

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 899 286**

51 Int. Cl.:

G10L 19/008 (2013.01)

G10L 19/02 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **15.08.2005 E 17193794 (9)**

97 Fecha y número de publicación de la concesión europea: **22.09.2021 EP 3279893**

54 Título: **Configuración de envoltente temporal para codificación espacial de audio usando filtrado de Wiener de dominio de frecuencia**

30 Prioridad:

25.08.2004 US 604836 P

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

10.03.2022

73 Titular/es:

**DOLBY LABORATORIES LICENSING CORPORATION (100.0%)
1275 Market Street
San Francisco, CA 94103-1410, US**

72 Inventor/es:

**VINTON, MARK STUART y
SEEFELDT, ALAN JEFFREY**

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 899 286 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

Configuración de envolvente temporal para codificación espacial de audio usando filtrado de Wiener de dominio de frecuencia

Campo de la invención

5 La presente invención se refiere a codificadores de audio basados en bloques, en los que cuando es decodificada la información de audio, ésta tiene una resolución de envolvente temporal limitada por la velocidad de bloque, que incluye codificadores, decodificadores y sistemas perceptuales y paramétricos de audio, que corresponden a métodos, a programas informáticos para implementar tales métodos y a un flujo binario producido por tales codificadores.

Antecedentes de la invención

10 Muchas técnicas de codificación de audio de velocidad binaria reducida están “basadas en bloques” ya que la codificación incluye procesamiento que divide cada una de la una o más señales de audio que están siendo codificadas en bloques de tiempo y actualiza al menos parte de la información secundaria asociada al audio codificado según una frecuencia no superior a la velocidad de bloque. Como resultado de ello, la información de audio, cuando se decodifica, tiene una resolución de envolvente temporal limitada por la velocidad de bloque. En consecuencia, la estructura detallada de las señales decodificadas de audio con respecto al tiempo no se mantiene durante períodos de tiempo más pequeños que la granularidad de la técnica de codificación (normalmente, en el intervalo de 8 a 50 milisegundos por bloque).

15 Tales técnicas de codificación de audio basadas en bloques incluyen no solamente técnicas perceptuales de codificación bien establecidas conocidas como AC-3, AAC y varias formas de MPEG en las que generalmente se mantienen canales discretos a través del proceso de codificación/decodificación, sino que también incluyen técnicas de codificación de velocidad binaria limitada recientemente introducidas, denominadas en algunas ocasiones “Codificación de pista binaural” y “Codificación paramétrica estéreo” en las que múltiples canales de entrada son de mezcla descendente y de mezcla ascendente a partir de un canal único a través del proceso de codificación/decodificación. Detalles de tales sistemas de codificación se encuentran en varios documentos que incluyen aquellos que se citan más adelante de acuerdo con el título “técnica anterior”. Como consecuencia del uso de un canal único en tales sistemas de codificación, las señales de salida reconstruidas son, forzosamente, versiones a escala de amplitud entre sí para un bloque particular, teniendo las diferentes señales de salida sustancialmente la misma estructura fina envolvente.

20 Aunque todas las técnicas de codificación de audio basadas en bloques pueden beneficiarse de una resolución de envolvente temporal mejorada de sus señales decodificadas de audio, la necesidad de tal mejora es particularmente grande en técnicas de codificación basadas en bloques que no mantienen canales discretos a través de todo el proceso de codificación/decodificación. Algunos tipos de señales de entrada, tales como, por ejemplo, un aplauso, son particularmente problemáticas para tales sistemas, haciendo que se estreche o colapse la imagen espacial percibida que es reproducida.

Breve descripción de los dibujos

La figura 1 es un diagrama de bloques funcional esquemático de un codificador o función de codificación que incluye aspectos de la presente invención.

La figura 2 es un diagrama de bloques funcional esquemático de un decodificador o función de decodificación que incluye aspectos de la presente invención.

Sumario de la invención

De acuerdo con un aspecto de la invención, se proporciona un método de decodificación de señal de audio en el que una o más señales de audio de entrada han sido codificadas en un flujo binario que comprende información de audio e información secundaria que se refiere a la información de audio y que es útil para decodificar flujo binario, incluyendo la codificación procesamiento que divide cada una de la una o más señales de audio de entrada en bloques de tiempo y actualiza al menos parte de la información secundaria según una frecuencia no superior a la velocidad de bloque, de manera que la información de audio, una vez decodificada, usando la información secundaria, tenga una resolución de envolvente temporal limitada por la velocidad de bloque, la decodificación incluyendo, adicionalmente, comparar la envolvente temporal de al menos una señal de audio de entrada y la envolvente temporal de una reconstrucción decodificada estimada de cada una de al menos una señal de audio de entrada mencionada, empleando dicha reconstrucción estimada al menos parte de la información de audio y al menos parte de la información secundaria, y la decodificación incluyendo además recibir el flujo binario, emplear la decodificación la información de audio y la información secundaria, y reformatear las reconstrucciones de audio decodificadas de las señales de audio de entrada usa la información de reformateo de envolvente temporal.

Otro aspecto incluye un decodificador de señal de audio adaptado para realizar el método explicado anteriormente.

Descripción detallada de la invención

La figura 1 muestra un ejemplo de un proceso codificador o de codificación que no es parte de la invención. Una pluralidad de señales de entrada de audio tales como señales PCM, muestras de tiempo de señales analógicas de audio respectivas, 1-n, son aplicadas a respectivos convertidores o funciones de conversión de dominio de tiempo a dominio de frecuencia ("T/F") 2-1 a 2-n. Las señales de audio pueden representar, por ejemplo, direcciones espaciales tales como izquierda, central, derecha, etcétera. Cada T/F puede implementarse, por ejemplo, dividiendo las muestras de entrada de audio en bloques, formando ventanas de los bloques, superponiendo los bloques, transformando cada uno de los bloques formados como ventanas y superpuestos en el dominio de frecuencia mediante el cálculo de una transformada discreta de frecuencia (DFT) y dividiendo los espectros de frecuencia resultantes en bandas que simulan las bandas críticas del oído, por ejemplo, veintiuno bandas que utilizan, por ejemplo, la escala de banda equivalente-rectangular (ERB). Tales procesos DFT son bien conocidos en la técnica. Pueden emplearse otros parámetros y técnicas de conversión de dominio de tiempo a dominio de frecuencia. Ni los parámetros particulares ni la técnica particular son fundamentales para la invención. Sin embargo, con el fin de facilitar su explicación, la siguiente descripción acepta que se emplee tal técnica de conversión DFT.

