



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 362 920**

51 Int. Cl.:  
**H04S 1/00** (2006.01)  
**H04S 3/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06742984 .5**  
96 Fecha de presentación : **18.05.2006**  
97 Número de publicación de la solicitud: **1999997**  
97 Fecha de publicación de la solicitud: **10.12.2008**

54 Título: **Método mejorado para la conformación de señales en reconstrucción de audio multicanal.**

30 Prioridad: **28.03.2006 US 787096 P**

45 Fecha de publicación de la mención BOPI:  
**15.07.2011**

45 Fecha de la publicación del folleto de la patente:  
**15.07.2011**

73 Titular/es: **FRAUNHOFER-GESELLSCHAFT ZUR  
FÖRDERUNG DER ANGEWANDTEN FORSCHUNG  
E.V.**  
**Hansastrasse 27C**  
**80686 München, DE**

72 Inventor/es: **Disch, Sascha;**  
**Linzmeier, Karsten;**  
**Herre, Jürgen y**  
**Popp, Harald**

74 Agente: **Arizti Acha, Mónica**

ES 2 362 920 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

Método mejorado para la conformación de señales en reconstrucción de audio multicanal

5 Campo de la invención

La presente invención se refiere a un concepto de conformación de señal mejorada en reconstrucción de audio multicanal y en particular a un nuevo enfoque de conformación de envolvente.

10 Antecedentes de la invención y técnica anterior

15 El desarrollo reciente en la codificación de audio permite la recreación de una representación multicanal de una señal de audio basándose en una señal estéreo (o mono) y datos de control correspondientes. Estos métodos difieren sustancialmente de las soluciones basadas en matrices anteriores, tales como Dolby Prologic, puesto que se transmiten datos de control adicionales para controlar la recreación, a lo que también se hace referencia como mezcla ascendente, de los canales envolventes basándose en los canales estéreo o mono transmitidos. Tales decodificadores de audio multicanal paramétricos reconstruyen N canales basándose en M canales transmitidos, donde  $N > M$ , y los datos de control adicionales. El uso de los datos de control adicionales provoca una tasa de transmisión de datos significativamente inferior que la transmisión de todos los N canales, haciendo que la codificación sea muy eficaz, mientras que al mismo tiempo se garantiza la compatibilidad con los dispositivos de M canales y los dispositivos de N canales. Los M canales pueden ser una representación de canal mono único, de canal estéreo o de 5.1 canales. De este modo, es posible tener una señal original de 7.2 canales, mezclada de manera descendente para dar una señal de 5.1 canales compatible hacia atrás, y parámetros de audio espaciales que permiten a un decodificador de audio espacial reproducir una versión muy similar de los 7.2 canales originales, con una sobrecarga de tasa de transmisión de bits adicional pequeña.

25 Estos métodos de codificación de envolvente paramétrica comprenden usualmente una parametrización de la señal envolvente basándose en parámetros ILD (diferencia de nivel entre canales) e ICC (coherencia entre canales), variables en tiempo y frecuencia. Estos parámetros describen, por ejemplo, correlaciones y tasas de energía entre pares de canales de la señal multicanal original. En el proceso de decodificación, la señal multicanal recreada se obtiene distribuyendo la energía de los canales mezclados de manera descendente recibidos entre todos los pares de canales como se describe por los parámetros ILD transmitidos. Sin embargo, puesto que una señal multicanal puede tener una distribución de energía equitativa entre todos los canales, mientras las señales en los diferentes canales son muy diferentes, dando así la impresión respecto a la escucha de un sonido muy amplio, la amplitud correcta se obtiene mezclando señales con versiones decorrelacionadas de las mismas, como se describe por el parámetro ICC.

30 La versión decorrelacionada de la señal, que a menudo se denomina también señal húmeda o difusa, se obtiene pasando la señal a través de un reverberador, tal como un filtro pasa todo. Una forma simple de decorrelación es aplicar un retardo específico a la señal. En general, existen una gran cantidad de reverberadores diferentes conocidos en la técnica, la implementación precisa del reverberador usado es de menor importancia.

35 La salida del decorrelacionador tiene una respuesta de tiempo que usualmente es muy plana. De este modo, una señal de Dirac de entrada proporciona una salida repentina de ruido descendente. Cuando se mezcla la señal decorrelacionada con la señal original, para algunos tipos de señales transitorias, como las señales de aplausos, es importante realizar un procesamiento posterior de la señal para evitar la percepción de artefactos introducidos adicionalmente que pueden dar como resultado la percepción de un tamaño de habitación más grande y artefactos de tipo pre-eco.

40 En general, la invención se refiere a un sistema que representa audio multicanal como una combinación de datos mezclados de manera descendente de audio (por ejemplo, uno o dos canales) y datos multicanal paramétricos relacionados. En tal esquema (por ejemplo, en la codificación de indicadores binaurales) se transmite un flujo de datos de mezcla descendente de audio, en el que puede indicarse que la forma más simple de mezcla descendente es simplemente añadir las señales diferentes de una señal multicanal. Tal señal (señal de suma) va acompañada de un flujo de datos multicanal paramétricos (información secundaria). La información secundaria comprende, por ejemplo, uno o más de los tipos de parámetros presentados anteriormente para describir la interrelación espacial de los canales originales de la señal multicanal. En un sentido, el esquema multicanal paramétrico actúa como un pre/postprocesador respecto al extremo de envío/recepción de los datos de mezcla descendente, por ejemplo, con la señal de suma y la información secundaria. Debe indicarse que la señal de suma de los datos de mezcla descendente puede codificarse de manera adicional usando cualquier codificador de audio o de voz.

45 Como la transmisión de señales multicanal sobre portadoras de bajo ancho de banda está pasando a ser cada vez más popular, estos sistemas, también conocidos como "codificación de audio espacial", "MPEG envolvente", se han desarrollado mejor recientemente.

Las siguientes publicaciones se conocen en el contexto de estas tecnologías:

50 [1] C. Faller y F. Baumgarte, "Efficient representation of spatial audio using perceptual parametrization," en Proc. IEEE WASPAA, Mohonk, NY, octubre de 2001.

60 [2] F. Baumgarte y C. Faller, "Estimation of auditory spatial cues for binaural cue coding," en Proc. ICASSP 2002, Orlando, FL, mayo de 2002.

[3] C. Faller y F. Baumgarte, "Binaural cue coding: a novel and efficient representation of spatial audio," en Proc. ICASSP 2002, Orlando, FL, mayo de 2002.

65 [4] F. Baumgarte y C. Faller, "Why binaural cue coding is better than intensity stereo coding," en Proc. AES 112th Conv., Munich, Alemania, mayo de 2002.

[5] C. Faller y F. Baumgarte, "Binaural cue coding applied to stereo and multi-channel audio compression," en Proc. AES

112th Conv., Munich, Alemania, mayo de 2002.

[6] F. Baumgarte y C. Faller, "Design and evaluation of binaural cue coding," en AES 113th Conv., Los Angeles, CA, octubre de 2002.

5 [7] C. Faller y F. Baumgarte, "Binaural cue coding applied to audio compression with flexible rendering," en Proc. AES 113th Conv., Los Angeles, CA, octubre de 2002.

[8] J. Breebaart, J. Herre, C. Faller, J. Rödén, F. Myburg, S. Disch, H. Purnhagen, G. Hoto, M. Neusinger, K. Kjörling, W. Oomen: "MPEG Spatial Audio Coding / MPEG Surround: Overview and Current Status", 119th AES Convention, Nueva York 2005, Preprint 6599

10 [9] J. Herre, H. Purnhagen, J. Breebaart, C. Faller, S. Disch, K. Kjörling, E. Schuijers, J. Hilpert, F. Myburg, "The Reference Model Architecture for MPEG Spatial Audio Coding", 118th AES Convention, Barcelona 2005, Preprint 6477

[10] J. Herre, C. Faller, S. Disch, C. Ertel, J. Hilpert, A. Hoelzer, K. Linzmeier, C. Spenger, P. Kroon: "Spatial Audio Coding: Next-Generation Efficient and Compatible Coding of Multi-Channel Audio", 117th AES Convention, San Francisco 2004, Preprint 6186

15 [11] J. Herre, C. Faller, C. Ertel, J. Hilpert, A. Hoelzer, C. Spenger: "MP3 Surround: Efficient and Compatible Coding of Multi-Channel Audio", 116th AES Convention, Berlín 2004, Preprint 6049.

Una técnica relacionada, que se centra en la transmisión de dos canales a través de una señal mono transmitida se denomina "estéreo paramétrico" y por ejemplo se describe más ampliamente en las siguientes publicaciones:

[12] J. Breebaart, S. van de Par, A. Kohlrausch, E. Schuijers, "High-Quality Parametric Spatial Audio Coding at Low Bitrates", AES 116th Convention, Berlín, Preprint 6072, mayo de 2004

20 [13] E. Schuijers, J. Breebaart, H. Purnhagen, J. Engdegard, "Low Complexity Parametric Stereo Coding", AES 116th Convention, Berlín, Preprint 6073, mayo de 2004.

