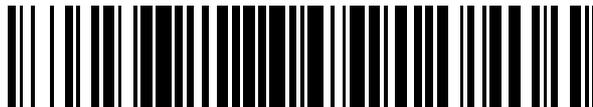


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 390 831**

51 Int. Cl.:  
**H04S 1/00**

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **07731711 .3**

96 Fecha de presentación: **08.03.2007**

97 Número de publicación de la solicitud: **1999998**

97 Fecha de publicación de la solicitud: **10.12.2008**

54 Título: **Procedimiento de síntesis binaural teniendo en cuenta un efecto de sala**

30 Prioridad:  
**28.03.2006 FR 0602694**

45 Fecha de publicación de la mención BOPI:  
**16.11.2012**

45 Fecha de la publicación del folleto de la patente:  
**16.11.2012**

73 Titular/es:  
**FRANCE TELECOM (100.0%)  
6, PLACE D'ALLERAY  
75015 PARIS, FR**

72 Inventor/es:  
**FAURE, JULIEN;  
GUERIN, ALEXANDRE;  
NICOL, ROZENN y  
PALLONE, GRÉGORY**

74 Agente/Representante:  
**LINAGE GONZÁLEZ, Rafael**

ES 2 390 831 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

Procedimiento de síntesis binaural teniendo en cuenta un efecto de sala

5 La invención se refiere a la espacialización sonora, denominada representación en 3D, de la señal de audio, integrando en particular un efecto de sala, particularmente en el dominio de las técnicas binaurales.

De ese modo, el término “binaural” pretende la restitución sobre unos cascos estereofónicos, o un par de auriculares, de una señal sonora con sin embargo unos efectos de espacialización. La invención no se limita en todo caso a la técnica antes citada y se aplica, particularmente, a unas técnicas derivadas de la “binaural” tales como las técnicas de restitución “transaural”, es decir sobre unos altavoces distantes. TRANSAURAL® es una marca comercial registrada por la sociedad COOPER BAUCK CORPORATION.

15 Una aplicación específica de la invención es, por ejemplo, el enriquecimiento de los contenidos de audio aplicando de manera eficaz unas funciones de transferencia acústica de la cabeza de un oyente a unas señales monofónicas, con el fin de sumergir a este último en una escena sonora en 3D, incluyendo en particular un efecto de sala.

Para la implementación de las técnicas “binaurales” sobre cascos o altavoces, se define la función de transferencia, o filtro, de una señal sonora entre una posición de una fuente sonora en el espacio y los dos oídos de un oyente. La función de transferencia acústica de la cabeza antes citada se designa como HRTF de “Head Related Transfer Function” en inglés en su forma frecuencial y HRIR de “Head Related Impulse Response” en inglés en su forma temporal. Para una dirección del espacio, se obtienen al final dos HRTF: una para el oído derecho y una para el oído izquierdo.

25 En particular, la técnica binaural consiste en aplicar dichas funciones de transferencia acústica de la cabeza a unas señales de audio monofónicas, con el fin de obtener una señal estereofónica que permita, durante una escucha en un casco, evitar la sensación de que las fuentes sonoras provienen de una dirección particular del espacio. La señal del oído derecho se obtiene filtrando la señal monofónica con la HRTF del oído derecho y la señal del oído izquierdo se obtiene filtrando igualmente la señal monofónica con la HRTF del oído izquierdo.

30 La patente de EE.UU. nº 5.438.623 describe un ejemplo de procedimiento de espacialización en 3D de canales de audio en los que tales funciones HRTF sintéticas se deducen a partir de datos experimentales y/o de un modelo de una cabeza esférica.

35 Los parámetros físicos esenciales que permiten caracterizar estas funciones de transferencia son:

- La ITD de “Interaural Time Difference” en inglés, definida como la diferencia interaural de tiempo de llegada de las ondas sonoras de una misma fuente sonora entre el oído izquierdo y el oído derecho del oyente. La ITD está ligada principalmente a la fase de las HRTF.

40 - El módulo espectral, que permite particularmente percibir unas diferencias de nivel entre el oído izquierdo y el oído derecho en función de la frecuencia.

45 - Cuando las HRTF, o las HRIR, de la cabeza del oyente no se consideran como correspondientes a unas condiciones de propagación sonora en campo libre (condición anecoica), las funciones de transferencia antes citadas pueden tener en cuenta unos fenómenos de reflexión, difusión, difracción, que corresponden a la respuesta acústica de la sala en la que estas funciones de transferencia se han medido o simulado. Las funciones de transferencia antes citadas se denominan entonces BRIR de “Binaural Room Impulse Response” en inglés en su forma temporal.

50 Las técnicas binaurales antes citadas pueden ser por ejemplo empleadas para simular una representación en 3D del tipo 5.1 en el casco de audición. En esta técnica, a cada posición del altavoz del sistema “surround” en inglés, o de múltiples altavoces, le corresponden un par de HRTF, una HRTF para el oído izquierdo y una HRTF para el oído derecho. La suma de los 5 canales de la señal del modo 5.1 convolucionados por los 5 filtros HRTF para cada oído de un oyente permiten obtener dos canales binaurales derecho e izquierdo, los cuales simulan el modo 5.1 para una escucha sobre un los cascos de audición de audio.

55 Se habla en esta situación de “binaural virtual surround” en inglés para la espacialización binaural que simula un sistema de altavoces múltiples.

60 Cuando, en la representación en 3D, se tiene en cuenta el hecho, para el oyente, de percibir las fuentes sonoras más o menos alejadas de la cabeza, fenómeno conocido bajo el nombre de externalización, y éstas de manera independiente de la dirección de procedencia de las fuentes sonoras, sucede frecuentemente, en una representación en 3D binaural, que las fuentes sean percibidas en el interior de la cabeza por el oyente. La fuente así percibida se denomina no externalizada.

65

Diferentes trabajos han mostrado que la adición de un efecto de sala en los métodos de representación en 3D binaurales permite incrementar considerablemente la externalización de las fuentes sonoras. Considérese, particularmente, D. R. Begault y E. M. Wenzel, "Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source". J. Audio Eng. Soc., vol. 49, nº 10, 2001.

Actualmente, existen dos métodos principales, que permiten integrar el efecto de sala en las HRIR.

- El primero, relativo al efecto de sala real, consiste en medir unas HRIR en una sala no anecoica, que incluya por lo tanto un efecto de sala. Las HRIR obtenidas, que no son otras que las BRIR, deben tener una duración suficientemente larga para integrar las primeras reflexiones sonoras, una duración superior a 500 muestras temporales para una frecuencia de muestreo de 44.100 Hz, pero esta duración debe ser incluso más importante, es decir superior a 20.000 muestras temporales a la misma frecuencia de muestreo, si se desea integrar el efecto de reverberación tardía. Se observa en cualquier caso que las BRIR antes mencionadas se pueden obtener de manera equivalente mediante la convolución de las HRIR medidas en un entorno anecoico con el efecto de sala deseado, representado por la respuesta impulsional de la sala.

- El segundo, relativo al efecto de sala artificial, es el resultante de la acústica virtual y consiste en integrar el efecto de sala en la HRIR, de manera sintética. Esta operación se realiza gracias a unos espacializadores que introducen unos efectos de reverberación artificial. El inconveniente de tales métodos es que la obtención de una representación realista necesita una potencia de cálculo importante.