Cada una de las salidas de dominio de frecuencia de T/F 2-1 a 2-n es un conjunto de coeficientes espectrales. Estos conjuntos pueden designarse $Y[k]_1$ a $Y[k]_n$, respectivamente. Todos estos conjuntos pueden aplicarse a un codificador o función de codificación basada en bloques ("codificador basado en bloques") 4. El codificador basado en bloques puede ser, por ejemplo, cualquiera de los codificadores basados en bloques conocidos mencionados anteriormente solos o en algunas ocasiones en combinación o cualquier codificador basado en bloques que incluya variaciones de estos codificadores mencionados anteriormente. Aunque los aspectos de la invención son particularmente beneficiosos para usar en combinación con codificadores basados en bloques que no mantienen canales discretos durante la codificación y decodificación, los aspectos de la invención son útiles prácticamente en combinación con cualquier codificador basado en bloques.

Las salidas de un codificador común basado en bloques 4 pueden caracterizarse como "información de audio" e "información secundaria". La información de audio puede comprender datos que representan múltiples canales de señal como también es posible en sistemas de codificación basados en bloques tales como, por ejemplo, AC-3, AAC y otros, o, pueden comprender solamente un canal único derivado mediante la mezcla descendente de múltiples canales de entrada, tal como los sistemas de codificación de referencia binaria y los sistemas paramétricos de codificación estéreo antes mencionados (el canal de mezcla descendente en un codificador de cifrado de referencia binaria o un sistema paramétrico de codificación estéreo también pueden ser codificados de manera perceptual, por ejemplo, con AAC o alguna otra codificación adecuada). También puede comprender un canal único o múltiples canales derivados mediante la mezcla descendente de múltiples canales de entrada, tal como se describe en la solicitud provisional de patente US S.N. 60/588.256, presentada el 14 de Julio de 2004 por Davis et al., titulada "Low Bit Rate Audio Encoding and Decoding in Which Multiple Channels are Represented By Monophonic Channel and Auxiliary Information". La información secundaria puede comprender datos que se refieren a la información de audio y es útil en su decodificación. En el caso de varios sistemas de codificación de mezcla descendente, la información secundaria puede comprender parámetros espaciales tales como, por ejemplo, diferencias de amplitud entre canales, diferencias de tiempo o fase entre canales y correlación cruzada entre canales.

La información de audio y la información secundaria procedentes del codificador basado en bloques 4 pueden aplicarse entonces a convertidores o funciones de conversión de dominio de frecuencia a dominio de tiempo correspondientes ("F/T") 6 y 8 de manera que cada uno realice generalmente las funciones de inversión de una T/F descrita anteriormente, a saber, una FFT inversa, seguido de la adición de formación de ventanas y superposición. La información de dominio de tiempo de F/T 6 y 8 se aplica a un agrupador o a una función de agrupación de flujo binario ("agrupador de flujo binario") 10 que proporciona una salida codificada de flujo binario. De manera alternativa, si el codificador va a proporcionar un flujo binario que representa información de dominio de frecuencia, F/T 6 y 8 pueden omitirse.

La información de audio de dominio de frecuencia y la información secundaria procedentes del codificador basado en bloques 4 se aplican también a un estimador o función de estimación de decodificación ("estimador de decodificación") 14. El estimador de decodificación 14 puede simular al menos una parte de un decodificador o función de decodificación diseñada para decodificar el flujo binario codificado proporcionado por el agrupador de flujo binario 10. Un ejemplo de tal decodificador o función de decodificación se describe más adelante con respecto a la figura 2. El estimador de decodificación 14 puede proporcionar conjuntos de coeficientes espectrales $X[k]_1$ a $X[k]_n$ que aproximan los conjuntos de coeficientes espectrales $Y[k]_1$ a $Y[k]_n$ de señales de entrada de audio correspondientes que se espera que sean obtenidas en el decodificador o función de decodificación. Alternativamente, puede proporcionar tales coeficientes espectrales en una cantidad menor que todas las señales de entrada de audio para una menor cantidad que todos los bloques de tiempo de las señales de entrada de audio y/o para menos que todas las bandas de frecuencia (es decir, no puede proporcionar todos los coeficientes espectrales). Esto puede darse, por ejemplo, si se desea mejorar solamente las señales de entrada que representan canales que se consideran más importantes que otros. Como otro ejemplo, esto puede darse si se desea mejorar únicamente las partes de frecuencia más baja de señales en las que el oído es más sensible a los detalles finos de las envolventes temporales de forma de onda.

Cada una de las salidas de dominio de frecuencia de T/F 2-1 a 2-n, cada uno de los conjuntos de coeficientes espectrales $Y[k]_1$ a $Y[k]_n$, también se aplica a respectivos dispositivos o funciones de comparación ("comparación") 12-1 a 12-n. Tales conjuntos se comparan con conjuntos correspondientes de bloques de tiempo correspondientes de

coeficientes espectrales estimados $X[k]_1$ a $X[K]_n$ en la comparación correspondiente 12-1 a 12-n. Cada resultado de la comparación en cada comparación 12-1 a 12-n se aplica a un calculador o función de cálculo de filtro ("cálculo de filtro") 15-1 a 15-n. Esta información debe ser suficiente para cada cálculo de filtro con el objeto de definir los coeficientes de un filtro para cada bloque de tiempo, generando este filtro, cuando se aplica a una reconstrucción decodificada de una señal de entrada, la señal que tiene una envolvente temporal con una resolución mejorada. Es decir, el filtro volvería a configurar la señal, de modo que esta reproduzca más fielmente la envolvente temporal de la señal original. La resolución mejorada es una resolución más fina que la velocidad de bloque. Otros detalles de un filtro preferido se señalan más adelante.

