

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 421 713**

51 Int. Cl.:

G10L 19/16 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **30.01.2004 E 09012227 (6)**

97 Fecha y número de publicación de la concesión europea: **22.05.2013 EP 2136361**

54 Título: **Transcodificación de audio de baja complejidad**

30 Prioridad:

06.02.2003 US 445931 P
09.06.2003 US 458798

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
05.09.2013

73 Titular/es:

DOLBY LABORATORIES LICENSING CORP.
(100.0%)
100 PORTRERO AVENUE
SAN FRANCISCO, CA 94103, US

72 Inventor/es:

LENNON, BRIAN TIMOTHY;
TRUMAN, MICHAEL MEAD y
ANDERSEN, ROBERT LORING

74 Agente/Representante:

PÉREZ BARQUÍN, Eliana

ES 2 421 713 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

Transcodificación de audio de baja complejidad

5 **Campo técnico**

La presente invención concierne, en general, a métodos y dispositivos de codificación de audio y, más específicamente, concierne a métodos y dispositivos mejorados para codificar y transcodificar información de audio.

10 **Técnica anterior**A. CODIFICACIÓN

15 Muchos sistemas de comunicaciones se enfrentan al problema de que la demanda de transmisión de información y de la capacidad de grabación excede a menudo de la capacidad disponible. Como resultado, existe un interés considerable entre los que se encuentran en los campos de retransmisión y grabación para reducir la cantidad de información requerida para transmitir o grabar una señal de audio destinada a la percepción humana sin degradar la calidad percibida.

20 Existe también un interés en mejorar la calidad percibida de la señal de salida para una anchura de banda o capacidad de almacenamiento dadas.

Los métodos tradicionales para reducir los requisitos de capacidad de información implican la transmisión o la grabación solamente de partes seleccionadas de la señal de entrada. Las partes restantes se descartan. Las técnicas conocidas como la codificación perceptiva convierten típicamente una señal de audio original en componentes espectrales o señales de sub-banda de frecuencias, de manera que aquellas partes de la señal que son redundantes o irrelevantes pueden ser identificadas y descartadas con mayor facilidad. Se estima que una parte de una señal es redundante si puede volverse a crear a partir de otras partes de la señal. Se estima que una parte de una señal es irrelevante si es perceptivamente insignificante o inaudible. Un descodificador perceptivo puede volver a crear las partes redundantes omitidas de una señal codificada, pero no puede crear ninguna información irrelevante omitida que no sea también redundante.

35 Sin embargo, la pérdida de información irrelevante es aceptable en muchas aplicaciones, porque su ausencia no tiene un efecto perceptible sobre la señal descodificada.

Una técnica de codificación de señales es perceptivamente transparente si descarta solamente aquellas partes de una señal que son redundantes o bien perceptivamente irrelevantes. Una manera en la cual las partes irrelevantes de una señal pueden ser descartadas, es representar componentes espectrales con menores niveles de precisión, lo que se denomina a menudo cuantificación. La diferencia entre un componente espectral original y su representación cuantificada es conocida como ruido de cuantificación. Las representaciones con una precisión menor tienen un nivel más alto de ruido de cuantificación. Las técnicas de codificación perceptiva intentan controlar el nivel del ruido de cuantificación, de manera que sea inaudible.

45 Si una técnica perceptivamente transparente no puede conseguir una reducción suficiente de los requisitos de capacidad de información, se necesita entonces una técnica perceptivamente no transparente para descartar las partes adicionales de la señal que no son redundantes y son perceptivamente relevantes. El resultado inevitable es que la fidelidad percibida de la señal transmitida o registrada se degrada. Preferiblemente, una técnica perceptivamente no transparente descarta solamente aquellas partes de la señal que se estima que tienen el menor significado perceptivo.

50 Se puede utilizar una técnica de codificación denominada "acoplamiento", que se considera a menudo como perceptivamente no transparente, para reducir los requisitos de capacidad de información. De acuerdo con esta técnica, los componentes espectrales en dos o más señales de audio de entrada son combinados para formar una señal de canal acoplado con una representación compuesta de estos componentes espectrales. Se genera también una información lateral, que representa una envolvente espectral de los componentes espectrales en cada una de las señales de audio de entrada, que son combinadas para formar la representación compuesta. Una señal codificada que incluye la señal de canal acoplado y la información lateral, es transmitida o registrada para su descodificación subsiguiente por un receptor. El receptor genera señales desacopladas, que son réplicas inexactas de las señales de entrada originales, generando copias de la señal de canal acoplado y utilizando la información lateral para efectuar una estimación de los componentes espectrales en las señales copiadas, de manera que las envolventes espectrales de las señales de entrada originales son sustancialmente restauradas. Una técnica típica de acoplamiento para un sistema estéreo de dos canales, combina componentes de alta frecuencia de las señales de los canales izquierdo y derecho, para formar una sola señal de componentes compuestos de alta frecuencia, y genera información lateral que representa las envolventes espectrales de los componentes de alta frecuencia en las señales originales de los canales izquierdo y derecho. Un ejemplo de una técnica de acoplamiento está descrito en el documento "Compresión Digital de Audio (AC-3)", que es el documento estándar A/52 (1994) del Comité de

Sistemas Avanzados de Televisión (ATSC), que en esta memoria se denomina documento A/52.

Una técnica de codificación conocida como regeneración espectral, es una técnica perceptivamente no transparente que puede ser utilizada para reducir los requisitos de capacidad de información. En muchas implementaciones, esta técnica se denomina “regeneración de alta frecuencia” (HFR), porque solamente se regeneran los componentes espectrales de alta frecuencia. De acuerdo con esta técnica, se transmite o almacena una señal de banda base que contenga solamente componentes de baja frecuencia de una señal de audio de entrada. Se proporciona también información lateral, que representa una envolvente espectral de los componentes originales de alta frecuencia. Se transmite o se registra una señal codificada que incluya la señal de banda base y la información lateral, para su posterior descodificación por un receptor. El receptor regenera los componentes de alta frecuencia omitidos con niveles espectrales basados en la información lateral, y combina la señal de banda base con los componentes de alta frecuencia regenerados, para producir una señal de salida. Se puede encontrar una descripción de métodos conocidos para la HFR en el artículo de Makhoul y Berouti “High-Frequency Regeneration in Speech Coding Systems” (“Regeneración de Alta Frecuencia en Sistemas de Codificación del Habla”), Proc. of the International Conf. on Acoust., Speech and Signal Proc., abril de 1979. Se divulgan técnicas mejoradas de regeneración espectral que son adecuadas para la codificación de música de alta calidad en los documentos US-A1-2003/0187663, US-A1-2003/0233234, US-A1-2003/0233236 y US-A1- 2004/0225505.

