

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 623 365**

51 Int. Cl.:

G10L 19/008 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **30.09.2005 PCT/EP2005/010595**

87 Fecha y número de publicación internacional: **13.07.2006 WO06072270**

96 Fecha de presentación y número de la solicitud europea: **30.09.2005 E 05796746 (5)**

97 Fecha y número de publicación de la concesión europea: **25.01.2017 EP 1829026**

54 Título: **Compactación de información secundaria para la codificación paramétrica de audio espacial**

30 Prioridad:

10.01.2005 US 32689

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

11.07.2017

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (50.0%)
Hansastraße 27c
80686 München, DE y
AGERE SYSTEMS LLC (50.0%)**

72 Inventor/es:

**HERRE, JÜRGEN y
FALLER, CHRISTOF**

74 Agente/Representante:

SALVA FERRER, Joan

ES 2 623 365 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

Compactación de información secundaria para la codificación paramétrica de audio espacial

5 **ANTECEDENTES DE LA INVENCION**

La materia de esta solicitud se relaciona a la materia de las siguientes solicitudes de E.U.A.:

- 10 ○ Solicitud de E.U.A., No. De Serie 09/848,877, presentada el 4 de Mayo del 2001, con número de expediente del apoderado Faller 5;
- 15 ○ Solicitud de E.U.A., No. De Serie 10/045,458, presentada el 7 de Noviembre del 2001, con número de expediente del apoderado Baumgarte 1-6-8, la cual reclama por sí misma el beneficio de la fecha de presentación de la solicitud provisional de E.U.A., No. 60/311,565, presentada el 10 de Agosto del 2001;
- 20 ○ Solicitud de E.U.A., No. De Serie 10/155,437, presentada el 24 de mayo del 3002, con número de expediente del apoderado Baumgarte 2-10;
- Solicitud de E.U.A., No. De Serie 10/246,570, presentada el 17 de Septiembre del 2002, con número de expediente del apoderado Baumgarte 3-11;
- 25 ○ Solicitud de E.U.A., No. De Serie 10/815,591, presentada el 1 de Abril del 2004, con número de expediente del apoderado Baumgarte 7-12;
- Solicitud de E.U.A., No. De Serie 10/936,.464, presentada el 8 de Septiembre del 2004, con número de expediente del apoderado Baumgarte 8-7-15;
- 30 ○ Solicitud de E.U.A., No. De Serie 10/762,100, presentada el 20 de Enero del 2005 (Faller 13-1);
- Solicitud de E.U.A., No. De Serie 11/006,492, presentada el 7 de Diciembre del 2001, con número de expediente del apoderado Allamanche 1-2-17-3; y
- 35 ○ Solicitud de E.U.A., No. De Serie 10/006,xxx, presentada el 7 de Diciembre del 2001, con número de expediente del apoderado Allamanche 2-3-18-4.

La materia de esta solicitud también se refiere a la materia descrita en los siguientes documentos:

- 40 ○ F. Baumgarte y C. Faller, "Binaural Cue Coding – Part I: Psychoacoustic fundamentals and design principles." IEEE Trans on Speech and Audio Proc., vol. 11, no. 6, Noviembre del 2003;
- C. Faller y F. Baumgarte, "Binaural Cue Coding – Parte II: Schemes and applications, IEEE Trans. On Speech and Audio Proc., vol. 11, no. 6, Noviembre del 2003; y
- 45 ○ . C.Faller, "Coding of spatial audio compatible with different playback formats," Preprint 117th Conv. Aud. Eng. Soc., Octubre del 2004.

Campo de la Invención

50 La presente invención se refiere a la codificación de señales de audio y a la síntesis subsiguiente de escenas de auditorio de los datos de audio codificados.

Descripción de la Técnica Relacionada

55 Cuando una persona oye una señal de audio (es decir, sonidos) generada por una fuente de audio particular, esta señal de audio típicamente llegará a .los oídos, izquierdo y derecho, de la persona, en dos diferentes momentos y con dos diferentes niveles de audio (es decir, decibeles), donde esos diferentes momentos y niveles son funcione de las diferencias en las trayectorias a través de las cuales la señal de audio viaja para llegar a los oídos, izquierdo y derecho, respectivamente. El cerebro de la persona interpreta estas diferencias en el tiempo y nivel para suministrar a la persona la percepción. Esa señal de audio recibida está siendo generada por una fuente de audio ubicada en una posición particular (por ejemplo, dirección y distancia) con relación a la persona. Una escena de auditorio es el efecto neto que una persona oye simultáneamente las

60

señales de audio generadas por una o más fuentes de audio ubicadas en una o más diferentes posiciones con relación a la persona.

5 La existencia de este proceso por el cerebro puede ser usado para sintetizar escenas de auditorio, donde las señales de audio de una o más diferentes fuentes de audio son modificadas a propósito, para generar señales de audio izquierdas y derechas, que dan la percepción que diferentes fuentes de audio se ubican en diferentes posiciones con relación al escucha.

10 La Figura 1 muestra un diagrama de bloques de alto nivel del sintetizador 100 de señal binaural convencional, el cual convierte una sola señal de una fuente de audio (por ejemplo, una mono-señal) en señales de audio, izquierdas y derechas de una señal binaural, donde una señal binaural se define será las dos señales recibidas en los tambores del oído de un escucha. Además de la señal de fuente de audio, el sintetizador 100 recibe un conjunto de señales espaciales, que corresponden a la posición deseada de la fuente de audio con relación al escucha, En formas de realización típicas, el conjunto de señales espaciales comprende un valor de diferencia del nivel Inter.-canales (ICLD) (el cual identifica la diferencia en el nivel de audio entre las señales de audio, izquierdas y derechas, conforme son recibidas en los oídos izquierdo y derecho, respectivamente) y un valor de diferencia de tiempo Inter.-canales (ICTD) (el cual identifica la diferencia en el tiempo de la llegada entre las señales de audio izquierdas y derechas, según se reciben en los oídos, izquierdo y derecho, respectivamente). Además, o como una alternativa, algunas técnicas de síntesis implican el modelo de una dirección dependiente de la función de transferencia para el sonido, desde la fuente de señal a los tambores del oído, también referida como la función de transferencia (HRTF). Véase, por ejemplo, B. Blauert, *The Psychophysics of Human Sound Localization* (La Psicofísica de la Localización del Sonido Humano) MIT Press, 1983.

20 Usando el sintetizador 100 de señal binaurales de la Figura 1, la señal de mono audio generada por una sola fuente de sonido, puede ser procesada, de modo que, cuando se escuche en los audífonos, la fuente de sonido es colocada espacialmente por la aplicación de un conjunto apropiado de indicadores espaciales (por ejemplo, ICLD, ICTD y/o HRTF) para generar la señal de audio para cada oído. Véase, D.R. Begault, *3-D Sound for Virtual Reality and Multimedia*, (Sonido en tres dimensiones para realidad virtual y multimedia), Academic Press, Cambridge, Mass, 1994.

30 El sintetizador 100 de señal binaural de la Figura 1, genera el tiempo más simple de escenas de auditorio, aquellas que tienen una sola fuente de audio colocada con relación al escucha. Escenas de auditorio más complejas, que comprenden dos o más fuentes de audio, ubicadas en diferentes posiciones con relación del escucha, pueden ser generadas usando un sintetizador de escena de auditorio, que es esencialmente realizado usando múltiples casos del sintetizador de señales binaurales, donde cada caso de sintetizador de la señal binaural genera la señal binaural que corresponde a una diferente fuente de audio. Puesto que cada diferente fuente de audio tiene una diferente ubicación con relación al escucha, un diferente conjunto de pistas espaciales se usa para generar la señal de audio binaural para cada diferente fuente de audio.

35 La tesis de doctorado PHD " Parametric coding of spatial audio" de Christof Faller describe diferentes escenarios BCC, en los que los canales originales se mezclan de manera descendente a uno o una pluralidad de canales de mezcla descendente, y donde se calcula la información paramétrica secundaria para los canales originales, se transmite a un decodificador y se utiliza para mezclar de manera ascendente el uno o más canales de mezcla descendente.

40 El documento WO2004/049309A describe un método para codificar una señal de audio, en el que se calculan primeros parámetros cuya característica representa un aspecto de la señal de audio en una primera instancia. Además, también se calculan valores de segundos parámetros que representan el aspecto de la señal de audio en una instancia posterior. Un subconjunto de los segundos parámetros está asociado con una porción particular de un rango de frecuencias de la señal de audio. Este rango de frecuencias se selecciona para cubrir todas las frecuencias presentes en la señal de audio.

50 El documento WO2005/069274A1 describe un aparato y un método para construir una señal de salida multicanal o para generar una señal de mezcla descendente, en que pistas ICC y pistas ICTD para ambos lados se calculan en un codificador. Estos dos valores pueden ser transmitidos desde un codificador a un decodificador. Alternativamente, el codificador puede calcular una pista resultante de ICC o ICTD introduciendo las pistas para ambos lados en una función matemática tal como una función de promedio, etc., para derivar el valor resultante de las dos medidas de coherencia. La fecha de prioridad del documento WO 2005/069274 A1 es anterior a la fecha de prioridad de la presente memoria descriptiva y la fecha de publicación del documento WO 2005/069274 A1 es posterior a la fecha de prioridad de la presente memoria descriptiva.

55 Es un objeto de la presente invención proporcionar un concepto de codificación/decodificación mejorado para canales de audio.

60 Este objeto se consigue mediante un método para codificar canales de audio de acuerdo con la reivindicación 1, un aparato para codificar canales de audio de acuerdo con la reivindicación 9, un medio legible por máquina de acuerdo con la reivindicación 12, una secuencia de bits de audio codificado de acuerdo con la reivindicación 13, un método para decodificar

de acuerdo con la reivindicación 14, un aparato para decodificar de acuerdo con la reivindicación 24, o un medio legible por máquina de acuerdo con la reivindicación 26.

DESCRIPCIÓN DE LA INVENCION

5

De acuerdo con una realización, la presente invención es un método, aparato, y medio leíble por máquina para codificar canales de audio. Uno o más códigos de pista se generan por dos o más canales de audio, en que al menos un código de pista es un código de pista combinado, generado mediante la combinación de dos o más códigos de pista estimados utilizando una función promedio ponderada considerando la importancia de los canales individuales, y cada código de pista estimado es estimado desde un grupo de dos más canales de audio.

10

15

De acuerdo con otra realización, la presente invención es un aparato para codificar canales de audio de entrada C, para generar E canales de audio transmitidos. El aparato comprende un estimador de código y un mezclador descendente. El estimador de código genera uno o más códigos de pista para dos o más canales de audio, en que al menos un código de pista es un código de pista combinado, generado mediante la combinación de dos o más códigos de pista estimados utilizando una función promedio ponderada considerando la importancia de los canales individuales, y cada código de pista estimado se estima de un grupo de dos o más de los canales de audio. El mezclador descendente mezcla hacia abajo los canales de entrada C, para generar los E canales transmitidos, donde $C > E \geq 1$, en que el aparato está adaptado para transmitir información acerca de los códigos de pista para habilitar a un decodificador a ejecutar un procesamiento de síntesis durante la decodificación de los E canales transmitidos.

20

25

De acuerdo con otra realización, la presente invención es una secuencia de bits de audio codificado, generador por codificar canales de audio, en que uno o más códigos de pista son generados por dos o más canales de audio, en que al menos un código de pista es un código de pista combinado, generado combinando dos o más códigos de pista estimados utilizando una función promedio ponderada considerando la importancia de los canales individuales, y cada código de pista estimado se estima de un grupo de dos o más de los canales de audio. Este uno o más códigos de pista y canales de audio transmitidos E corresponden a dos o más canales de audio, en que $E \geq 1$, se codifican en la secuencia de bits de audio codificado.

30

De acuerdo con otra realización, la presente invención es una secuencia de bits de audio codificado, que comprende uno o más códigos de pista y E canales de audio transmitidos. Este uno o más códigos de pista son generados por dos o más canales de audio, en que al menos un código de pista es un código de pista combinado, generado combinando dos o más códigos de pista estimados utilizando una función promedio ponderada considerando la importancia de los canales individuales, y cada código de pista estimado se estima de un grupo de dos o más de los canales de audio. Los E canales de audio transmitidos corresponden a dos o más canales de audio.