Aunque el ejemplo de la figura 1 muestra la comparación y el cálculo de filtro en el dominio de frecuencia, en principio, la comparación y el cálculo de filtro pueden realizarse en el dominio de tiempo. Ya fueran realizados en el dominio de frecuencia o en el dominio de tiempo, solo se determinaría una configuración de filtro por bloque de tiempo (aunque puede aplicarse la misma configuración de filtro a varios bloques consecutivos de tiempo). Aunque, en principio, una configuración de filtro puede determinarse en base a una banda por banda (tal como por banda de la escala ERB), al hacerlo de este modo se requeriría el envío de un gran número de bits de información secundaria, lo que perdería la ventaja de la invención, a saber, mejorar la resolución de envolvente temporal con un bajo incremento en la velocidad binaria.

Cada medición de la comparación en cada comparación 12-1 a 12-n se aplica a un dispositivo o función de decisión ("decisión") 16-1 a 16-n. Cada decisión compara la medición de la comparación frente a un umbral. Una medición de la comparación puede adoptar varias formas y no es fundamental. Por ejemplo, puede calcularse el valor absoluto de la diferencia de cada valor correspondiente de coeficiente y las diferencias sumadas a fin de proporcionar un número único cuyo valor indique el grado en el que las formas de onda de señal difieren entre si durante un bloque de tiempo. Este número puede ser comparado con un umbral tal que, si sobrepasa el umbral, se proporcionaría un indicador "si" al cálculo de filtro correspondiente. En ausencia de un indicador "si", pueden impedirse los cálculos de filtro para el bloque, o si se calculan, no pueden ser enviados por el cálculo de filtro. Tal información si/no para cada señal constituye un aviso que también puede aplicarse al agrupador de flujo binario 10 para su inclusión en el flujo binario (de esta manera, puede haber una pluralidad de avisos, uno para cada señal de entrada y cada uno de tales avisos puede representarse mediante un bit).

Alternativamente, cada decisión 16-1 a 16-n puede recibir información de un cálculo de filtro correspondiente 14-1 a 14-n en lugar o además de la información procedente de una comparación correspondiente 12-1 a 12-n. La decisión correspondiente 16 puede emplear las características calculadas de filtro (por ejemplo, sus magnitudes medias o pico) como base para la toma de una decisión o para ayudar en la toma de una decisión.

Como se menciona anteriormente, cada cálculo de filtro 14-1 a 14-n proporciona una representación de los resultados de la comparación, los cuales constituir los coeficientes de un filtro, filtro que, cuando se aplica a una reconstrucción decodificada de una señal de entrada, origina la señal que tiene una envolvente temporal con una resolución mejorada. Si los coeficientes espectrales estimados $X[k]_1$ a $X[K]_n$ fueran incompletos (en el caso de que el estimador de decodificación proporcionara coeficientes espectrales en una cantidad menor que todas las señales de entrada de audio para una menor cantidad que todos los bloques de tiempo de las señales de entrada de audio y/o para menos que todas las bandas de frecuencia), no puede haber salidas para cada comparación 12-1 a 12-n para todos los bloques de tiempo, bandas de frecuencia y señales de entrada. El lector debe observar que $X[k]_1$ a $X[K]_n$ se refieren a salidas reconstruidas, mientras que $Y[k]_1$ a $Y[K]_n$ se refieren a entradas.

La salida de cada cálculo de filtro 14-1 a 14-n puede aplicarse al ensamblador de flujo binario 10. Aunque la información de filtro puede emitirse por separado desde el flujo binario, se prefiere que sea enviada como parte del flujo binario y como parte de la información secundaria. Cuando los aspectos de la invención se aplican a sistemas existentes de codificación basada en bloques, la información adicional proporcionada por aspectos de la presente invención puede introducirse en partes de los flujos binarios de tales sistemas que se pretende que lleven información auxiliar.

En realizaciones prácticas, es probable que no solamente la información de audio sino también la información secundaria y los coeficientes de filtro sean cuantificados o codificados de algún modo para minimizar sus costes de transmisión. Sin embargo, la falta de cuantificación y descuantificación se muestra en las figuras con el fin de simplificar la presentación y debido a que tales detalles son bien conocidos y no ayudan a entender la invención.

50 Diseño de filtro de Wiener en el dominio de frecuencia

De preferencia, cada uno de los dispositivos o funciones de cálculo de filtro 14-1 a 14-n se caracteriza por un filtro FIR en el dominio de frecuencia que representa los cambios multiplicativos en el dominio de tiempo requeridos para obtener una reproducción más exacta de una envolvente temporal original del canal de señal. Este problema de filtro se puede formular al menos como un problema de cuadros, que a menudo se denomina diseño de filtro de Wiener. Véase, por ejemplo, X. Rong Li, Probability, Random Signals, and Statistics, CRC Press 1999, New York, págs. 423. La aplicación de técnicas de filtro de Wiener tiene la ventaja de reducir los bits adicionales requeridos para transmitir la información de reformateo de filtro a un decodificador. Normalmente, las aplicaciones convencionales del filtro de Wiener son diseñadas y aplicadas en dominio de tiempo.

El problema de diseño de filtro de menos cuadros de dominio de frecuencia puede definirse como sigue: dada la representación espectral DFT de una señal original $Y[k]$ y la representación espectral DFT de una aproximación de tal canal original $X[k]$, se calcula un conjunto de coeficientes de filtro (a_m) que minimiza la ecuación 1. Se observa que $Y[k]$ y $X[k]$ son valores complejos y, por tanto, en general, a_m también será un valor complejo.

