot \text{sen} 2\theta$$

$$\text{sen } 3\theta \equiv \cos \theta \cdot \text{sen} 2\theta + \text{sen} \theta \cdot \cos 2\theta$$

10

La señal X_2 puede ser obtenida a partir de cualquiera de las combinaciones ponderadas siguientes:

$$X_2 = \text{Señal} \cdot \cos 2\theta = W \cdot C_2 \quad (10a)$$

$$X_2 = \text{Señal} \cdot \cos 2\theta = \text{Señal} \cdot (\cos^2 \theta - \text{sen}^2 \theta) = X \cdot C_1 - Y \cdot S_1 \quad (10b)$$

$$X_2 = \frac{1}{2}(W \cdot C_2 + X \cdot C_1 - Y \cdot S_1) \quad (10c)$$

15 El valor calculado en la ecuación 10c es una media de las dos primeras expresiones. La señal Y_2 puede ser obtenida a partir de cualquiera de las combinaciones ponderadas siguientes:

$$Y_2 = \text{Señal} \cdot \text{sen } 2\theta = W \cdot S_2 \quad (11a)$$

$$Y_2 = \text{Señal} \cdot \text{sen } 2\theta = \text{Señal} \cdot (2 \cos \theta \cdot \text{sen} \theta) = X \cdot S_1 + Y \cdot C_1 \quad (11b)$$

$$Y_2 = \frac{1}{2}(W \cdot S_2 + X \cdot S_1 + Y \cdot C_1) \quad (11c)$$

20 El valor calculado en la ecuación 11c es una media de las dos primeras expresiones. Las señales de tercer orden pueden ser obtenidas a partir de la combinaciones ponderadas siguientes:

$$X_3 = \text{Señal} \cdot \cos 3\theta = X \cdot C_2 - Y \cdot S_2 \quad (12)$$

$$Y_3 = \text{Señal} \cdot \text{sen } 3\theta = X \cdot S_2 + Y \cdot C_2 \quad (13)$$

25 Otras combinaciones ponderadas pueden ser usadas para calcular las cuatro señales X_2 , Y_2 , X_3 , Y_3 . Las ecuaciones mostradas anteriormente son simplemente ejemplos de cálculos que pueden ser usados.

Otras técnicas pueden ser usadas para derivar las cuatro características estadísticas. Por ejemplo, si están disponibles recursos suficientes de procesamiento, puede ser práctico obtener C_1 de la ecuación siguiente:

$$C_1(n) = \frac{2 \sum_{k=0}^{K-1} W(n-k) \cdot X(n-k)}{\sum_{k=0}^{K-1} (W(n-k)^2 + X(n-k)^2 + Y(n-k)^2)} \quad (14a)$$

Esta ecuación calcula el valor de C_1 en la muestra n analizando las señales de canales W , X e Y sobre las K muestras anteriores.

5 Otra técnica que puede ser usada para obtener C_1 es un cálculo que usa un filtro de aplanamiento recurrente de primer orden en lugar de las sumas finitas en la ecuación 14a, como se muestra en la ecuación siguiente:

$$C_1(n) = \frac{2 \sum_{k=0}^{\infty} W(n-k) \cdot X(n-k) \cdot (1-\alpha)^k}{\sum_{k=0}^{\infty} (W(n-k)^2 + X(n-k)^2 + Y(n-k)^2) \cdot (1-\alpha)^k} \quad (14b)$$

10 La constante de tiempo del filtro de aplanamiento es determinada por el factor α . Este cálculo puede ser realizado como se muestra en el esquema de bloques ilustrado en la Figura 10. Los errores de dividir por cero que ocurrirían cuando el denominador de la expresión en la ecuación 14b es igual a cero pueden ser evitados sumando un valor pequeño ϵ al denominador como se muestra en la figura. Esto modifica la ecuación ligeramente como sigue:

$$C_1(n) = \frac{2 \sum_{k=0}^{\infty} W(n-k) \cdot X(n-k) \cdot (1-\alpha)^k}{\sum_{k=0}^{\infty} (W(n-k)^2 + X(n-k)^2 + Y(n-k)^2 + \epsilon) \cdot (1-\alpha)^k} \quad (14c)$$

El error de dividir por cero también puede ser evitado usando un bucle de realimentación como se muestra en la Figura 11. Esta técnica usa la estimación anterior $C_1(n-1)$ para calcular la función de error siguiente:

$$Err(n) = 2 W(n) \cdot X(n) - C_1(n-1) \cdot (W(n)^2 + X(n)^2 + Y(n)^2 + \epsilon) \quad (15)$$

15 Si el valor de la función de error es mayor que cero, la estimación previa de C_1 es demasiado pequeña, el valor de signo ($Err(n)$) es igual a uno y la estimación es aumentada en una cantidad de ajuste igual a α_1 . Si el valor de la función de error es menor que cero, la estimación previa de C_1 es demasiado grande, la función signo ($Err(n)$) es igual a uno negativo y la estimación es reducida en una cantidad de ajuste igual a α_1 . Si el valor de la función de error es cero, la estimación previa de C_1 es correcta, la función signo ($Err(n)$) es igual a cero y la estimación no es cambiada. Una
20 versión aproximada de la estimación de C_1 es generada en el elemento de retardo o almacenamiento mostrado en la porción inferior izquierda del esquema de bloques ilustrado en la Figura 11, y una versión aplanada de esta estimación es generada en la salida etiquetada C_1 en la porción inferior derecha del esquema de bloques. La constante de tiempo del filtro de aplanamiento es determinada por el factor α_2 .