En lo que concierne a la espacialización sonora "binaural" un método actual consiste en modelizar los filtros binarios, descomponiendo las HRTF, o las HRIR, en una componente de fase mínima (filtro de fase mínima determinado por el módulo espectral de la HRTF) y un retardo puro. Para una descripción más detallada de un método de ese tipo, podrá referirse de modo útil a los artículos de D. J. Kistler y F. L. Wightman, "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction" J. Acoustic Soc. Am. 91 (3) págs. 1637-1647, 1992 y de Kulkarni A. et ál. "On the minimum-phase approximation of head-related functions" 1995 IEEE ASSP Workshop on Applications of Signal Processing Audio and Acoustics (IEEE catalog number: 95TH8144).

La diferencia de retardo observada entre las HRTF o las HRIR del oído izquierdo y del oído derecho corresponde entonces al índice de localización IRD. Existen diferentes métodos para extraer los retardos de las HRIR o HRTF. Los principales métodos se describen por S. Busson "Individualisation d'indices acoustiques pour la synthèse binaurale" Thèse de doctorat de l'Université de la Méditerranée Aix-Marseille II, 2006.

El módulo espectral se obtiene tomando el módulo de la transformada de Fourier de las HRIR. El número de coeficientes se puede reducir entonces, por ejemplo promediando la energía sobre un número reducido de bandas de frecuencias, por ejemplo de acuerdo con unas técnicas de alisado frecuencial basadas en las propiedades de integración del sistema auditivo.

Independientemente de la manera en la que se modelizan los filtros HRTF o HRIR, llegado al caso de la BRIR, existen varios métodos de implementación de la espacialización sonora binaural.

Entre estos, el método más simple y el más directo es la implementación bicanal del binaural, representado en la figura 1.

De acuerdo con este método, la espacialización de las fuentes se realiza de manera independiente unas de las otras. Se asocia a cada fuente un par de filtros HRTF. El filtrado se puede realizar o bien en el dominio temporal, en la forma de un producto de convolución, o bien en el dominio de la frecuencia, bajo la forma de una multiplicación compleja, o incluso en cualquier otro dominio de transformada, tal como el dominio PQMF de Pseudo Quadrature Mirror Filter en inglés por ejemplo.

La implementación de canal múltiple del binaural es una alternativa a la implementación bicanal que ofrece una implementación más eficaz que consiste en una descomposición lineal de las HRTF, en la forma de una suma de productos de funciones de la dirección (ganancias de codificación) y de filtros elementales (filtros de decodificación). Esta descomposición permite separar las etapas de codificación y de decodificación, siendo el número de filtros entonces independiente del número de fuentes a espacializar. Los filtros elementales pueden a su vez ser modelizados por un filtro de fase mínima y un retardo puro para simplificar su implementación. Es posible igualmente extraer los retardos de las HRTF de origen e integrarlos por separado en la codificación.

Las técnicas anteriores previamente mencionadas presentan unos inconvenientes principales, durante la implementación de los filtros BRIR, teniendo en cuenta el efecto de sala, en particular:

- La complejidad: las BRIR, debido al hecho de la importante duración de las respuestas de la sala, contienen un número de muestras temporales que puede ser muy elevado, superior a 20.000 muestras para unas salas de tamaño medio, estando ligado este número al retardo de los ecos de la sala y por tanto a las dimensiones de esta

última. En consecuencia, los filtros BRIR correspondientes requieren una potencia de cálculo y un tamaño de memoria muy importantes.

5 - La externalización: la modelización en la forma de un filtro de fase mínima, asociado a un retardo puro, permite reducir el tamaño de los filtros. No obstante, el hecho de extraer un único retardo interaural para cada filtro BRIR no permite tener en cuenta las primeras reflexiones.

En este caso, del timbre sonoro se respeta completamente pero el efecto de externalización no se reproduce.

10 La presente invención tiene por objetivo solucionar los inconvenientes antes citados de la técnica anterior y está definida por el procedimiento de la reivindicación 1 y el programa de ordenador de la reivindicación 11.

15 En particular, un objetivo de la presente invención es proponer un procedimiento de cálculo de parámetros de modelización de filtros BRIR, filtros HRIR, teniendo en cuenta un efecto de sala de la técnica anterior, estos parámetros comprenden uno o varios retardos asociados eventualmente a unas ganancias y al menos a un espectro de amplitud, con el fin de permitir una implementación de alto rendimiento o bien en el dominio temporal, o bien en el dominio frecuencial o de transformada.

20 Otro objetivo de la presente invención es la implementación de un procedimiento de cálculo de filtros BRIR específicos, los cuales, aunque equivalentes en términos de calidad a unos filtros BRIR clásicos u originales permiten un posicionamiento y una externalización que satisface unos orígenes, reduciendo fuertemente la potencia de cálculo y el tamaño de memoria necesario para la implementación del filtrado correspondiente.

25 El procedimiento de espacialización en 3D de canales de audio, a partir de al menos una función de transferencia del filtro acústico que incorpore un efecto de sala, objetivo de la presente invención, se caracteriza porque consiste en al menos, para un número de muestras específico que corresponde al tamaño de la respuesta de impulsos de la función de transferencia, en descomponer esta función de transferencia en al menos un conjunto de valores de retardo y de amplitud asociados a unos picos de amplitud, en extraer sobre este número de muestras al menos un módulo espectral de la función de transferencia, en constituir a partir de cada retardo sucesivo de su amplitud y de su módulo espectral asociados una función de transferencia elemental directamente aplicada a los canales de audio en el dominio temporal, frecuencial o de transformada.

35 El procedimiento de acuerdo con la invención además se caracteriza porque los valores de retardo y de amplitud asociados a unos picos de amplitud corresponden a los instantes de llegada de las reflexiones.

40 El procedimiento objetivo de la invención por otro lado se caracteriza porque la descomposición de la función de transferencia se ejecuta por un proceso de detección de los retardos mediante detección de los picos de amplitud, estando asociado el primer pico de amplitud al retardo correspondiente al instante de llegada de la onda sonora directa.

El procedimiento objetivo de la invención igualmente se caracteriza porque la extracción de cada módulo espectral se ejecuta mediante la transformación tiempo-frecuencia.

45 El procedimiento objetivo de la invención igualmente se caracteriza porque, para un número de muestras que corresponde a la respuesta de impulsos de la función de transferencia descompuesta en sub-bandas de frecuencias de intervalo  $k$  determinado, el valor del módulo espectral de la función de transferencia se define como un valor real de ganancia representativo de la energía de la función de transferencia en cada sub-banda.

50 El procedimiento objetivo de la invención igualmente se caracteriza porque a cada retardo se asocia un módulo espectral y porque el módulo espectral de la función de transferencia se define en cada sub-banda como un valor real de ganancia representativo de la energía de la función de transferencia parcial en dicha sub-banda, siendo este valor de ganancia una función del retardo asociado.

55 Esta modulación del módulo espectral en función del retardo aplicado permite implementar una reconstrucción del filtro BRIR mucho más próxima al filtro BRIR de origen.

60 Finalmente, el procedimiento objetivo de la invención se caracteriza porque cada función de transferencia elemental en cada sub-banda de frecuencias de intervalo  $k$  está formada por una multiplicación compleja, función o no del retardo asociado a cada pico de amplitud que incluye un valor real de ganancia, y por un retardo puro, incrementado en el intervalo de retardo en comparación con el retardo asignado a la primera muestra correspondiente al instante de llegada de la onda sonora directa.