En un decodificador de audio espacial, la mezcla ascendente multicanal se calcula a partir de una parte de señal directa y una parte de señal difusa, que se deriva por medio de la decorrelación de la parte directa, como ya se mencionó anteriormente. De este modo, en general, la parte difusa tiene una envolvente temporal diferente que la parte directa. El término "envolvente temporal" describe en este contexto la variación de la energía o amplitud de la señal en el tiempo. La envolvente temporal diferente lleva a artefactos (pre y post-ecos, "manchado" temporal) en las señales mezcladas de manera ascendente para señales de entrada que tienen una imagen estéreo amplia y, al mismo tiempo, una estructura de envolvente transitoria. Las señales transitorias son generalmente señales que varían intensamente en un corto periodo de tiempo.

30 Los ejemplos probablemente más importantes para esta clase de señales son las señales de tipo aplauso, que con frecuencia están presentes en las grabaciones en tiempo real.

Para evitar artefactos causados por la introducción de sonido difuso/decorrelacionado con una envolvente temporal inapropiada en la señal de mezcla ascendente, se han propuesto varias técnicas:

35 La solicitud de estadounidense 11/006.492 ("Diffuse Sound Shaping for BCC Schemes and The Like") muestra que la calidad perceptual de señales transitorias críticas puede mejorarse conformando la envolvente temporal de la señal difusa para coincidir con la envolvente temporal de la señal directa.

Este enfoque ya se ha introducido en la tecnológica de MPEG envolvente por medio de diferentes herramientas, tales como la "conformación de envolvente temporal" (TES) y el "procesamiento temporal" (TP). Puesto que la envolvente temporal objetivo de la señal difusa se deriva a partir de la envolvente de la señal de mezcla descendente transmitida, este método no requiere transmitir información secundaria adicional. Sin embargo, como consecuencia, la estructura fina temporal del sonido difuso es la misma para todos los canales de salida. Como la parte de señal directa, que se deriva directamente de la señal de mezcla descendente transmitida, también tiene una envolvente temporal similar, este método puede mejorar la calidad perceptual de las señales de tipo aplauso en términos de "nitidez", por ejemplo. Sin embargo, como entonces la señal directa y la señal difusa tienen envolventes temporales similares para todos los canales, tales técnicas pueden mejorar la calidad subjetiva de las señales de tipo aplauso, pero no pueden mejorar la distribución espacial de eventos de aplauso únicos en la señal, ya que esto sería posible solamente si un canal reconstruido fuese mucho más intenso al producirse la señal transitoria que en los otros canales, lo cual es imposible teniendo señales que comparten básicamente la misma envolvente temporal.

50 Un método alternativo para superar el problema se describe en la solicitud estadounidense 11/006.482 ("individual Channel Shaping for BCC Schemes and The Like"). Este enfoque emplea información secundaria de banda ancha temporal de grano fino que se transmite por medio del codificador para realizar una conformación temporal fina de la señal directa y la difusa. Evidentemente, este enfoque permite una estructura fina temporal que es individual para cada canal de salida y así puede albergar también señales para las que se producen eventos transitorios solamente en un subconjunto de los canales de salida. Una variación adicional de este enfoque se describe en el documento US 60/726.389 ("Methods for Improved Temporal and Spatial Shaping of Multi-Channel Audio Signals"). Ambos enfoques descritos para mejorar la calidad perceptual de las señales codificadas transitorias comprenden una conformación temporal de la envolvente de la señal difusa prevista para coincidir con una envolvente temporal de señales directas correspondientes.

60 Aunque ambos métodos previamente descritos de la técnica anterior pueden mejorar la calidad subjetiva de las señales de tipo aplauso en términos de nitidez, sólo el último enfoque puede mejorar también la redistribución espacial de la señal reconstruida. Entonces, la calidad subjetiva de las señales de aplauso sintetizadas sigue siendo insatisfactoria, debido a que la conformación temporal de la combinación de sonido tanto seco como difuso lleva a distorsiones características (los ataques de los aplausos individuales se perciben o bien como no "cerrados", cuando sólo se realiza una conformación temporal flexible, o bien las distorsiones se introducen si se aplica conformación a la señal con una resolución temporal muy alta). Esto se vuelve evidente, cuando una señal difusa es simplemente una copia retardada de la señal directa. Entonces, la señal difusa mezclada con la señal directa es más susceptible de tener una composición

espectral diferente que la señal directa. De este modo, aun cuando la envolvente se ajusta a escala para coincidir con la envolvente de la señal directa, las diferentes contribuciones espaciales, que no se originan directamente de la señal original, estarán presentes en la señal reconstruida. Las distorsiones introducidas pueden volverse aún peores, cuando la parte de señal difusa se mejora (se hace más fuerte) durante la reconstrucción, cuando la señal difusa se ajusta a escala para coincidir con la envolvente de la señal directa.

Numerosas publicaciones se refieren al problema de codificar y decodificar apropiadamente señales multicanal.

La solicitud de patente internacional WO 2004/097794 A2 se refiere al procesamiento avanzado de señales de audio multicanal basadas en un banco de filtros modulado exponencialmente complejo y métodos de señalización de tiempo adaptiva. Un sintetizador para generar una señal de decorrelación basándose en una señal de entrada es operativo sobre una pluralidad de señales de subbanda, en el que una señal de subbanda incluye una secuencia de al menos dos muestras de subbanda. El sintetizador incluye fases de filtrado para filtrar cada señal de subbanda usando un filtro de reverberación para obtener una pluralidad de señales reverberadas, en el que una pluralidad de señales de subbanda reverberadas juntas representan una señal de decorrelación. Esta señal de decorrelación se usa para reconstruir una señal basándose en una señal estéreo paramétricamente codificada que consiste en una señal mono y una medida de coherencia.

La publicación "Parametric multi-channel audio coding: synthesis of coherence cues" Faller C., enero de 2006, IEEE transactions on audio, speech and language processing, IEEE service center, N.Y., EE.UU., páginas 299 a 310, XP007900793, página 303 a página 305, se refiere a formas de sintetizar indicadores de coherencia. Para dicho propósito, se usan filtros de decorrelación que modelan la reverberación tardía con respuestas de impulso que corresponden a varios 100 ms, dando como resultado la habilidad del esquema para generar sonido difuso que suena de forma natural.

El documento "MPEG4-EXT2: CE on low complexity parametric stereo", OOMEN W *et al.*, diciembre de 2003, norma internacional ISO/IEC JTC1/SC 29/WG11, describe un concepto de cómo puede disminuirse la complejidad en el análisis y síntesis de estéreo paramétrico mediante el uso de bancos de filtros QMF en lugar de filtros FFT.

La solicitud de patente estadounidense 2005/00583004 A1 se refiere a la codificación BCC y en particular a esquemas de codificación, en los que uno o más de los canales de entrada se transmiten como canales no modificados que no se mezclan de manera descendente en el codificador BCC y no se mezclan de manera ascendente en el decodificador BCC.

### 30 **Sumario de la invención**

Es el objetivo de la presente invención proporcionar un concepto de conformación de señal mejorada en reconstrucción multicanal.

Este objetivo se logra por medio de un aparato según las reivindicaciones 1 ó 29, un método según la reivindicación 28 y un programa informático según la reivindicación 30.

La presente invención se basa en el descubrimiento de que un canal de salida reconstruido, reconstruido con un reconstructor multicanal usando al menos un canal de mezcla descendente derivado de la mezcla descendente de una pluralidad de canales originales y usando una representación paramétrica que incluye información adicional sobre una estructura (fina) temporal de un canal original, puede reconstruirse de manera eficaz con alta calidad, cuando se usa un generador para generar una componente de señal directa y una componente de señal difusa basándose en el canal de mezcla descendente. La calidad puede mejorarse esencialmente si sólo se modifica la componente de señal directa de modo que la estructura fina temporal del canal de salida reconstruido se ajusta a una estructura fina temporal deseada, indicada por medio de la información adicional transmitida sobre la estructura fina temporal.

En otras palabras, el ajuste a escala de las partes de señal directa derivadas directamente de la señal de mezcla descendente, apenas introduce artefactos adicionales en el momento en que se produce una señal transitoria. Cuando, como en la técnica anterior, la parte de señal húmeda se ajusta a escala para coincidir con una envolvente deseada, puede darse perfectamente el caso de que la señal transitoria original en el canal reconstruido se enmascare por una señal difusa mejorada mezclada con la señal directa, lo cual se describirá de manera más extensa a continuación.

La presente invención resuelve este problema al ajustar a escala sólo la componente de señal directa, sin dar así oportunidad para introducir artefactos adicionales, a costa de transmitir parámetros adicionales para describir la envolvente temporal dentro de la información secundaria.