25 Las cuatro características estadísticas C_1 , S_1 , C_2 , S_2 pueden ser obtenidas usando circuitos y procesos correspondientes a los esquemas de bloques mostrados en la Figura 12. Las señales X_2 , Y_2 , X_3 , Y_3 con términos de orden superior pueden ser obtenidas según las ecuaciones 10c, 11c, 12 y 13 usando circuitos y procesos correspondientes a los esquemas de bloques mostrados en la Figura 13.

30 Los procesos usados para derivar las cuatro características estadísticas de las señales de entrada de los canales W , X e Y incurrirán en algún retardo si estos procesos usan técnicas de promediado en el tiempo. En un sistema en tiempo real, puede ser ventajoso añadir algún retardo a los trayectos de señales de entrada como se muestra en la Figura 9 para compensar el retardo en la derivación estadística. En muchas implementaciones, un valor típico de retardo para análisis estadístico está entre 10 ms y 50 ms. El retardo insertado en el trayecto de señal de entrada debería ser generalmente menor o igual que el retardo de análisis estadístico. En muchas implementaciones, el retardo de trayecto de señal puede ser omitido sin degradación significativa en el rendimiento funcional global del sistema.

35 2. Procedimiento multibanda

Las técnicas tratadas anteriormente derivan características estadísticas de banda ancha que pueden ser expresadas como valores escalares que varían con el tiempo pero no varían con la frecuencia. Las técnicas de derivación pueden ser ampliadas para derivar características estadísticas dependientes de la banda de frecuencias que pueden ser

expresadas como vectores con elementos correspondientes a un número de frecuencias diferentes o subbandas de frecuencias diferentes. Alternativamente, cada una de las características estadísticas C_1 , S_1 , C_2 y S_2 dependientes de la frecuencia puede ser expresada como una respuesta de impulsos.

5 Si los elementos en cada uno de los vectores C_1 , S_1 , C_2 y S_2 son tratados como valores de ganancia dependientes de la frecuencia, combinaciones ponderadas de las señales X_2 , Y_2 , X_3 e Y_3 pueden ser generadas aplicando un filtro apropiado a la señales de canales W, X, e Y que tienen respuestas de frecuencia basadas en los valores de ganancia en estos vectores. Las operaciones de multiplicar mostradas en las ecuaciones y los esquemas anteriores son sustituidas por una operación de filtración tal como convolución.

10 El análisis estadístico de las señales de canales W, X e Y puede ser realizado en el dominio de frecuencia o en el dominio de tiempo. Si el análisis es realizado en el dominio de frecuencia, las señales de entrada pueden ser transformadas en un dominio de frecuencia de tiempo breve usando una transformada de Fourier de bloque o similar para generar coeficientes en el dominio de frecuencia y las cuatro características estadísticas pueden ser calculadas para cada coeficiente en el dominio de frecuencia o para grupos de coeficientes en el dominio de frecuencia que definen subbandas de frecuencias. El proceso usado para generar las señales X_2 , Y_2 , X_3 e Y_3 puede efectuar este procesamiento sobre una base de coeficiente por coeficiente o sobre una base de banda por banda.

F. Implementación en un sistema de micrófonos

Las técnicas tratadas anteriormente pueden ser incorporadas en una disposición de transductor/procesador para formar un sistema 15 de micrófonos que puede proporcionar señales de salida con exactitud espacial mejorada. En una implementación mostrada esquemáticamente en la Figura 14, el sistema 15 de micrófonos comprende tres transductores acústicos A, B, C coincidentes o casi coincidentes que tienen modelos direccionales cardioides de sensibilidad que están dispuestos en los vértices de un triángulo equilátero con cada transductor enfrentado hacia fuera alejado del centro del triángulo. Los modelos de ganancias direccionales de transductor pueden ser expresados como

$$\text{Ganancia}_A(\theta) = \frac{1}{2} + \frac{1}{2} \cos \theta \quad (16a)$$

$$\text{Ganancia}_B(\theta) = \frac{1}{2} + \frac{1}{2} \cos(\theta - 120^\circ) \quad (16b)$$

$$25 \quad \text{Ganancia}_C(\theta) = \frac{1}{2} + \frac{1}{2} \cos(\theta + 120^\circ) \quad (16c)$$

donde el transductor A mira hacia delante a lo largo del eje X, el transductor B mira hacia atrás y a la izquierda en un ángulo de 120 grados respecto al eje X y el transductor C mira hacia atrás y a la derecha en un ángulo de 120 grados respecto al eje X.

30 Las señales de salida de estos transductores pueden ser convertidas en señales en formato B de primer orden de tres canales (W, X, Y) como sigue:

$$\begin{aligned} W &= \frac{2}{3} [\text{Ganancia}_A(\theta) + \text{Ganancia}_B(\theta) + \text{Ganancia}_C(\theta)] \\ &= \frac{2}{3} \left[\frac{1}{2} + \frac{1}{2} \cos \theta + \frac{1}{2} + \frac{1}{2} \cos(\theta - 120^\circ) + \frac{1}{2} + \frac{1}{2} \cos(\theta + 120^\circ) \right] = 1 \end{aligned} \quad (17a)$$

$$\begin{aligned} X &= \frac{4}{3} \text{Ganancia}_A(\theta) - \frac{2}{3} \text{Ganancia}_B(\theta) - \frac{2}{3} \text{Ganancia}_C(\theta) \\ &= \frac{4}{3} \left[\frac{1}{2} + \frac{1}{2} \cos \theta \right] - \frac{2}{3} \left[\frac{1}{2} + \frac{1}{2} \cos(\theta - 120^\circ) \right] - \frac{2}{3} \left[\frac{1}{2} + \frac{1}{2} \cos(\theta + 120^\circ) \right] = \cos \theta \end{aligned} \quad (17b)$$