5

$$\min_{a_m} \left[E \left| Y[k] - \sum_{m=0}^{M-1} a_m X[k-m] \right|^2 \right], \quad (1)$$

donde k es el índice espectral, E es el operador de expectativa y M es la longitud del filtro que está siendo diseñado.

La ecuación 1 puede ser nuevamente expresada utilizando expresiones de matriz como se muestra en la ecuación 2:

$$\min_{a_m} \left[E \left| \bar{Y}_k - \bar{A}^T \bar{X}_k \right|^2 \right], \quad (2)$$

10 donde

$$\begin{aligned} \bar{Y}_k &= [Y[k]] \\ \bar{X}_k^T &= [X[k] \quad X[k-1] \quad \dots \quad X[k-M+1]] \end{aligned}$$

y

$$\bar{A}^T = [a_0 \quad a_1 \quad \dots \quad a_{M-1}].$$

15 De este modo, mediante el ajuste de las derivadas parciales en la ecuación 2 con respecto a cada uno de los coeficientes de filtro en cero, es simple mostrar la solución al problema de minimización, el cual se da en la ecuación 3.

$$\bar{A} = \bar{R}_{XX}^{-1} \bar{R}_{XY}, \quad (3)$$

donde

$$\bar{R}_{XX} = \begin{bmatrix} E(X_K X_K^*) & E(X_K X_{K-1}^*) & \dots & E(X_K X_{K-M+1}^*) \\ E(X_{K-1} X_K^*) & E(X_{K-1} X_{K-1}^*) & \dots & E(X_{K-1} X_{K-M+2}^*) \\ \vdots & \vdots & \ddots & \vdots \\ E(X_{K-M+1} X_K^*) & E(X_{K-M+1} X_{K-1}^*) & \dots & E(X_{K-M+1} X_{K-M+1}^*) \end{bmatrix}$$

20

y

$$\bar{R}_{YX}^T = [E(Y_K X_K^*) \quad E(Y_K X_{K-1}^*) \quad \dots \quad E(Y_K X_{K-M+1}^*)].$$

25

La ecuación 3 define el cálculo de los coeficientes óptimos de filtro que minimizan el error entre el espectro original ($Y[k]$) y el espectro reconstruido ($X[k]$) de un canal particular. Generalmente, se calcula un conjunto de coeficientes de filtro para cada bloque de tiempo de cada señal de entrada.

30

En una realización práctica de aspectos de la invención, se emplea un filtro de Wiener de 12ª orden, aunque la invención no se limita al uso de un filtro de Wiener de tal tamaño. Tal realización práctica emplea procesamiento en el dominio de frecuencia después de una DFT. En consecuencia, los coeficientes de filtro de Wiener son números complejos y cada filtro requiere la transmisión de veinticuatro números reales. Para transmitir de manera eficiente tal información de filtro a un decodificador, se puede utilizar cuantificación de vector (VQ) para codificar los coeficientes de cada filtro. Se puede utilizar un libro de códigos, de manera que solo tenga que emitirse un índice al decodificador

para transmitir la información compleja de filtro de 12ª orden. En una realización práctica, se ha encontrado útil un libro de códigos de tabla VQ que tiene 24 dimensiones y 16.536 entradas. La invención no se limita al uso de la cuantificación de vector, tampoco al uso de un libro de códigos.

5 Aunque la descripción anterior adopta el uso de una DFT para estimar el contenido espectral y para diseñar el filtro de Wiener, en general se puede utilizar cualquier transformación.

10 La figura 2 muestra un ejemplo de un entorno de decodificador o proceso de decodificación en el que pueden emplearse aspectos de la presente invención. Tal codificador o proceso de codificación puede ser adecuado para el funcionamiento en cooperación con un codificador o proceso de codificación, como se describe con respecto al ejemplo de la figura 1. Un flujo binario codificado, tal como el que se produce mediante la disposición de la figura 1, es recibido en cualquier modo adecuado de transmisión o almacenamiento de señal y es aplicado a un desagrupador de flujo binario 30 que desagrupa el flujo binario según sea necesario para separar la información codificada de audio de la información secundaria y avisos de sí/no (si estuvieran incluidos en el flujo binario). De preferencia, la información secundaria incluye un conjunto de coeficientes de filtro para usar en la mejora de la reconstrucción de cada una de la una o más de las señales de entrada que se aplicaron a la disposición de codificación de la figura 1.

15 En este ejemplo, se supone que existe una señal reproducida que corresponde a cada señal de entrada y que la información de reformateo de filtro de envolvente temporal se proporciona para cada señal reproducida, aunque esto no tiene por qué ser así, como se menciona anteriormente. De este modo, los conjuntos 1-n de información secundaria de coeficientes de filtro se muestran como salidas del desagrupador de flujo binario 30. La información de coeficientes de filtro para cada señal de entrada se aplica a los respectivos filtros de reformateo 36-1 a 36-n, cuyo funcionamiento se explica más adelante. Cada uno de los filtros también puede recibir un aviso correspondiente de sí/no 31-1 a 31-n indicando si el filtro tiene que ser activo durante un bloque particular de tiempo.

20 La información secundaria procedente del agrupador de flujo binario 30 también puede incluir otra información tal como, por ejemplo, diferencias de amplitud entre canales, diferencias de tiempo o de fase entre canales y correlación cruzada entre canales en el caso de una codificación de referencia binaural o sistema paramétrico estéreo. Un decodificador basado en bloques 42 recibe la información secundaria procedente del desagrupador de flujo binario 30 junto con la información convertida de audio de dominio de tiempo-a-frecuencia del desagrupador de flujo binario 30. La información de audio del desagrupador 30 se aplica mediante un convertidor o función de conversión de dominio de tiempo a dominio de frecuencia ("T/F") 46, que puede ser la misma que cualquiera de uno de los convertidores o funciones de conversión de dominio de frecuencia ("T/F") 2-1 a 2-n de la figura 1.