B. TRANSCODIFICACIÓN

Las técnicas de codificación conocidas han reducido los requisitos de capacidad de información de las señales de audio para un nivel dado de calidad percibida o, a la inversa, han mejorado la calidad percibida de las señales de audio que tienen una capacidad de información especificada. A pesar de este éxito, existe una demanda para un mayor avance y continúa la investigación sobre la codificación para descubrir nuevas técnicas de codificación y para descubrir nuevas formas de utilizar las técnicas conocidas.

Una consecuencia de los avances adicionales es una potencial incompatibilidad entre señales que son codificadas por técnicas de codificación más modernas y por equipos existentes que implementan técnicas de codificación más antiguas. Aunque se ha hecho un gran esfuerzo por organizaciones de estándares y por fabricantes de equipos para impedir la obsolescencia prematura, los receptores más antiguos no pueden descodificar siempre correctamente las señales que están descodificadas por técnicas de codificación más modernas. A la inversa, los receptores más modernos no siempre pueden descodificar correctamente las señales que están codificadas por técnicas de codificación más antiguas. Como resultado, tanto los profesionales como los consumidores adquieren y mantienen muchos equipos si desean asegurar la compatibilidad con las señales codificadas por técnicas de codificación antiguas y modernas.

Una manera con la que puede aliviarse o evitarse esta carga es adquirir un transcodificador que pueda convertir señales codificadas de un formato a otro. Un transcodificador puede servir como puente entre diferentes técnicas de codificación. Por ejemplo, un transcodificador puede convertir una señal que está codificada con una técnica de codificación moderna en otra señal que sea compatible con los receptores que pueden descodificar solamente aquellas señales que están codificadas por una técnica más antigua.

La transcodificación convencional implementa procesos completos de descodificación y codificación. Haciendo referencia al ejemplo de transcodificación mencionado anteriormente, una señal de entrada codificada se descodifica utilizando una técnica de descodificación más moderna para obtener componentes espectrales que son convertidos después en una señal digital de audio mediante el filtrado por síntesis. La señal digital de audio es convertida después en componentes espectrales nuevamente mediante el filtrado por análisis, y estos componentes espectrales son codificados después utilizando una técnica de codificación más antigua. El resultado es una señal codificada que es compatible con equipos de recepción más antiguos. La transcodificación puede ser utilizada también para convertir formatos antiguos en modernos, para convertir entre formatos contemporáneos diferentes y para convertir entre velocidades binarias diferentes del mismo formato.

Las técnicas convencionales de transcodificación tienen serias desventajas cuando se utilizan para convertir señales que están codificadas por sistemas de codificación perceptivos. Una desventaja es que los equipos de transcodificación convencional son relativamente costosos, porque deben implementar procesos completos de descodificación y codificación.

Una segunda desventaja es que la calidad percibida de la señal transcodificada tras la descodificación, es casi siempre degradada con respecto a la calidad percibida de la señal de entrada codificada tras la descodificación.

El documento de Mat Hans y otros colaboradores, titulado “An MPEG Audio Layered Transcoder” (“Un transcodificador estratificado de audio MPEG”), borradores de artículos presentados en la Convención AES, de septiembre de 1998 (1998-09) páginas 1-18, XP001014304, divulga un método de transcodificación de una primera señal codificada en una segunda señal codificada, donde la primera señal codificada transporta primeros valores estimados cuantificados y primeros factores de estimación que representan componentes espectrales de una señal de audio, donde cada primer factor de estimación está asociado con uno o más primeros valores estimados

cuantificados, estando estimado cada primer valor cuantificado de acuerdo con su primer factor de estimación asociado, y cada primer valor estimado cuantificado y primer factor de estimación asociado representan un respectivo componente espectral. Los bits están asignados de acuerdo con un primer proceso de asignación de bits, como respuesta a uno o más primeros parámetros de control, y se obtienen valores estimados descuantificados a partir de los primeros valores estimados cuantificados, descuantificando de acuerdo con resoluciones de cuantificación basadas en números de bits asignados por el primer proceso de asignación de bits. Los bits se asignan de acuerdo con un segundo proceso de asignación de bits, como respuesta a uno o más segundos parámetros de control y se obtienen segundos valores estimados cuantificados cuantificando los valores estimados descuantificados, utilizando resoluciones de cuantificación basadas en números de bits asignados por el segundo proceso de asignación de bits, donde cada segundo factor de estimación está asociado con uno o más de los segundos valores estimados cuantificados, siendo estimado el segundo valor cuantificado de acuerdo con su segundo factor de estimación asociado, representando cada uno de los segundos valores estimados cuantificados y su factor de estimación asociado, un respectivo componente espectral. Los segundos valores estimados cuantificados y los segundos factores de estimación son ensamblados en la segunda señal codificada.

En un caso, esta técnica anterior produce una cadena de bits de salida que tiene los mismos factores de estimación que la cadena de entrada original. En otro caso, esta técnica anterior produce una cadena de bits de salida que codifica componentes de error o de diferencia y elige factores de estimación basándose en estos componentes de diferencia.

Divulgación de la invención

Es un objeto de la presente invención proporcionar técnicas de codificación que puedan ser utilizadas para mejorar la calidad de las señales transcodificadas y permitir implementar los equipos de transcodificación de una manera menos costosa.

Este objeto se consigue por medio de la presente invención, como se establece en las reivindicaciones. Una técnica de transcodificación descodifica una señal de entrada codificada para obtener componentes espectrales, y después codifica los componentes espectrales en una señal de salida codificada. Se evitan los costes de implementación y la degradación de la señal en que incurrir los filtrados por síntesis y por análisis. Los costes de implementación del transcodificador pueden reducirse aún más proporcionando parámetros de control en la señal codificada, en lugar de hacer que el transcodificador determine estos parámetros de control por sí mismo.