$$\begin{aligned} Y &= \frac{2}{\sqrt{3}} \text{Ganancia}_B(\theta) - \frac{2}{\sqrt{3}} \text{Ganancia}_C(\theta) \\ &= \frac{2}{\sqrt{3}} \left[\frac{1}{2} + \frac{1}{2} \cos(\theta - 120^\circ) \right] - \frac{2}{\sqrt{3}} \left[\frac{1}{2} + \frac{1}{2} \cos(\theta + 120^\circ) \right] = \text{sen} \theta \end{aligned} \quad (17c)$$

35 Un mínimo de tres transductores es necesario para captar las señales en formato B de tres canales. En la práctica, cuando se usan transductores de bajo coste, puede ser preferible usar cuatro transductores. Los diagramas esquemáticos mostrados en las Figuras 15A y 15B ilustran dos disposiciones alternativas. Un conjunto de tres transductores puede ser dispuesto con los transductores enfrentados en ángulos diferentes tales como 60, -60 y 180 grados. Un conjunto de cuatro transductores puede ser dispuesto en una denominada configuración en "T" con los transductores enfrentados en 0,90, -90 y 180 grados, o dispuesto en una denominada configuración en "Cruz" con los

transductores enfrentados en 45, -45, 135 y -135 grados. Los modelos de ganancias para la configuración en "Cruz" son:

$$Ganancia_{LF}(\theta) = \frac{1}{2} + \frac{1}{2} \cos(\theta - 45^\circ) \quad (18a)$$

$$Ganancia_{RF}(\theta) = \frac{1}{2} + \frac{1}{2} \cos(\theta + 45^\circ) \quad (18b)$$

$$Ganancia_{LB}(\theta) = \frac{1}{2} + \frac{1}{2} \cos(\theta - 135^\circ) \quad (18c)$$

$$Ganancia_{RB}(\theta) = \frac{1}{2} + \frac{1}{2} \cos(\theta + 135^\circ) \quad (18d)$$

donde los subíndices LF, RF, LB y RB designan ganancias para los transductores enfrentados en las direcciones izquierda hacia delante, derecha hacia delante, izquierda hacia atrás y derecha hacia atrás.

Las señales de salida de la configuración en cruz de transductores pueden ser convertidas en las señales en formato B de primer orden de tres canales (W, X, Y) como sigue:

$$X = \frac{1}{\sqrt{2}} [Ganancia_{LF}(\theta) + Ganancia_{RF}(\theta) - Ganancia_{LB}(\theta) - Ganancia_{RB}(\theta)] = \cos \theta \quad (19b)$$

$$W = \frac{1}{2} [Ganancia_{LF}(\theta) + Ganancia_{RF}(\theta) + Ganancia_{LB}(\theta) + Ganancia_{RB}(\theta)] = 1 \quad (19a)$$

$$Y = \frac{1}{\sqrt{2}} [Ganancia_{LF}(\theta) - Ganancia_{RF}(\theta) + Ganancia_{LB}(\theta) - Ganancia_{RB}(\theta)] = \sin \theta \quad (19c)$$

En la práctica real, los modelos de ganancia direccionales para cada transductor se desvían del modelo cardioide ideal. Las ecuaciones de conversión mostradas anteriormente pueden ser ajustadas para tener en cuenta estas desviaciones. Además, los transductores pueden tener peor sensibilidad direccional a frecuencias más bajas; sin embargo, esta propiedad puede ser tolerada en muchas aplicaciones porque los oyentes son generalmente menos sensibles a los errores direccionales en frecuencias más bajas.

G. Ecuaciones de mezcla

El conjunto de siete señales de primero, segundo y tercer orden (W, X, Y, X₂, Y₂, X₃, Y₃) puede ser mezclado o combinado por una matriz para excitar un número deseado de altavoces. El conjunto siguiente de ecuaciones de mezcla definen una matriz de 7x5 que puede ser usada para excitar cinco altavoces en una configuración típica de sonido envolvente que incluye canales izquierdo (L), derecho (R), central (C), izquierdo envolvente (LS), y derecho envolvente (RS).

$$\begin{bmatrix} S_L \\ S_C \\ S_R \\ S_{LS} \\ S_{RS} \end{bmatrix} = \begin{bmatrix} 0,2144 & 0,1533 & 0,3498 & -0,1758 & 0,1971 & -0,1266 & -0,0310 \\ 0,1838 & 0,3378 & 0,0000 & 0,2594 & 0,0000 & 0,1598 & 0,0000 \\ 0,2144 & 0,1533 & -0,3498 & -0,1758 & -0,1971 & -0,1266 & 0,0310 \\ 0,2451 & -0,3227 & 0,2708 & 0,0448 & -0,2539 & 0,0467 & 0,0809 \\ 0,2451 & -0,3227 & -0,2708 & 0,0448 & 0,2539 & 0,0467 & -0,0809 \end{bmatrix} \cdot \begin{bmatrix} W \\ X \\ Y \\ X_2 \\ Y_2 \\ X_3 \\ Y_3 \end{bmatrix}$$

Las funciones de ganancia de altavoces que son provistas por estas ecuaciones de mezcla son ilustradas gráficamente en la Figura 16. Estas funciones de ganancia suponen que la matriz de mezcla es alimentada con un conjunto ideal de señales de entrada.