35

40

De acuerdo con otra realización, la presente invención es un método, aparato y medio que se puede leer por máquina para decodificar los E canales de audio transmitidos, para generar canales de audio de reproducción C, donde $C > E \geq 1$. Los códigos de pista que corresponden a los E canales transmitidos son recibidos, en que al menos un código de pista es un código de pista combinado, generado combinando dos o más códigos de pista estimados, y cada código de pista estimado se estima de un grupo de dos o más canales de audio, que corresponden a los E canales transmitidos. Uno o más de los E canales transmitidos se mezclan en forma ascendente, para generar uno o más canales mezclados ascendentemente. Uno o más de los canales de reproducción C se sintetizan aplicando los códigos de pista a este uno o más canales mezclados en forma ascendente, en que dos o más códigos de pista derivados se derivan del código de pista combinado y cada código de pista derivado es aplicado para generar dos o más canales sintetizados.

45

BREVE DESCRIPCIÓN DE LOS DIBUJOS

50

Otros aspectos, características y ventajas de la presente invención llegarán a ser más completamente evidentes de la siguiente descripción detallada, las reivindicaciones anexas y los dibujos acompañantes, en los cuales números de referencia similares identifican elementos similares o idénticos.

La Figura 1 muestra un diagrama de bloques de alto nivel del sintetizador de señal binaural;

55

La Figura 2 es un diagrama de bloques de una codificación de pista binaural genérica (BCC) de un sistema de procesamiento de audio;

La Figura 3 muestra un diagrama de bloques de un mezclador descendente, que puede ser usado como el mezclador descendente de la Figura 2;

60

La Figura 4 muestra un diagrama de bloques de un sintetizador de BCC, que puede ser usado para el decodificador de la Figura 2;

La Figura 5 muestra un diagrama de bloques del estimador BCC de la Figura 2, de acuerdo con una realización de la presente invención;

5 La Figura 6 ilustra la generación de los datos de la ICTD y la ICLD para el audio de cinco canales;

La Figura 7 ilustra la generación de datos ICC para el audio de cinco canales;

10 La Figura 8 muestra un diagrama de bloques de una forma de realización del sintetizador BCC de la Figura 4, que puede ser usado en el decodificador BCC, para generar una señal de audio estéreo o de múltiples canales, dada una señal de suma transmitida sencilla $s(n)$ más las pistas espaciales;

15 La Figura 9 ilustra cómo son variados los ICTL y las ICLD dentro de una sub-banda, como una función de la frecuencia;

La Figura 10 muestra un diagrama de bloques de un sintetizador de BCC que puede ser usados para el decodificador de la Figura 2 para un esquema 5 a 2 de BCC; y

20 la Figura 11 muestra un diagrama de flujo del proceso de un sistema BCC, tal como aquel mostrado en la Figura 2, con relación a una realización de la invención.

DESCRIPCIÓN DETALLADA

25 En la codificación de pista binaural (BCC), un codificador codifica los canales de audio de entrada C para generar E canales de audio transmitidos, donde $C > E \geq 1$. En particular, dos o más de los canales de entrada C son provistos en un dominio de frecuencia, y uno o más de los códigos de pista son generados para cada cual de uno o más bandas de frecuencia diferentes, en los dos o más canales de entrada en el dominio de frecuencia. Además, los canales de entrada C son mezclados en forma descendente para generar los E canales transmitidos. En algunas formas de realización de la mezcla descendente, al menos uno de los E canales transmitidos se basa en dos o más de los canales de entrada C y al menos uno de los E canales transmitidos se basa en solamente uno de los canales de entrada.

35 En una realización, un codificador de BCC tiene dos o más bancos de filtros, un estimador de código y un mezclador descendente. Los dos o más bancos de filtros convierten dos o más de los canales de entrada C desde un dominio de tiempo en un dominio de frecuencia. El estimador de código genera uno o más códigos de pista para cada cual de una o más diferentes bandas de frecuencia, en dos o más canales de entrada convertidos. El mezclador descendente, mezcla los canales de entrada C para generar los E canales transmitidos, donde $C > E \geq 1$.

40 En la decodificación de BCC, los E canales de audio transmitidos son decodificados para generar canales de audio de reproducción en particular, para cada uno o más bandas de frecuencia diferente, uno o más de los E canales transmitidos se mezclan en forma ascendente en un dominio de frecuencia para generar dos o más de los canales de reproducción C en el dominio de frecuencia, donde $C > E \geq 1$. Uno o más de los códigos de pista se aplican a cada cual de uno o más bandas de frecuencia diferentes en dos o más canales de reproducción en el dominio de frecuencia para generar dos o más canales modificados, y los dos o más canales modificados son convertidos desde el dominio de frecuencia en un dominio de tiempo. En algunas formas de realización de mezcla ascendente, al menos uno de los canales de reproducción C se basa en al menos uno de los E canales transmitidos y al menos un código de pista y al menos uno de los canales de reproducción C se basa en solamente uno de los E canales transmitidos, sencillo e independiente de cualquier código de pista.

50 En una realización, un decodificador de BCC tiene un mezclador ascendente, un sintetizador y uno o más bancos de filtro inversos. Para cada una o más diferentes bandas de frecuencia, el mezclador ascendente mezcla uno o más de los E canales transmitidos en un dominio de frecuencia para generar dos o más de los canales de reproducción c en el dominio de frecuencia, donde $C > E \geq 1$. El sintetizador aplica uno o más códigos de pista a cada una o más bandas de frecuencia diferentes en dos o más canales de reproducción en el dominio de frecuencia, para generar dos o más canales modificados. Este uno o más bancos de filtro inversos convierten los dos o más canales modificados del dominio de frecuencia en un dominio de tiempo.

55 Dependiendo de la forma de realización particular, un canal de reproducción dado puede ser basado en un canal transmitido sencillo, más bien que una combinación de dos o más canales transmitidos. Por ejemplo, cuando hay solamente un canal transmitido, cada uno de los C canales de reproducción se basa en ese canal transmitido. En estas situaciones, la mezcla ascendente corresponde a la copia del canal transmitido correspondiente. Como tal, para aplicaciones en que hay solamente un canal transmitido, el mezclador ascendente puede ser realizado usando un replicador, que copia el canal transmitido para cada canal de reproducción.

60

Los codificadores de BCC y/o decodificadores pueden ser incorporados en un número de sistemas o aplicaciones, que incluyen, por ejemplo grabadoras / reproductoras de video digital, grabadoras/reproductoras de audio digital, computadoras, transmisores/receptores de satélite, transmisores/receptores de cable, transmisores/receptores de comunicaciones terrestres, sistemas de entretenimiento del hogar y sistemas de películas.

5

Proceso BCC Genérico

La Figura 2 es un diagrama de bloques de un sistema 200 de procesamiento de audio de codificación de pista binaural genérico (BCC), que comprende un codificador 202 y un decodificador 204. El codificador 202 incluye el mezclador descendente 206 y un estimador 208 del BCC.

10

El mezclador descendente 206 convierte los C canales de audio de entrada $x_i(n)$ en E canales de audio transmitidos $y_i(n)$, donde $C > E \geq 1$. En este documento, las señales expresadas usando la variable n son señales de dominio de tiempo, mientras las señales expresadas usando la variable k son señales de dominio de frecuencia. Dependiendo de la forma de realización particular, la mezcla descendente puede ser realizada en cualquiera del dominio de tiempo o el dominio de frecuencia. El estimador de BCC 208 genera códigos de BCC de los canales de audio de entrada C y transmite sus códigos de BCC como información secundaria en banda o fuera de banda, con relación a los E canales de audio transmitidos. Los códigos de BCC típicos incluyen uno o más de diferencia de tiempo de Inter.-canales (ICTD), diferencia de nivel de Inter.-canales (ICLD) y datos de correlación Inter.-canales (ICC) estimados entre ciertas parejas de canales de entrada, como una función de la frecuencia y el tiempo. La forma de realización particular dictará entre cuáles parejas particulares de canales de entrada y códigos de BCC se estima.

15

20

Los datos de ICC que corresponden a la coherencia de una señal binaural, que se relaciona al ancho percibido de la fuente de audio. Cuanto más ancha es la fuente de audio, menor será la coherencia entre los canales izquierdo y derecho de la señal binaural resultante. Por ejemplo, la coherencia de la señal binaural, que corresponde a una orquesta se difunde sobre una etapa de auditorio que es típicamente menor que la coherencia de la señal binaural que corresponde a un solo violín, que se toca solo. En general, una señal de audio con coherencia menor, es usualmente percibida como mayor difundida en el espacio del auditorio. Como tal, los datos de ICC se relacionan típicamente al ancho de fuente aparente y un grado del escucha encerrado. Véase, por ejemplo, en J. Blauert, The Psychophysics of Human Sound Localización (La Psicofísica de la Ubicación del Sonido Humano, por MIT Press, 1983).

25

30

Dependiendo de la aplicación particular, los E canales de audio transmitidos y los códigos de B correspondientes, pueden ser transmitidos directamente al decodificador 204 o almacenados en algún tipo adecuado de dispositivo de almacenamiento, para el acceso subsiguiente por el decodificador 204. Dependiendo de la situación, el término de "transmisión" puede referirse a cualquiera de la transmisión directa a un decodificador o almacenamiento para la provisión subsiguiente a un decodificador. En cualquier caso, el decodificador 204 recibe los canales de audio transmitidos y la información secundaria y realiza la mezcla ascendente y la síntesis de BCC usando los códigos de BCC para convertir los E canales de audio transmitidos, los canales de audio de reproducción), $x_i(n)$ para la reproducción de audio. Dependiendo de la forma de realización particular, la mezcla ascendente puede ser realizada en cualquiera del dominio de tiempo o el dominio de frecuencia.

35

40

Además del proceso BCC mostrado en la Figura 2, un sistema de procesamiento de audio de BCC genérico puede incluir codificación adicional y etapas de decodificación para comprimir ulteriormente las señales de audio en el codificador y luego descomprimir las señales de audio en el decodificador, respectivamente Estas codificaciones de audio se pueden basar en las técnicas de compresión / descompresión de audio convencionales, tal como aquéllas basadas en la modulación del código de pulso (PCM), PCM diferencial (DPCM) o DPCM adaptivo (ADPCM).

45

Cuando el mezclador descendente 206 genera una simple señal de suma (es decir, $E = 1$), la codificación BCC es capaz de representar señales de audio de múltiples canales en un régimen de bits sólo ligeramente mayor del requerido para representar una señal de mono-audio. Esto es así, debido a que la ICTD, ICLD estimadas y los datos de ICC entre una pareja de canales, contienen alrededor de dos órdenes de magnitud menor de información que una forma de onda de audio.

50

No sólo el régimen de bits bajo de la codificación de BCC, sino también su aspecto de compatibilidad anterior, son de interés. Una señal de suma transmitida sencilla corresponde a una mono mezcla descendente de la señal del estéreo original o de múltiples canales. Para receptores que no soportan la reproducción de sonido estéreo o de múltiples canales. Escuchar la señal de suma transmitida es un método válido de presentar el material de audio en el equipo de reproducción único de bajo perfil. La codificación BCC puede, por lo tanto, ser usado para aumentar los servicios existentes que implican la entrega de material de mono-audio hacia el audio de múltiples canales. Por ejemplo, los sistemas de radio-transmisión de mono-audio existentes, pueden ser aumentados por la reproducción del estéreo o de múltiples canales si la información secundaria de BCC puede ser incrustada en el canal de transmisión existente. Capacidades análogas existen cuando el audio de multi-canales de mezcla descendente de dos señales de suma, corresponden al audio de estéreo.

55

60

La BCC procesa las señales de audio con un cierto tiempo y resolución de frecuencia. La resolución de frecuencia usada es en su mayor parte motivada por la resolución de frecuencia del sistema del auditorio humano. La psicoacústica sugiere que la percepción espacial es más probablemente basada en la representación crítica de banda de la señal de entrada acústica. Esta resolución de frecuencia está considerada usando un banco de filtro invertible (por ejemplo, basado en una transformación rápida de Fourier (FFT) o un filtro de espejo de cuadratura (QMF)) con sub-bandas con anchos de banda iguales o proporcionales al ancho de banda crítico del sistema de auditorio humano.

Mezcla Descendente Genérica

En formas de realización preferidas, las señales de suma transmitidas contienen todos los componentes de señal de dicha señal de audio de entrada. La meta es que cada componente de señal se mantenga completamente. La suma sencilla de los canales de entrada de audio a menudo resulta en la amplificación o atenuación de los componentes de señal. En otras palabras, la potencia de los componentes de señal en una suma "simple" es a menudo mayor o menor que la suma de la potencia del componente de señal correspondiente de cada canal. Una técnica de mezcla descendente puede ser usada, que iguala la señal de suma de manera que la potencia de los componentes de señal sea la señal de suma de aproximadamente la misma como la potencia correspondiente de todos los canales de entrada.