Las diversas características de la presente invención y sus modos de realización preferidos pueden comprenderse mejor haciendo referencia a la discusión siguiente y a los dibujos que se acompañan, en los cuales las referencias numéricas similares hacen referencia a elementos similares en las diversas figuras. El contenido de la siguiente discusión y de los dibujos se establece solamente como ejemplos y no debe entenderse que representan limitaciones en el alcance de la presente invención.

Breve descripción de los dibujos

La figura 1 es un diagrama esquemático de un transmisor de codificación de audio.

La figura 2 es un diagrama esquemático de un receptor de descodificación de audio.

La figura 3 es un diagrama esquemático de un transcodificador.

Las figuras 4 y 5 son diagramas esquemáticos de transmisores de codificación de audio.

La figura 6 es un diagrama esquemático de bloques de un aparato que puede implementar diversos aspectos de la presente invención.

Modos de llevar a cabo la invención

A. DESCRIPCIÓN GLOBAL

Un sistema básico de codificación de audio incluye un transmisor de codificación, un receptor de descodificación y un camino de comunicación o medio de grabación. El transmisor recibe una señal de entrada que representa uno o más canales de audio y genera una señal codificada que representa el audio. El transmisor transmite entonces la señal codificada al camino de comunicación para su transporte, o bien al medio de grabación para su almacenamiento. El receptor recibe la señal codificada desde el camino de comunicaciones o medio de grabación y genera una señal de salida que puede ser una réplica exacta o aproximada del audio original. Si la señal de salida no es una réplica exacta, muchos sistemas de codificación intentan proporcionar una réplica que es perceptivamente indistinguible del audio de entrada original.

Un requisito inherente y obvio para el funcionamiento correcto de cualquier sistema de codificación, es que el

receptor debe ser capaz de descodificar correctamente la señal codificada. Sin embargo, debido a los avances en técnicas de codificación, surgen situaciones en las que es deseable utilizar un receptor para descodificar una señal que ha sido codificada mediante técnicas de codificación que el receptor no puede descodificar correctamente. Por ejemplo, una señal codificada puede haber sido generada por una técnica de codificación que espera que el descodificador realice la regeneración espectral, pero donde el receptor no puede realizar la regeneración espectral. A la inversa, una señal codificada puede haber sido generada por una técnica de codificación que no espera que el descodificador realice la regeneración espectral, pero donde el receptor espera y requiere una señal codificada que necesita la regeneración espectral.

La presente invención está dirigida hacia la transcodificación que pueda proporcionar un puente entre técnicas de codificación y equipos de codificación incompatibles.

A continuación se describen unas pocas técnicas de codificación, como introducción a una descripción detallada de algunas maneras en las cuales puede implementarse la presente invención.

1. Sistema básico

a) TRANSMISOR DE CODIFICACIÓN

La figura 1 es una ilustración esquemática de una implementación de un transmisor de codificación de audio de banda repartida, que recibe desde el camino 11 una señal de audio de entrada. El banco de filtros 12 de análisis reparte la señal de audio de entrada en componentes espectrales que representan el contenido espectral de la señal de audio.

El codificador 13 realiza un proceso que codifica al menos algunos de los componentes espectrales en información espectral codificada. Los componentes espectrales que no están codificados por el codificador 13 son cuantificados por el cuantificador 15, utilizando una resolución de cuantificación que está adaptada como respuesta a los parámetros de control recibidos desde el controlador 14 de cuantificación. Opcionalmente, alguna o toda la información espectral codificada puede ser cuantificada también. El controlador 14 de cuantificación deduce los parámetros de control a partir de las características detectadas en la señal de audio de entrada. En la implementación ilustrada, las características detectadas se obtienen a partir de la información proporcionada por el codificador 13. El controlador 14 de cuantificación puede deducir también los parámetros de control como respuesta a otras características de la señal de audio, incluyendo las características temporales. Estas características pueden ser obtenidas a partir del análisis de la señal de audio antes, durante o después del proceso realizado por el banco de filtros 12 de análisis. Los datos que representan la información espectral cuantificada, la información espectral codificada y los datos que representan los parámetros de control, son ensamblados por el formateador 16 para formar una señal codificada, que se hace pasar a lo largo del camino 17 para su transmisión o almacenamiento. El formateador 16 puede ensamblar también otros datos en la señal codificada como palabras de sincronización, paridad o códigos de detección de errores, claves de recuperación de base de datos y señales auxiliares, que no son pertinentes para la comprensión de la presente invención, y no son estudiadas con más detalle.

La señal codificada puede ser transmitida por caminos de comunicación en banda base o modulada, en todo el espectro, incluyendo desde frecuencias supersónicas a ultravioletas, o puede ser grabada en un medio que utilice esencialmente cualquier tecnología de grabación, incluyendo la cinta magnética, tarjetas o disco, tarjetas o discos ópticos, y marcas detectables en medios como el papel.

(1) Banco de filtros de análisis

El banco de filtros 12 de análisis y el banco de filtros 25 de síntesis, que se estudian a continuación, pueden ser implementados esencialmente de cualquier forma que se desee, incluyendo una amplia gama de tecnologías digitales de filtrado, transformadas de bloques y transformadas de pequeñas ondulaciones. En un sistema de codificación de audio, el banco de filtros 12 de análisis se implementa mediante una Transformada Modificada Discreta del Coseno (MDCT) y el banco de filtros 25 de síntesis se implementa mediante una Transformada Inversa Modificada Discreta del Coseno (IMDCT) que están descritas en el artículo de Princen y otros colaboradores "Subband/Transform Coding Using Filter Bank Designs Based on Time Domain Aliasing Cancellation" (Codificación por Transformada de Subbanda utilizando diseños de bancos de filtros basados en la cancelación del repliegue en el dominio del tiempo), Proc. of the International Conf. on Acoust, Speech and Signal Proc., mayo de 1987, páginas 2161-64. En principio, no es importante ninguna implementación particular de un banco de filtros.