30 El decodificador basado en bloques 42 proporciona una o más salidas, cada una de las cuales es una aproximación de una señal correspondiente de entrada en la figura 1. Aunque algunas señales de entrada no pueden tener una señal de salida correspondiente, el ejemplo de la figura 2 muestra señales de salida 1-n, cada una de las cuales es una aproximación que corresponde a la señal correspondiente de las señales de entrada 1-n de la figura 1. En este ejemplo, cada una de las señales de salida 1-n del decodificador 42 se aplica a un filtro de reformateo correspondiente 36-1 a 36-n, cada uno de los cuales se puede implementar como un filtro FIR. Los coeficientes de cada filtro FIR son controlados, en una base de bloques, a través de la información de filtro respectiva que se refiere a un canal particular de entrada cuya salida reconstruida será mejorada. De preferencia, el reformateo multiplicativo de la envolvente en el dominio de tiempo se consigue mediante la convolución de cada filtro FIR con una salida de decodificador basado en bloques en cada uno de los filtros 36-1 a 36-n. De esta manera, la configuración de envolvente temporal de acuerdo con los aspectos de la presente invención se aprovecha de la dualidad de frecuencia de tiempo, la convolución en el dominio de tiempo es equivalente a la multiplicación en el dominio de frecuencia y viceversa. Cada una de las señales de salida decodificadas y filtradas, se aplica después a los convertidores o funciones de conversión de dominio de frecuencia a dominio de tiempo respectivos ("F/T") 44-1 a 44-n de manera que cada una realice las funciones inversas de una T/F descrita anteriormente, a saber, una FFT inversa, seguida de la adición de formación de ventanas y de superposición. Alternativamente, se puede emplear un filtro adecuado de reformateo de dominio de tiempo después de cada uno de los convertidores de dominio de frecuencia a tiempo. Por ejemplo, los n coeficientes polinomiales de una curva polinomial de enésimo orden pueden enviarse como información secundaria en lugar de coeficientes de filtro FIR y la curva aplicada mediante multiplicación en el dominio de tiempo. Aunque es preferible emplear técnicas de filtro de Wiener para transmitir la información de reformateo de filtro al decodificador, se pueden emplear otras técnicas de dominio de frecuencia y dominio de tiempo, tales como aquellas señaladas en la solicitud de Patente US S. N. 10/113.858 de Truman y Vinton, titulada "Broadband Frequency Translation for High Frequency Regeneration", presentada el 28 de marzo de 2002 y publicada como US 2003/0187663 A1 el 2 de octubre de 2003.

Implementación

55 La invención puede implementarse en hardware o software, o en una combinación de ambos (por ejemplo, en series lógicas programables). A menos que se especifique de otro modo, los algoritmos incluidos como parte de la invención no están inherentemente relacionados con ningún ordenador concreto u otro aparato. En particular, pueden emplearse varias máquinas de uso general con programas escritos de acuerdo con las enseñanzas de este documento, o puede resultar más conveniente construir aparatos más especializados (por ejemplo, circuitos integrados) para realizar las etapas de método requeridas. Por tanto, la invención puede implementarse en uno o más programas informáticos que se ejecuten en uno o más sistemas informáticos programables, comprendiendo cada uno al menos un procesador, al

menos un sistema de almacenamiento de datos (que incluye una memoria volátil y no volátil y/o elementos de almacenamiento), al menos un dispositivo o puerto de entrada y al menos un dispositivo o puerto de salida. El código de programa se aplica a datos de entrada para realizar las funciones descritas en el presente documento y generar la información de salida. La información de salida se aplica a uno o más dispositivos de salida, en un modo conocido.

- 5 Cada programa puede implementarse en cualquier lenguaje informático deseado (que incluye lenguajes de programación de máquina, ensamble o procedimiento de alto nivel, lógico u orientado por objeto) para comunicarse con un sistema informático. En cualquier caso, el lenguaje puede ser un lenguaje compilado o interpretado.

De preferencia, cada programa informático se almacena o descarga en un dispositivo o medio de almacenamiento (por ejemplo, en una memoria o medio de estado sólido o medio magnético u óptico) que puede ser leído a través de un ordenador programable de uso general o especial, para configurar y hacer funcionar el ordenador cuando el medio o dispositivo de almacenamiento sea leído por el sistema informático para realizar los procedimientos descritos en el presente documento. También puede considerarse la implementación del sistema de la invención como medio de almacenamiento legible por ordenador, configurado con un programa informático, en el que el medio de almacenamiento configurado de este modo hace que el sistema informático funcione en un modo específico y predefinido para realizar las funciones descritas en el presente documento.

Se han descrito varias realizaciones de la invención. Sin embargo, se entenderá que pueden realizarse varias modificaciones sin apartarse del espíritu y ámbito de aplicación de la invención. Por ejemplo, algunas de las etapas descritas en el presente documento pueden ser independientes del orden y, por tanto, pueden realizarse en un orden diferente del descrito.

20 **Técnica anterior**

Las siguientes patentes, solicitudes de patente y publicaciones son relevantes para la invención.

AC-3

Estándar ATSC A52/A: Estándar de compresión de audio digital (AC-3), Revisión A, Advanced Television Systems Committee, 20 de agosto de 2001. El documento A/52A se encuentra disponible en la red mundial (WWW) en <http://www.atsc.org/standards.html>.

“Design and Implementation of AC-3 Coders,” de Steve Vernon, IEEE Trans. Consumer Electronics, vol. 41, núm. 3, agosto de 1995.

“The AC-3 Multichannel Coder” de Mark Davis, Audio Engineering Society Preprint 3774, Convención AES 95.º, octubre de 1993.

30 “High Quality, Low-Rate Audio Transform Coding for Transmission and Multimedia Applications” de Bosi et al, Audio Engineering Society Preprint 3365, Convención AES 93.º, octubre de 1992.

Patentes US 5.583.962; 5.632.005; 5.633.981; 5.727.119; y 6.021.386.