H. Implementación

Dispositivos que incorporan diversos aspectos de la presente invención pueden ser implementados de modos diversos que incluyen software para ejecución por un ordenador o algún otro dispositivo que incluye componentes más especializados tales como circuitos de procesador de señales digitales (DSP: digital signal processor) acoplados a componentes similares a los hallados en un ordenador de uso general. La Figura 17 es un diagrama de bloques esquemático de un dispositivo 70 que puede ser usado para implementar aspectos de la presente invención. El

- 5 procesador 72 proporciona recursos de cómputo. La memoria RAM 73 es la memoria de acceso aleatorio (RAM) del sistema usada por el procesador 72 para procesamiento. La memoria ROM 74 representa alguna forma de almacenamiento persistente tal como la memoria de solo lectura (ROM: read only memory) o memoria flash para almacenar programas necesarios para hacer funcionar el dispositivo 70 y posiblemente para llevar a cabo diversos aspectos de la presente invención. El control 75 de entrada/salida (I/O: input/output) representa circuitos de interfaz para recibir y transmitir señales por vía de los canales 76, 77 de comunicación. En la realización mostrada, todos los componentes principales del sistema se conectan al bus 71 que puede representar más de un bus físico o lógico; una arquitectura de bus no es necesaria para implementar la presente invención.
- 10 El dispositivo 78 de almacenamiento es opcional. Programas que implementan diversos aspectos de la presente invención pueden ser grabados en un dispositivo 78 de almacenamiento que tiene un soporte de almacenamiento tal como una cinta o disco magnético, o un soporte óptico. El soporte de almacenamiento también puede ser usado para grabar programas de instrucciones para sistemas operativos, utilidades y aplicaciones.
- 15 Las funciones necesarias para poner en práctica diversos aspectos de la presente invención pueden ser realizados por componentes que son implementados en una extensa variedad de modos que incluyen componentes lógicos discretos, circuitos integrados, uno o más circuitos integrados de aplicaciones específicas (ASICs: application-specific integrated circuits) y/o procesadores controlados por programas. La manera en la que estos componentes son implementados no es importante para la presente invención.
- 20 Implementaciones de software de la presente invención pueden ser transportadas por una variedad de soportes legibles por máquina tales como trayectos de comunicación en banda base o modulados en todo el espectro que incluye desde frecuencias supersónicas a ultravioletas, o soportes de almacenamiento que transportan información usando esencialmente cualquier tecnología de grabación que incluye cinta, tarjetas o disco magnético, tarjetas o disco óptico, y marcaciones detectables sobre soportes incluyendo papel.

REIVINDICACIONES

1. Un método para aumentar la resolución espacial de señales de audio que representan un campo sonoro, comprendiendo el método:
- 5 recibir tres o más señales de audio de entrada que representan el campo sonoro en función de la dirección angular con términos angulares de orden cero y de primer orden;
- analizar las tres o más señales de audio de entrada para derivar características estadísticas del campo sonoro expresadas como funciones seno y coseno de primer orden de direcciones angulares de energía acústica en el campo sonoro;
- 10 derivar dos o más señales procesadas de combinaciones ponderadas de las tres o más señales de audio de entrada en el que las tres o más señales de audio son ponderadas según las características estadísticas, en el que las dos o más señales procesadas representan el campo sonoro en función de la dirección angular con términos angulares de uno o más órdenes mayores que uno;
- proporcionar cinco o más señales de audio de salida que representan el campo sonoro en función de la dirección angular con términos angulares de órdenes cero, uno y mayor que uno, en el que las cinco o más señales de audio de salida comprenden las tres o más señales de audio de entrada y las dos o más señales procesadas.
- 15 2. El método según la reivindicación 1, en el que las tres o más señales de audio de entrada son recibidas desde una pluralidad de transductores acústicos, teniendo cada uno sensibilidades direccionales con términos angulares de un orden no mayor que el primer orden.
3. El método según la reivindicación 1 o 2, que deriva de las características estadísticas dos o más señales que representan el campo sonoro en función de la dirección angular con términos angulares de segundo orden.
- 20 4. El método según la reivindicación 1 o 2, que deriva de las características estadísticas cuatro o más señales procesadas que representan el campo sonoro en función de la dirección angular con términos angulares de segundo orden y tercer orden.
5. El método según la reivindicación 1 o 2, que deriva de las características estadísticas cuatro o más señales procesadas que representan el campo sonoro en función de la dirección angular con términos angulares de dos o más órdenes mayores que uno.
- 25 6. El método según una cualquiera de las reivindicaciones 1 a 5, en el que las características estadísticas son derivadas al menos en parte de medias de las tres o más señales de audio de entrada calculadas sobre intervalos de tiempo.
7. El método según una cualquiera de las reivindicaciones 1 a 5, en el que cada una de las señales de audio de entrada es representada por muestras y las características estadísticas son derivadas al menos en parte de una suma de una pluralidad de las muestras para una señal de audio de entrada respectiva.
- 30 8. El método según una cualquiera de las reivindicaciones 1 a 5, en el que las características estadísticas son derivadas al menos en parte aplicando un filtro de aplanamiento a valores derivados de las tres o más señales de audio de entrada.
9. El método según una cualquiera de las reivindicaciones 1 a 8, que deriva características estadísticas dependientes de la frecuencia para las tres o más señales de audio de entrada.
- 35 10. El método según la reivindicación 9, que comprende:
- aplicar una transformada de bloque a las tres o más señales de audio de entrada para generar coeficientes en el dominio de frecuencia; derivar las características estadísticas dependientes de la frecuencia a partir de coeficientes individuales en el dominio de la frecuencia o grupos de coeficientes en el dominio de la frecuencia, y
- 40 derivar las dos o más señales procesadas aplicando filtros a las tres o más señales de audio de entrada que tienen respuestas de frecuencia basadas en las características estadísticas dependientes de la frecuencia.
11. El método según la reivindicación 9, que comprende derivar las dos o más señales procesadas aplicando filtros a las tres o más señales de audio de entrada que tienen respuestas de impulsos basadas en las características estadísticas dependientes de la frecuencia.
- 45 12. Un aparato (70) para aumentar la resolución espacial de señales de audio que representan un campo sonoro, comprendiendo el aparato medios para realizar el método según una cualquiera de las reivindicaciones 1 a 11.
13. Un soporte (78) de almacenamiento que graba un programa de instrucciones ejecutable por un dispositivo (70), en el que la ejecución del programa de instrucciones causa que el dispositivo realice el método según una cualquiera de las reivindicaciones 1 a 11.
- 50