La Figura 3 muestra un diagrama de bloques de un mezclador descendente 300, que puede ser usado para el mezclador descendente 206 de la Figura 2, de acuerdo con ciertas formas de realización del sistema BCC 200. El mezclador descendente 300 tiene un banco de filtros (FB) 302 para cada canal de entrada $x_i(n)$, un bloque de mezcla descendente 304, un bloque opcional de escala/retardo y un FB inverso (IFB) para cada canal codificado $y_i(n)$.

Cada banco 302 de filtros convierte cada marco (por ejemplo, de 20 mseg) de un canal de entrada digital correspondiente $x_i(n)$, en el dominio de tiempo en un conjunto de coeficientes de entrada $\tilde{x}_i(k)$ en el dominio de frecuencia. El bloque 304 de mezcla descendente mezcla cada sub-banda de coeficientes de entrada correspondiente C en una sub-banda de coeficientes de dominio de frecuencia de mezcla descendente E. La Ecuación (1) representa la mezcla descendente de la sub-banda enésima de coeficientes de entrada $(\tilde{x}_1(k), \tilde{x}_2(k), \dots, \tilde{x}_C(k))$ para generar la enésima sub-banda de coeficientes de mezcla descendente $(\hat{y}_1(k), \hat{y}_2(k), \dots, \hat{y}_E(k))$, como sigue:

$$\begin{bmatrix} \hat{y}_1(k) \\ \hat{y}_2(k) \\ \vdots \\ \hat{y}_E(k) \end{bmatrix} = D_{CE} \begin{bmatrix} \tilde{x}_1(k) \\ \tilde{x}_2(k) \\ \vdots \\ \tilde{x}_C(k) \end{bmatrix}, \tag{1}$$

donde D_{CE} es una matriz de mezcla descendente C por E de valor real.

El bloque 306 opcional de escala/retardo comprende un conjunto de multiplicadores 310, cada uno de los cuales multiplica un coeficiente de mezcla descendente correspondiente $\hat{y}_i(k)$ por un factor de escala $e_i(k)$, para generar un coeficiente de escala correspondiente $\tilde{y}_i(k)$. La motivación para la operación de escala es equivalente a la igualación generalizada para la mezcla descendente con factores de ponderación arbitrarios para cada canal. Si los canales de entrada son independientes, luego la potencia $\tilde{p}_{\tilde{y}_i(k)}$ de la señal de mezcla descendente en cada sub-banda es dada por la Ecuación (2) como sigue:

$$\begin{bmatrix} \tilde{p}_{\tilde{y}_1(k)} \\ \tilde{p}_{\tilde{y}_2(k)} \\ \vdots \\ \tilde{p}_{\tilde{y}_E(k)} \end{bmatrix} = \overline{D}_{CE} \begin{bmatrix} p_{\tilde{x}_1(k)} \\ p_{\tilde{x}_2(k)} \\ \vdots \\ p_{\tilde{x}_C(k)} \end{bmatrix}, \tag{2}$$

donde \bar{D}_{CE} se deriva elevando al cuadrado cada elemento de matriz en la matriz de mezcla descendente C por E, D_{CE} y $P_{x_i(k)}$ es la potencia de la sub-banda k del canal de entrada i.

5 Si las sub-bandas no son independientes, entonces los valores de potencia $P_{y_i(k)}$ de la señal de mezcla descendente serán mayores o menores que esa calculada usando la Ecuación (2), debido a las amplificaciones o cancelaciones cuando los componentes de señal son en fase o fuera de fase, respectivamente. Para prevenir esto, la operación de mezcla descendente de la Ecuación (1) se aplica en las sub-bandas, seguido por la operación de escala de los multiplicadores 310. Los factores de escala $e_i(k)$ ($1 \leq i \leq E$) pueden derivarse usando la Ecuación (3) como sigue:

$$e_i(k) = \sqrt{\frac{P_{\hat{y}_i(k)}}{P_{y_i(k)}}}, \quad (3)$$

10 donde $P_{\hat{y}_i(k)}$ es la potencia de sub-banda como se calculó de la Ecuación (2) y $P_{y_i(k)}$ es la potencia de la señal de sub-banda $\hat{y}_i(k)$ de mezcla descendente correspondiente.

Además de, o en lugar de, proporcionar la escala opcional, el bloque 306 de escala/retardo puede, opcionalmente, aplicar retardos a las señales.

15 Cada banco 308 de filtros inversos convierte un conjunto de coeficientes escalados correspondientes $\tilde{y}_i(k)$ en el dominio de frecuencia en un marco de un canal $y_i(n)$ transmitido digital correspondiente.

20 Aunque la Figura 3 muestra todos los canales de entrada C siendo convertidos en el dominio de frecuencia para la mezcla descendente subsiguiente, en formas de realización alternativas, uno o más (pero menor de C-1) de los canales de entrada C pueden desviarse algo o todo el proceso mostrado en la Figura 3 y ser transmitidos como un número equivalente de canales de audio sin modificar. Dependiendo de la forma de realización particular, estos canales de audio o modificados pueden o no ser usados por el estimador 208 de BCC de la Figura 2, en generar los códigos de BCC transmitidos.

25 En una forma de realización del mezclador descendente 300, que genera una señal de suma sencilla $y(n)$, $E = 1$ y las señales $\tilde{x}_c(k)$, (k) de cada sub-banda de cada canal de entrada c, son agregados y luego multiplicados con un factor $e(k)$, de acuerdo con la Ecuación (4) como sigue:

$$\tilde{y}(k) = e(k) \sum_{c=1}^C \tilde{x}_c(k). \quad (4)$$

30 el factor $e(k)$ se da por la Ecuación (5), como sigue:

$$e(k) = \sqrt{\frac{\sum_{c=1}^C p_{\tilde{x}_c}(k)}{p_{\tilde{x}}(k)}}, \quad (5)$$

35 donde $p_{\tilde{x}_c}(k)$ es una estimación de tiempo corto de la potencia de $\tilde{x}_c(k)$ en el índice k de tiempo y $p_{\tilde{x}}(k)$ es una estimación de tiempo breve de la potencia de $\sum_{c=1}^C \tilde{x}_c(k)$. Las sub-bandas igualadas se transforman de nuevo al dominio de tiempo que resulta en la señal de suma $y(n)$ que es transmitida al decodificador de BCC.

Síntesis de BCC Genérica

La Figura 4 muestra un diagrama de bloques de un sintetizador de B 400, que puede ser usado para el decodificador 204 de la Figura 2, de acuerdo a ciertas formas de realización del sistema de BCC 200. Este sintetizador de BCC 400 tiene un banco 402 de filtros para cada canal transmitido $y_i(n)$, un bloque mezclador ascendente 404, los retardos 406, multiplicadores 408, bloque 410 de correlación y un banco 412 de filtros inversos, para cada canal de reproducción $\hat{x}_i(n)$.

Cada banco 402 de filtros convierte cada marco de un canal $y_i(n)$ transmitido digital correspondiente en el dominio del marco en un conjunto de coeficientes de entrada $\tilde{y}_i(k)$ en el dominio de frecuencia. El bloque de mezcla ascendente 404 mezcla cada sub-banda de E, que corresponde a los coeficientes de canal transmitidos, en una sub-banda correspondiente de coeficientes de dominio de frecuencia de mezcla ascendente C. La Ecuación (4) representa la mezcla ascendente de la enésima subbanda de coeficientes de canal transmitidos $(\tilde{y}_1(k), \tilde{y}_2(k), \dots, \tilde{y}_E(k))$ para generar la enésima sub-banda de coeficientes de mezcla ascendente $(\tilde{s}_1(k), \tilde{s}_2(k), \dots, \tilde{s}_C(k))$, como sigue:

$$\begin{bmatrix} \tilde{s}_1(k) \\ \tilde{s}_2(k) \\ \vdots \\ \tilde{s}_C(k) \end{bmatrix} = U_{EC} \begin{bmatrix} \tilde{y}_1(k) \\ \tilde{y}_2(k) \\ \vdots \\ \tilde{y}_E(k) \end{bmatrix}, \quad (6)$$

donde U_{EC} es una matriz de mezcla ascendente E por C de valor real. Ejecutando la mezcla ascendente en el dominio de frecuencia habilita que esta mezcla ascendente sea aplicada individualmente en cada diferente subbanda.

Cada retardo 407 aplica un valor $d_i(k)$ de retardo basado en un código BCC correspondiente para datos de ICTD, con el fin de asegurar que los valores de ICTD deseados aparezcan entre ciertas parejas de canales de reproducción. Cada multiplicador 408 aplica un factor de escala $a_i(k)$ con base en un código de BCC correspondiente para datos ICLD con el fin de asegurar que los valores de ICLD deseados aparezcan entre ciertas parejas de canales de reproducción. El bloque 410 de correlación realiza una operación de descorrelación A basada en códigos de BCC correspondientes para datos de ICC con el fin de asegurar que los valores de ICC deseados aparezcan entre ciertas parejas de canales de reproducción. Una descripción ulterior de las operaciones del bloque 410 de correlación puede ser encontrada en la solicitud de patente de E.U.A., No. De Serie 10/155,437, presentada el 24 de Mayo del 2003, como Baumgarte 2-10.

La síntesis de los valores de ICLD pueden ser menos problemática que la síntesis de los valores de ICTD y de ICC, puesto que la síntesis de ICLD implica escalar meramente las señales de sub-bandas. Puesto que las pistas de ICLD son las pistas direccionales muy usadas comúnmente es usualmente más importante que los valores de ICLD se aproximen a aquellos de la señal de audio original. Como tal, los datos de ICLD pueden ser estimados entre todas las parejas de canales. Los factores de escala $a_i(k)$ ($1 \leq i \leq C$) para cada sub-banda se escogen preferiblemente de modo que la potencia de sub-banda de cada canal de reproducción se aproxime a la potencia correspondiente del canal de audio de entrada original.

Una meta puede ser aplica modificaciones de señales relativamente pocas para sintetizar los valores de ICTD y de ICC. Como tal, los datos de BCC pueden no incluir los valores de ICTD y de ICC, para todas las parejas de canales. En ese caso, el sintetizador de BCC 400 sintetizará los valores de ICTD y de ICC solamente entre ciertas parejas de canales.

Cada banco de filtro inverso 412 convierte un conjunto de coeficientes sintetizados correspondientes $\tilde{x}_i(k)$ en el dominio de frecuencia en un marco de un canal $\hat{x}_i(n)$ de reproducción digital correspondiente.

Aunque la Figura 4 muestra todos los E de los canales transmitidos siendo convertidos en el dominio de frecuencia para la mezcla ascendente subsiguiente y el proceso de BCC, en formas de realización alternativas, uno o más (pero no todos) los E canales transmitidos pueden desviarse algo o todo del proceso mostrado en la Figura 4. Por ejemplo, uno o más de los canales transmitidos pueden ser canales sin modificar, que no se someten a alguna mezcla ascendente. Además de ser uno o más de los canales de reproducción C, estos canales sin modificar, a su vez, pueden ser, pero no tienen que ser, usados como canales de referencia a los cuales el proceso de BCC se aplica para sintetizar uno o más de los otros canales de reproducción.

En cualquier caso, tales canales sin modificar pueden ser sometidos a retardos para compensar el tiempo de proceso implicado en la mezcla ascendente y/o el proceso de BCC usado para generar el resto de los canales de reproducción.

Nótese que, aunque la Figura 4 muestra canales de producción C siendo sintetizados de e canales transmitidos, donde C es también el número de canales de entrada originales, la síntesis de BCC no se limita a ese número de canales de reproducción. En general, el número de canales de reproducción puede ser cualquier número de canales, que incluyen números mayores de o menores de, y posiblemente aún situaciones donde el número de canales de reproducción es igual a, o menor que, el número de canales transmitidos.

10 “Diferencias Relevantes Perceptualmente” entre Canales de Audio

Suponiendo una señal simple de suma, la BCC sintetiza una señal de audio de canal estéreo o de múltiples canales, de manera que ICTD, ICLD e ICC se aproximen a las pistas o indicaciones correspondientes de la señal de audio original. En lo siguiente, el papel de ICTD, ICLD e ICC, en relación con los atributos de imagen espacial del auditorio, se discute.

El conocimiento acerca del oído espacial implica que un evento de auditorio, ICTD e ICLD se relacionen a la dirección percibida. Cuando se consideran respuestas de impulso de espacio binaural (BRIRx), de una fuente, hay una relación entre el ancho del evento de auditorio y el envolvente del escucha y los datos de ICC estimados para las partes temprana y tardía de las BTIRs. Sin embargo, la relación entre ICC y estas propiedades para señales generales (y no justamente las BRIRs) no son directas.