Se comprenderá mejor con la lectura de la descripción y la observación de los dibujos posteriores en los que, además de la figura 1 relativa a una técnica de espacialización sonora binaural de la técnica anterior:

- la figura 2 representa, a título puramente ilustrativo, un organigrama de las etapas esenciales de implementación del procedimiento de espacialización 3D de canales de audio a partir de al menos un filtro BRIR que incorpora un efecto de sala, de acuerdo con el objetivo de la presente invención;

5 - la figura 3 representa un detalle de la implementación de la etapa de descomposición ejecutada en la etapa A de la figura 2a;

10 - la figura 3b representa un cronograma de muestras que permiten explicar el modo de funcionamiento de una sub etapa A<sub>0</sub> de constitución de un primer vector l<sub>i</sub> y de un primer vector desfasado l<sub>i+1</sub> de picos de amplitud de la figura 3a;

15 - la figura 3c representa a título ilustrativo un cronograma de las muestras de picos de amplitud que explican un proceso de construcción de un segundo vector a partir de un vector de desviación entre el primer vector desfasado y el primer vector ilustrado en la figura 3b, reagrupando este segundo vector los índices de intervalos de los picos de amplitud aislados;

20 - la figura 3d representa un cronograma de los picos de amplitud representativos de las primeras reflexiones debidas al efecto de sala obtenido a partir del segundo vector ilustrado en la figura 3c, estando asignado a cada una de las primeras reflexiones un retardo correspondiente al parámetro correspondiente al instante de llegada de la onda sonora directa, tras unos retardos sucesivos específicos añadidos al parámetro de retardo de la onda sonora directa.

25 El procedimiento de espacialización en 3D del canal de audio a partir de al menos un filtro BRIR que incorpore un efecto de sala, de acuerdo con el objetivo de la invención se describirá ahora en relación con la figura 2 y las figuras siguientes.

30 El procedimiento objetivo de la invención consiste para un número N dado de muestras específico, que corresponde al tamaño de la respuesta de impulsos del filtro BRIR, en descomponer en una etapa A este filtro BRIR en al menos un conjunto de valores de amplitud y de valores de retardo que describen un conjunto de picos de amplitud.

En la etapa A de la figura 2, la operación de descomposición se indica:

$$\left[ A_n, n \right]_{n=1}^{n=N} A_{Mx} | \Delta x = \Delta_0 + \delta x$$

35 En esta relación, A<sub>n</sub> indica la amplitud de la muestra de intervalo n y A<sub>Mx</sub> indica la amplitud de cada pico de amplitud, Δx designa el retardo asociado a cada uno de los picos de amplitud correspondiente.

40 Este retardo es función del retardo Δ<sub>0</sub> que corresponde al instante de llegada de la onda directa tal como se describirá a continuación en la descripción. La etapa A viene seguida de una etapa B que consiste en extraer, sobre el número N de muestras al menos un módulo espectral medio del filtro BRIR, estando descrito cada módulo espectral:

$$\overline{BRIR}_N = G_N$$

45 La etapa B va seguida entonces por una etapa C que consiste en constituir a partir de cada retardo sucesivo, de la amplitud y del módulo espectral asociados a este retardo establecido en la etapa B un filtro BRIR elemental indicado por BRIR<sub>e</sub> directamente aplicado al canal de audio en el dominio temporal, frecuencial o transformado, tal como se describirá a continuación en la descripción.

50 De manera más específica, se comprende que la descomposición del filtro BRIR en la etapa A se ejecuta por un proceso de detección de los retardos mediante la detección de los picos de amplitud, estando asociado el primer pico de amplitud al retardo Δ<sub>0</sub> que corresponde al instante de llegada de la onda sonora directa.

De ese modo, el primer pico de amplitud se define por los parámetros A<sub>M0</sub>|Δ<sub>0</sub>.

55 Se comprende igualmente que a los otros picos de amplitud se asocian entonces sucesivamente, además del retardo Δ<sub>0</sub>, un valor δx que depende de la posición del pico de amplitud en las N muestras, el retardo asignado a cada pico de amplitud A<sub>Mx</sub> viene dado por Δx = Δ<sub>0</sub> + δx.

60 Se pueden utilizar también otros métodos de detección del primer pico, tal como es bien conocido en la técnica actual, en particular para determinar el valor del retardo Δ<sub>0</sub>, que puede por ejemplo ser tomado igual al retardo interaural.

La etapa B, de extracción de al menos un módulo espectral del filtro BRIR de una duración de N muestras permite asegurar una correspondencia del timbre entre cada filtro BRIR original y el filtro BRIR reconstruido a partir de los filtros BRIR<sub>e</sub> elementales, tal como se describirá posteriormente en la descripción.

5 En particular, y de manera no limitativa, la extracción del módulo espectral se puede ejecutar mediante una transformación tiempo-frecuencia tal como una transformada de Fourier, tal como se describirá posteriormente en la descripción.

10 La implementación de los filtros BRIR elementales BRIR<sub>e</sub> formados cada uno a partir del valor de cada módulo espectral del filtro BRIR y por supuesto de la amplitud y del retardo  $\Delta x$  considerado, permite asegurar una reducción de los costes de cálculo.

15 Todos los métodos de filtrado a partir de un filtro de fase mínima o no, asociados a todos los métodos de implementación de los retardos pueden ser apropiados para la descomposición propuesta. En particular, el procedimiento objetivo de la invención puede por ejemplo combinarse con una implementación de canales múltiples de la espacialización 3D binaural.

20 Un modo de implementación particular preferido no limitativo del procedimiento objetivo de la invención se describirá ahora en relación con las figuras 3a a 3d.

El modo de implementación antes citado se implementa en el marco de la descomposición de filtros BRIR para una implementación eficaz en el dominio de las sub-bandas temporales complejas más particularmente, aunque de manera no limitativa, el dominio PQMF complejo.

25 Una implementación así se puede utilizar mediante un decodificador definido por la norma MPEG surround, con el fin de obtener una representación en 3D binaural del tipo 5.1. El modo 5.1 se define por la norma MPEG spatial audio coding ISO/IEC 23003-1 (doc. N7947).

Con referencia a la solicitud de patente francesa titulada:

30 "Procédé et dispositif de spatialisation sonore binaurale efficace dans le domaine transformé"

35 presentada el mismo día en nombre del solicitante, se indica que el filtrado binaural se puede efectuar directamente en el dominio de las sub-bandas, es decir en el dominio codificado, con el fin de reducir los costes de decodificación que incluye la implementación del procedimiento.

El modo de realización antes citado se puede trasponer al dominio temporal es decir al dominio no transformado en sub-bandas o a cualquier otro dominio transformado.

40 El procedimiento objetivo de la invención permite obtener de manera general y en particular en su modo de implementación preferido:

45 - unos retardos que corresponden al retardo  $\Delta_0$  del instante de llegada de la onda sonora directa y a los retardos de las primeras reflexiones de la sala, siendo estos retardos implementados a continuación en el dominio de las sub-bandas;

- unos valores de ganancia, valores reales, siendo una ganancia por ejemplo atribuida a cada sub-banda y para cada reflexión a partir del contenido espectral de los filtros BRIR, tal como se explicará a continuación.