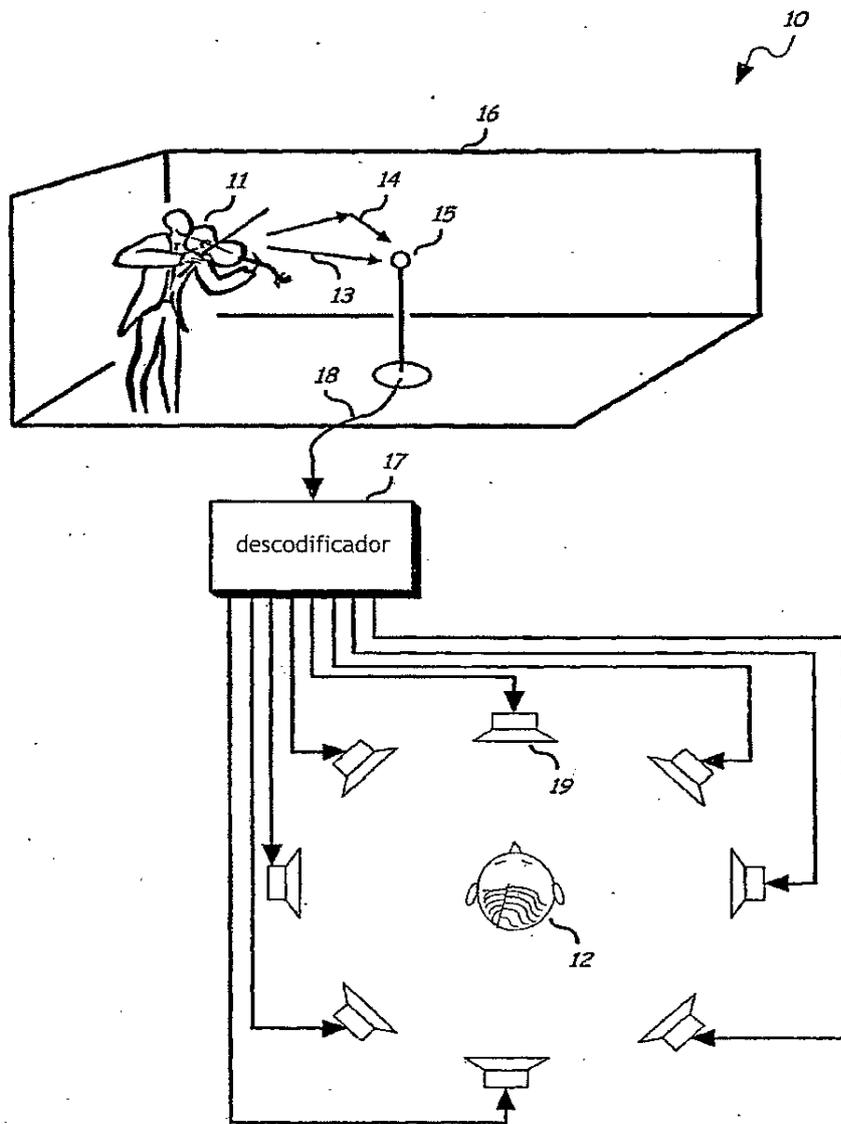


Fig. 1

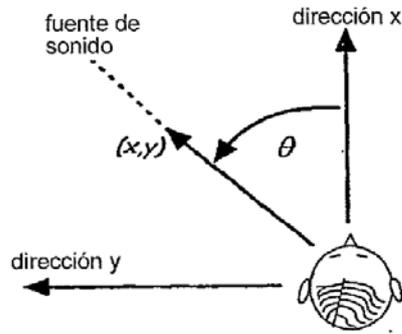


Fig. 2

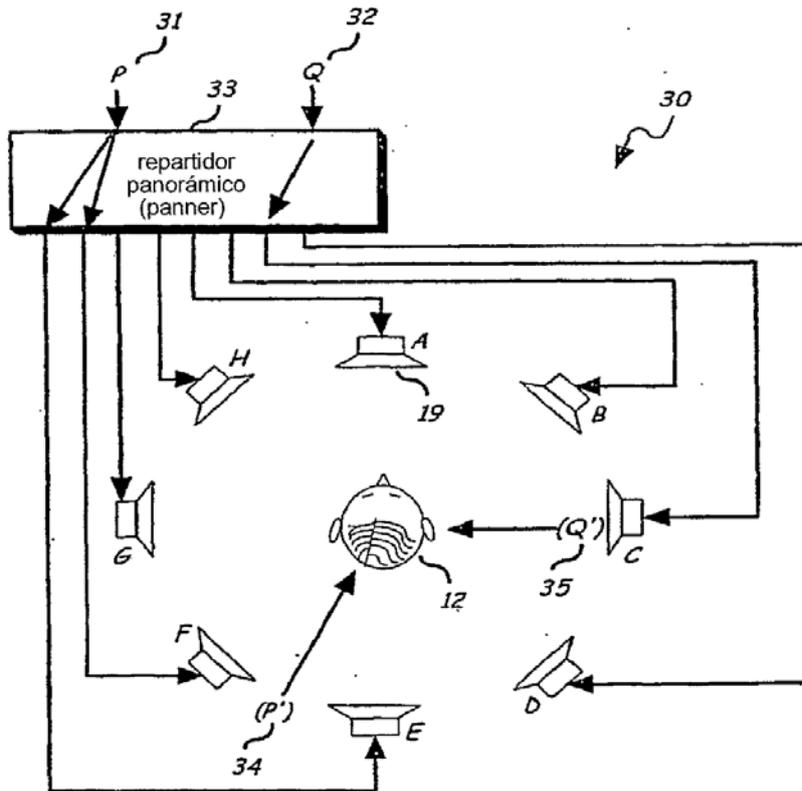


Fig. 3

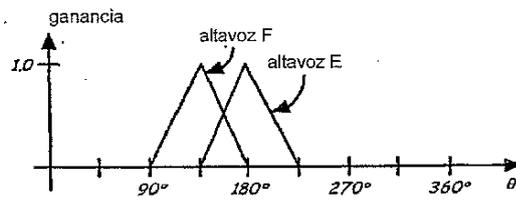


Fig. 4

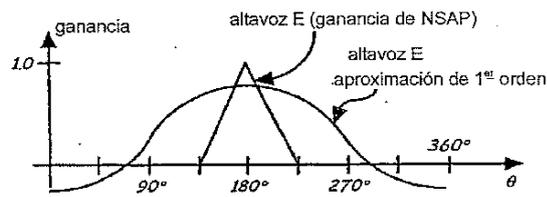


Fig. 5

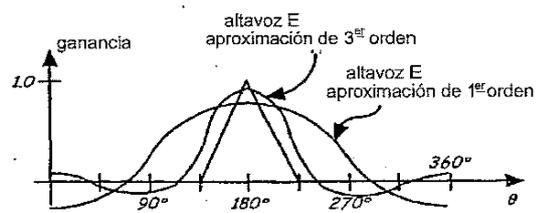


Fig. 6

sistema "ideal":
todos los 8 altavoces excitados
individualmente por la fuente

audio discreto
de 8 canales

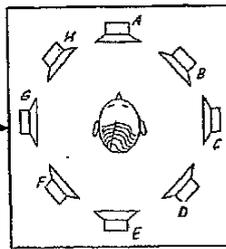


Fig. 7A

sistema "Ambisonic":
3 canales desde el micrófono de
campo sonoro descodificados
a los altavoces

formato B
de primer orden

descodifica-
dor en forma-
to B de pri-
mer orden

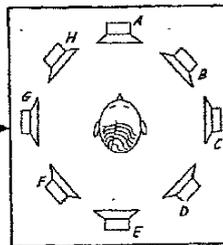


Fig. 7B

sistema "Ambisonic de orden
superior": información direccional
de mayor calidad usada para
hacer descodificador de mejor calidad

formato B
de tercer orden

descodifica-
dor en forma-
to B de ter-
cer orden

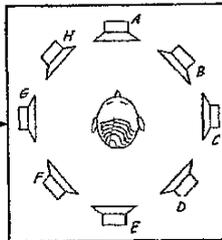


Fig. 7C

sistema de "Extrapolación de Armónicos
Esféricos": entrada de 3 canales es
extrapolada a 7 canales para proporcionar
descodificación de resolución superior