Las señales de audio de estéreo y de múltiples canales, usualmente contienen una mezcla compleja de señales de fuentes activas concurrentemente superpuestas de componentes de señal reflejadas, que resultan de grabar en espacios encerrados o agregados por el ingeniero de grabación para crear artificialmente una impresión espacial. Diferentes señales de fuente y sus reflexiones ocupan diferentes regiones en el plano de tiempo-frecuencia. Esto es reflejado por ICTD, ICLD e ICC, que varía como una función de tiempo y frecuencia. En este caso, la relación entre el ICTD, ICLD e ICC instantáneos y las direcciones de eventos de auditorio y la impresión espacial no son obvias. La estrategia de ciertas realizaciones de BCC es sintetizar a ciegas estas pistas o indicaciones de modo que se aproximen a las pistas correspondientes de la señal de audio original.

Los bancos de filtros con sub-bandas de anchos de banda iguales a do veces el ancho de banda rectangular equivalente (ERB) se usan. La escucha informal revela que la calidad de audio de BCC no mejora notablemente cuando se escoge la resolución de frecuencia mayor. Una resolución de frecuencia menor puede ser conveniente, puesto que resulta en valores menores de ICTD, ICLD e ICC, de los necesarios para ser transmitidos al decodificador y sí en un régimen menor de bits.

Considerando el tiempo de resolución, ICTD, ICLD e ICC se consideran típicamente en intervalos de tiempo regulares. Se obtiene el alto desempeño cuando ICTD, ICLD e ICC se consideran aproximadamente cada 4 a 16 ms. Nótese que, a no ser que las pistas se consideren en intervalos de tiempo muy cortos, el efecto de precedencia no se considera directamente. Suponiendo una pareja clásica de guía retardada de estímulos de sonido, si esta guía y retardo caen en un intervalo de tiempo donde solamente un conjunto de pistas se sintetiza, entonces el dominio de ubicación de la guía no se considera. A pesar de esto, BCC logra calidad de audio reflejada en una clasificación promedio de MUSHRA de aproximadamente 87 (es decir, “excelente” calidad de audio) en promedio hasta casi 100 para ciertas señales de audio.

La pequeña diferencia perceptualmente, lograda a menudo, entre la señal de referencia y la señal sintetizada, implica que las pistas relacionadas a un amplio intervalo de atributos de imagen especial de auditorio son consideradas implícitamente sintetizando ICTD, ICLD e ICC en intervalos de tiempo regulares. En lo siguiente, se dan algunos argumentos de cómo ICTD, ICLD e ICC pueden relacionarse a un intervalo de atributos de imagen espaciales de auditorio.

50 Estimación de Pistas Espaciales

En lo siguiente, se describirá cómo ICTD, ICLD e ICC se estimaron. La tasa de bits para la transmisión de estas pistas espaciales (cuantificadas y codificadas) puede ser justamente de unos cuantos kb/s y así, con BCC, es posible transmitir señales de audio estéreo y de múltiples canales a regímenes de bits cercanos a los que se requieren para un solo canal de audio.

La Figura 5 muestra un diagrama de bloques del estimador de BCC 208 de la Figura 2, de acuerdo con una realización de la presente invención. El estimador 208 de BCC comprende bancos de filtros (FB) 502 que pueden ser los mismos que los bancos de filtros 302 de la Figura 3, y el bloque 504 de estimación, el cual genera las pistas espaciales ICTD, ICLD e ICC para cada sub-banda de frecuencia, generada por los bancos de filtro 502.

Estimación de ICTD, ICLD e ICC para señales Estéreo

Se usan las siguientes medidas para ICTD, ICLD e ICC para las señales de sub-bandas correspondientes $\tilde{x}_1(k)$ y $\tilde{x}_2(k)$ de dos canales de audio (por ejemplo estéreo).

- 5 o ICTD (muestras)

$$\tau_{12}(k) = \arg \max_d \{ \Phi_{12}(d, k) \}, \quad (7)$$

con una estimación de tiempo breve de la función de correlación transversal normalizada, dada por la ecuación (8) como sigue:

$$\Phi_{12}(d, k) = \frac{p_{\tilde{x}_1\tilde{x}_2}(d, k)}{\sqrt{p_{\tilde{x}_1}(k-d_1)p_{\tilde{x}_2}(k-d_2)}}, \quad (8)$$

10 donde

$$\begin{aligned} d_1 &= \max\{-d, 0\} \\ d_2 &= \max\{d, 0\}, \end{aligned} \quad (9)$$

15 y $p_{\tilde{x}_1\tilde{x}_2}(d, k)$ es una estimación de tiempo breve del medio de $\tilde{x}_1(k-d_1)\tilde{x}_2(k-d_2)$.

- o ICLD [dB]:

$$\Delta L_{12}(k) = 10 \log_{10} \left(\frac{p_{\tilde{x}_2}(k)}{p_{\tilde{x}_1}(k)} \right). \quad (10)$$

20 o ICC:

$$c_{12}(k) = \max_d |\Phi_{12}(d, k)|. \quad (11)$$

Nótese que el valor absoluto de la correlación transversal normalizada se considera y $c_{12}(k)$ tiene un intervalo de [0,1].

Estimación de ICTD, ICLD e ICC para Señales de Audio de Múltiples canales

25

Cuando hay más de dos canales de entrada, es típicamente suficiente definir ICTD e ICLD entre un canal de referencia por ejemplo el número 1 de canal) y los otros canales, como se ilustró en la Figura, para el caso de canales $C = 5$, donde $\tau_{1c}(k)$ y $\Delta L_{1c}(k)$ denotan ICTD e ICLD, respectivamente, entre el canal de referencia 1 y el canal c .

30

En formas opuestas a ICTD e ICLD, ICC típicamente tiene más grados de libertad. La ICC según se definió, puede tener diferentes valores entre todas las posibles parejas de canal de entrada. Para los canales C , hay $C(C-1)/2$ parejas de canales posibles, p.ej. para 5 canales hay 10 parejas de canales como se ilustra en la Figura 7(a). Sin embargo, tal esquema requiere que para cada sub-banda en cada índice de tiempo, $C(C-1)/2$ valores de ICC se estimen y transmitan, lo que resulta en alta complejidad computacional y alto régimen de bits.

35

5 Alternativamente, para cada sub-banda, ICTD e ICLD determinan la dirección en la cual el evento de auditorio del componente de señal correspondiente en la sub-banda se realizó. Un parámetro de ICC simple por sub-banda, puede luego ser usado para describir la coherencia general entre todos los canales de audio. Buenos resultados pueden ser obtenidos estimando y transmitiendo pistas de ICC sólo entre los dos canales con más energía en cada sub-banda en cada índice de tiempo. Esto se ilustra en la Figura 7(b), donde para los instantes de tiempo $k-1$ y k , las parejas de canales (3, 4) y (1, 2) son más fuertes respectivamente. Una regla heurística puede ser usada para determinar ICC entre las otras parejas de canales.

Síntesis de Pistas Espaciales

10 La Figura 8 muestra un diagrama de bloques de una forma de realización del sintetizador de BCC 400 de la Figura 4, que puede ser usado en un decodificador de BCC para generar una señal de audio estéreo o de múltiples canales, dada una señal $s(n)$ de suma transmitida sencilla, más las pistas o indicaciones espaciales. La señal de suma $s(n)$ se descompone en sub-bandas, donde $\tilde{s}(k)$ denota una de tales sub-bandas. Para generar las sub-bandas correspondientes de cada canal de salida, se aplican los retardos d_c , factores de escala a_c y filtros h_c a la sub-banda correspondiente de la señal de suma. (Para simplicidad de notación, el índice de tiempo k es ignorado en los retardos, factores de escala y filtros.) Las ICTD se sintetizan por imponer retardos, las ICLD por escala, y las ICLD por aplicar filtros de descorrelación. El proceso mostrado en la Figura 8 se aplicó independientemente a cada sub-banda.

Síntesis de ICTD

20 Los retardos d_c se determinaron de los ICTD $\tau_{1c}(k)$, de acuerdo con la Ecuación (12), como sigue:

$$d_c = -\frac{1}{2}(\max_{2 \leq l \leq C} \tau_{1l}(k) + \min_{2 \leq l \leq C} \tau_{1l}(k)), \text{ for } c = 1 \quad (12)$$

$$d_c = \tau_{1c}(k) + d_1, \text{ for } 2 \leq c \leq C.$$

25 El retardo para el canal de referencia d_1 es calculado de modo que la magnitud máxima de los retardos d_c sea mínima. Cuanto menos señales de sub-banda sean modificados, menor será la ocurrencia de un peligro para los artefactos. Si el régimen de muestreo de la sub-banda no proporciona una resolución de tiempo suficientemente alta para la síntesis de la ICTD, los retardos se pueden imponer más precisamente por usar filtros adecuados de todos los pasos.

Síntesis de la ICLD

30 Con el fin que las señales de sub-banda de salida tengan las ICLDs $\Delta L_{12}(k)$ deseadas entre el canal c y el canal 1 de referencia, los factores de ganancia a_c deben satisfacer la Ecuación (13) como sigue:

$$\frac{a_c}{a_1} = 10^{\frac{\Delta L_{1c}(k)}{20}}. \quad (13)$$

35 Adicionalmente, las sub-bandas de salida son normalizadas preferiblemente, de modo que la suma de la potencia de todos los canales de salida sea igual a la potencia de la señal de suma de entrada. Puesto que la potencia de la señal original total en cada sub-banda se preserva en la señal de suma, esta normalización resulta en la potencia de sub-banda absoluta para cada canal de salida, que se aproxima a la potencia correspondiente de la señal de audio de entrada del codificador original. Dadas estas restricciones, los factores de escala a_c son dados estas restricciones, los factores de escala a_c son dados por la Ecuación 14 como sigue:

$$a_c = 1/\sqrt{1 + \sum_{i=2}^C 10^{\Delta L_{1i}/10}}, \text{ para } c = 1 \quad (14)$$

$$a_c = 10^{\Delta L_{1c}/20} a_1, \text{ De otra manera } (c \neq 1).$$

40

Síntesis de la ICC

En ciertas realizaciones, el objeto de la síntesis de la ICC es reducir la correlación entre las sub-bandas después que los retardos y el escalado se han logrado. Sin afectar la ICTD y la ICLD. Esto puede lograrse por diseñar los filtros h_c en la Figura 6, de modo que ICTD e ICLD sean variadas efectivamente como una función de frecuencia, de modo que la variación promedio sea de cero, en cada sub-banda (banda crítica de auditorio).

La Figura 9 ilustra cómo ICTD e ICLD son variadas dentro de una subbanda como una función de la frecuencia. La amplitud de la variación de ICTD e ICLD determina el grado de la descorrelación y se controla como una función de la ICC, mientras la ICLD varían aleatoriamente (como en la Figura 9(b)). Se puede variar ICLD tan suavemente como ICTD, pero esto resultaría en mayor correlación de las señales de audio resultantes.

Otro método de sintetizar la ICC, particularmente adecuado para la síntesis de ICC de múltiples canales, se describe en mayor detalle en C. Faller, "Codificación de audio de múltiples canales paramétricos. Síntesis de pistas de coherencia" IEEE Trans on Speech Síntesis and Audio Proc., 2003, cuyas enseñanzas se incorporan aquí como referencia. Como una función de tiempo y frecuencia, cantidades específicas de reverberación tardía artificial se agregan a cada uno de los canales de salida para lograr una ICC deseada. Adicionalmente, la modificación espectral puede ser aplicada, tal como la envoltura espectral de los acercamientos de señales resultantes, la envoltura espectral de la señal de audio original.

Otras técnicas de síntesis de la ICC relacionadas y no relacionadas para las señales estéreo (o parejas de canales de audio) se han presentado en E. Chuijiers, W. Oomenm B den Brinker y J. Breebaart. "Avances en la codificación paramétrica para audio de alta calidad", en Preprint 114° Conv. Aud. Eng. Soc., Marzo del 2003 y J. Engdegard, H. Purnhagen, J. Roden y L. Liljeryd, "Ambiente Sintético en la codificación estéreo paramétrica! En Preprint 117° Conv. Aud. Eng. Soc. Mayo del 2004, las enseñanzas de ambos se incorporan aquí como referencia.

C a E BCC

Como se describió previamente, BCC se puede llevar a cabo con más de un canal de transmisión. Una variación de BCC se ha descrito, la cual representa canales de audio C no como un canal simple (transmitido), sino como E canales denotados C a E BCC. Hay (al menos) dos motivaciones para C a E BCC.

- BCC con un canal de transmisión proporciona una trayectoria compatible hacia atrás para mejorar los mono-sistemas existentes, para la reproducción de audio estéreo o de múltiples canales. Los sistemas mejorados transmiten la señal de suma mezclada en forma descendente de BCC a través de la mono-infraestructura existente, mientras transmite adicionalmente la información secundaria de BCC, C a E BCC es aplicable a la codificación compatible hacia atrás del canal E de audio del canal C.
- C a E BCC introduce la capacidad de escalabilidad en términos de diferentes grados de reducción del número de canales transmitidos. Se espera que cuantos más canales de audio que se transmitan, mejor será la calidad del audio.