50 De ese modo, para una ejecución descrita a modo de ejemplo no limitativo en el dominio de las sub-bandas temporales complejas, la extracción de los retardos consiste en al menos para todo filtro BRIR correspondiente a una posición del espacio, tal como se representa en la figura 3a y a partir de la envolvente temporal del filtro establecida en el número de muestras N que corresponden al tamaño de la respuesta de impulsos del filtro BRIR, estando indicada esta envolvente temporal como  $[A_{n,i}]_{i=1}^{n=N}$ , en ejecutar en una primera subetapa, indicada como A<sub>0</sub>,

55 que consiste en identificar los índices del intervalo de muestras temporal cuyo valor de amplitud es superior a un valor de umbral indicado como V en la etapa A<sub>01</sub> de la figura 3a. Se comprende, en particular, que la comparación  $A_n > V$  se efectúa para cada muestra de las N muestras sucesivamente mediante la vuelta a la etapa A<sub>01</sub> por intermedio de la subetapa A<sub>02</sub> sucesivamente en las N muestras.

60 Esta operación permite engendrar un primer vector indicado por I<sub>i</sub> en la subetapa A<sub>03</sub>, y un primer vector desfasado indicado por I<sub>i+1</sub> en la subetapa A<sub>04</sub>. El primer vector I<sub>i</sub> corresponde a los índices del intervalo de muestras temporales cuyo valor de amplitud es superior al valor de umbral V. El primer vector desfasado I<sub>i+1</sub> se deduce del primer vector por desfase de un índice. El primer vector y el primer vector desfasado son representativos de la posición de los picos de amplitud en el número N de muestras.

La etapa A<sub>0</sub> va seguida de una etapa A<sub>1</sub> que consiste en determinar si las muestras temporales cuya amplitud es superior al valor de umbral V corresponden a unos picos de amplitud aislados mediante el cálculo de un vector de desviación l' que representa la diferencia entre el primer vector desfasado l<sub>i+1</sub> y el primer vector l<sub>i</sub>.

5 En efecto, se comprende que, si los valores contenidos en el vector de desviación l' son importantes, entonces se está en presencia del comienzo de un pico distinto al pico precedente, tal como se describirá posteriormente en la descripción.

10 La etapa A<sub>1</sub> viene seguida entonces de una etapa A<sub>2</sub> que consiste en calcular un segundo vector P que reagrupa los índices de picos de amplitud aislados en el número N de muestras para un umbral de desviación definido por un valor específico W.

15 Finalmente, la etapa A<sub>2</sub> va seguida de la etapa A<sub>3</sub> que consiste en identificar, a partir de las muestras del segundo vector, para cada pico aislado identificado, el índice de muestra de amplitud máxima entre un número determinado de muestras, tomado igual al valor W anteriormente citado siguiendo la muestra identificada por el segundo vector. Este valor W se puede determinar experimentalmente.

20 El índice y la amplitud de cualquier otra nueva muestra de amplitud máxima se memorizan en la forma de un vector de índice de retardo y de un vector de amplitud.

De ese modo al final de la etapa A<sub>3</sub>, se dispone en la forma de un vector de índice D'(i) y un vector de amplitud A'(i) por ejemplo del conjunto de los valores de índice de retardo y de amplitud de los picos de amplitud antes citados.

25 Se realizará ahora una descripción específica de la implementación de las etapas A<sub>0</sub>, A<sub>1</sub>, A<sub>2</sub> y A<sub>3</sub> representadas en la figura 2 en relación con las figuras 3b, 3c y 3d.

Con referencia a la figura 3b, para un filtro BRIR temporal que corresponde a una posición del espacio, la envolvente temporal de este último viene dada por:

30 
$$BRIR_{env}(t) = |BRIR(t)|$$

La etapa A<sub>0</sub> consiste entonces en encontrar todos los índices de las muestras cuyo valor de envolvente es superior al valor de umbral V.

35 De manera particularmente ventajosa y de acuerdo con un aspecto notable del procedimiento objeto de la invención, el valor de umbral V es en sí mismo función de la energía de la envolvente temporal del filtro BRIR.

De ese modo, el valor de umbral V verifica ventajosamente la relación:

40 
$$V = C \sqrt{\frac{\sum_N BRIR(t)^2}{N}}$$

En la relación anterior, además de N que representa el número de muestras temporales, C se es una constante fijada en 1, por ejemplo.

45 A continuación de las comparaciones ejecutadas en las etapas A<sub>01</sub> y A<sub>02</sub>, se almacena sobre la comparación lograda, en un vector l<sub>i</sub> de tamaño K, siendo K el número de muestras cuyo valor absoluto de la amplitud sobrepasa el valor de umbral V para constituir el primer vector.

50 A título de ejemplo no limitativo, en la figura 3b, se ha representado la envolvente temporal de un filtro BRIR para el que el umbral V se fija en el valor real 0,037.

El vector l<sub>i</sub> representado en la etapa A<sub>03</sub> de la figura 3a se escribe:

55 
$$l_i = [89 \ 90 \ 91 \ 92 \ 93 \ 94 \ 95 \ 96 \ 97 \ 98 \ 101 \ 104 \ 108 \ 110 \ 116 \ 422 \ 423 \ 424 \ 427...].$$

A partir de la memorización del vector l<sub>i</sub>, se procede igualmente por desfase del índice del primer pico de amplitud, el índice 89, a la memorización del vector desfasado l<sub>i+1</sub>, correspondiendo el vector l<sub>i+1</sub> por ejemplo al vector l<sub>i</sub> en el que se ha suprimido el primer pico de amplitud.

60 Se dispone de ese modo del primer vector l<sub>i</sub> y del primer vector desfasado l<sub>i+1</sub>.

En la etapa A<sub>1</sub> se procede entonces al cálculo del vector l', vector de desviación, diferencia entre el primer vector desfasado l<sub>i+1</sub> y el primer vector l<sub>i</sub>.

El vector de desviación l' verifica la relación en el ejemplo dado:

$$l' = [1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 3 \ 3 \ 4 \ 2 \ 6 \ 306 \ 1 \ 1 \ 3 \ \dots].$$

Los valores importantes contenidos en el vector l' indican la presencia de un pico de amplitud distinto al pico de amplitud precedente.

La etapa A<sub>2</sub> consiste entonces en calcular el segundo vector P que reagrupa los índices de los picos distintos.

En el ejemplo dado el primer pico P(1) viene dado naturalmente por P(1) = l(1) = 89 es decir por el primer pico de amplitud mencionado anteriormente. El índice de los picos siguientes corresponde a los índices incrementados en 1 de los valores de l' que sobrepasan un umbral de desviación definido por un valor W. A modo de ejemplo no limitativo y de manera experimental W se puede fijar en el valor 20. En estas condiciones, el valor l'(15) = 306 > W determina un segundo pico aislado. El valor del índice del intervalo de este segundo pico P(2) viene dado entonces por l(15+1) = 422.

De ese modo el segundo vector P se puede escribir en la forma:

$$P = [89 \ 422 \ \dots].$$

La etapa A<sub>3</sub> de la figura 3a puede, tal como se representa en la figura 3c, consistir, a partir de cada una de las muestras P(i) del segundo vector representativo de la envolvente temporal, en encontrar la muestra que tiene el valor de amplitud máximo entre las W = 20 muestras siguientes.