formato B
de primer orden

convertidor
de primer
orden a tercer
orden

descodifica-
dor en forma-
to B de ter-
cer orden

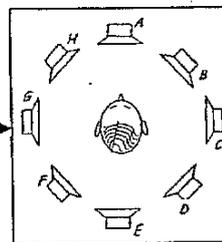


Fig. 7D

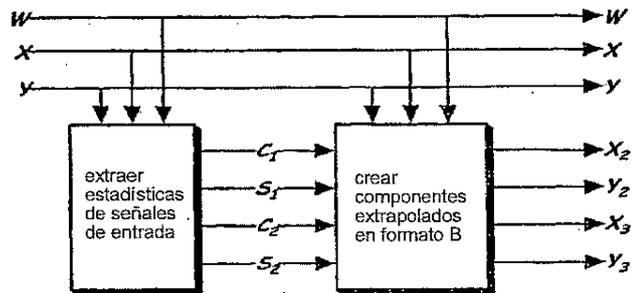


Fig. 8

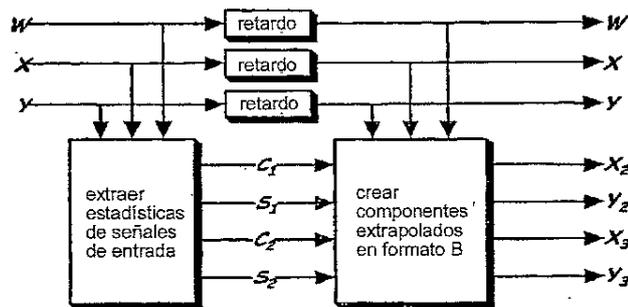


Fig. 9

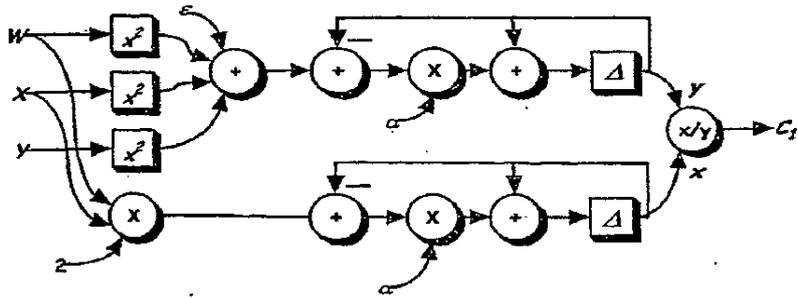


Fig. 10

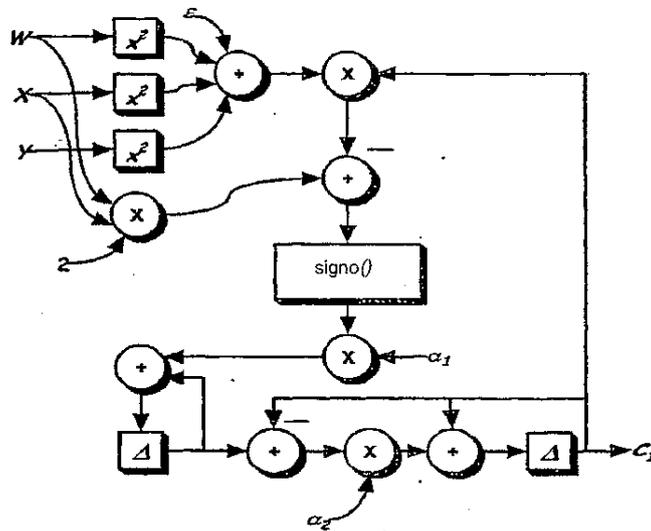