Los detalles del proceso de señal para C a E BCC, tal como cómo definir las pistas ICTD, ICLD e ICC, se describen en la solicitud de E.U.A, No. De Serie 10/762,100, presentada el 20 de enero del 2004 (Faller 13-1).

Compactar Información Secundaria

Como se describió antes, en un esquema de BCC típico, el codificador transmite al decodificador ICTD, ICLD y/o ICC códigos estimados entre parejas o grupos diferentes de canales de audio. Esta información secundaria se transmite además a las señales de mezcla descendente (por ejemplo, mono o estéreo), con el fin de obtener una señal de audio de múltiples canales después de la decodificación de BCC. Así, es conveniente minimizar la cantidad de la información secundaria mientras no degrade la calidad subjetiva del sonido decodificado.

Puesto que los valores de ICLD e ICTD típicamente se relacionan a un canal de referencia, los valores C-1 de ICLD e ICTD son suficientes para describir las características de los C canales codificados. Por otra parte, las ICC se definen entre parejas arbitrarias de canales. Como tales, para C canales codificados, hay $C(C-1)/2$ parejas ICC posibles. Para 5 canales codificados, esto corresponderá a 10 parejas de ICC. En la práctica, con el fin de limitar la cantidad de información de ICC transmitida, sólo la información de ICC para ciertas parejas será transmitida.

La Figura 10 muestra un diagrama de bloques de un sintetizador de BCC 1000, que puede ser usado para el decodificador 204 de la Figura 2 para un esquema de BCC 5 a 2. Como se muestra en la Figura 10, el sintetizador de BCC 1000 recibe dos

señales de entrada $y_1(n)$ e $y_2(n)$ y la información secundaria BCC (no mostrada) y genera cinco señales de salida sintetizadas $\hat{x}_1(n), \dots, \hat{x}_5(n)$, donde la primera, segunda, tercera, cuarta y quinta señales de salida corresponden a las señales envolventes izquierda, derecha, centro, izquierda posterior y derecha posterior, respectivamente, mostradas en las Figuras 6 y 7.

5 Los parámetros de retardo, escala y descorrelación, derivados de la información secundaria de ICTD, ICLD e ICC transmitidas, se aplican a los elementos 1004, 1006 y 1008, respectivamente, para sintetizar las cinco señales de salida $\hat{x}_i(n)$ desde las cinco señales de "mezcla ascendente" $\tilde{s}_i(k)$, generadas por el elemento 1002 de mezcla ascendente. Como se muestra en la Figura 10, la descorrelación se realiza solamente entre los canales izquierdo y posterior izquierdo (es decir los canales 1 y 4) y entre los canales derecho y posterior derecho, (es decir, los canales 2 y 5). Como tal, no más de dos conjuntos de datos de ICC necesitan ser transmitidos al sintetizador 1000 de BCC, donde estos dos conjuntos caracterizan los valores de ICC entre las dos parejas de canales para cada sub-banda. Mientras esto ya es una reducción considerable en la cantidad de la información lateral de ICC, es conveniente una reducción ulterior.

15 De acuerdo con una realización de la presente invención, en el contexto del esquema BCC de 5 a 2 de la Figura 10 para cada sub-banda, el codificador de BCC correspondiente combina el valor de ICC estimado a la pareja de canal "izquierdo / posterior izquierdo" con el valor de ICC estimado para la pareja del canal "derecho / posterior derecho con el fin de generar un valor de ICC combinado sencillo, que indique efectivamente una cantidad global de descorrelación frontal/posterior y que se transmite al decodificador de BCC como la información secundaria de ICC. Experimentos informales indicaron que esta simplificación resulta en virtualmente ninguna pérdida en la calidad de audio, mientras reduce la información de ICC transmitida por un factor de dos.

20 En general, realizaciones de la presente invención se dirigen a esquemas de BCC en los cuales dos o más diferentes ICC estimadas entre diferentes parejas de canales, o grupos de canales, se combinan para la transmisión, como se indica por la Ecuación (15) como sigue:

25

$$ICC_{transmitida} = f(ICC_1, ICC_2, \dots, ICC_N), \quad (15)$$

donde f es una función que combina N diferentes ICCs.

30 Con el fin de obtener una medida de la ICC, que sea representativa de la imagen espacial, puede ser ventajoso usar un promedio ponderado para la función f , que considere la importancia de canales individuales, donde la importancia del canal puede ser basad en las potencias del canal, como se representa por la Ecuación (16) como sigue:

$$ICC_{transmitida} = \frac{\sum_{i=1}^N p_i ICC_i}{\sum_{i=1}^N p_i}, \quad (16)$$

35 donde p_i es la potencia de la pareja de canales correspondiente en la sub-banda. En este caso, las ICC estimadas de las parejas de Canales más fuertes son ponderadas más que las ICC estimadas de las parejas de canales más débiles. La potencia combinada p_i de una pareja de canales puede ser calculada como la suma de potencias de canales individuales para cada sub-banda.

40 En el decodificador, dada la $ICC_{transmitida}$, las ICC pueden ser derivadas para cada pareja de canales. En una posible forma de realización, el decodificador usa simplemente la $ICC_{transmitida}$ como el código de ICC derivado para cada pareja de canales. Por ejemplo, en el contexto del esquema de BCC 5 a 2 de la Figura 10, la $ICC_{transmitida}$ puede ser usada directamente para la descorrelación de la pareja de canales tanto izquierdo/posterior izquierdo como la pareja de canales derecho/posterior derecho.

45

En otra posible forma de realización, si el decodificador estima potencias de parejas de canales de las señales sintetizadas, entonces la ponderación de la Ecuación (16) puede ser estimada y el proceso del decodificador puede usar opcionalmente esta información y otros argumentos preceptuales y estadísticos de señal para generar una regla para derivar dos códigos de ICC optimizados perceptualmente, individuales.

50

Aunque la combinación de los valores de ICC se ha descrito en el contexto de un esquema de BCC de 5 a 2 particular, la presente invención puede ser realizada en el contexto de cualquier esquema de BCC de C a E, que incluye aquellos en donde $E = 1$.

5 La Figura 11 muestra un diagrama de flujo del procesamiento de un sistema BCC, tal como el mostrado en la Figura 2, relacionado a una realización de la presente invención. La Figura 11 muestra sólo esas etapas asociadas con el procesamiento relacionado con la ICC.

10 En particular, el codificador de BCC estima los valores de ICC entre dos o más grupos de canales (Etapas 1102), combina dos o más de estos valores de ICC estimados, para generar uno o más valores de ICC combinados (etapa 1104) y transmite los valores de ICC combinados (posiblemente junto con uno o más valores de ICC "no combinados") como la información secundaria de BCC a un decodificador de BCC (etapa 1106). El decodificador de BCC deriva dos o más valores de ICC desde los valores de ICC combinados, recibidos (etapa 1108) y descorrelaciona grupos de canales usando los valores de ICC derivados (y posiblemente uno o más valores de ICC no combinados, recibidos) (etapa 1110).

15 Otras Realizaciones Alternativas

20 La presente invención se ha descrito en el contexto del esquema de BCC de 5 a 2 de la Figura 10. En ese ejemplo, un codificador de BCC (1) estima dos códigos de ICC para dos parejas de canales, que consisten de cuatro canales diferentes (es decir, izquierdo / posterior izquierdo y derecho / posterior derecho) y (2) promedios de esos dos códigos de ICC, para generar un código de IC combinado, el cual se transmite al decodificador de BCC. Este decodificador (1) de BCC deriva dos códigos de ICC deseados del código de ICC combinado, transmitido (nótese que el código de ICC combinado puede simplemente ser usado para ambos de los códigos de ICC derivados) y (2) aplica cada uno de los dos códigos de ICC derivados a una pareja diferente de canales sintetizados, para generar cuatro canales descorrelacionados (es decir, canales izquierdo / posterior izquierdo, derecho y posterior derecho, sintetizados).

30 La presente invención puede también ser llevada a cabo en otros contextos. Por ejemplo, un codificador de BCC puede estimar dos códigos de ICC de tres canales de entrada, A, B y C, donde un código de ICC estimado corresponde a los canales A y B y el otro código de ICC estimado corresponde a los canales A y C. En ese caso, el codificador puede ser el que estima dos códigos de ICC de dos parejas de canales de entrada, donde las dos parejas de canales de entrada comparten un canal común (es decir el canal de entrada A). El codificador puede luego generar y transmitir un código de ICC combinado, sencillo, basado en los dos códigos de ICC estimados. Un decodificador de BCC puede luego derivar dos códigos de ICC del código de ICC combinado, transmitido, y aplicar esos dos códigos de ICC derivados para sintetizar tres canales descorrelacionados (es decir, canales sintetizados A, B y C). En este caso, cada código de ICC derivado puede decirse que será aplicado para generar una pareja de canales descorrelacionados, donde las dos parejas de canales descorrelacionados comparten un canal común (es decir, el canal A sintetizado).

40 Aunque la presente invención se ha descrito en el contexto de los esquemas de codificación de BCC, que emplean códigos de ICC combinados, la presente invención puede también ser realizada en el contexto de esquemas de codificación de BCC que emplean códigos de pista de BCC combinados, que se generan combinando dos o más códigos de pista de BCC, además de los códigos de ICC, tal como códigos de ICTD y/o de ICLD, en lugar de, o además de, emplear los códigos de ICC combinados.

45 Aunque la presente invención se ha descrito en el contexto de esquemas de codificación de BCC, que implican los códigos ICTD, ICLD e ICC, la presente invención puede también ser realizada en el contexto de otros esquemas de codificación de BCC, que implican solamente uno o dos de estos tres tipos de códigos (es decir, ICLD e ICC, pero no ICTD) y/o uno o más tipos adicional de códigos.

50 En el esquema de BCC de 5 a 2, representado en la Figura 10, los dos canales transmitidos $y_1(n)$ e $y_2(n)$ son típicamente generados aplicando un esquema de mezcla descendente de una etapa, particular, para los cinco canales mostrados en las Figuras 6 y 7, donde el canal y_1 es generado como una suma ponderada de canales 1, 3 y 4, y el canal y_2 es generado como una suma ponderada de canales 2, 3 y 5, donde, por ejemplo, en cada suma ponderada, el factor de ponderación para el canal 3 es la mitad del factor de ponderación usado para cada uno de los otros dos canales. En este esquema de BCC de una etapa, los códigos de pista de BCC estimados corresponden a diferentes parejas de los cinco canales de entrada originales. Por ejemplo, un conjunto de códigos de ICC estimados se basa en los canales 1 y 4, y otro conjunto de códigos de ICC estimados se basa en los canales 2 y 5.

60 En una alternativa, relativa a un esquema de BCC de múltiples etapas, los canales son mezclados en forma descendente en secuencia, con los códigos de pista BCC que corresponden potencialmente a diferentes grupos de canales en cada etapa, en la secuencia de mezcla descendente. Por ejemplo, para los cinco canales en las Figuras 6 y 7, en un codificador de BCC, los canales originales izquierdo y posterior izquierdo pueden ser mezclados en forma descendente, para formar un canal izquierdo

primero mezclado en forma ascendente con un primer conjunto de códigos de pista de BCC generados, que corresponden a aquellos dos canales originales. Similarmente, los canales originales derecho y posterior derecho pueden ser mezclados en forma descendente para formar el canal derecho primero mezclado en forma descendente, con un segundo conjunto de códigos de pista de BCC generados, que corresponden a aquellos dos canales originales. En una segunda etapa de mezcla descendente, el canal izquierdo primero mezclado en forma descendente, puede ser mezclado hacia abajo con el canal central original para formar un canal izquierdo segundo mezclado en forma descendente, con un tercer conjunto de códigos de pista de BCC generados, que corresponde al canal izquierdo primero mezclado en forma descendente y el canal central original. Similarmente, el canal derecho primero mezclado en forma descendente, puede ser mezclado hacia abajo con el canal central original, para formar un canal derecho segundo mezclado en forma descendente, con un cuarto conjunto de códigos de pista de BCC, generados, que corresponden al canal derecho primero mezclado en forma descendente y el canal central original. Los canales izquierdo y derecho, segundos mezclados en forma descendente, pueden ser luego transmitidos con todos los cuatro conjuntos de códigos de pista de BCC como la información secundaria. En una manera análoga, un decodificador de BCC correspondiente puede luego aplicar subsiguientemente estos cuatro conjuntos de códigos de pista de BCC, en diferentes etapas de un esquema de mezcla ascendente en secuencia, de dos etapas, para sintetizar cinco canales de salida desde los dos canales "estéreo" transmitidos.