Los bancos de filtros de análisis que están implementados por medio de transformadas en bloques, reparten un bloque o intervalo de una señal de entrada, en un conjunto de coeficientes de transformada que representan el contenido espectral de ese intervalo de la señal. Un grupo de uno o más coeficientes contiguos de la transformada representa el contenido espectral dentro de una sub-banda particular de frecuencias que tiene una anchura de banda equiparable al número de coeficientes del grupo.

Los bancos de filtros de análisis que se implementan mediante algún tipo de filtro digital, tal como un filtro de

múltiples fases, en lugar de una transformada por bloques, reparten una señal de entrada en un conjunto de señales de sub-banda. Cada señal de sub-banda es una representación basada en tiempos del contenido espectral de la señal de entrada, dentro de una sub-banda particular de frecuencias. Preferiblemente, la señal de sub-banda está diezmada, de manera que cada una de las señales de sub-banda tiene una anchura de banda que es equiparable al número de muestras de la señal de sub-banda para un intervalo unitario de tiempo.

La discusión siguiente hace referencia de manera más particular a implementaciones que utilizan transformadas por bloques como la transformada de Cancelación del Repliegue en el Dominio del Tiempo (TDAC) mencionada anteriormente. En esta discusión, el término “componentes espectrales” se refiere a coeficientes de la transformada, y los términos “sub-banda de frecuencias” y “señal de sub-banda” conciernen a grupos de uno o más coeficientes contiguos de la transformada. Sin embargo, los principios de la presente invención pueden ser aplicados a otros tipos de implementaciones, de manera que los términos “sub-banda de frecuencias” y “señal de sub-banda” conciernen también a una señal que representa el contenido espectral de una parte o de toda la anchura de banda de una señal, y el término “componentes espectrales” puede entenderse en general que se refiere a muestras o elementos de la señal de sub-banda. Los sistemas de codificación perceptiva implementan normalmente el banco de filtros de análisis para proporcionar sub-bandas de frecuencia que tienen anchuras de banda equiparables a las denominadas anchuras de banda críticas del sistema auditivo humano.

(2) Codificación

El codificador 13 puede realizar esencialmente cualquier tipo de proceso de codificación que se desee. En una implementación, el proceso de codificación convierte los componentes espectrales en una representación estimada que comprende valores estimados y factores de estimación asociados, que se estudian a continuación. En otras implementaciones, pueden utilizarse también los procesos de codificación como la formación de matrices o la generación de información lateral para la regeneración espectral o el acoplamiento. Algunas de estas técnicas se estudian con más detalle a continuación.

El transmisor 10 puede incluir otros procesos de codificación, que no están sugeridos en la figura 1. Por ejemplo, los componentes espectrales cuantificados pueden estar sometidos a un proceso de codificación de entropía, tal como la codificación aritmética o la codificación de Huffman. No es necesaria una descripción detallada de los procesos de codificación como estos para comprender la presente invención.

(3) Cuantificación

La resolución de la cuantificación proporcionada por el cuantificador 15 se adapta como respuesta a los parámetros de control recibidos desde el controlador 14 de cuantificación. Estos parámetros de control pueden ser deducidos de cualquier manera que se desee; sin embargo, en un codificador perceptivo, se utiliza algún tipo de modelo perceptivo para estimar cuánto ruido de cuantificación se puede enmascarar por la señal de audio a codificar. En muchas aplicaciones, el controlador de cuantificación responde también a restricciones impuestas en la capacidad de información de la señal codificada. La restricción se expresa algunas veces en términos de una velocidad binaria máxima permisible para la señal codificada o para una parte especificada de la señal codificada.

En implementaciones preferidas de los sistemas de codificación perceptiva, los parámetros de control son utilizados por un proceso de asignación de bits, para determinar el número de bits a asignar a cada componente espectral, y para determinar las resoluciones de cuantificación que el cuantificador 15 utiliza para cuantificar cada componente espectral, de manera que la audición del ruido de cuantificación se hace mínima, sujeta a las restricciones de la capacidad de información o de la velocidad binaria. No hay una implementación particular del controlador 14 de cuantificación que sea crítica para la presente invención.

Un ejemplo de controlador de cuantificación se divulga en el Documento A/52, que describe un sistema de codificación denominado algunas veces como Dolby AC-3. En esta implementación, los componentes espectrales de una señal de audio están representados por una representación estimada, en la cual los factores de estimación proporcionan una estimación de la forma espectral de la señal de audio. Un modelo perceptivo utiliza los factores de estimación para calcular una curva de enmascaramiento que estima los efectos de enmascaramiento de la señal de audio. El controlador de cuantificación determina entonces un umbral de ruido permisible, el cual controla cómo son cuantificados los componentes espectrales, de manera que el ruido de cuantificación se distribuye de alguna manera óptima para cumplir con un límite o velocidad binaria impuestos en la capacidad de información. El umbral de ruido permisible es una réplica de la curva de enmascaramiento y está desplazada de la curva de enmascaramiento en una cantidad determinada por el controlador de cuantificación. En esta implementación, los parámetros de control son los valores que definen el umbral de ruido permisible. Estos parámetros pueden ser expresados de diversas maneras, tales como una expresión directa del propio umbral, o como valores tales como los factores de estimación y de un desplazamiento del cual puede deducirse el umbral de ruido permitido.

b) RECEPTOR DE DESCODIFICACIÓN

La figura 2 es una ilustración esquemática de una implementación de un receptor 20 de decodificación de audio de

banda repartida, que recibe desde el camino 21 una señal codificada que representa una señal de audio. El desformateador 22 obtiene información espectral cuantificada, información espectral codificada y parámetros de control a partir de la señal codificada. La información espectral cuantificada es descuantificada por el descuantificador 23, utilizando una resolución que se adapta como respuesta a los parámetros de control.