AAC

35 ISO/IEC JTC1/SC29, “Information technology - very low bitrate audio-visual coding” ISO/IEC IS-14496 (parte 3, audio), 1996

1) ISO/IEC 13818-7. “MPEG-2 advanced audio coding, AAC”. Estándar Internacional, 1997;

M. Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagiri, H. Fuchs, M. Dietz, J. Herre, G. Davidson, e Y. Oikawa: “ISO/IEC MPEG-2 Advanced Audio Coding”. Proc. de la Convención AES 101.º, 1996;

40 M. Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagiri, H. Fuchs, M. Dietz, J. Herre, G. Davidson, Y. Oikawa: “ISO/IEC MPEG-2 Advanced Audio Coding”, Diario de AES, vol. 45, núm. 10, octubre de 1997, págs. 789-814;

Karlheinz Brandenburg: “MP3 and AAC explained”. Proc. Of the AES 17th International Conference on High Quality Audio Coding, Florencia, Italia, 1999; y

G.A. Soulodre et al.: “Subjective Evaluation of State-of-the-Art Two-Channel Audio Codecs” J. Audio Eng. Soc., vol. 46, núm. 3, págs., 164-177, marzo de 1998.

45 **Estéreo de Intensidad MPEG**

Patentes US 5.323.396; 5.539.829; 5.606.618 y 5.621.855.

Solicitud de patente publicada US 2001/0044713, publicada.

Codificación espacial y paramétrica

- Solicitud de patente provisional US S.N. 60/588.256, presentada el 14 de julio de 2004 de Davis et al, titulada' "Low Bit Rate Audio Encoding and Decoding in Which Multiple Channels are Represented By Monophonic Channel and Auxiliary Information".
- Solicitud de patente publicada US 2003/0026441, publicada el 6 de febrero de 2003.
- 5 Solicitud de patente publicada US 2003/0035553, publicada el 20 de febrero de 2003,
- Solicitud de patente publicada US 2003/0219130 (Baurngarte & Faller) publicada el 27 de noviembre de 2003, Audio Engineering Society Paper 5852, marzo de 2003.
- Solicitud de patente internacional publicada WO 03/090206, publicada el 30 de octubre de 2003
- Solicitud de patente internacional publicada WO 03/090207, publicada el 30 de octubre de 2003
- 10 Solicitud de patente internacional publicada WO 03/090208, publicada el 30 de octubre de 2003
- Solicitud de patente internacional publicada WO 03/007656, publicada el 22 de enero de 2003
- Publicación de Solicitud de patente publicada US 2003/0236583 A1, Baumgarte et al, publicada el 25 de diciembre de 2003, "Hybrid Multi-Channel/Cue Coding/Decoding of Audio Signals" Solicitud S.N. 10/246.570.
- 15 "Binaural Cue Coding Applied to Stereo and Multi-Channel Audio Compression", de Faller et al, Audio Engineering Society Convention Paper 5574, Convención 112.º, Munich, mayo de 2002.
- "Why Binaural Cue Coding is Better than Intensity Stereo Coding" by Baumgarte et al, Audio Engineering Society Convention Paper 5575, Convención 112.º Munich, mayo de 2002.
- "Design and Evaluatin of Binaural Cue Coding Schemes" de Baurngarte et al, Audio Engineering Society Convention Paper 5706, Convención 113.º, Los Ángeles, octubre de 2002.
- 20 "Efficient Representation of Spatial Audio Using Perceptual Parametrization", de Faller et al, IEEE Workshop on Applications of Signal Processing to Audio and Acoustics 2001, New Paltz, Nueva York, octubre de 2001, págs. 199-202.
- "Estimation of Auditory Spatial Cues for Binaural Cue Coding" de Baurngarte et al, Proc. ICASSP 2002, Orlando, Florida, mayo de 2002, págs. II-1801-1804.
- 25 "Binaural Cue Coding: A Novel and Efficient Representation of Spatial Audio" de Faller et al, Proc. ICASSP 2002, Orlando, Florida, mayo de 2002, págs. II-1841-II-1844.
- "High-quality parametric spatial audio coding at low bitrates" de Breebaart et al, Audio Engineering Society Convention Paper 6072, Convención 116.º, Berlín, mayo de 2004.
- 30 "Audio Coder Enhancement using Scalable Binaural Cue Coding with Equalized Mixing" de Baurngarte et al, Audio Engineering Society Convention Paper 6060, Convención 116.º, Berlín, mayo de 2004.
- "Low complexity parametric stereo coding" de Schuijers et al, Audio Engineering Society Convention Paper 6073, Convención 116.º, Berlín, mayo de 2004.
- "Synthetic Ambience in Parametric Stereo Coding" de Engdegard et al, Audio Engineering Society Convention Paper 6074, Convención 116.º, Berlín, mayo de 2004.
- 35 Otras
- Patente US 5.812.971, de Herre, "Enhanced Joint Stereo Coding Method Using Temporal Envelope Shaping" 22 de septiembre de 1998
- "Intensity Stereo Coding" de Herre et al, Audio Engineering Society Preprint 3799, Convención 96.º, Amsterdam, 1994.