Aunque la presente invención se ha descrito en el contexto de los esquemas de codificación de BCC, en los cuales los códigos de pista de ICC combinados se transmiten con uno o más canales de audio (es decir, los E canales transmitidos) junto con otros códigos de BCC, en realizaciones alternativas, los códigos de pista de ICC combinados pueden ser transmitidos, o solos o con otros códigos de BCC, a un sitio (por ejemplo un decodificador o un dispositivo de almacenamiento) que ya tiene canales transmitidos y posiblemente otros códigos de BCC.

Aunque la presente invención se ha descrito en el contexto de los esquemas de codificación de BCC, la presente invención puede también ser realizada en el contexto de otros sistemas de procesamiento de audio, en los cuales las señales de audio son descorrelacionadas u otros procesamientos de audio que necesitan señales descorrelacionadas.

Aunque la presente invención se ha descrito en el contexto de formas de realización en las cuales el codificador recibe la señal de audio de entrada en el dominio de tiempo y genera señales de audio transmitidas en el dominio de tiempo y el decodificador recibe las señales de audio transmitidas en el dominio de tiempo y genera señales de audio de reproducción en el dominio de tiempo, la presente invención no se limita así. Por ejemplo, en otras formas de realización cualquiera de una o más de las señales de audio de la entrada, transmitidas y de reproducción pueden ser representadas en un dominio de frecuencia.

Los codificadores y/o decodificadores de BCC pueden ser usados en conjunto con o incorporados en una variedad de diferentes aplicaciones o sistemas, que incluyen los sistemas para la televisión o distribución de música electrónica, películas, radiodifusión, flujo y/o recepción. Estos incluyen sistemas para codificar/decodificar transmisiones por vía, por ejemplo, terrestre, de satélite, cable, internet, intranets o medios físicos (por ejemplo discos compactos, discos versátiles digitales, microcircuitos semiconductores, discos duros, tarjetas de memoria y similares). Los codificadores y/o decodificadores pueden también ser empleados en juegos y sistemas de juegos, que incluyen, por ejemplo, productos de software interactivos, intentados para actuar con un usuario para entretenimiento (acción, juegos de funciones, estrategia, aventuras, simulaciones, carreras, deportes, galerías, tarjetas y juegos de tablero) y/o educación, que puede ser publicada para múltiples máquinas, plataformas u otros medios. Además, los codificadores y/o decodificadores de BCC pueden ser incorporados en grabadoras/reproductoras de audio o sistemas de CD-ROM/DVD. Los codificadores y/o decodificadores pueden también ser incorporados en aplicaciones de software de computadoras personales (PC), que incorporan un decodificador digital (por ejemplo, reproductor, decodificador) y aplicaciones de software que incorporan capacidades de codificaciones digitales (por ejemplo, un codificador, grabadora y tocadiscos automático).

La presente invención puede también ser llevada a cabo como procesos a base de circuitos, que incluyen la posible realización como un solo circuito integrado, (tal como un ASIC o un FPGA) un módulo de múltiples microcircuitos, una tarjeta sencilla o un paquete de circuitos de múltiples tarjetas. Como será evidente a un experto en la técnica, varias funciones de elementos de circuito pueden también ser realizadas como etapas de proceso en un programa de software. Tal software puede ser empleado en, por ejemplo, un procesador de señales digitales, microcontrolador o un ordenador para propósitos generales.

La presente invención puede también ser llevada a cabo en la forma de métodos y aparatos para practicar esos métodos. La presente invención puede también ser incorporada en la forma de códigos de programas incorporados en medios tangibles, tal como en disquetes, CD-ROMs, discos duros o cualquier otro medio de almacenamiento que se pueda leer por una máquina, en que, cuando el código del programa se carga dentro y se ejecuta por una máquina, tal como por un ordenador, esta máquina llega a ser un aparato para practicar la invención. La presente invención puede también ser incorporada en la forma de un código de programa, por ejemplo, cuando se almacena en un medio de almacenamiento, cargado en, y/o ejecutado por, una máquina o transmitido sobre algún medio de transmisión o portador, por ejemplo sobre un alambrado o cableado eléctrico, a través de fibras ópticas o por medio de la radiación electromagnética, en que, cuando el código de programa se carga en y se

ejecuta por una máquina, tal como por un ordenador, esta máquina llega a ser un aparato para practicar la invención. Cuando se incorpora en un procesador para propósitos generales, los segmentos del código del programa se combinan con el procesador para proporcionar un dispositivo único que opera análogamente a los circuitos lógicos específicos.

5 Asimismo se entenderá que varios cambios en los detalles, materiales y arreglos de las partes, que se han descrito e ilustrado, con el fin de explicar la naturaleza de esta invención, se pueden hacer por los expertos en la técnica, sin apartarse del ámbito de la invención, como se expresa en las siguientes reivindicaciones.

10 Aunque las etapas en las siguientes reivindicaciones del método, si las hay, se exponen en una secuencia particular, con el rotulado correspondiente, a no ser que las exposiciones de las reivindicaciones expresen de otra manera una secuencia particular para llevar a cabo algunas o todas estas etapas. Las etapas no necesariamente se intentan para estar limitadas a las formas de realización expuestas en esa secuencia particular.

REIVINDICACIONES

- 1.- Un método para codificar canales de audio, el método comprende:
- 5 generar (1102, 1104) uno o más códigos de pista a ser transmitidos para dos o más canales de audio, en que:
- al menos un código de pista de los códigos de pista a ser transmitidos es un código de pista combinado generado mediante la combinación (1104) de dos o más códigos de pista estimados utilizando una función promedio ponderada; la función promedio ponderada considerando la importancia de los canales individuales; y
- 10 cada código de pista estimado se estima de un grupo diferente de dos o más de los canales de audio; y
- transmitir (1106) o almacenar los uno o más códigos de pista a ser transmitidos.
- 15 2.- El método de la reivindicación 1, que además comprende transmitir E canales de audio transmitidos, que corresponden a los dos o más canales de audio, donde $E \leq 1$.
- 3.- El método de la reivindicación 2, en que
- 20 los dos o más canales de audio comprenden C canales de audio de entrada, donde $C > E$, y los C canales de entrada se mezclan en forma descendente para generar los E canales transmitidos.
- 4.- El método de la reivindicación 1, en que los uno o más códigos de pista a ser transmitidos se transmiten para habilitar a un decodificador a ejecutar un procesamiento de síntesis durante la decodificación de los E canales transmitidos basado en los códigos de pista combinados, en que los E canales de audio transmitidos corresponden a los dos o más canales de audio, donde $E \geq 1$.
- 25 5.- El método de la reivindicación 1, en que los uno o más códigos de pista a ser transmitidos comprenden uno o más de los siguientes códigos: un código de correlación de inter-canales combinados (ICC), un código de una diferencia de nivel de inter-canales combinados (ICLD), y un código de una diferencia de tiempo inter-canales combinados (ICTD).
- 30 6.- El método de la reivindicación 1, en que:
- 35 cada código de pista estimado usado para generar el código de pista combinado se asocia con un factor de ponderación, usado en generar el promedio ponderado; y
- el factor de ponderación para cada código de pista estimado se basa en la potencia en el grupo de canales que corresponde al código de pista estimado.
- 40 7.- El método de la reivindicación 1, en que el código de pista combinado es un código de ICC combinado.
- 8.- El método de la reivindicación 7, en que
- 45 los dos o más canales de audio comprenden un canal izquierdo, un canal posterior izquierdo, un canal derecho y un canal posterior derecho;
- se genera un primer código de ICC estimado desde los canales izquierdo y posterior izquierdo;
- 50 se genera un segundo código de ICC estimado desde los canales derecho y posterior derecho; y
- se genera el código de ICC combinado, combinando el primero y segundo códigos de ICC estimados.
- 55 9.- Aparato (202) para codificar canales de audio, el aparato comprende:
- medios (208) para generar uno o más códigos de pista a ser transmitidos para dos o más canales de audio, en que
- 60 al menos un código de pista de los uno o más códigos de pista a ser transmitidos es un código de pista combinado generado mediante la combinación de dos o más códigos de pista estimados utilizando una función promedio ponderada; la función promedio ponderada considerando la importancia de los canales individuales; y

cada código de pista estimado se estima de un grupo diferente de dos o más de los canales de audio; y medios para transmitir o almacenar los uno o más códigos de pista a ser transmitidos.

5

10.- El aparato de la reivindicación 9,

en que el aparato es para codificar C canales de audio de entrada, para generar E canales de audio transmitidos:

10

donde los medios para generar incluyen un estimador de códigos,

donde el aparato comprende adicionalmente un mezclador descendente (206) adaptado para mezclar, en forma descendente, los C canales de entrada, para generar los E canales transmitidos, donde $C > E \geq 1$, y

15

en que el aparato está adaptado para transmitir información acerca de los uno o más códigos de pista a ser transmitidos, para habilitar a un decodificador a realizar un procesamiento de síntesis durante la decodificación de los E canales transmitidos.

11.- El aparato de la reivindicación 10, en que

20

el aparato es un sistema seleccionado del grupo que consiste de una grabadora digital de video, una grabadora digital de audio, un ordenador, un transmisor de satélite, un transmisor de cable, un transmisor de radiodifusión terrestre, un sistema de entretenimiento del hogar, y un sistema para reproducir películas; y

el sistema comprende el estimador de códigos y el mezclador de tipo descendente.

25

12.- Un medio que se puede leer por máquina, que tiene codificado un código de programa, en que, cuando el código de programa se ejecuta en una máquina, la máquina implementa un método para codificar canales de audio, el método comprende:

30

generar (1102, 1104) uno o más códigos de pista a ser transmitidos para dos o más canales de audio, en que:

al menos un código de pista de los uno o más códigos de pista a ser transmitidos es un código de pista combinado generado mediante la combinación (1104) de dos o más códigos de pista estimados utilizando una función promedio ponderada, la función promedio ponderada considerando la importancia de los canales individuales; y

35

cada código de pista estimado se estima de un grupo diferente de dos o más de los canales de audio; y

transmitir (1106) o almacenar los uno o más códigos de pista a ser transmitidos.

40

13.- Una secuencia de bits de audio codificado generada mediante la codificación de canales de audio, en que:

se generan uno o más códigos de pista a ser transmitidos para dos o más canales de audio, en que:

45

al menos un código de pista de los uno o más códigos de pista a ser transmitidos es un código de pista combinado generado mediante la combinación de dos o más códigos de pista estimados utilizando una función promedio ponderada, la función promedio ponderada considerando la importancia de los canales individuales; y

cada código de pista estimado se estima de un grupo diferente de dos o más de los canales de audio; y

50

en que los uno o más códigos de pista a ser transmitidos y E canales de audio transmitidos que corresponden a los dos o más canales de audio, donde $E \geq 1$, se codifican en la secuencia de bits de audio codificado.

14.- Un método para decodificar E canales de audio transmitidos, para generar C canales de audio de reproducción, donde $C > E \geq 1$, el método comprende:

55

recibir códigos de pista que corresponden a los E canales transmitidos, en que:

al menos un código de pista de los códigos de pista recibidos es un código de pista combinado generado mediante la combinación de dos o más códigos de pista estimados; y

60

cada código de pista estimado se estima de un grupo diferente de dos o más canales de audio de entrada de codificación;

5 realizar una mezcla ascendente de uno o más de los E canales transmitidos para generar uno o más canales con mezcla ascendente, y

synetizar uno o más de los C canales de reproducción mediante la aplicación de los códigos de pista recibidos al uno o más canales de mezcla ascendente, en que:

10 dos o más códigos de pista derivados son derivados del código de pista combinado (1108); y

cada código de pista derivado es aplicado (1110) para generar dos o más canales sintetizados.

15 15.- El método de la reivindicación 14, en que los códigos de pista comprenden uno o más de un código de ICC combinado, un código de ICLD combinado, y un código de ICTD combinado.

16.- El método de la reivindicación 14, en que el código de pista combinado es un promedio de dos o más códigos de pista estimados.

20 17.- El método de la reivindicación 16, en que el código de pista combinado es un promedio ponderado de los dos o más códigos de pista estimados.

18.- El método de la reivindicación 17, en que

25 cada código de pista estimado usado para generar el código de pista combinado se asocia con un factor de ponderación usado para generar el promedio ponderado; y

el factor de ponderación para cada código de pista estimado se basa en la potencia en el grupo de canales que corresponde al código de pista estimado.

30 19.- El método de la reivindicación 14, en que los dos o más códigos de pista derivados son derivados mediante:

la derivación de un factor de ponderación para cada grupo de dos o más canales asociados con un código de pista estimado, y

35 la derivación de los dos o más códigos de pista derivados como una función del código de pista combinado y dos o más factores de ponderación derivados.