- 5 Opcionalmente, alguna o toda la información espectral codificada puede ser descuantificada también. La información espectral codificada es descodificada por el descodificador 24 y combinada con los componentes espectrales descuantificados, que son convertidos en una señal de audio por el banco de filtros 25 de síntesis y pasados a lo largo del camino 26.
- 10 Los procesos realizados en el receptor son complementarios a los correspondientes procesos realizados en el transmisor. El desformateador 22 desensambla lo que fue ensamblado por el formateador 16. El descodificador 24 realiza un proceso de descodificación que es una inversa exacta o bien una cuasi-inversa del proceso de codificación realizado por el codificador 13, y el descuantificador 23 realiza un proceso que es un cuasi-inverso del proceso realizado por el cuantificador 15. El banco de filtros 25 de síntesis lleva a cabo un proceso de filtrado que es
- 15 inverso al llevado a cabo por el banco de filtros 12 de análisis. Los procesos de descodificación y descuantificación se dice que son procesos cuasiinversos porque no pueden proporcionar una inversión perfecta de los procesos complementarios en el transmisor.

20 En algunas implementaciones, el ruido sintetizado o pseudo-aleatorio puede ser insertado en algunos de los bits menos significativos de componentes espectrales descuantificados, o utilizados como un sustituto para uno o más componentes espectrales. El receptor puede realizar también procesos adicionales de descodificación para tener en cuenta cualquier otra codificación que pueda haber sido realizada en el transmisor.

c) TRANSCODIFICADOR

25 La figura 3 es una ilustración esquemática de una implementación de un transcodificador 30, que recibe desde el camino 31 una señal codificada que representa una señal de audio. El desformateador 32 obtiene información espectral cuantificada, información espectral codificada, uno o más primeros parámetros de control y uno o más segundos parámetros de control a partir de la señal codificada. La información espectral cuantificada es

30 descuantificada por el descuantificador 33 utilizando una resolución que se adapta como respuesta al uno o más parámetros de control recibidos desde la señal codificada. Opcionalmente, alguna o toda la información espectral codificada puede ser también descuantificada. Si fuera necesario, toda o parte de la información espectral puede ser descodificada por el descodificador 34 para su transcodificación.

35 El codificador 35 es un componente opcional que puede no ser necesario para una aplicación particular de transcodificación. Si fuera necesario, el codificador 35 realiza un proceso que codifica al menos parte de la información espectral descuantificada, o información espectral codificada y/o descodificada, como una información espectral recodificada. Los componentes espectrales que no son codificados por el codificador 35, son re-

40 cuantificados por el cuantificador 36, utilizando una resolución de cuantificación que se adapta como respuesta al uno o más segundos parámetros de control recibidos desde la señal codificada. Opcionalmente, alguna o toda la información espectral re-codificada puede ser cuantificada también. Los datos que representan la información espectral re-cuantificada, la información espectral re-codificada y los datos que representan los uno o más segundos parámetros de control son ensamblados por el formateador 37 como una señal codificada, que se hace pasar a lo

45 largo del camino 38 para la transmisión o el almacenamiento. El formateador 37 puede ensamblar también otros datos como señal codificada, como se ha estudiado anteriormente para el formateador 16.

El transcodificador 30 es capaz de realizar sus operaciones más eficientemente debido a que no se requieren recursos informáticos para implementar un controlador de cuantificación para determinar los primeros y segundos parámetros de control. El transcodificador 30 puede incluir uno o más controladores cuantificadores como el

50 controlador 14 de cuantificación descrito anteriormente, para obtener los uno o más segundos parámetros de control y/o los uno o más primeros parámetros de control, en lugar de obtener estos parámetros a partir de la señal codificada. Las características del transmisor 10 de codificación que son necesarias para determinar los primeros y segundos parámetros de control, se estudian a continuación.

55 2. Representación de valores

(1) Estimación

60 Los sistemas de codificación de audio deben representar típicamente señales de audio con una gama dinámica que excede de 100 dB. El número de bits necesarios para una representación binaria de una señal de audio o de sus componentes espectrales, que puede expresar esta gama dinámica es proporcional a la precisión de la representación.

65 En aplicaciones como el disco compacto convencional, el audio modulado por código de impulsos (PCM) está representado por dieciséis bits. Muchas aplicaciones profesionales utilizan aún más bits, 20 o 24 bits por ejemplo, para representar el audio PCM con una gama dinámica mayor y con mayor precisión.

Una representación entera de una señal de audio o de sus componentes espectrales es muy ineficiente y muchos sistemas de codificación utilizan otro tipo de representación, que incluye un valor estimado y un factor de estimación asociado de la forma

5

$$s = v \cdot f \tag{1}$$

donde: s = el valor de un componente de audio; v = un valor estimado; y f = el factor de estimación asociado.

10 El valor v estimado puede ser expresado esencialmente de cualquier manera que pueda desearse, incluyendo las representaciones fraccionarias y las representaciones enteras. Los valores positivos y negativos pueden ser representados de una diversidad de maneras, incluyendo magnitudes con signo y diversas representaciones de complementos, como el complemento a uno y el complemento a dos de los números binarios. El factor f de estimación puede ser un simple número o puede ser esencialmente cualquier función tal como una función exponencial g^f o una función logarítmica $\log_g f$, donde g es la base de las funciones exponencial y logarítmica.

15

En una implementación preferida adecuada por ser utilizada en muchos ordenadores digitales, se utiliza una representación particular de coma flotante en la cual la "mantisa" m es el valor estimado, expresado como una fracción binaria que utiliza una representación de complemento a dos, y un "exponente" x representa el factor de estimación, que es la función exponencial 2^{-x} . El resto de esta divulgación hace referencia a mantisas y a exponentes de coma flotante; sin embargo, debe entenderse que esta representación particular es meramente una manera en la cual puede aplicarse la presente invención a una información de audio representada por valores estimados y factores de estimación.

20

El valor de un componente de la señal de audio está expresado en esta representación particular de coma flotante como sigue:

25

$$s = m \cdot 2^{-x} \tag{2}$$

Por ejemplo, supóngase que un componente espectral tiene un valor igual a 0,17578125₁₀, que es igual a la fracción binaria 0,00101101₂. Este valor puede ser representado por muchas parejas de mantisas y exponentes como se ilustra en la tabla I.