El índice de esta nueva muestra se almacena en el vector D' y su amplitud se almacena en el vector A' tal como se ha mencionado en relación con la etapa A<sub>3</sub> de la figura 3a de acuerdo con las relaciones:

$$D'(i) = \text{índice}(\max(\text{BRIR}_{\text{env}}([P(i); P(i+W)]))),$$

$$A'(i) = \text{BRIR}(D'(i)) * \text{signo}(\text{BRIR}(D'(1))).$$

De manera no limitativa para el ejemplo dado en relación con la figura 3:

$$D' = [92 \ 423 \ \dots],$$

$$A' = [0,1878 \ 0,0924 \ \dots].$$

Si la amplitud de la primera muestra de la amplitud máxima indicada por A(1) es negativa, entonces se tiene en cuenta el valor absoluto de esta última.

Las amplitudes A de las amplitudes máximas se pueden normalizar entonces en energía por la relación:

$$A = \frac{A'}{\sqrt{\sum_{l=1:L} A'(l)^2}}$$

En la relación anterior L es el número de elementos de D' y de A, es decir de los vectores de índice y de amplitud representativos de cada pico. Este número depende naturalmente del valor de umbral V y del valor de la constante W antes citados.

Una representación de las amplitudes normalizadas, de los picos de amplitud y de su posición de retardo sucesivo con relación al primer pico de amplitud al que está asignado el retardo Δ<sub>0</sub>, se representa en la figura 3d.

Una descripción más detallada de un primer y un segundo modo de puesta en práctica de los filtros BRIR elementales directamente aplicables y aplicados a los canales de audio en el dominio de la transformada, en particular en el dominio PQMF complejo descompuesto en sub-bandas SB<sub>k</sub>, se dará a título de ejemplo no limitativo a continuación en la descripción.

Se recuerda que la descomposición en sub-bandas en el dominio antes citado permite descomponer las N muestras de la respuesta de impulsos del filtro BRIR en M sub-bandas de frecuencias, por ejemplo M = 64, para una aplicación en el MPEG surround antes citado.

La ventaja de una transformación de ese tipo es poder aplicar unas ganancias reales a cada sub-banda, mientras se sortean unos problemas de repliegue del espectro generados por el sub muestreo inherente al banco de filtros.

5 En el dominio de las sub-bandas antes citadas, los retardos y las ganancias se aplican sobre las muestras complejas, tal como se describirá posteriormente en la descripción.

De acuerdo con un modo de puesta en práctica no limitativo, el valor de cada módulo espectral del filtro BRIR se define en cada sub-banda como al menos un valor real de ganancia representativo de la energía del filtro BRIR en dicha sub-banda.

10 En este primer modo de puesta en práctica, los valores de las ganancias correspondientes indicados por  $G(k, n)$  en la que  $k$  designa el intervalo de la sub-banda considerado y  $n$  el intervalo de la muestra entre las  $N$  muestras, se obtienen promediando la energía de la amplitud espectral de cada filtro BRIR en cada sub-banda.

15 Para un filtro BRIR frecuencial  $BRIR^*(f)$  correspondiente a la transformada de Fourier a 8192 muestras del filtro  $BRIR(t)$  temporal, completado desde 0 para obtener las 8192 muestras, el valor de las ganancias  $G(k, n)$  viene dado por la relación:

$$G(k, n) = \sqrt{\frac{\sum_{f=f_1}^{f=f_1+M'} (H(f)BRIR^*(f))^2}{M'}}$$

20 En la relación precedente, se indica que  $H$  es una ventana de ponderación, ventana rectangular por ejemplo de longitud  $M'$  superior o igual a la longitud de la sub-banda  $SB_k$ , por ejemplo  $M' = 64$ . La ventana de ponderación se centra en la frecuencia central de la sub-banda  $k$  y la frecuencia  $f_1$  es inferior o igual a la frecuencia de partida de la sub-banda  $k$ .

25 De acuerdo con un segundo modo de puesta en práctica preferente del procedimiento objetivo de la invención, a cada retardo se asocia un módulo espectral. El valor de cada módulo espectral se define en cada sub-banda como al menos un valor de ganancia representativo de la energía del filtro BRIR parcial en dicha sub-banda, siendo este valor de ganancia función del retardo aplicado en función del índice de cada muestra del pico de amplitud, a partir del vector de índice y de amplitud.

30 De ese modo en este segundo modo de realización, las ganancias  $G(k, n)$  se modulan y pueden por lo tanto variar en cada nuevo retardo  $l$  aplicado. Los valores de ganancias vienen dados entonces por la relación:

$$35 \quad G(k, n, l) = \sqrt{\frac{\sum_{f=f_1}^{f=f_1+M'} (H(f)BRIR^*(f, l))^2}{M'}}$$

40 En la relación precedente,  $BRIR^*(f, l)$  es la transformada de Fourier del filtro  $BRIR(t)$  temporal abierto entre las muestras  $D'(1)-Z$  y  $D'(1+1)$ , siendo la energía espectral calculada la del filtro BRIR así abierto, y completado por unos 0 para obtener 8192 muestras.  $Z$  depende de la frecuencia de muestreo y puede tomar el valor  $Z = 10$  para una frecuencia de muestreo de 44,1 kHz.

45 El segundo modo de realización antes citado permite de manera notable una reconstrucción mucho más próxima de la función de transferencia o filtro BRIR de origen y en particular tener en cuenta cada uno de los retardos provocados por las reflexiones sucesivas en la sala, lo que permite obtener una representación del efecto de la sala particularmente bueno y realista.

50 Se comprende entonces que cada filtro BRIR elemental, en cada banda de frecuencia  $k$  puede entonces ventajosamente estar formado por una multiplicación compleja, que incluye un valor real de ganancia, función o no del retardo aplicado función del índice de cada muestra del pico de amplitud, de acuerdo con el primer o el segundo modo de realización que permanece descrito anteriormente en la descripción.

La operación de multiplicación compleja viene dada por la relación:

$$55 \quad S'(k, n) = G(k, n) A(l) e^{-j\pi \frac{(k+0,5)d(l)}{M}} E(k, n).$$

El filtro BRIR elemental se forma igualmente por un retardo puro incrementado por la desviación del retardo con respecto al retardo  $\Delta_0$  asignado al primer pico de amplitud. Este retardo puede implementarse entonces por intermedio de una línea de retardo aplicada al producto obtenido por la rotación bajo la forma de multiplicación compleja antes citada.

5 La muestra obtenida verifica entonces la relación:  

$$S(k, n) = S'(k, n - D(l)).$$

10 En las relaciones anteriores  $E(k, n)$  designa la  $n$ -ésima muestra compleja de la sub-banda  $k$  considerada,  $S(k, n)$  designa la  $n$ -ésima muestra de la sub-banda  $k$  después de la aplicación de las ganancias y de los retardos,  $M$  es el número de la sub-banda y  $d(l)$  y  $D(l)$  son tales que corresponden a la aplicación del  $l$ -ésimo retardo de  $D(l)M+d(l)$  muestras en el dominio temporal no sub muestreadas.

15 El retardo  $D(l)M+d(l)$  corresponde a los valores de  $D'(l)$  calculados de acuerdo con el proceso de detección de picos de amplitud anteriormente descrito en relación con las figuras 3a a 3d.

Además,  $A(l)$  designa la amplitud del pico asociada al retardo correspondiente y  $G(k, n)$  designa la ganancia real aplicada a la  $n$ -ésima muestra compleja de la sub-banda  $SB_k$  del intervalo  $k$  considerado.