Según una realización de la presente invención, los parámetros de ajuste a escala de la envolvente se derivan usando una presentación de la señal directa y difusa con un espectro blanqueado, es decir, donde diferentes partes espectrales de la señal tienen energías casi idénticas. Las ventajas de usar espectros blanqueados son dobles. Por un lado, el uso de un espectro blanqueado como base del cálculo de un factor de ajuste a escala usado para ajustar a escala la señal directa permite la transmisión de sólo un parámetro por ranura de tiempo que incluye la información sobre la estructura temporal. Como es usual en la codificación de audio multicanal que las señales se procesan dentro de numerosas bandas de frecuencia, esta característica ayuda a disminuir el número de información secundaria necesaria adicionalmente y, de este modo, el incremento de la tasa de transmisión de bits para la transmisión del parámetro adicional. Típicamente, otros parámetros tales como ICLD e ICC se transmiten una vez por trama de tiempo y banda de parámetro. Puesto que la cantidad de bandas de parámetro puede ser mayor que 20, es una ventaja principal el tener que transmitir sólo un único parámetro por canal. Generalmente, en la codificación multicanal, la señales se procesan en una estructura de trama, es decir, en entidades que tienen varios valores de muestreo, por ejemplo 1024 por trama. Adicionalmente, como ya se ha mencionado, las señales se dividen en varias partes espectrales antes de procesarse, de modo que finalmente de manera típica se transmite un parámetro ICC e ICLD por trama y parte espectral de la señal. La segunda ventaja del uso de sólo un parámetro tiene una causa física, puesto que las señales transitorias en cuestión tienen naturalmente espectros amplios. Por consiguiente, para tener en cuenta la energía de las señales transitorias

dentro de los canales individuales de manera correcta, es más apropiado usar espectros blanqueados para el cálculo de los factores de ajuste a escala de energía.

En una realización adicional de la presente invención, el concepto de la invención de modificar la componente de señal directa se aplica solamente a una parte espectral de la señal por encima de cierto límite espectral en presencia de señales residuales adicionales. Esto se debe a que las señales residuales junto con la señal de mezcla descendente permiten una reproducción de alta calidad de los canales originales.

Resumiendo, el concepto de la invención está diseñado para proporcionar calidad espacial y temporal mejoradas con respecto a los enfoques de la técnica anterior, evitando los problemas asociados con aquellas técnicas. Por lo tanto, la información secundaria se transmite para describir la estructura de envolvente de tiempo fina de los canales individuales y permitir así la conformación temporal/espacial fina de las señales de canal de mezcla ascendente en el lado del decodificador. El método de la invención descrito en este documento se basa en los siguientes descubrimientos/consideraciones:

- Las señales de tipo aplauso pueden verse como compuestas por palmadas cercanas individuales y diferenciadas y un ambiente de tipo ruido que se origina a partir de palmadas lejanas muy densas.

- En un decodificador de audio espacial, la mejor aproximación de las palmadas cercanas en cuanto a la envolvente temporal es la señal directa. Por lo tanto, sólo se procesa la señal directa por el método de la invención.

- Puesto que la señal difusa representa principalmente la parte de ambiente de la señal, cualquier procesamiento sobre una resolución temporal fina es probable que introduzca artefactos de distorsión y modulación (aunque puede lograrse una cierta mejora subjetiva de la "nitidez" del aplauso por medio de tal técnica). Como una consecuencia de estas consideraciones, la señal difusa por tanto no se toca (es decir, no se somete a una conformación de tiempo fina) mediante el procesamiento de la invención).

- No obstante, la señal difusa contribuye al equilibrio de energía de la señal de mezcla ascendente. El método de la invención tiene esto en cuenta al calcular un factor de ajuste a escala de banda ancha modificado a partir de la información transmitida que debe aplicarse solamente a la parte de señal directa. Este factor modificado se elige de modo que la energía total en un intervalo de tiempo dado es la misma dentro de ciertos límites como si se hubiera aplicado el factor original tanto a la parte directa como a la parte difusa de la señal en este intervalo.

- Usando el método de la invención, se obtiene mejor calidad subjetiva de audio si la resolución espectral de los indicadores espaciales se elige como baja (por ejemplo, "ancho de banda completo") para garantizar la conservación de la integridad espectral de las señales transitorias contenidas en la señal. En este caso, el método propuesto no incrementa necesariamente la tasa de transmisión de bits de información espacial secundaria promedio, puesto que la resolución espectral se intercambia de manera segura por la resolución temporal.

La mejora subjetiva de la calidad se logra por medio de la amplificación o amortiguación ("conformación") de la parte seca de la señal a lo largo del tiempo solamente y, de este modo,

- Mejorando la calidad de la señal transitoria al reforzar la parte de señal directa en el sitio de la señal transitoria, mientras se evita una distorsión adicional que se origina de una señal difusa con envolvente temporal inapropiada

- Mejorar la localización espacial mejorando la parte directa con respecto a la parte difusa en el origen espacial de un evento transitorio y amortiguándola con respecto a la parte difusa en posiciones alejadas panorámicas.

#### **Breve descripción de los dibujos**

La figura 1 muestra un diagrama de bloques de un codificador multicanal y un decodificador correspondiente;

la figura 1b muestra un diagrama esquemático de la reconstrucción de señal usando señales decorrelacionadas;

la figura 2 muestra un ejemplo de un reconstructor multicanal de la invención;

la figura 3 muestra un ejemplo adicional de un reconstructor multicanal de la invención;

la figura 4 muestra un ejemplo de representaciones de banda de parámetro usadas para identificar bandas de parámetro diferentes dentro de un esquema de decodificación multicanal;

la figura 5 muestra un ejemplo de un decodificador multicanal de la invención; y

La figura 6 muestra un diagrama de bloques que detalla un ejemplo de un método de la invención de reconstrucción de un canal de salida.

**Descripción detallada de las realizaciones adicionales**

La figura 1 muestra un ejemplo de codificación de datos de audio multicanal según la técnica anterior, para ilustrar más claramente el problema resuelto por el concepto de la invención.

En general, en el lado de un codificador, una señal 10 multicanal original se introduce en el codificador 12 multicanal, derivando información 14 secundaria que indica la distribución espacial de los diversos canales de la señal multicanal original unos respecto a otros. Además de la generación de información 14 secundaria, un codificador 12 multicanal genera una o más señales 16 de suma, mezclándose de manera descendente a partir de la señal multicanal original. Configuraciones populares usadas ampliamente son las denominadas configuraciones 5-1-5 y 5-2-5. En la configuración 5-1-5 el codificador genera una única señal 16 de suma monofónica a partir de cinco canales de entrada y, por tanto, un correspondiente decodificador 18 tiene que generar cinco señales reconstruidas de una señal 20 multicanal reconstruida. En la configuración 5-2-5, el codificador genera dos canales mezclados de manera descendente a partir de cinco canales de entrada, conteniendo típicamente el primer canal de los canales mezclados de manera descendente información sobre un lado izquierdo o un lado derecho y conteniendo el segundo canal de los canales mezclados de manera descendente información sobre el otro lado.

Los parámetros de muestreo que describen la distribución espacial de los canales originales son, como se indica por ejemplo en la figura 1, los parámetros ICLD e ICC previamente introducidos.

Puede observarse que dentro del análisis que deriva la información 14 secundaria, las muestras de los canales originales de la señal 10 multicanal se procesan típicamente en dominios de subbanda que representan un intervalo específico de frecuencias de los canales originales. Un intervalo de frecuencias individual se indica mediante  $K$ . En algunas aplicaciones, los canales de entrada pueden filtrarse por medio de un banco de filtros híbrido antes del procesamiento, es decir, las bandas de parámetro  $K$  pueden subdividirse adicionalmente, indicándose cada subdivisión como  $k$ .

Adicionalmente, el procesamiento de los valores de muestreo que describen un canal original se hace por tramas dentro de cada banda de parámetro individual, es decir, varias muestras consecutivas forman una trama de duración finita. Los parámetros BCC mencionados anteriormente describen normalmente una trama completa.

Un parámetro relacionado de alguna manera con la presente invención y ya conocido en la técnica es el parámetro ICLD, que describe la energía contenida dentro de una trama de señal de un canal con respecto a las tramas correspondientes de otros canales de la señal multicanal original.

Comúnmente, la generación de canales adicionales para derivar una reconstrucción de una señal multicanal a partir de una señal de suma transmitida solamente se logra con la ayuda de señales decorrelacionadas, que se derivan de la señal de suma usando decorreladores o reverberadores. En una aplicación típica, la frecuencia de muestreo discreta puede ser de 44,100 kHz, de modo que una única muestra representa un intervalo de longitud finita de aproximadamente 0,02 ms de un canal original. Puede observarse que, usando bancos de filtros, la señal se divide en numerosas partes de señal, representando cada una un intervalo finito de frecuencias de la señal original. Para compensar un posible incremento de los parámetros que describen el canal, la resolución de tiempo normalmente se reduce, de manera que una parte de tiempo de longitud finita descrita por una única muestra dentro de un dominio de banco de filtros puede incrementarse a más de 0,5 ms. La longitud de trama típica puede variar entre 10 y 15 ms.