Fig. 11

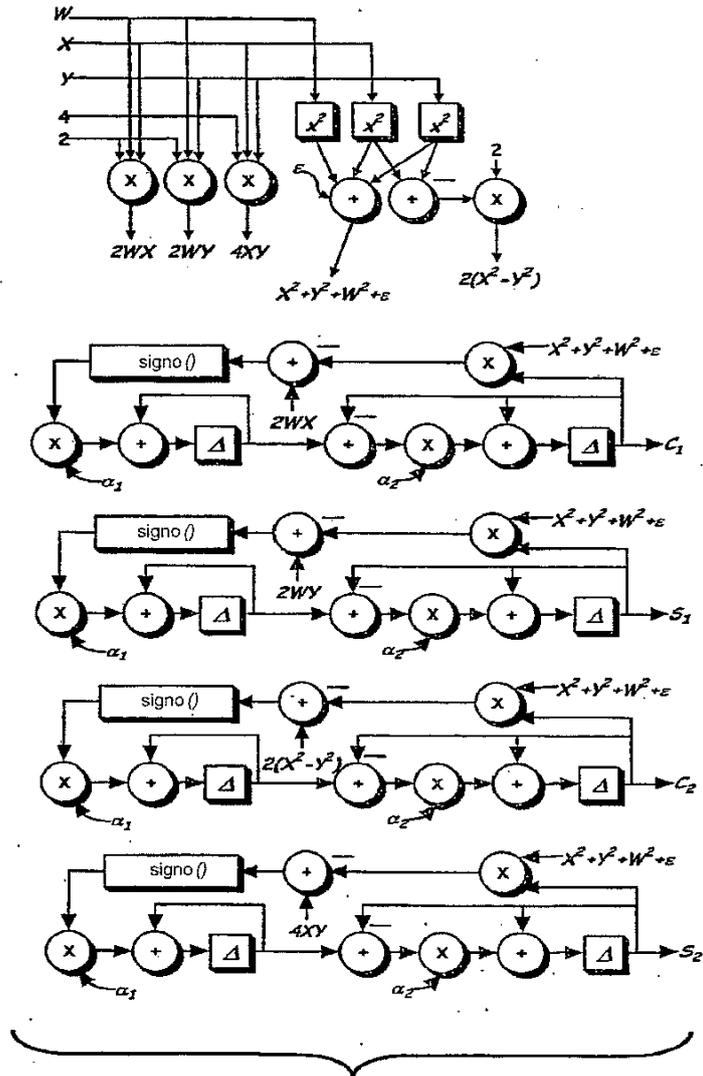


Fig. 12

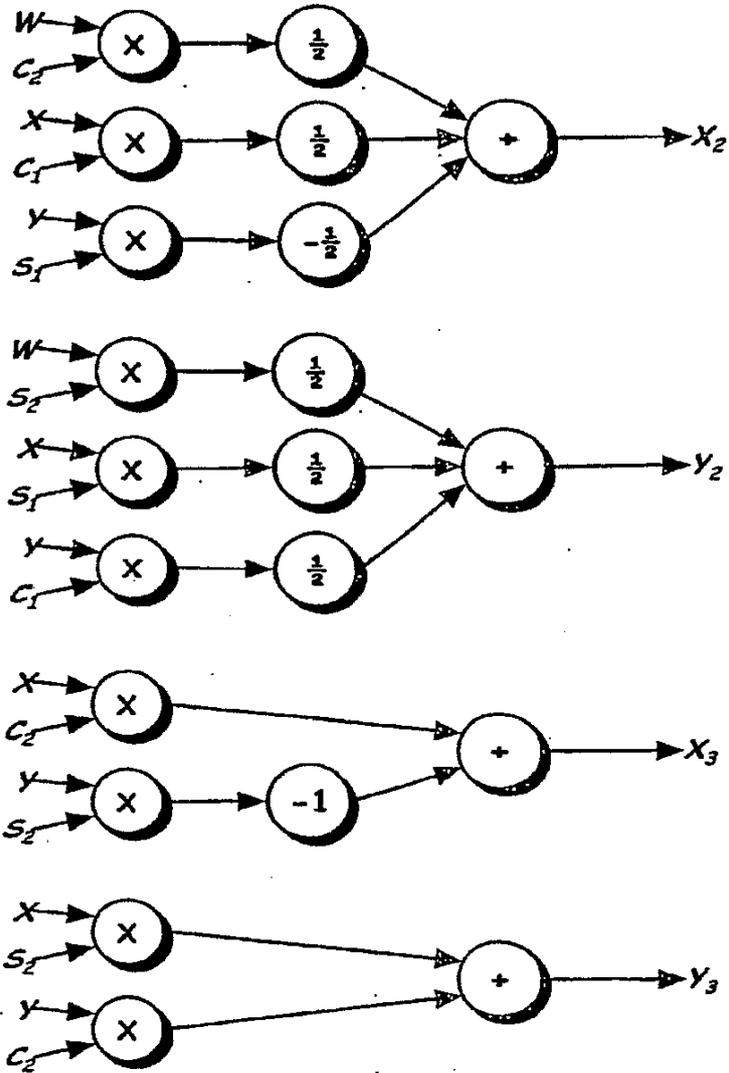


Fig. 13

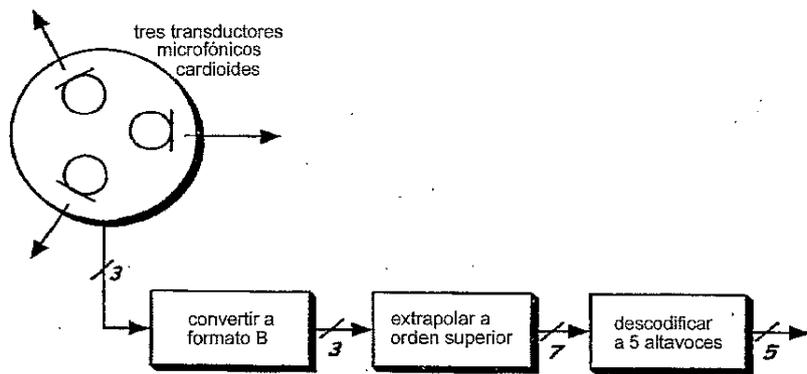


Fig. 14

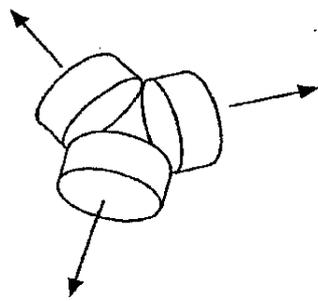


Fig. 15A

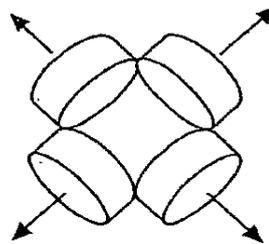


Fig. 15B

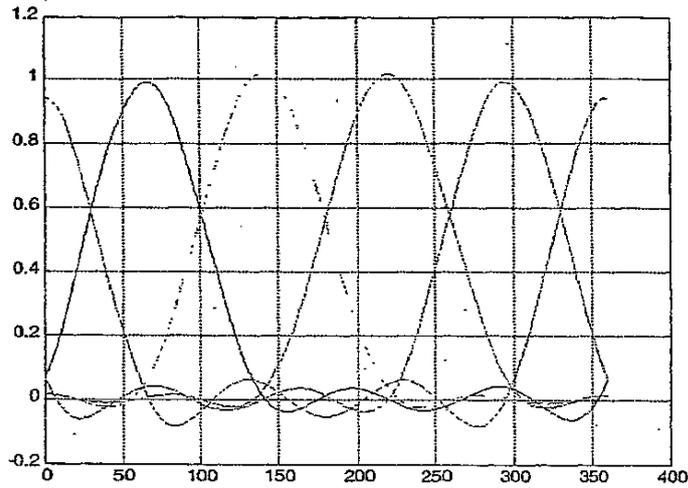


Fig. 16

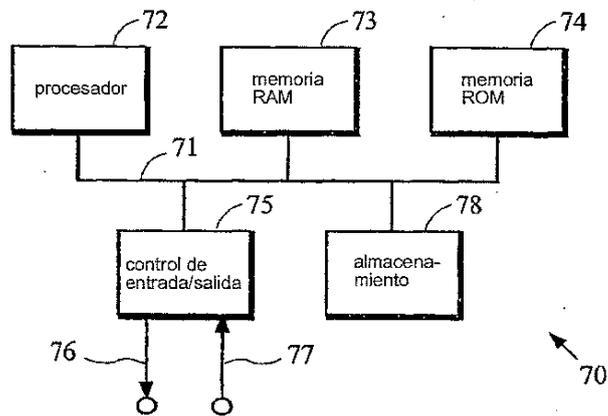


Fig. 17