20 Finalmente, el procedimiento objetivo de la invención permite el tratamiento de la reverberación tardía. Se recuerda que la reverberación tardía corresponde a la parte de la respuesta de una sala para la que el campo acústico se difundirá y en consecuencia no se pueden discernir las reflexiones. Sin embargo es posible tratar los efectos de sala incluyendo una reverberación tardía, de acuerdo con el procedimiento objetivo de la invención. En este caso, el  
 25 procedimiento de acuerdo con la invención consiste en volver a añadir a los valores de los picos de amplitud detectados, una pluralidad de valores de amplitud arbitrarios y repartidos más allá de un instante arbitrario a partir del que se considera que las reflexiones discretas han finalizado y en el que comienza el fenómeno de reverberación tardía. Estos valores de amplitud se calculan y reparten más allá de la duración arbitraria, la cual se puede tomar  
 30 igual a 200 milisegundos por ejemplo, hasta la última muestra del número de muestras correspondientes al tamaño de la respuesta de impulsos BRIR.

De ese modo, de acuerdo con el procedimiento objetivo de la invención, los picos de amplitud de las primeras reflexiones se determinan tal como se ha descrito anteriormente en relación con las figuras 2 y siguientes y, a partir  
 35 de una muestra  $t_1$  que corresponde a 200 milisegundos, determinada experimentalmente y que corresponde al comienzo de la reverberación tardía, hasta una muestra  $t_2$  que corresponde al final de la reverberación o, llegado el caso, al final de las  $N$  muestras de la respuesta de impulsos del filtro BRIR, se vuelven añadir  $R$  valores a los vectores  $D'$  y  $A'$  tales como:

40 
$$D'(L+r) = t_1 + (t_2 - t_1) / (R - 1),$$
  

$$A(L+r) = 1.$$

En la relación anterior,  $L$  es el número de picos detectado,  $r$  es un entero comprendido entre 1 y  $R$ .

45 La toma en consideración del segundo modo de realización antes citado en el que los valores de las ganancias se modulan en función del retardo de cada pico de amplitud, permite introducir entonces de manera eficaz la reverberación tardía en el dominio de las sub-bandas.

50 El fenómeno de reverberación tardía se puede tratar igualmente mediante una línea de retardo añadida al tratamiento de las primeras reflexiones.

La invención cubre finalmente un programa de ordenador que comprende un conjunto de instrucciones memorizadas sobre un soporte de almacenamiento de un ordenador o de un dispositivo dedicado de espacialización sonora en 3D de señales de audio caracterizado porque, durante su ejecución, este programa de ordenador ejecuta el  
 55 procedimiento de espacialización sonora en 3D a partir de al menos un filtro BRIR que incluye un efecto de sala descrito anteriormente en la descripción en relación con las figuras 2 y 3a a 3d.

60 Se comprende, en particular, que el programa de ordenador antes citado puede ser un programa ejecutable directamente implantado en la memoria permanente de un ordenador o de un dispositivo de síntesis binaural de un efecto de sala en espacialización sonora.

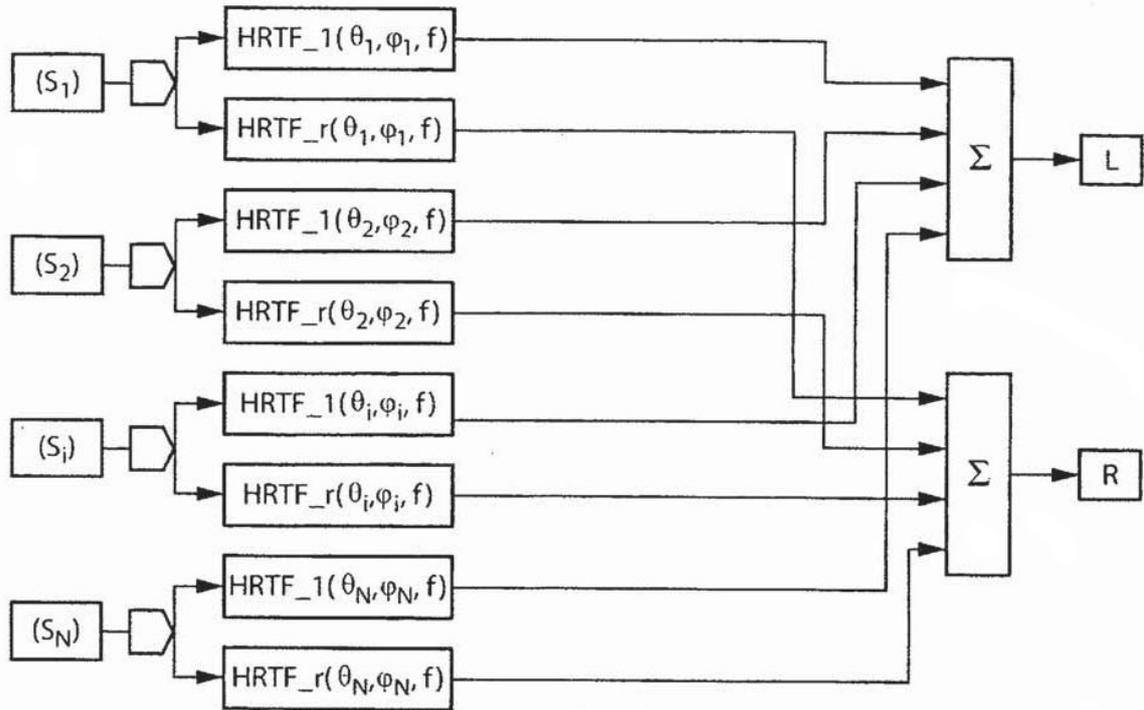
La puesta en práctica de la invención se puede ejecutar entonces de manera totalmente digital.