La derivación de la señal decorrelacionada puede hacer uso de diferentes estructuras de filtro y/o retardos o combinaciones de los mismos sin limitar el alcance de la invención. Puede observarse adicionalmente que no hay que usar necesariamente el espectro completo para derivar las señales decorrelacionadas. Por ejemplo, sólo las partes espectrales por encima de un límite inferior espectral (el valor específico de  $K$ ) de la señal de suma (señal de mezcla descendente) pueden usarse para derivar las señales decorrelacionadas usando retardos y/o filtros. Una señal decorrelacionada describe de este modo generalmente una señal derivada de la señal de mezcla descendente (canal de mezcla descendente) de modo que un coeficiente de correlación, cuando se deriva usando la señal decorrelacionada y el canal de mezcla descendente, se desvía significativamente de la unidad, por ejemplo, en 0,2.

La figura 1b da un ejemplo extremadamente simplificado del proceso de mezcla descendente y reconstrucción durante la codificación de audio multicanal para explicar el gran beneficio del concepto de la invención de ajustar a escala solamente la componente de señal directa durante la reconstrucción de un canal de una señal multicanal. Para la siguiente descripción se asumen algunas simplificaciones. La primera simplificación es que la mezcla descendente de un canal izquierdo y uno derecho es una simple adición de las amplitudes dentro de los canales. La segunda gran simplificación es que la correlación se asume como un simple retardo de la señal completa.

Bajo estos supuestos, van a codificarse una trama de un canal 21a izquierdo y un canal 21b derecho. Como se indica sobre el eje x de las ventanas mostradas, en la codificación de audio multicanal, el procesamiento se realiza típicamente sobre valores de muestreo, muestreados con una frecuencia de muestreo fija. Para facilidad de la explicación, esto se ignorará en el siguiente resumen breve.

Como ya se mencionó, en el lado del codificador, un canal izquierdo y uno derecho se combinan (mezclan de manera descendente) en un canal 22 de mezcla descendente que se transmite al decodificador. En el lado del decodificador, una señal 23 decorrelacionada se deriva del canal 22 de mezcla descendente transmitido, que es la suma del canal 21a izquierdo y del canal 21b derecho en este ejemplo. Como ya se explicó, la reconstrucción del canal izquierdo se realiza entonces a partir de las tramas de señal derivadas del canal 22 de mezcla descendente y la señal 23 decorrelacionada.

Puede observarse que cada trama individual se somete a un ajuste a escala global antes de la combinación, como se indica por el parámetro ICLD, que relaciona las energías dentro de las tramas individuales de canales individuales con la energía de las tramas correspondientes de los demás canales de una señal multicanal.

Puesto que en el presente ejemplo se asume que energías iguales están contenidas dentro de la trama del canal 21a izquierdo y la trama del canal 21b derecho, el canal 22 de mezcla descendente transmitido y la señal 23 decorrelacionada se ajustan a escala aproximadamente por un factor de 0,5 antes de la combinación. Esto es, cuando

la mezcla ascendente es igual de simple que la mezcla descendente, es decir, la suma de las dos señales, la reconstrucción del canal 21a izquierdo original es la suma del canal 24a de mezcla descendente ajustado a escala y la señal 24b decorrelacionada ajustada a escala.

Debido a la suma para la transmisión y al ajuste a escala debido al parámetro ICLD, la relación señal a fondo de la señal transitoria disminuirá en un factor de 2 aproximadamente. Adicionalmente, cuando simplemente se suman las dos señales, se introducirá un artefacto adicional de tipo eco en la posición de la estructura transitoria retardada en la señal 24b decorrelacionada ajustada a escala.

Como se indica en la figura 1b, la técnica anterior trata de resolver el problema del eco al ajustar a escala la amplitud de la señal 24b decorrelacionada ajustada a escala, para hacerla coincidir con la envolvente del canal 24b ajustado a escala transmitido, tal como se indica mediante las líneas discontinuas en la trama 24b. Debido al ajuste a escala, la amplitud en la posición de la señal transitoria original en el canal 21a izquierdo puede incrementarse. Sin embargo, la composición espectral de la señal decorrelacionada en la posición del ajuste a escala en la trama 24b es diferente de la composición espectral de la señal transitoria original. Por consiguiente, se introducen artefactos audibles en la señal, aunque la intensidad general de la señal puede reproducirse bien.

La gran ventaja de la presente invención es que la presente invención sólo ajusta a escala una componente de señal directa del canal reconstruido. Puesto que este canal tiene una componente de señal que corresponde a la señal transitoria original con la composición espectral correcta y el tiempo correcto, ajustar a escala solamente el canal de mezcla descendente proporcionará una señal reconstruida que reconstruye el evento transitorio original con alta precisión. Éste es el caso puesto que sólo se mejoran mediante el ajuste a escala partes de señal que tienen la misma composición espectral que la señal transitoria original.

La figura 2 muestra un diagrama de bloques de un ejemplo de un reconstructor multicanal de la invención, para detallar los aspectos principales del concepto de la invención.

La figura 2 muestra un reconstructor 30 multicanal, con un generador 32, un modificador de señal directa y un combinador 36. El generador 32 recibe un canal 38 de mezcla descendente mezclado de manera descendente a partir de una pluralidad de canales originales y una representación 40 de parámetros que incluye información sobre una estructura temporal de un canal original.

El generador genera una componente 42 de señal directa y una componente 44 de señal difusa basándose en el canal de mezcla descendente.

El modificador 34 de señal directa recibe tanto la componente 42 de señal directa como la componente 44 de señal difusa y adicionalmente la representación 40 de parámetros con la información sobre una estructura temporal del canal original. Según la presente invención, el modificador 34 de señal directa modifica solamente la componente 42 de señal directa usando la representación de parámetros para derivar una componente 46 de señal directa modificada.

La componente 46 de señal directa modificada y la componente 44 de señal difusa, que no se altera por el modificador 34 de señal directa, se introducen en el combinador 36 que combina la componente 46 de señal directa modificada y la componente 44 de señal difusa para obtener un canal 50 de salida reconstruido.

Al modificar solamente la componente 42 de señal directa derivada del canal 38 de mezcla descendente transmitido sin reverberación (decorrelación), es posible reconstruir una envolvente de tiempo para el canal de salida reconstruido que coincide en gran medida con una envolvente de tiempo del canal original subyacente sin introducir artefactos adicionales y distorsiones audibles, como en las técnicas de la técnica anterior.

Como se comentará con mayor detalle en la descripción de la figura 3, la conformación de envolvente de la invención restaura la envolvente de banda ancha de la señal de salida sintetizada. Comprende un procedimiento modificado de mezcla ascendente, seguido del aplanamiento de la envolvente y la reconformación de la parte de señal directa de cada canal de salida. Para la reconformación, se usa información secundaria de envolvente de banda ancha paramétrica contenida en el flujo de bits de la representación de parámetros. Esta información consiste, según una realización de la presente invención, en tasas (envRatio) que relacionan la envolvente de la señal de mezcla descendente transmitida con la envolvente de la señal de canal de entrada original. En el decodificador, se derivan factores de ganancia a partir de estas tasas para su aplicación a la señal directa en cada ranura de tiempo en una trama de un canal de salida dado. La parte de sonido difuso de cada canal no se altera según el concepto de la invención.

La realización preferida de la presente invención mostrada en el diagrama de bloques de la figura 3 es un reconstructor 60 multicanal modificado para ajustarse al flujo de señal del decodificador de un decodificador MPEG espacial.

El reconstructor 60 multicanal comprende un generador 62 para generar una componente 64 de señal directa y una componente 66 de señal difusa usando un canal 68 de mezcla descendente derivado mediante la mezcla descendente de una pluralidad de canales originales y una representación 70 de parámetros con información sobre las propiedades espaciales de los canales originales de la señal multicanal, tal como se usa en la codificación MPEG. El reconstructor 60 multicanal comprende adicionalmente un modificador 68 de señal directa, que recibe la componente 64 de señal directa, la componente 66 de señal difusa, la señal 69 de mezcla descendente e información 72 secundaria de envolvente adicional como entradas.

El modificador de señal directa proporciona en su salida 73 del modificador la componente de señal directa modificada, modificada tal como se describe con mayor detalle a continuación.

El combinador 74 recibe la componente de señal directa modificada y la componente de señal difusa para obtener el canal 76 de salida reconstruido.

Como se muestra en la figura, la presente invención puede implementarse fácilmente en entornos multicanal ya existentes. La aplicación general del concepto de la invención en un esquema de codificación de este tipo podría encenderse y apagarse según algunos parámetros transmitidos adicionalmente dentro del flujo de bits de parámetros. Por ejemplo, podría introducirse una bandera adicional **bsTempShapeEnable**, que indica, cuando se establece en 1, que se requiere del uso del concepto de la invención.

Adicionalmente, podría introducirse una bandera adicional especificando de manera específica la necesidad de la aplicación del concepto de la invención sobre un canal basándose en el canal. Por lo tanto, puede usarse una bandera adicional, denominada por ejemplo **bsEnvShapeChannel**. Esta bandera, disponible para cada canal individual, puede entonces indicar el uso del concepto de la invención, cuando se establece en 1.