REIVINDICACIONES

- 5 1.- Método para decodificar una de señal de audio en el que una o más señales de audio de salida se han codificado en un flujo binario que comprende información de audio e información secundaria relacionada con la información de audio y útil para decodificar el flujo binario, en el que cada una de la una o más señales de audio de entrada se ha dividido en bloques de tiempo y al menos se ha actualizado parte de la información secundaria según una frecuencia no superior a una velocidad de bloque, de modo que la información de audio, cuando se decodifica, usando la información secundaria, tiene una resolución de envolvente temporal limitada por la velocidad de bloque,
- 10 en el que el flujo binario incluye además información de reformateo de envolvente temporal indicativa de una comparación entre la envolvente temporal de al menos una señal de audio de entrada y la envolvente temporal de una reconstrucción decodificada estimada de cada una de al menos una señal de audio de entrada mencionada, empleando dicha reconstrucción estimada al menos parte de la información de audio y al menos parte de la información secundaria, el método comprendiendo
- recibir (30) el flujo binario, y
- 15 decodificar (42) el flujo binario en una o más reconstrucciones decodificadas de la una o más señales de audio de entrada, la decodificación empleando la información de audio y la información secundaria; y
- reformatear (36-1, 36-n) la una o más reconstrucciones de la una o más señales de audio de entrada usando dicha información de reformateo de envolvente temporal.
- 2.- Método según la reivindicación 1, en el que la reconstrucción estimada es la misma que al menos una porción de la decodificación.
- 20 3.- Método según la reivindicación 1, en el que dicha comparación se realiza para uno o más bloques de tiempo de una o más frecuencias de banda de una o más señales de audio de entrada.
- 4.- Método según la reivindicación 3, en el que dicha comparación se realiza para menos de todas las señales de audio.
- 25 5.- Método según la reivindicación 3, en el que dicha comparación se realiza para menos de todos los bloques de tiempo de una señal de audio.
- 6.- Método según la reivindicación 3, en el que dicha comparación se realiza para menos de todas las bandas de frecuencia de una señal de audio.
- 7.- Método según cualquiera de las reivindicaciones 1 a 6, en el que parte de la información de reformateo de envolvente temporal no está incluida en el flujo binario.
- 30 8.- Método según la reivindicación 7, en el que la información de reformateo de envolvente temporal no está incluida en el flujo binario cuando una medida de la comparación es mayor o menor de un umbral.
- 9.- Método según las reivindicaciones 1 a 8, en el que la información secundaria incluye la información de reformateo de envolvente temporal.
- 35 10.- Método de cualquiera de las reivindicaciones 1 a 9, en el que reformatear la una o más reconstrucciones decodificadas de la una o más señales de audio de entrada usando información de reformateo de envolvente temporal origina una señal que tiene una envolvente temporal con una resolución mejorada.
- 11.- Método según la reivindicación 10, en el que se reformatea la señal de manera que reproduzca más fielmente la envolvente temporal de la señal original.
- 40 12.- Método según la reivindicación 10 u 11, en el que la resolución mejorada es una resolución más fina que la velocidad de bloque.
- 13.- Método de cualquiera de las reivindicaciones 1 a 12, en el que una pluralidad de señales de audio de entrada se combina en una señal compuesta, dicha información de audio es una versión codificada de codificación perceptual de la señal compuesta, y la información secundaria incluye uno o más parámetros indicativos de propiedades espaciales de la pluralidad de señales de audio de entrada.
- 45 14.- Método de cualquiera de las reivindicaciones 1 a 12, en el que una pluralidad de señales de audio de entrada se combina en una o más señales derivadas, cada una siendo derivada desde una o una combinación de unas de al menos dos señales de audio de entrada, dicha información de audio es una versión codificada de codificación perceptual de las señales derivadas, y la información secundaria incluye uno o más parámetros indicativos de propiedades espaciales de parte o toda la pluralidad de señales de audio de entrada.
- 50 15.- Decodificador de señal de audio en el que una o más señales de audio de entrada han sido codificadas en un flujo binario que comprende información de audio e información secundaria que se refiere a la información de audio y

- que es útil para decodificar el flujo binario, en el que cada una de la una o más señales de audio de entrada se ha dividido en bloques de tiempo y al menos parte de la información secundaria ha sido actualizada según una frecuencia no superior a la velocidad de bloque, de manera que la información de audio, una vez decodificada usando la información secundaria, tenga una resolución de envolvente temporal limitada por la velocidad de bloque, en el que el
- 5 flujo binario incluye además información de reformateo de envolvente temporal indicativa de una comparación entre la envolvente temporal de al menos una señal de audio de entrada y la envolvente temporal de una reconstrucción decodificada estimada de cada una de al menos una señal de audio de entrada mencionada, empleando dicha reconstrucción estimada al menos parte de la información de audio y al menos parte de la información secundaria, comprendiendo el decodificador de señal de audio
- 10 medios (30) para recibir el flujo binario, y
- medios (42) para decodificar el flujo binario en una o más reconstrucciones decodificadas de la una o más señales de audio de entrada, empleando la decodificación la información de audio y la información secundaria; y
- medios (36-1, 36-n) para reformatear una o más reconstrucciones decodificadas de la una o más señales de audio de entrada usando dicha información de reformateo de envolvente temporal.