**REIVINDICACIONES**

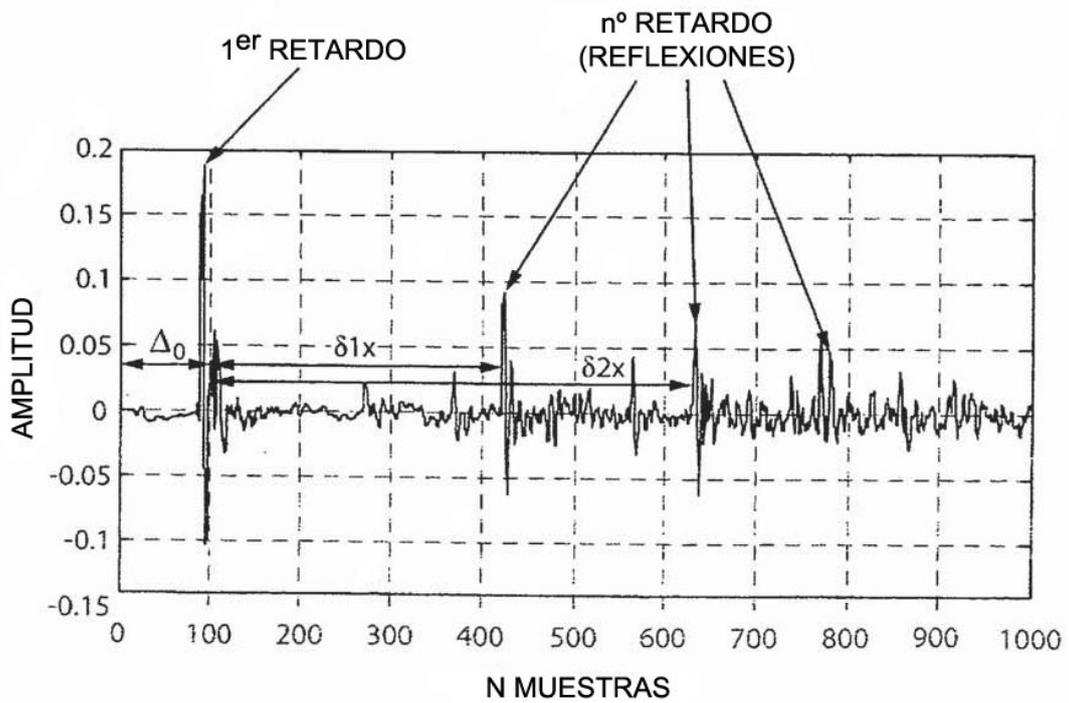
1. Procedimiento de espacialización en 3D de canales de audio, a partir de al menos un filtro BRIR que incorpore un efecto de sala, caracterizado porque consiste al menos, para un número de muestras específico que corresponde al tamaño de la respuesta de impulsos del filtro BRIR, en:
- descomponer (A) este filtro BRIR en al menos un conjunto de valores de retardo de amplitud asociados a los instantes de llegada de las reflexiones;
  - extraer (B) sobre dicho número de muestras al menos un módulo espectral medio del filtro BRIR;
  - constituir (C), a partir de cada retardo sucesivo, de su amplitud y de su módulo espectral medio asociados, un filtro BRIR elemental directamente aplicado a dichos canales de audio en el dominio temporal, frecuencial o transformado.
2. Procedimiento de acuerdo con la reivindicación 1, caracterizado porque los valores de retardo y de amplitud asociados a unos picos de amplitud corresponden a los instantes de llegada de las reflexiones.
3. Procedimiento de acuerdo con la reivindicación 1, caracterizado porque dicha descomposición del filtro BRIR se ejecuta por un proceso de detección de los retardos mediante la detección de los picos de amplitud, estando asociado el primer pico de amplitud al retardo correspondiente al instante de llegada de la onda sonora directa.
4. Procedimiento de acuerdo con una de las reivindicaciones 1 a 3, caracterizado porque la extracción de cada módulo espectral se ejecuta mediante una transformación tiempo-frecuencia.
5. Procedimiento de acuerdo con una de las reivindicaciones 1 a 4, caracterizado porque la extracción de los retardos consiste al menos, para cualquier filtro BRIR correspondiente a la posición del espacio, a partir de la envolvente temporal del filtro establecida sobre dicho número de muestras correspondientes al tamaño de la respuesta de impulsos del filtro BRIR, en:
- identificar ( $A_0$ ) los índices del intervalo de muestras temporales cuyo valor de amplitud es superior a un valor de umbral, para engendrar un primer vector y un primer vector desfasado representativos de la posición de los picos de amplitud en dicho número de muestras;
  - determinar ( $A_1$ ) la existencia de picos de amplitud aislados mediante el cálculo de un vector de desviación entre el primer vector desfasado y el primer vector;
  - calcular ( $A_2$ ) un segundo vector que reagrupa los índices de los picos de amplitud aislados sobre dicho número de muestras;
  - discriminar ( $A_3$ ) a partir de las muestras de dicho segundo vector los índices sucesivos de muestras de amplitud máxima entre un número determinado de muestras sucesivas, estando memorizados el índice y la amplitud de dichas muestras de amplitud máxima en la forma de un vector de índice de retardo y de amplitud.
6. Procedimiento de acuerdo con una de las reivindicaciones 1 a 5, caracterizado porque, para un número de muestras correspondiente a la respuesta de impulsos del filtro BRIR descompuesto en sub-bandas de frecuencias de intervalo k determinado, dicho valor del módulo espectral del filtro BRIR se define como un valor real de ganancia representativo de la energía del filtro BRIR en cada sub-banda.
7. Procedimiento de acuerdo con la reivindicación 6, caracterizado porque el valor del módulo espectral del filtro BRIR en cada sub-banda se calcula mediante la aplicación de una ventana de ponderación centrada en la frecuencia central de la sub-banda de frecuencias de intervalo k y de longitud igual o superior a la longitud de la sub-banda de frecuencias.
8. Procedimiento de acuerdo con una de las reivindicaciones 6 ó 7, caracterizado porque, a cada retardo se asocia un módulo espectral y porque dicho módulo espectral se define en cada sub-banda como un valor real de ganancia representativo de la energía del filtro BRIR parcial en dicha sub-banda, siendo este valor de ganancia función del retardo asociado.
9. Procedimiento de acuerdo con una de las reivindicaciones 6 a 8, caracterizado porque cada filtro BRIR elemental en cada sub-banda de frecuencias de intervalo k está formado por:
- una multiplicación compleja, función o no del retardo aplicado en función del índice de cada muestra de pico de amplitud que incluye el valor real de ganancia;

- un retardo puro, incrementado en la desviación del retardo con respecto al retardo asignado a la primera muestra que corresponde al instante de llegada de la onda sonora directa.

- 5 10. Procedimiento de acuerdo con una de las reivindicaciones 1 a 9, caracterizado porque, para el tratamiento de la reverberación tardía, éste consiste en volver a añadir a los valores típicos de amplitud detectados una pluralidad de valores de amplitud es arbitrarias, repartidas, desde un instante arbitrario, hasta una última muestra de los números de muestras que corresponden al tamaño de la respuesta de impulsos del filtro BRIR.
- 10 11. Programa de ordenador que incluye un conjunto de instrucciones memorizadas sobre un soporte de almacenamiento de un ordenador o de un dispositivo dedicado de espacialización sonora en 3D de señales de audio, caracterizado porque, durante su ejecución, dicho programa ejecuta el procedimiento de espacialización sonora en 3D a partir de al menos un filtro BRIR que incluye un efecto de sala, de acuerdo con una de las reivindicaciones 1 a 10.