5 Puede observarse además que para facilidad en la presentación, sólo se describe una configuración de dos canales en la figura 3. Por su puesto, la presente invención no pretende limitarse a una configuración de dos canales solamente. Por otra parte, cualquier configuración de canales puede usarse en conexión con el concepto de la invención. Por ejemplo, cinco o siete canales de entrada pueden usarse en conexión con la conformación de envolvente avanzada de la invención.

10 Cuando se aplica el concepto de la invención en el esquema de codificación MPEG, como se indica en la figura 3, y la aplicación del concepto de la invención se señala al establecer **bsTempShapeEnable** igual a 1, las componentes de la señal difusa y la señal directa se sintetizan por separado por el generador 62 usando una mezcla posterior modificada en el dominio de subbanda híbrido según la siguiente fórmula:

$$y_{directa}^{n,k} = M^{n,k} w_{directa}^{n,k} \quad 0 \leq k < K$$

$$y_{difusa}^{n,k} = M^{n,k} w_{difusa}^{n,k} \quad 0 \leq k < K$$

15 Aquí y en los siguiente párrafos, el vector  $W_{m,k}$  describe el vector de  $n$  parámetros de subbanda híbridos para la  $k$ -ésima subbanda del dominio de subbanda. Como se indica por la ecuación anterior, los parámetros de señal directa y difusa  $y$  se derivan por separado en la mezcla ascendente. Las salidas directas contienen la componente de señal directa y la señal residual, que es una señal que puede estar presente adicionalmente en la codificación MPEG. Las salidas difusas suministran solamente la señal difusa. Según el concepto de la invención, sólo se procesa adicionalmente la componente de señal directa mediante la conformación de envolvente guiada (la conformación de envolvente de la invención).

20 El proceso de conformación de envolvente emplea una operación de extracción de envolvente sobre diferentes señales. El proceso de extracción de envolventes que tiene lugar dentro del modificador 68 de señal directa se describe con mayor detalle en los siguientes párrafos ya que es un paso obligado antes de la aplicación de la modificación de la invención a la componente de señal directa.

25 Como ya se mencionó, dentro del dominio de subbanda híbrido, las subbandas se indican como  $k$ . Varias subbandas  $k$  también pueden organizarse en bandas de parámetro  $k$ .

30 La asociación de subbandas con las bandas de parámetro que subyace a la realización de la presente invención comentada más adelante, se da en la tabla de la figura 4.

En primer lugar, para cada ranura en un trama, las energías  $E_{ranura}^k$  de ciertas bandas de parámetro  $k$  se calculan siendo  $y^{n,k}$  una señal de entrada de subbanda híbrida.

$$E_{ranura}^k(n) = \sum_I y^{n,k} (y^{n,k})^* \quad \bar{k} = \{k | \bar{\kappa}(k) = \kappa\} \quad \forall \kappa_{inicio} < \kappa < \kappa_{detencion}$$

35 con  $\kappa_{inicio}=10$  y  $\kappa_{detencion}=18$ .

La suma incluye atribuir todas las  $\bar{k}$  a una banda de parámetro  $\kappa$  según la Tabla A.1.

A continuación, se calcula un promedio de energía a largo plazo  $\bar{E}_{ranura}^k$  para cada banda de parámetro como:

$$\bar{E}_{ranura}^k(n) = (1 - \alpha) E_{ranura}^k(n) + \alpha \bar{E}_{ranura}^k(n-1)$$

$$\alpha = \exp\left(-\frac{64}{0.4 \cdot 44100}\right)$$

40 Con  $\alpha$  siendo un factor de equilibrio que corresponde a un paso bajo IIR de primer orden (con una constante de tiempo de 400 ms aproximadamente) y  $n$  indica el índice de la ranura de tiempo. La energía (banda ancha) promedio total suavizada

$\bar{E}_{total}$  se calcula como

$$\bar{E}_{total}(n) = (1 - \alpha) E_{total}(n) + \alpha \bar{E}_{total}(n-1)$$

45 con

$$E_{total}(n) = \frac{1}{\kappa_{detencion} - \kappa_{inicio} + 1} \sum_{\kappa=\kappa_{inicio}}^{\kappa_{max}} E_{ranura}^k(n)$$

$$\alpha = \exp\left(-\frac{64}{0.4 \cdot 44100}\right)$$

Como puede verse a partir de las fórmulas anteriores, la envolvente temporal se suaviza antes de que los factores de ganancia se deriven a partir de la representación suavizada de los canales. El suavizado significa generalmente la derivación de una representación suavizada a partir de un canal original con gradientes disminuidos.

- 5 Como puede verse a partir de las fórmulas anteriores, la operación de blanqueado descrita a continuación se basa en estimaciones de la energía total temporalmente suavizada y en estimaciones de la energía suavizada en las subbandas, garantizando así mayor estabilidad en las estimaciones de envolvente finales.  
La relación de estas energías se determina para obtener los pesos para una operación de blanqueado espectral:

$$w^k(n) = \frac{\bar{E}_{total}(n)}{\bar{E}_{ranura}^k(n) + \varepsilon}$$

- 10 La estimación de envolvente de banda ancha se obtiene por medio de la suma de las contribuciones ponderadas de las bandas de parámetro, normalización sobre un promedio de energía a largo plazo y cálculo de la raíz cuadrada

$$Env(n) = \sqrt{\frac{EnvAbs(n)}{\bar{Env}(n)}}$$

con

$$EnvAbs(n) = \sum_{k=k_{inicio}}^{k_{detencion}} w^k(n) \cdot E_{ranura}^k(n)$$

- 15 
$$\bar{Env}(n) = (1 - \beta) EnvAbs(n) + \beta \bar{Env}(n - 1)$$

$$\beta = \exp\left(-\frac{64}{0.04 \cdot 44100}\right)$$

donde  $\beta$  es un factor de equilibrio que corresponde a un paso bajo IIR de primer orden (una constante de tiempo de 40 ms aproximadamente).

- 20 Las medidas de energía o amplitud blanqueadas espectralmente se usan como la base del cálculo de los factores de ajuste a escala. Como puede verse a partir de las fórmulas anteriores, el blanqueado espectral significa la alteración del espectro de modo que la misma energía o amplitud media está contenida dentro de cada banda espectral de la representación de los canales de audio. Esto es enormemente ventajoso puesto que las señales transitorias en cuestión tienen espectros muy amplios, de modo que es necesario usar la información completa sobre la totalidad del espectro disponible para el cálculo de los factores de ganancia para no suprimir las señales transitorias con respecto a otras señales no transitorias. En otras palabras, las señales blanqueadas espectralmente son señales que tienen una energía aproximadamente igual en diferentes bandas espectrales de su representación espectral.

- 25 El modificador de señal directa de la invención modifica la componente de señal directa. Como ya se ha mencionado, el procesamiento puede restringirse a algunos índices de subbanda, comenzando con un índice de inicio, en presencia de señales residuales transmitidas. Adicionalmente, el procesamiento puede restringirse generalmente a índices de subbanda por encima de un índice umbral.

- 30 El proceso de conformación de envolvente consiste en un aplanamiento de la envolvente de sonido directo para cada canal de salida seguido de una reconformación hacia una envolvente objetivo. Esto da como resultado la aplicación de una curva de ganancia a la señal directa de cada canal de salida, si se señala **bsEnvShapeChannel=1** para este canal en la información secundaria.

- 35 El procesamiento se realiza para ciertas subbandas híbridas  $k$  solamente:  
 $k > 7$

En presencia de señales residuales transmitidas,  $k$  se elige para comenzar por encima de la banda residual más alta implicada en la mezcla ascendente del canal en cuestión.

- 40 Para la configuración 5-1-5, la envolvente objetivo se obtiene estimando la envolvente de la mezcla descendente transmitida  $Env_{Dmx}$ , como se describió en la sección previa, y ajustarla a escala posteriormente con las relaciones de envolvente recuadradas y transmitidas del codificador  $envRatio_{ch}$ .

- 45 Entonces, una curva de ganancia  $g_{ch}(n)$  para todas las ranuras en un trama se calcula para cada canal de salida estimando su envolvente  $Env_{ch}$  y relacionándola con la envolvente objetivo. Finalmente, esta curva de ganancia se convierte en una curva de ganancia efectiva para ajustar a escala solamente la parte directa del canal convertido de manera ascendente:

$$relacion_{ch}(n) = \min\left(4, \max\left(0.25, g_{ch} + ampRatio_{ch}(n) \cdot (g_{ch} - 1)\right)\right)$$

con

$$g_{ch}(n) = \frac{envRatio_{ch}(n) \cdot Env_{Dmx}(n)}{Env_{ch}(n)}$$

$$ampRatio_{ch}(n) = \frac{\sum_k |y_{ch,difusa}^{n,k}|}{\sum_k |y_{ch,directa}^{n,k}| + \varepsilon}$$

$$ch \in \{L, Ls, C, R, Rs\}$$

5 Para la configuración 5-2-5 la envolvente objetivo para  $L$  y  $Ls$  se deriva a partir de la envolvente de la señal de mezcla descendente transmitida del canal izquierdo  $Env_{DmxL}$ , para  $R$  y  $Rs$  se utiliza la envolvente de mezcla descendente transmitida del canal derecho  $Env_{DmxR}$ . El canal central se deriva a partir de la suma de las envolventes de la señal de mezcla descendente transmitida izquierda y derecha.