20.- El método de la reivindicación 19, en que cada factor de ponderación derivado se deriva mediante:

40 la estimación de la potencia en el grupo de canales que corresponden a un código de pista estimado; y

la derivación del factor de ponderación basada en las potencias estimadas para los diferentes grupos de canales que corresponden a diferentes códigos de pista estimadas.

45 21.- El método de la reivindicación 14, en que el código de pista combinado es un código de ICC combinado.

22.- El método de la reivindicación 21, en que

50 los dos o más canales de audio comprenden un canal izquierdo, un canal posterior izquierdo, un canal derecho, y un canal posterior derecho;

un primer código de ICC estimado se genera a partir de los canales izquierdo y posterior izquierdo;

55 un segundo código de ICC estimado se genera a partir de los canales derecho y posterior derecho; y

el código de ICC combinado se genera combinando el primero y segundo códigos de ICC estimados.

23.- El método de la reivindicación 22, en que

60 el código de ICC combinado se usa para descorrelacionar los canales izquierdo y posterior izquierdo sintetizados; y

el código de ICC combinado se usa para descorrelacionar los canales derecho y posterior derecho sintetizados.

24.- Aparato (204) para decodificar E canales de audio transmitidos para generar C canales de audio de reproducción, donde $C > E \geq 1$, el aparato comprende:

medios para recibir códigos de pista que corresponden a los E canales transmitidos, en que:

al menos un código de pista de los códigos de pista recibidos es un código de pista combinado generado mediante la combinación de dos o más códigos de pista estimados; y

cada código de pista estimado se estima de un grupo diferente de dos o más canales de audio de entrada de codificación;

medios para la mezcla ascendente de uno o más de los E canales transmitidos para generar uno o más canales de mezcla ascendente; y

medios para sintetizar uno o más de los C canales de reproducción mediante la aplicación de los códigos de pista recibidos a los uno o más canales de mezcla ascendente, en que:

los dos o más códigos de pista derivados se derivan del código de pista combinado; y

cada código de pista derivado se aplica para generar dos o más canales sintetizados.

25.- El aparato de la reivindicación 24, en que:

el aparato es un sistema seleccionado del grupo que consiste de un reproductor digital de video, un reproductor digital de audio, un ordenador, un receptor de satélite, un receptor de cable, un receptor de radiodifusión terrestre, un sistema de entretenimiento del hogar, y un sistema para reproducir películas; y

el sistema comprende el receptor, el mezclador de tipo ascendente, y el sintetizador.

26.- Un medio que se puede leer por una máquina, que tiene codificado un código de programa, en que, cuando este código de programa se ejecuta por una máquina, la máquina implementa un método para decodificar E canales de audio transmitidos para generar C canales de audio de reproducción, donde $C > E \geq 1$, el método comprende:

recibir unos códigos de pista que corresponden a los E canales transmitidos, en que:

al menos un código de pista de los códigos de pista recibidos es un código de pista combinado generado mediante la combinación de dos o más códigos de pista estimados; y

cada código de pista estimado se estima de un grupo diferente de dos o más canales de audio de entrada de codificación;

realizar una mezcla ascendente de uno o más de los E canales transmitidos para generar uno o más canales de mezcla ascendente; y

sintetizar uno o más de los C canales de reproducción mediante la aplicación de los códigos de pista recibidos a los uno o más de canales de mezcla ascendente; en que

dos o más códigos de pista derivados son derivados del código de pista combinado; y

cada código de pista derivado se aplica para generar dos o más canales sintetizados.

FIGURA 1
(TÉCNICA ANTERIOR)