**FIG. 1**  
(TÉCNICA ANTERIOR)



**FIG. 3d**

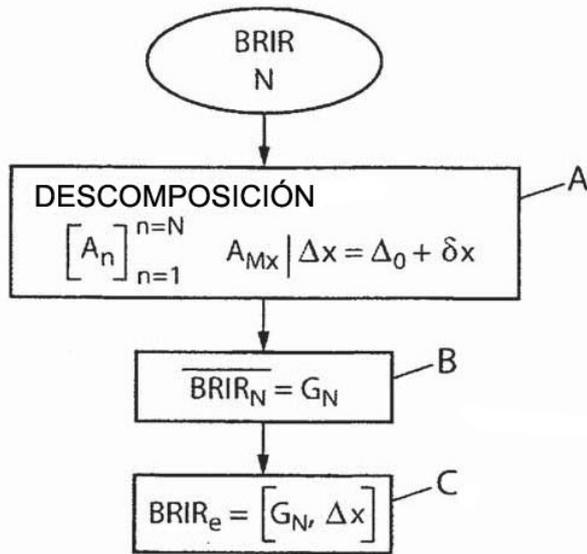


FIG. 2

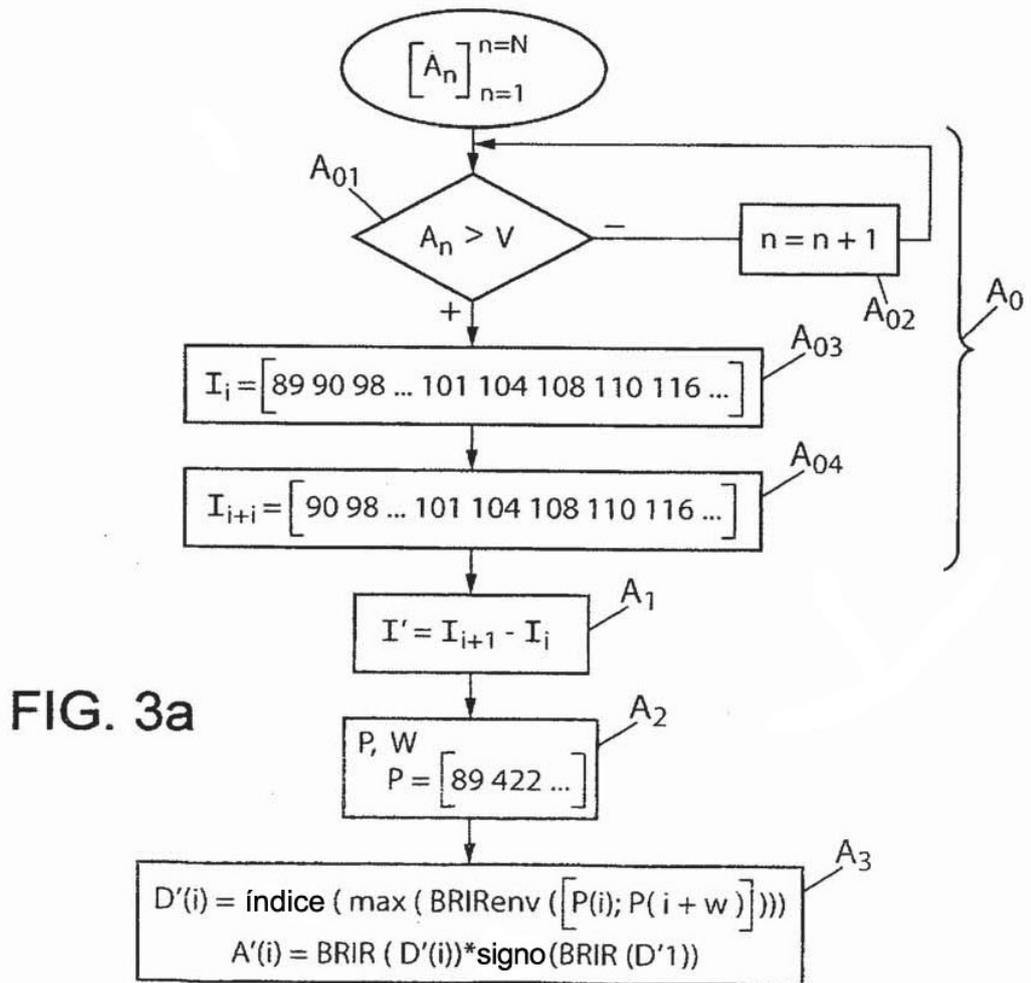


FIG. 3a

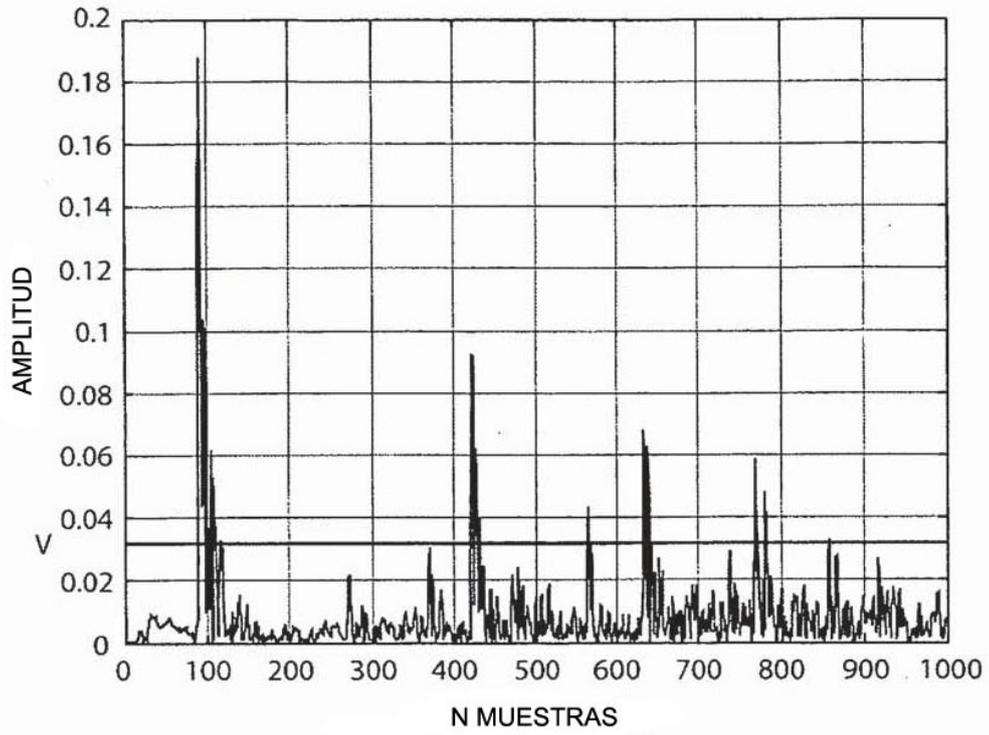


FIG. 3b

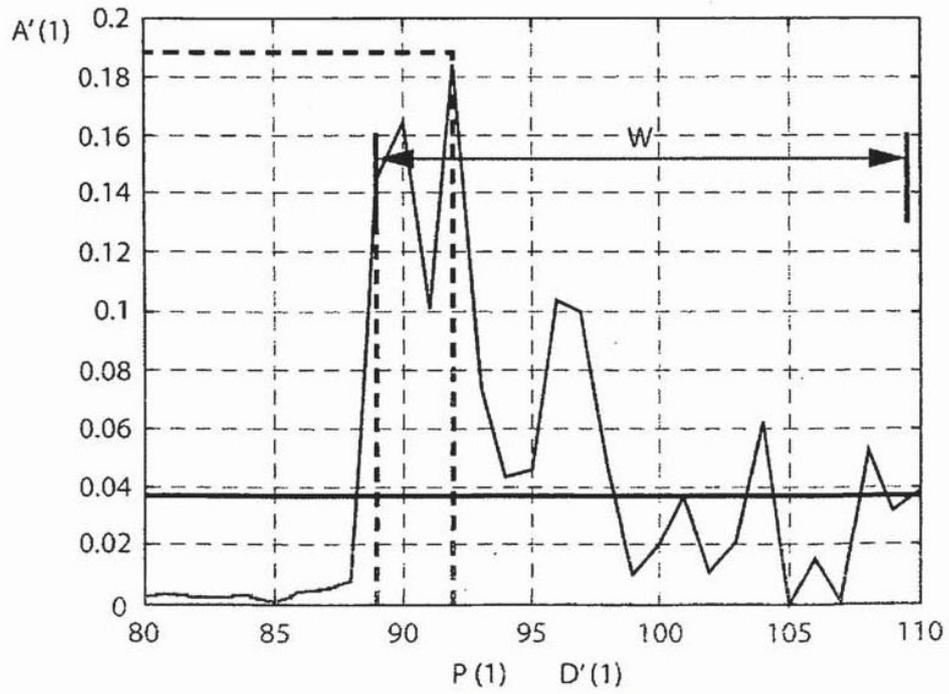


FIG. 3c