10 La curva de ganancia se calcula para cada canal de salida estimando su envolvente  $Env^{L,Ls,C,R,Rs}$  y relacionarla con la envolvente objetivo. En un segundo paso esta curva de ganancia se convierte en una curva de ganancia efectiva para ajustar a escala solamente la parte directa del canal mezclado de manera ascendente:

$$relación_{ch}(n) = \min(4, \max(0.25, g_{ch} + ampRatio_{ch}(n) \cdot (g_{ch} - 1)))$$

con

$$ampRatio_{ch}(n) = \frac{\sum_k |y_{ch,difusa}^{n,k}|}{\sum_k |y_{ch,directa}^{n,k}| + \varepsilon}, \quad ch \in \{L, Ls, C, R, Rs\}$$

15

$$g_{ch}(n) = \frac{envRatio_{ch}(n) \cdot Env_{DmxL}(n)}{Env_{ch}(n)}, \quad ch \in \{L, Ls\}$$

$$g_{ch}(n) = \frac{envRatio_{ch}(n) \cdot Env_{DmxR}(n)}{Env_{ch}(n)}, \quad ch \in \{R, Rs\}$$

$$g_{ch}(n) = \frac{envRatio_{ch}(n) \cdot 0.5(Env_{DmxL}(n) + Env_{DmxR}(n))}{Env_{ch}(n)}, \quad ch \in \{C\}$$

Para todos los canales, la curva de ganancia de ajuste de envolvente se aplica si  $bsEnvShapeChannel=1$ .

$$\bar{y}_{ch,directa}^k(n) = relación_{ch}(n) \cdot y_{ch,directa}^k(n), \quad ch \in \{L, Ls, C, R, Rs\}$$

20 Si no, la señal directa simplemente se copia

$$\bar{y}_{ch,directa}^k(n) = y_{ch,directa}^k(n), \quad ch \in \{L, Ls, C, R, Rs\}$$

Finalmente, la componente de señal directa modificada de cada canal individual tiene que combinarse con la componente de señal difusa del canal individual correspondiente dentro del dominio de subbanda híbrido según la siguiente ecuación:

$$y_{ch}^{n,k} = \bar{y}_{ch,directa}^{n,k} + y_{ch,difusa}^{n,k}, \quad ch \in \{L, Ls, C, R, Rs\}$$

25

Como puede verse a partir de los párrafos anteriores, el concepto de la invención enseña a mejorar la calidad perceptual y la distribución espacial de señales de tipo aplauso en un decodificador de audio espacial. La mejora se complementa mediante la derivación de factores de ganancia con una granularidad temporal de ajuste a escala fina para ajustar a escala solamente la parte directa de la señal de mezcla ascendente espacial. Estos factores de ganancia se derivan esencialmente de información secundaria transmitida y mediciones de nivel o energía de la señal directa y difusa en el codificador.

30

Puesto que el ejemplo anterior describe particularmente el cálculo basado en mediciones de amplitud, debe observarse que el método de la invención no está restringido a esto, sino que puede calcularse también con, por ejemplo, mediciones de energía u otras magnitudes apropiadas para describir una envolvente temporal de una señal.

35

El ejemplo anterior describe el cálculo para las configuraciones de 5-1-5 y 5-2-5 canales. Naturalmente, el principio expuesto anteriormente podría aplicarse de manera análoga para, por ejemplo, configuraciones de 7-2-7 y 7-5-7 canales.

- La figura 5 muestra un ejemplo de un decodificador 100 de audio multicanal de la invención, que recibe un canal 102 de mezcla descendente derivado mezclando de manera descendente una pluralidad de canales de una señal multicanal original y una representación 104 de parámetros que incluye información sobre una estructura temporal de los canales originales (izquierdo frontal, derecho frontal, izquierdo posterior y derecho posterior) de la señal multicanal original. El decodificador 100 multicanal tiene un generador 106 para generar una componente de señal directa y una componente de señal difusa para cada uno de los canales originales que subyacen al canal de mezcla 102 descendente. El decodificador 100 multicanal comprende además cuatro modificadores 108a a 108d de señal directa de la invención para cada uno de los canales que van a reconstruirse, de modo que el decodificador multicanal emite cuatro canales de salida (izquierdo frontal, derecho frontal, izquierdo posterior y derecho posterior) en sus salidas 112.
- 5 Aunque el decodificador multicanal de la invención se ha detallado usando una configuración de ejemplo de cuatro canales originales para su reconstrucción, el concepto de la invención puede implementarse en esquemas de audio multicanal con cantidades arbitrarias de canales.
- 10 La figura 6 muestra un diagrama de bloques, que detalla el método de la invención para generar un canal de salida reconstruido.
- 15 En una etapa de generación 110, una componente de señal directa y una componente de señal difusa se derivan del canal de mezcla descendente. En una etapa de modificación 112 la componente de señal directa se modifica usando parámetros de la representación de parámetros con información sobre una estructura temporal de un canal original. En una etapa de combinación 114, la componente de señal directa modificada y la componente de señal difusa se combinan para obtener un canal de salida reconstruido.
- 20 Dependiendo de ciertos requisitos de implementación de los métodos de la invención, los métodos de la invención pueden implementarse en hardware o software. La implementación puede llevarse a cabo usando un medio de almacenamiento digital, en particular un disco, un DVD o un CD, con señales de control legibles electrónicamente almacenadas en los mismos, que actúan conjuntamente con un sistema informático programable de manera que se realizan los métodos de la invención. En general, la presente invención es, por lo tanto, un producto de programa informático con un código de programa almacenado en un soporte legible por máquina, siendo el código de programa operativo para realizar los métodos de la invención cuando el producto de programa informático se ejecuta en un ordenador. En otras palabras, los métodos de la invención son, por consiguiente, un programa informático con un código de programa para llevar a cabo al menos uno de los métodos de la invención cuando el programa informático se ejecuta en un ordenador.
- 25 Aunque lo anterior se ha mostrado y descrito particularmente en relación a las realizaciones particulares de la misma, los expertos en la técnica entenderán que pueden realizarse otros diversos cambios en la forma y detalle sin alejarse del alcance que comprenden las reivindicaciones que siguen.
- 30

## REIVINDICACIONES

- 5 1. Reconstructor (30; 60) multicanal para generar un canal (50; 76) de salida reconstruido usando al menos un canal (38; 68) de mezcla descendente derivado mezclando de manera descendente una pluralidad de canales originales y usando una representación (40; 72) de parámetros, incluyendo la representación (40; 72) de parámetros información sobre una estructura temporal de un canal original, que comprende:
- 10 un generador (32; 62) para generar una componente (42; 64) de señal directa y una componente (44; 66) de señal difusa para el canal (50; 76) de salida reconstruido, basándose en el canal (38; 68) de mezcla descendente;
- 15 un modificador (34; 69) de señal directa para modificar la componente (42; 64) de señal directa usando la representación (40; 72) de parámetros, usando la información sobre la estructura temporal del canal original; y
- un combinador (36; 74) para combinar la componente (46) de señal directa modificada y la componente (44; 66) de señal difusa para obtener el canal (50; 76) de salida reconstruido, en el que el modificador de señal directa no altera la componente de señal difusa.
2. Reconstructor multicanal según la reivindicación 1, en el que el generador (32; 62) es operativo para generar la componente (42; 64) de señal directa usando solamente componentes del canal (38; 68) de mezcla descendente.
- 20 3. Reconstructor (30; 60) multicanal según las reivindicaciones 1 ó 2, en el que el generador (32; 62) es operativo para generar la componente (44; 66) de señal difusa usando una parte filtrada y/o retardada del canal (38; 68) de mezcla descendente.
4. Reconstructor (30; 60) multicanal según cualquiera de las reivindicaciones 1 a 3, en el que el modificador (34; 69) de señal directa es operativo para usar información sobre la estructura temporal del canal original que indica la energía contenida en el canal original dentro de una parte de tiempo de longitud finita del canal original.
- 25 5. Reconstructor (30; 60) multicanal según cualquiera de las reivindicaciones 1 a 3, en el que el modificador (34; 69) de señal directa es operativo para usar información sobre la estructura temporal del canal original que indica una amplitud media del canal original dentro de una parte de tiempo de longitud finita del canal original.
- 30 6. Reconstructor (30; 60) multicanal según cualquiera de las reivindicaciones 1 a 5, en el que el combinador (36; 74) es operativo para añadir la componente (46) de señal directa modificada y la componente (44; 66) de señal difusa para obtener la señal reconstruida.
- 35 7. Reconstructor multicanal según cualquiera de las reivindicaciones 1 a 6, en el que el reconstructor multicanal es operativo para usar un primer canal de mezcla descendente que tiene información sobre un lado izquierdo de la pluralidad de canales originales y un segundo canal (38; 68) de mezcla descendente que tiene información sobre un lado derecho de la pluralidad de canales originales, en el que un primer canal (50; 76) de salida reconstruido para un lado izquierdo se combina usando solamente componentes de señal directa y difusa generadas a partir del primer canal de mezcla descendente y en el que un segundo canal de salida reconstruido para un lado derecho se combina usando componentes de señal directa y difusa generadas solamente a partir de la segunda señal de mezcla descendente.
- 40 8. Generador (30; 60) multicanal según cualquiera de las reivindicaciones 1 a 7, en el que el modificador (34; 68) de señal directa es operativo para modificar la señal directa para partes de tiempo de longitud finita que son menores que partes de tiempo de trama de información paramétrica adicional dentro de la representación (40; 72) de parámetros, en el que la información paramétrica adicional se usa por el generador (32; 62) para generar las componentes de señal directa y difusa.
- 45 9. Generador (30; 60) multicanal según la reivindicación 8, en el que el generador (32; 62) es operativo para usar información paramétrica adicional que tiene información sobre la energía del canal original con respecto a otros canales de la pluralidad de canales originales.
- 50 10. Reconstructor (30; 60) multicanal según cualquiera de las reivindicaciones anteriores, en el que el modificador (34; 68) de señal directa es operativo para usar información sobre una estructura temporal del canal original que relaciona una estructura temporal del canal original con una estructura temporal del canal (38; 68) de mezcla descendente.
- 55 11. Reconstructor (30; 60) multicanal según cualquiera de las reivindicaciones anteriores, en el que la información sobre la estructura temporal del canal original y la información sobre la estructura temporal del canal de mezcla descendente tiene una medida de energía o de amplitud.
12. Reconstructor (30; 60) multicanal según cualquiera de las reivindicaciones anteriores, en el que el modificador (34; 68) de señal directa es operativo adicionalmente para derivar información temporal de mezcla descendente sobre la estructura temporal del canal (38; 68) de mezcla descendente.
- 60 13. Reconstructor (30; 60) multicanal según la reivindicación 12, en el que el modificador (34; 68) de señal directa es operativo para derivar información temporal de mezcla descendente que indica la energía contenida en el canal (38; 68) de mezcla descendente dentro de un intervalo de tiempo de longitud finita o una medida de amplitud para el intervalo de tiempo de longitud finita.
- 65 14. Reconstructor (30; 60) multicanal según las reivindicaciones 12 ó 13, en el que el modificador (34; 68) de señal directa es operativo adicionalmente para derivar una estructura temporal objetivo para el canal (38; 68) de mezcla descendente reconstruido usando la información temporal de mezcla descendente y la información sobre la estructura temporal del canal original.