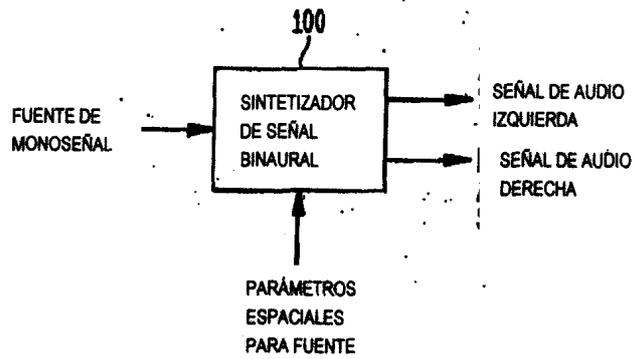


FIGURA 2

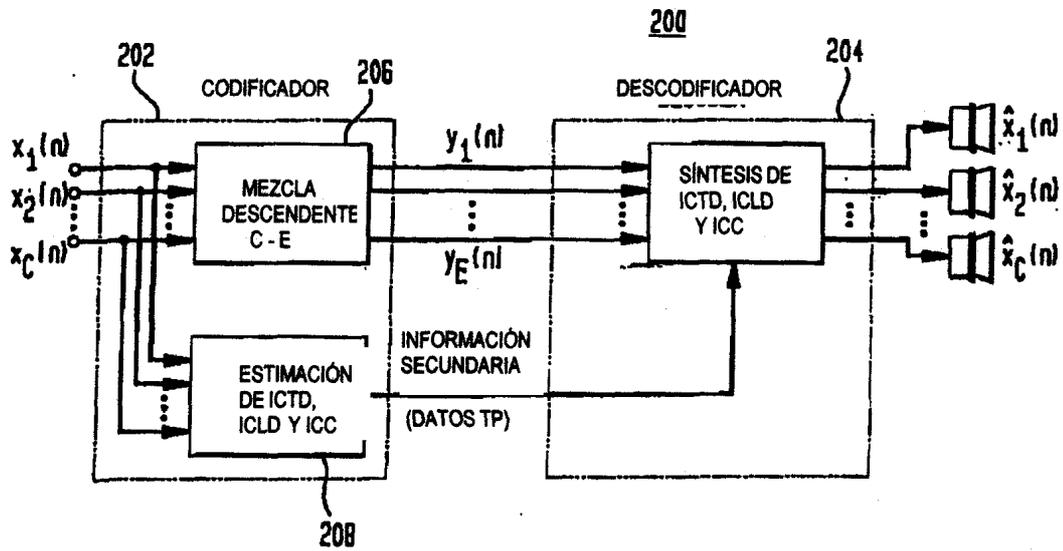


FIGURA 3

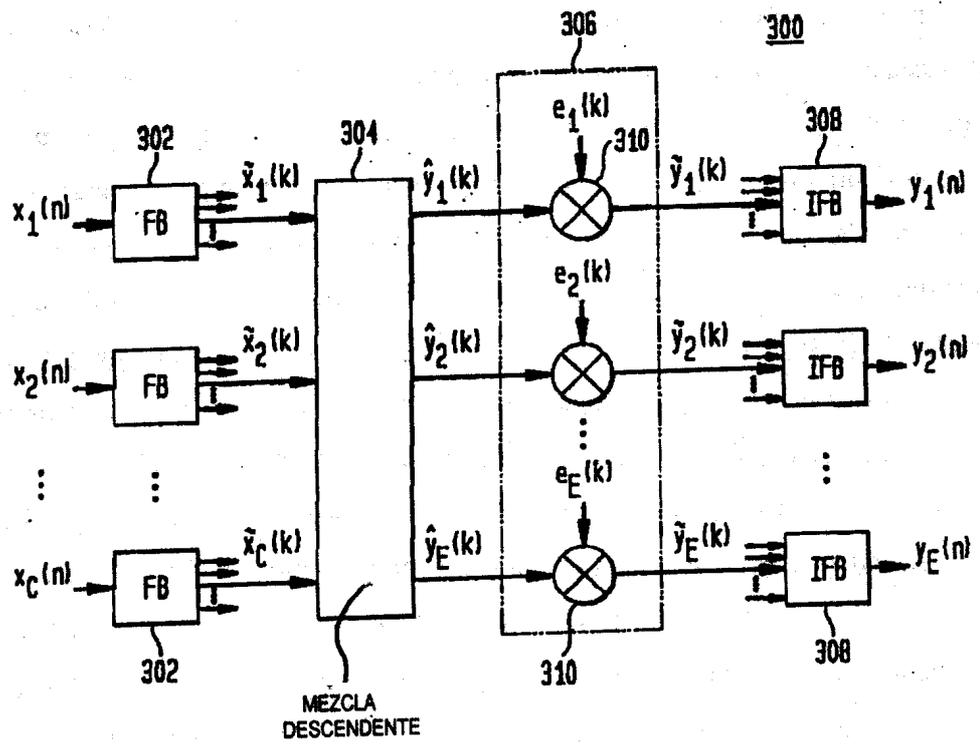


FIGURA 4

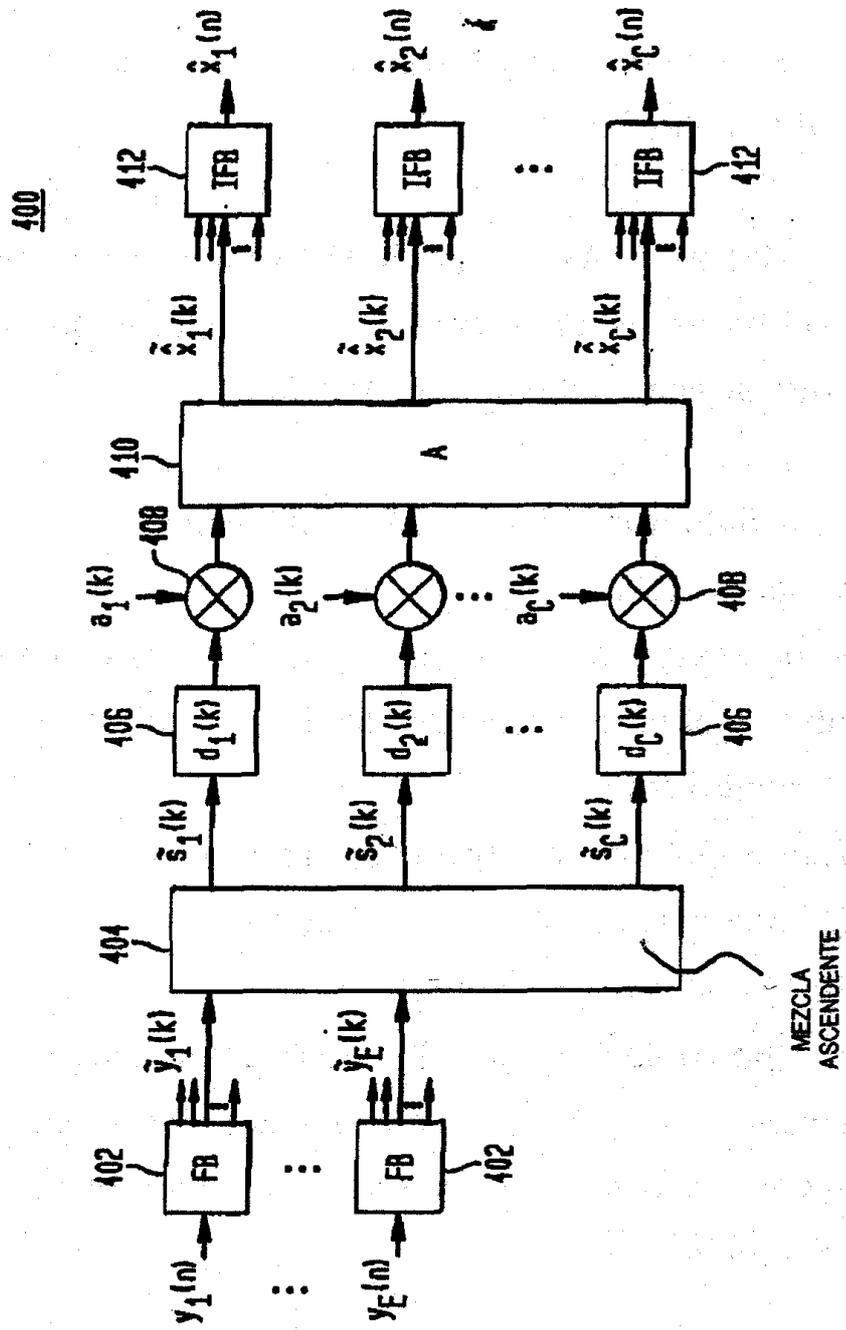


FIGURA 5

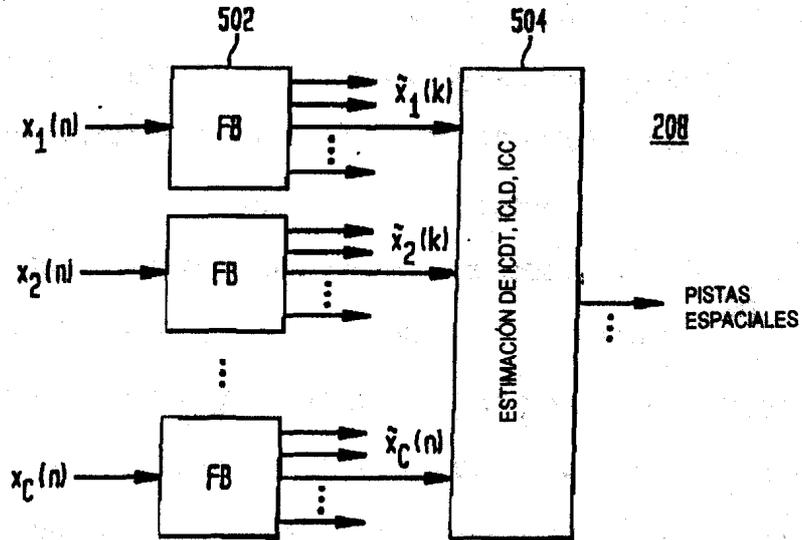


FIGURA 6

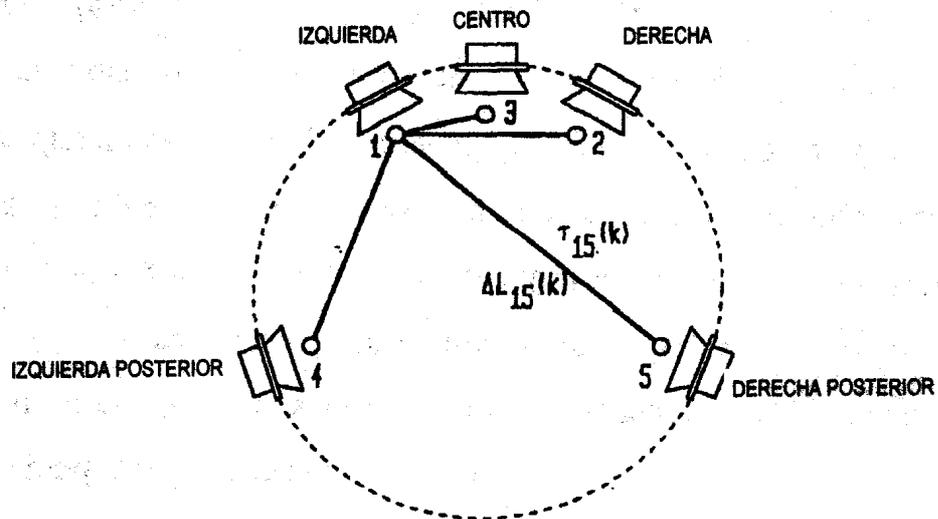


FIGURA 7A

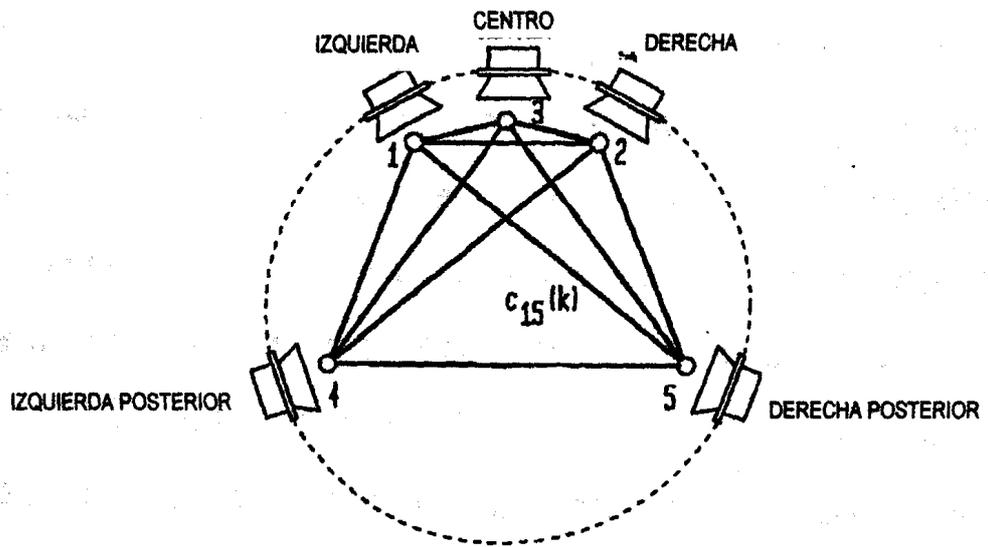


FIGURA 7B

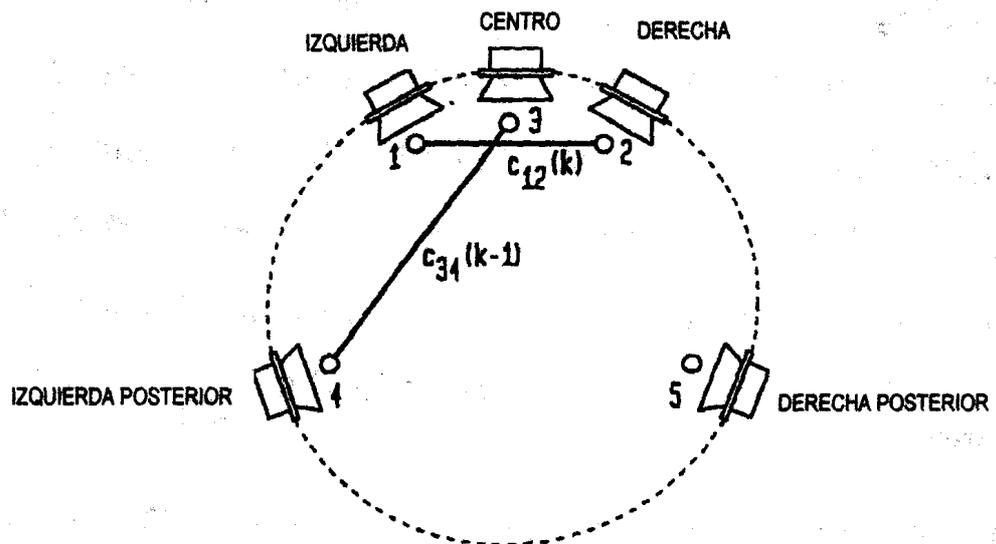


FIGURA 8

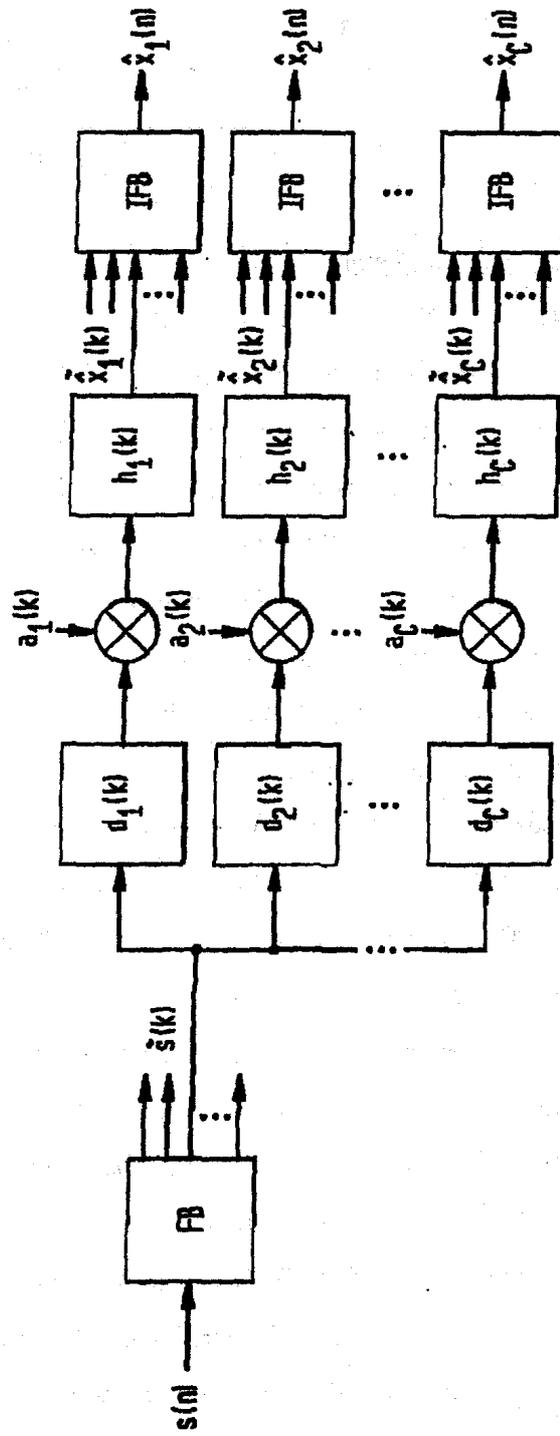


FIGURA 9

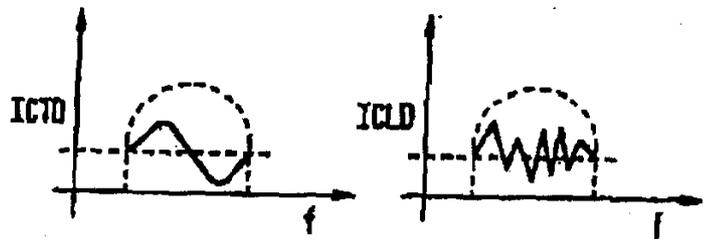


FIGURA 10

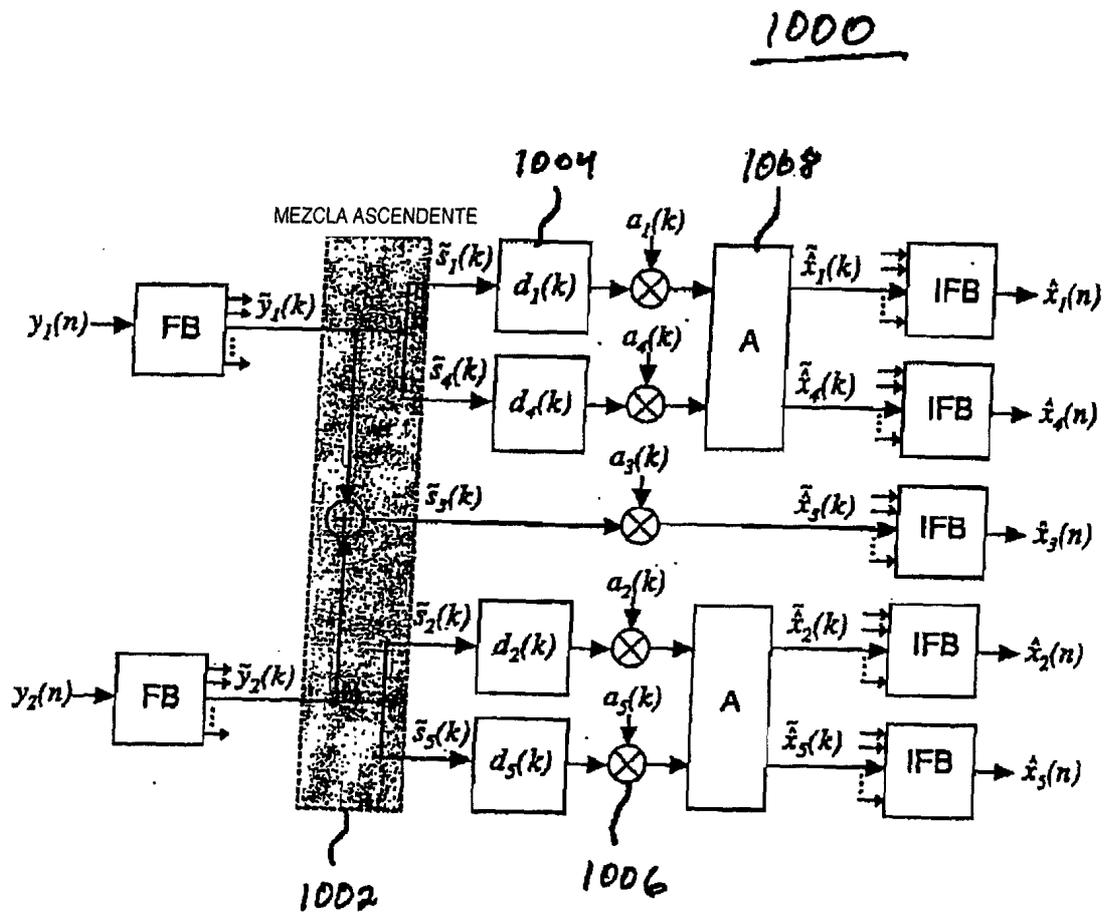


FIGURA 11