30

TABLA I

Mantisa (m)	Exponente (x)	Expresión
0,00101101 ₂	0	0,00101101 ₂ x 2 ⁰ = 0,17578125 x 1 = 0,17578125
0,0101101 ₂	1	0,0101101 ₂ x 2 ⁻¹ = 0,3515625 x 0,5 = 0,17578125
0,101101 ₂	2	0,101101 ₂ x 2 ⁻² = 0,703125 x 0,25 = 0,17578125
1,01101 ₂	3	1,01101 ₂ x 2 ⁻³ = 1,40625 x 0,125 = 0,17578125

35 En esta representación particular en coma flotante, un número negativo viene expresado por una mantisa que tiene un valor que es el complemento a dos de la magnitud del número negativo. Haciendo referencia a la última fila ilustrada en la tabla I, por ejemplo, la fracción binaria 1,01101₂ en una representación de complemento a dos, expresa el valor decimal -0,59375. Como resultado, el valor realmente representado por el número de coma flotante ilustrado en la última fila de la tabla es -0,59375 x 2⁻³ = -0,07421875, que difiere del valor pretendido ilustrado en la tabla. La significación de este aspecto se estudia a continuación.

40

(2) Normalización

45 El valor de un número en coma flotante puede ser expresado con menos bits si la representación de coma flotante se "normaliza". Una representación de coma flotante distinta de cero se dice que está normalizada si los bits de una expresión binaria de la mantisa han sido desplazados hacia las posiciones de los bits más significativos, tanto como sea posible, sin perder ninguna información sobre el valor. En una representación por complemento a dos, las mantisas positivas normalizadas son siempre mayores o iguales a +0,5 y menores que +1, y las mantisas negativa normalizadas son siempre menores que -0,5 y mayores o iguales a -1. Esto es equivalente a tener el bit más significativo distinto al bit de signo. En la tabla I, la representación en coma flotante de la tercera fila está normalizada. El exponente x para la mantisa normalizada es igual a 2, que es el número de desplazamiento de bits requeridos para desplazar un bit-uno a la posición de bit más significativa.

50

Supóngase un componente espectral que tiene un valor igual a la fracción decimal -0,17578125, que es igual al número binario 1,11010011₂. El bit-uno inicial en la representación del complemento a dos indica que el valor del número es negativo. Este valor puede ser representado como un número de coma flotante que tiene una mantisa

55

normalizada $m = 1,010011_2$. El exponente x para esta mantisa normalizada es igual a 2, que es el número de desplazamientos de bits requeridos para desplazar un bit-cero a la posición de bit más significativa.

5 La representación en coma flotante ilustrada en la primera, segunda y tercera filas de la tabla I son representaciones no normalizadas. Las representaciones ilustradas en las dos primeras filas de la tabla están "infra-normalizadas" y la representación ilustrada en la última fila de la tabla está "sobre-normalizada".

10 Para fines de codificación, el valor exacto de una mantisa de un número en coma flotante normalizado puede ser representado con menos bits. Por ejemplo, el valor de la mantisa no normalizada $m = 0,00101101_2$ puede ser representado por nueve bits. Se necesitan ocho bits para representar el valor fraccionario y se necesita un bit para representar el signo. El valor de la mantisa normalizada $m = 0,101101_2$ puede ser representado con solamente siete bits.

15 El valor de la mantisa sobre-normalizada $m = 1,01101_2$ ilustrado en la última fila de la tabla I, puede ser representado por menos bits aún; sin embargo, como se ha explicado anteriormente, un número en coma flotante con una mantisa sobre-normalizada, ya no representa el valor correcto.

20 Estos ejemplos ayudan a ilustrar por qué es normalmente deseable evitar mantisas infra-normalizadas y por qué es normalmente crítico evitar mantisas sobre-normalizadas. La existencia de mantisas infra-normalizadas puede significar que se usan bits ineficientemente en una señal codificada o que se representa un valor con menos precisión, pero la existencia de mantisas sobre-normalizadas significa normalmente que los valores se distorsionan seriamente.

25 (3) Otras consideraciones para la normalización

En muchas implementaciones, el exponente está representado por un número fijo de bits o, alternativamente, está restringido a un valor dentro de una gama prescrita. Si la longitud en bits de la mantisa es mayor que el valor máximo posible del exponente, la mantisa es capaz de expresar un valor que no puede ser normalizado. Por ejemplo, si el exponente está representado por tres bits, puede expresar cualquier valor de cero a siete. Si se representa la mantisa con dieciséis bits, el menor valor distinto de cero que es capaz de representar requiere catorce desplazamientos de bits para la normalización. El exponente de 3 bits no puede expresar claramente el valor necesario para normalizar este valor de la mantisa. Esta situación no afecta los principios básicos sobre los que está basada la presente invención, pero las implementaciones prácticas deben asegurar que las operaciones aritméticas no desplazan la mantisa más allá de la gama que el exponente asociado es capaz de representar.

35 Generalmente, es muy ineficiente representar cada componente espectral de una señal codificada con su propia mantisa y su propio exponente. Se necesitan menos exponentes si múltiples mantisas comparten un exponente común.

40 Esta disposición es denominada a menudo representación de coma flotante por bloques (BFP). El valor del exponente para el bloque se establece de manera que el valor con magnitud mayor en el bloque se representa por una mantisa normalizada.

45 Se necesitan menos exponentes, y como resultado menos bits para expresar los exponentes, si se utilizan bloques más grandes. Sin embargo, el uso de bloques más grandes originará normalmente que haya más valores en el bloque que estén infra-normalizados. Por tanto, el tamaño del bloque se elige normalmente para equilibrar el dilema entre el número de bits necesarios para transportar los exponentes y las imprecisiones e ineficiencias resultantes de representar mantisas infra-normalizadas.

50 La elección de un tamaño de bloque puede afectar también a otros aspectos de la codificación, tales como la precisión de la curva de enmascaramiento calculada por un modelo perceptivo utilizado en el controlador 14 de cuantificación.

55 En algunas implementaciones, el modelo perceptivo utiliza exponentes BFP como estimación de la forma espectral, para calcular una curva de enmascaramiento. Si se utilizan bloques muy grandes para la BFP, la resolución espectral del exponente BFP se reduce, y la precisión de la curva de enmascaramiento calculada por el modelo perceptivo se degrada. Se pueden obtener detalles adicionales en el Documento A/52.