15. Reconstructor (30; 60) multicanal según cualquiera de las reivindicaciones 12 a 14, en el que el modificador (34; 68) de señal directa es operativo para derivar la información temporal de mezcla descendente para una parte espectral del canal (38; 68) de mezcla descendente por encima de un límite inferior espectral.
- 5 16. Reconstructor (30; 60) multicanal según cualquiera de las reivindicaciones 12 a 15, en el que el modificador (34; 68) de señal directa es operativo adicionalmente para blanquear espectralmente el canal (38; 68) de mezcla descendente y para derivar la información temporal de mezcla descendente usando el canal (38; 68) de mezcla descendente blanqueado espectralmente.
- 10 17. Reconstructor (30; 60) multicanal según cualquiera de las reivindicaciones 12 a 16, en el que el modificador (34; 68) de señal directa es operativo adicionalmente para derivar una representación suavizada del canal (38; 68) de mezcla descendente y para derivar la información temporal de mezcla descendente a partir de la representación suavizada del canal de mezcla descendente.
- 15 18. Reconstructor (30; 60) multicanal según la reivindicación 17, en el que el modificador (34; 68) de señal directa es operativo para derivar la representación suavizada filtrando el canal (38; 68) de mezcla descendente con un filtro paso bajo de primer orden.
- 20 19. Reconstructor (30; 60) multicanal según cualquiera de las reivindicaciones anteriores, en el que el modificador (34; 68) de señal directa es operativo adicionalmente para derivar información sobre una estructura temporal de una combinación de la componente de señal directa y la componente de señal difusa.
- 25 20. Reconstructor (30; 60) multicanal según la reivindicación 19, en el que el modificador (34; 68) de señal directa es operativo para blanquear espectralmente la combinación de las componentes de señal directa y señal difusa y para derivar la información sobre la estructura temporal de la combinación de las componentes de señal directa y señal difusa usando las componentes de señales directa y difusa blanqueadas espectralmente.
- 30 21. Reconstructor (30; 60) multicanal según las reivindicaciones 19 ó 20, en el que el modificador (34; 68) de señal directa es operativo adicionalmente para derivar una representación suavizada de la combinación de las componentes de señal directa y difusa y para derivar la información sobre la estructura temporal de la combinación de las componentes de señal directa y difusa a partir de la representación suavizada de la combinación de las componentes de señal directa y difusa.
- 35 22. Reconstructor (30; 60) multicanal según la reivindicación 21, en el que el modificador (34; 68) de señal directa es operativo para derivar la representación suavizada de la combinación de las componentes de señal directa y difusa filtrando las componentes de señal directa y difusa con un filtro paso bajo de primer orden.
- 40 23. Reconstructor (30; 60) multicanal según cualquiera de las reivindicaciones anteriores, en el que el modificador (34; 68) de señal directa es operativo para usar información sobre la estructura temporal del canal original que representa una relación de la energía o amplitud para un intervalo de tiempo de longitud finita del canal original y la energía o amplitud para el intervalo de tiempo de longitud finita del canal (38; 68) de mezcla descendente.
- 45 24. Reconstructor (30; 60) multicanal según cualquiera de las reivindicaciones anteriores, en el que el modificador (34; 68) de señal directa es operativo para derivar una estructura temporal objetivo para el canal (50; 76) de salida reconstruido usando el canal (38; 68) de mezcla descendente y la información sobre la estructura temporal.
- 50 25. Reconstructor (30; 60) multicanal según la reivindicación 23, en el que el modificador (34; 68) de señal directa es operativo para modificar la componente de señal directa de tal modo que una estructura temporal del canal (50; 76) de salida reconstruido iguale la estructura temporal objetivo dentro de un rango de tolerancia.
- 55 26. Reconstructor (30; 60) multicanal según la reivindicación 24, en el que el modificador (34; 68) de señal directa es operativo para derivar un factor de ajuste a ajuste a escala intermedio, siendo el factor de ajuste a escala intermedio tal que la estructura temporal del canal (50; 76) de salida reconstruido iguale la estructura temporal objetivo dentro del rango de tolerancia, cuando el canal (50; 76) de salida reconstruido se combina usando las componentes de señal directa ajustadas a escala con el factor de ajuste a escala intermedio y la componente de señal difusa ajustada a escala con el factor de ajuste a escala intermedio.
- 60 27. Reconstructor (30; 60) multicanal según la reivindicación 25, en el que el modificador (34; 68) de señal directa es operativo adicionalmente para derivar un factor de ajuste a escala final usando el factor de ajuste a escala intermedio y las componentes de señal directa y difusa de tal modo que la estructura temporal del canal (50; 76) de salida reconstruido iguale la estructura temporal objetivo dentro del rango de tolerancia, cuando el canal (50; 76) de salida reconstruido se combina usando la componente de señal difusa y la componente de señal directa ajustadas a escala usando el factor de ajuste a escala final.
- 65 28. Método para generar un canal (50; 76) de salida reconstruido usando al menos un canal (38; 68) de mezcla descendente derivado mezclando de manera descendente una pluralidad de canales originales y usando una representación (40; 72) de parámetros, incluyendo la representación (40; 72) de parámetros información sobre una estructura temporal de un canal original, comprendiendo el método:  
generar una componente de señal directa y una componente de señal difusa para el canal (50; 76) de salida reconstruido, basándose en el canal (38; 68) de mezcla descendente;  
modificar la componente de señal directa usando la representación (40; 72) de parámetros, usando la información sobre la estructura temporal del canal original; y  
combinar la componente (46) de señal directa modificada y la componente de señal difusa para obtener el canal (50; 76) de salida reconstruido, en el que la etapa de modificar no altera la componente de señal difusa.
29. Decodificador de audio multicanal para generar una reconstrucción de una señal multicanal usando al

5

menos un canal (38; 68) de mezcla descendente derivado mezclando de manera descendente una pluralidad de canales originales y usando una representación (40; 72) de parámetros, incluyendo la representación (40; 72) de parámetros información sobre una estructura temporal de un canal original, el decodificador de audio multicanal, que comprende un reconstructor multicanal según las reivindicaciones 1 a 27.  
30. Programa informático con un código de programa para ejecutar el método según la reivindicación 28, cuando se ejecuta en un ordenador.