15

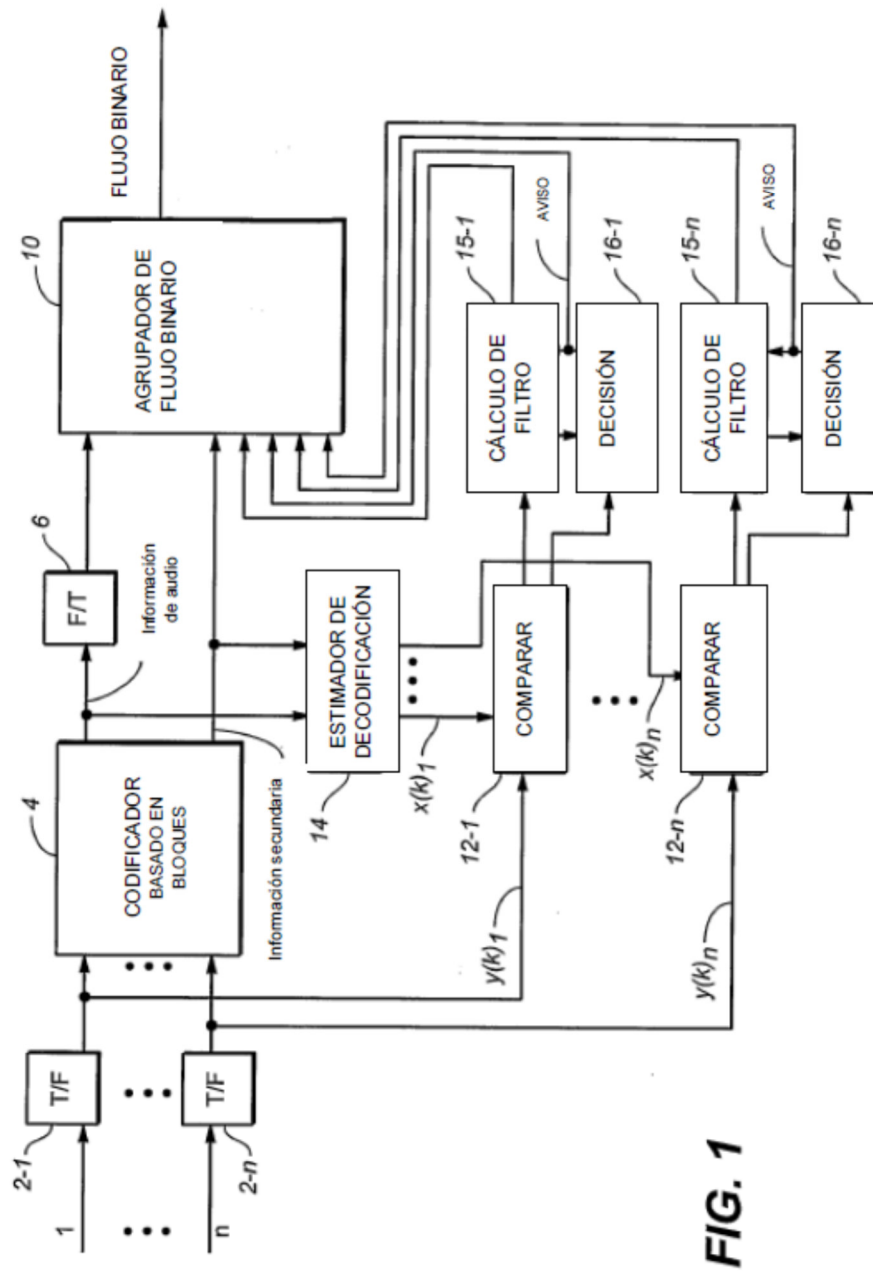


FIG. 1

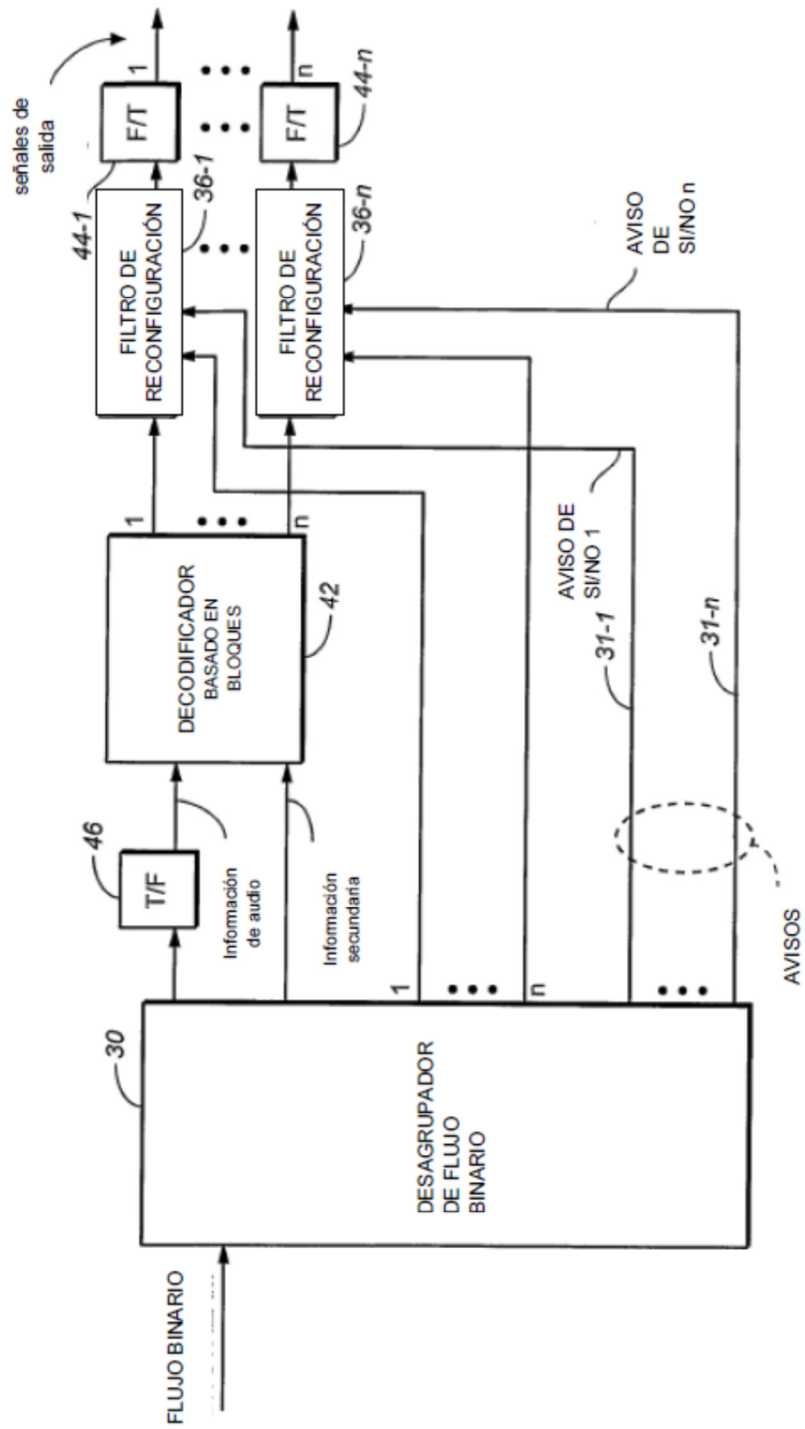


FIG. 2