60 Las consecuencias de utilizar representaciones BFP no se estudian en la descripción siguiente. Es suficiente comprender que cuando se utilizan representaciones BFP, es muy probable que algunos componentes espectrales queden siempre infra-normalizados.

(4) Cuantificación

65 La cuantificación del componente espectral representado en forma de coma flotante, se refiere generalmente a una cuantificación de la mantisa. El exponente no se cuantifica generalmente, sino que se representa con un número fijo

de bits o, alternativamente, se restringe a tener un valor dentro de una gama prescrita.

Si la mantisa normalizada $m = 0,101101$ ilustrada en la tabla I, se cuantifica con una resolución de $0,0625 = 0,0001_2$, la mantisa cuantificada $q(m)$ es igual a la fracción binaria $0,1011_2$, que puede ser representada por cinco bits y es igual a la fracción decimal $0,6875$. El valor representado por la representación en coma flotante, tras haberse cuantificado con esta resolución particular, es $q(m) \cdot 2^{-x} = 0,6875 \times 0,25 = 0,171875$.

Si la mantisa normalizada ilustrada en la tabla se cuantifica con una resolución de $0,25 = 0,01_2$, la mantisa cuantificada es igual a la fracción binaria $0,10_2$, que puede ser representada por tres bits y es igual a la fracción decimal $0,5$. El valor representado por la representación en coma flotante, tras haberse cuantificado con esta resolución más imprecisa es $q(s) = 0,5 \times 0,25 = 0,125$.

Estos ejemplos particulares se ofrecen meramente por conveniencia de la explicación. En principio, para la presente invención no es importante ninguna forma particular de cuantificación ni ninguna relación particular entre la resolución de cuantificación y el número de bits requerido para representar una mantisa cuantificada.

(5) Operaciones aritméticas

Muchos procesadores y otra lógica por hardware implementan un conjunto especial de operaciones aritméticas que pueden ser aplicadas directamente a una representación de números en coma flotante. Algunos procesadores y lógica de proceso no implementan tales operaciones y algunas veces es atractivo utilizar estos tipos de procesadores, porque son normalmente mucho menos costosos. Cuando se utilizan tales procesadores, un método de simular operaciones de coma flotante es convertir la representación de coma flotante en representaciones fraccionarias de coma fija y precisión extendida, realizar operaciones aritméticas enteras sobre los valores convertidos, y volver a convertir las representaciones a coma flotante. Un método más eficaz es realizar operaciones aritméticas con enteros sobre las mantisas y exponentes separadamente.

Al considerar los efectos que estas operaciones aritméticas pueden tener sobre las mantisas, un transmisor de codificación puede ser capaz de modificar sus procesos de codificación, de manera que la sobre-normalización o la infra-normalización en un proceso de descodificación subsiguiente puede ser controlado o impedido si se desea.

Si tiene lugar una sobre-normalización o una infra-normalización de una mantisa de un componente espectral en un proceso de descodificación, el descodificador no puede corregir esta situación sin cambiar también el valor del exponente asociado.

Esto es particularmente problemático para el transcodificador 30, porque un cambio en un exponente significa que se necesita el complejo proceso de un controlador de cuantificación para determinar los parámetros de control para la transcodificación. Si se cambia el exponente de un componente espectral, uno o más de los parámetros de control que son transportados en la señal codificada pueden no ser ya válidos y puede ser necesario determinarlos nuevamente, a menos que el proceso de codificación que determinó estos parámetros de control fuera capaz de anticipar el cambio.

Los efectos de la adición, sustracción y multiplicación son de un interés particular porque estas operaciones aritméticas se utilizan en técnicas de codificación como las descritas anteriormente.

(a) Adición

La adición de dos números de coma flotante puede ser realizada en dos pasos. En el primer paso la estimación de los dos números es armonizada si fuera necesario. Si los exponentes de los dos números no son iguales, los bits de la mantisa asociada con el mayor exponente son desplazados hacia la derecha en un número igual a la diferencia entre los dos exponentes. En el segundo paso, se calcula una "mantisa suma" sumando las mantisas de los dos números, utilizando la aritmética de complemento a dos. La suma de los dos números originales se representa entonces por la mantisa suma y el menor exponente de los dos exponentes originales.

Al concluir la operación de adición, la mantisa suma puede ser sobre-normalizada o infra-normalizada. Si la suma de las dos mantisas originales es igual o mayor que $+1$ o es inferior a -1 , la mantisa suma será sobre-normalizada. Si la suma de las dos mantisas originales es inferior a $+0,5$ y mayor o igual a $-0,5$, la mantisa suma será infra-normalizada.

Esta última situación puede surgir si las dos mantisas originales tienen signos opuestos.

(b) Substracción

La substracción de dos números de coma flotante puede realizarse en dos pasos, de una manera que es análoga a la descrita anteriormente para la adición. En el segundo paso, se calcula una "mantisa diferencia" restando una mantisa original de la otra mantisa original, utilizando la aritmética de complemento a dos. La diferencia de los dos

números originales se representa entonces por la mantisa diferencia y el menor de los dos exponentes originales.

Al concluir la operación de substracción, la mantisa diferencia puede ser sobre-normalizada o infra-normalizada.

- 5 Si la diferencia de las dos mantisas originales es inferior a +0,5 y mayor o igual a -0,5, la mantisa diferencia será infra-normalizada. Si la diferencia de las dos mantisas originales es igual o mayor que +1 o es inferior a -1, la mantisa diferencia será sobre-normalizada. Esta última situación puede surgir si las dos mantisas originales tienen signos opuestos.

10 (c) Multiplicación

La multiplicación de dos números de coma flotante puede ser realizada en dos pasos. En el primer paso, se calcula un "exponente suma" sumando los exponentes de los dos números originales. En el segundo paso, se calcula una "mantisa producto" multiplicando las mantisas de dos números, utilizando la aritmética del complemento a dos. El producto de los dos números originales se representa después por la mantisa producto y el exponente suma.

Al concluir la operación de la multiplicación, la mantisa producto puede estar infra-normalizada pero, con una excepción, nunca puede estar sobre-normalizada porque la magnitud de la mantisa producto nunca puede ser mayor o igual que +1 o menor que -1. Si el producto de las dos mantisas originales es menor que +0,5 y mayor o igual que -0,5, la mantisa suma quedará infra-normalizada.