FIG 1

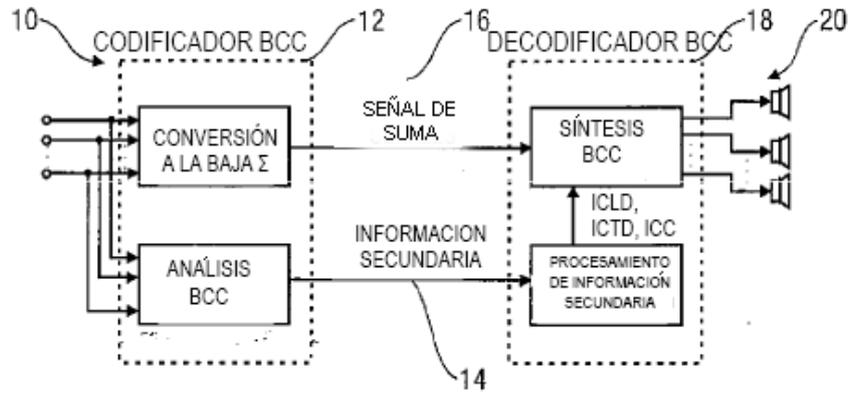


FIG 1b

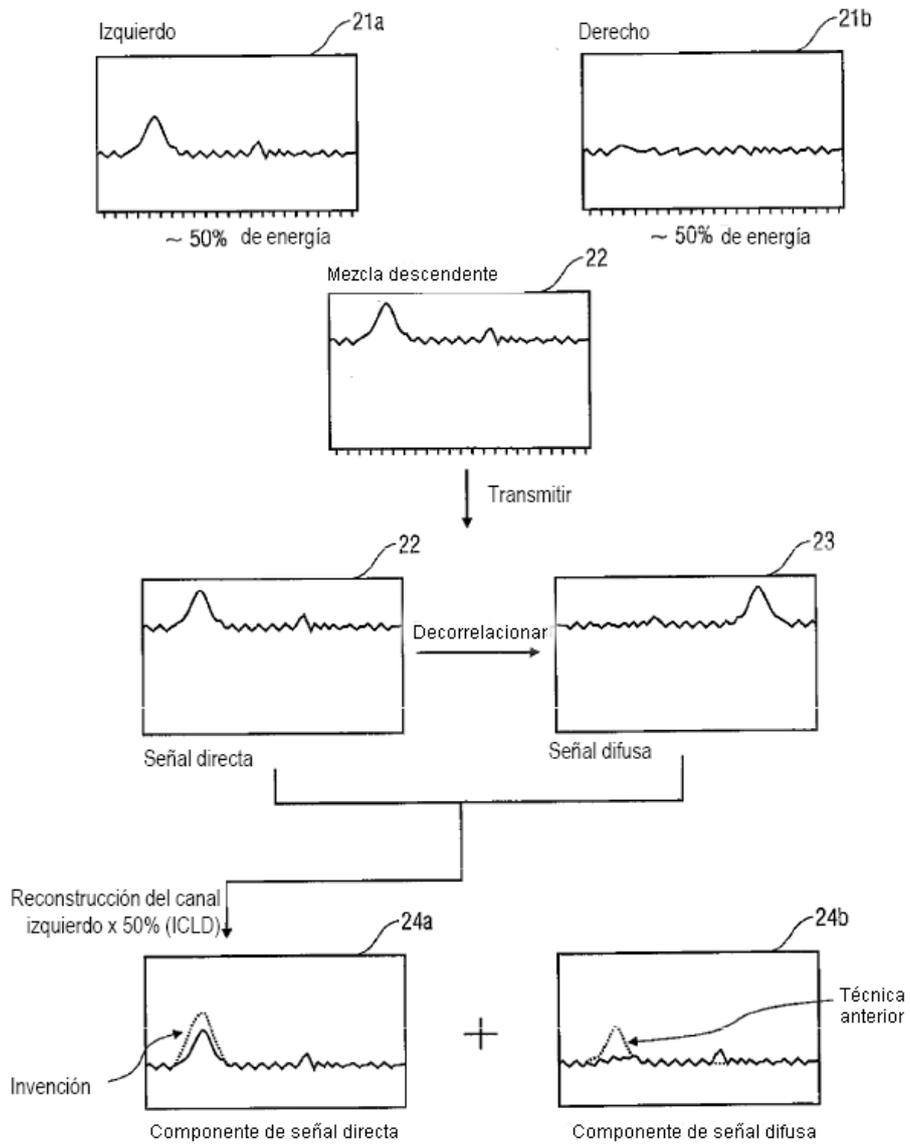


FIG 2

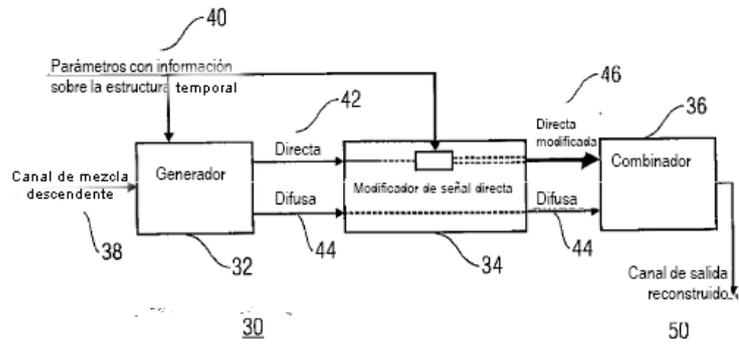


FIG 3

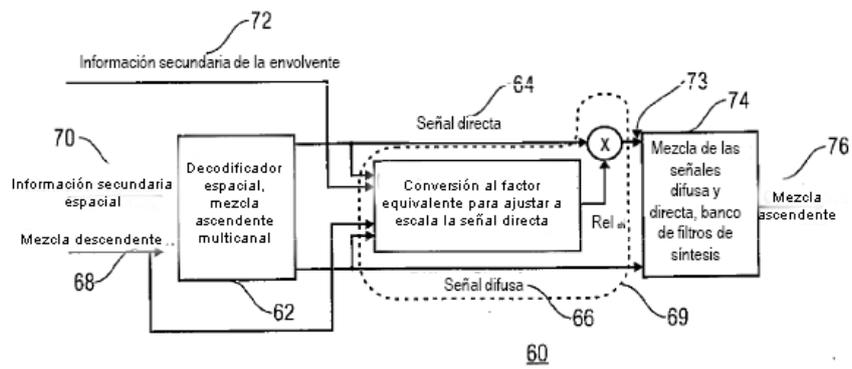


FIG 4

k	0	1	2	3	4	5	6	7	8	9	10
$\bar{k}(k)$	1	0	0	1	2	3	4	5	6	7	8
k	11	12	13	14	15	16	17	18	19	20	21
$\bar{k}(k)$	9	10	11	12	13	14	14	15	15	15	16
k	22	23	24	25	26	27	28	29	30	31	32
$\bar{k}(k)$	16	16	16	17	17	17	17	17	18	18	18
k	33	34	35	36	37	38	39	40	41	42	43
$\bar{k}(k)$	18	18	18	18	18	18	18	18	18	19	19
k	44	45	46	47	48	49	50	51	52	53	54
$\bar{k}(k)$	19	19	19	19	19	19	19	19	19	19	19
k	55	56	57	58	59	60	61	62	63	64	65
$\bar{k}(k)$	19	19	19	19	19	19	19	19	19	19	19
k	66	67	68	69	70						
$\bar{k}(k)$	19	19	19	19	19						

FIG 5

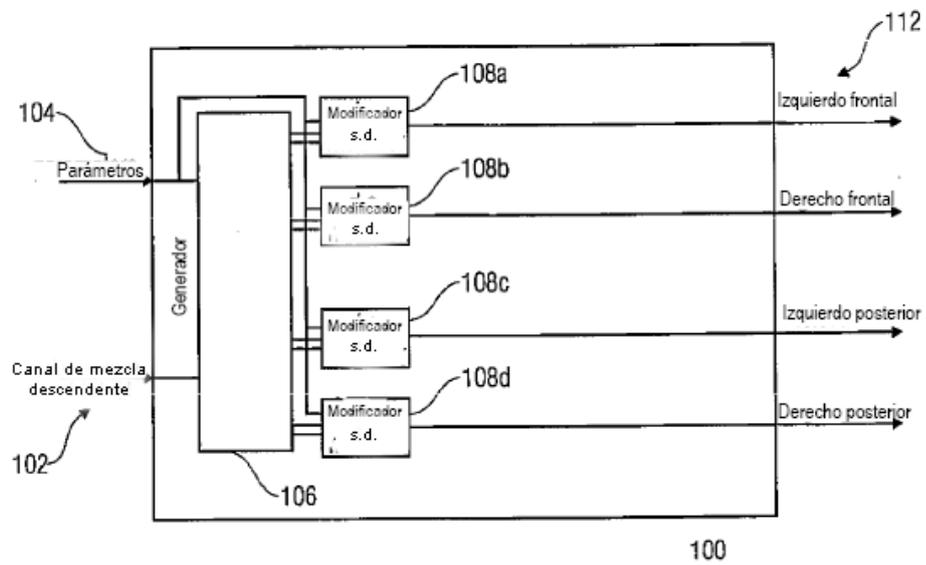


FIG 6