La única excepción a la regla para la sobre-normalización ocurre cuando ambos números de coma flotante a multiplicar tienen mantisas igual a -1. En este caso, la multiplicación produce una mantisa producto igual a +1, que está sobre-normalizada. Sin embargo, esta situación se puede impedir asegurando que al menos uno de los valores a multiplicar no es nunca negativo. Para las técnicas de síntesis estudiadas a continuación, la multiplicación se utiliza solamente para sintetizar señales de canales acoplados y para la regeneración espectral. La condición excepcional se evita en el acoplamiento requiriendo que el coeficiente de acoplamiento sea un valor no negativo, y se evita para la regeneración espectral requiriendo que la información de estimación de la envolvente, el parámetro de mezcla de componentes convertidos y el parámetro de mezcla de componentes similares al ruido sean valores no negativos.

El resto de la discusión supone que las técnicas de codificación están implementadas para evitar esta condición excepcional. Si esta condición no puede evitarse, deben adoptarse pasos para evitar también la sobre-normalización cuando se utiliza la multiplicación.

35 (d) Resumen

El efecto de estas operaciones con mantisas pueden ser resumidas como sigue: (1) la suma de dos números normalizados puede conducir a una suma que puede ser normalizada, infra-normalizada o sobre-normalizada; (2) la substracción de dos números normalizados puede conducir a una diferencia que puede ser normalizada, infranormalizada o sobre-normalizada; y (3) la multiplicación de dos números normalizados puede conducir a un producto que puede ser normalizado o infra-normalizado, pero a la vista de las limitaciones estudiadas anteriormente, no puede ser sobre-normalizado.

45 El valor obtenido a partir de estas operaciones aritméticas puede ser expresado con menos bits si es normalizado.

Las mantisas que son infra-normalizadas están asociadas con un exponente que es menor que el valor ideal para una mantisa normalizada; una expresión entera de la mantisa infra-normalizada perderá precisión porque se pierden bits significativos de las posiciones de bits menos significativos. Las mantisas que están sobre-normalizadas están asociadas con un exponentes que es mayor que el valor ideal para una mantisa normalizada; una expresión entera de la mantisa sobre-normalizada introducirá una distorsión porque los bits significativos se desplazan desde las posiciones más significativas a la posición del bit de signo. A continuación se estudia la manera en la que algunas técnicas de codificación afectan a la normalización.

55 3. Técnicas de codificación

Algunas aplicaciones imponen varios límites a la capacidad de información de una señal codificada que no pueden cumplirse por medio de técnicas básicas de codificación perceptiva, sin insertar niveles inaceptables de ruido de cuantificación en la señal descodificada. Se pueden utilizar técnicas de codificación adicionales que degradan también la calidad de la señal descodificada, pero lo hacen de una manera que reduce el ruido de cuantificación a un nivel aceptable. A continuación se estudian algunas de estas técnicas de codificación.

a) FORMACIÓN DE MATRICES

65 La formación de matrices puede ser utilizada para reducir los requisitos de capacidad de información en los sistemas de codificación de dos canales, si las señales en los dos canales tienen una alta correlación. Al formar matrices con

las dos señales con correlación en señales suma y diferencia, una de las dos señales que se ha formado como matriz tendrá un requisito de capacidad de información que es aproximadamente el mismo que una de las dos señales originales, pero la otra señal formada como matriz tendrá un requisito de capacidad de información mucho menor. Si las dos señales originales tienen una correlación perfecta, por ejemplo, el requisito de capacidad de información para una de las señales de matriz se aproximará a cero.

En principio, las dos señales originales pueden ser recuperadas perfectamente a partir de dos señales de matrices suma y diferencia; sin embargo, el ruido de cuantificación insertado por otras técnicas de codificación impedirá la recuperación perfecta. Los problemas con las matrices que pueden ser originados por el ruido de cuantificación no son pertinentes para la comprensión de la presente invención y no se van a estudiar con más detalle. Pueden obtenerse detalles adicionales a partir de referencias tales como la patente de Estados Unidos 5.291.557 y del artículo de Vernon, "Dolby Digital: Audio Coding for Digital Televisión and Storage Applications", (Dolby Digital: Codificación de audio para aplicaciones de televisión digital y de almacenamiento), de Audio Eng. Soc. 17ª Conferencia Internacional, agosto 1999, páginas 40-57. Véanse especialmente las páginas 50-51.

A continuación se ilustra una matriz típica para codificar un programa estereofónico de dos canales. Preferiblemente, se aplica la formación de matrices de forma adaptable a componentes espectrales solamente en las señales de sub-banda, si se estima que las dos señales originales de sub-banda tienen una correlación alta. La matriz combina los componentes espectrales de los canales de entrada de la izquierda y de la derecha en componentes espectrales de las señales de canal suma y diferencia, de la siguiente manera:

$$M_i = \frac{1}{2} (L_i + R_i) \quad (3a)$$

$$D_i = \frac{1}{2} (L_i - R_i) \quad (3b)$$

donde: M_i = componente espectral i en la salida del canal suma de la matriz; D_i = componente espectral i en la salida del canal diferencia de la matriz; L_i = componente espectral i en la entrada del canal izquierdo de la matriz; y R_i = componente espectral i en la entrada del canal derecho de la matriz.

Los componentes espectrales en las señales del canal suma y diferencia son codificados de una manera similar a la utilizada para componentes espectrales en las señales de las que están en forma de matriz. En situaciones en las que las señales de sub-banda para los canales izquierdo y derecho tienen una alta correlación y están en fase, los componentes espectrales en la señal del canal suma tienen magnitudes que son aproximadamente las mismas que las magnitudes de los componentes espectrales en los canales izquierdo y derecho, y los componentes espectrales en la señal del canal diferencia serán sustancialmente igual a cero. Si las señales de sub-banda para los canales izquierdo y derecho tienen una correlación alta y están invertidas en fase una con respecto a la otra, esta relación entre magnitudes de componentes espectrales en las señales del canal suma y diferencia se invierte.

Si se aplica la formación de matrices a señales de sub-banda de manera adaptable, se incluye una indicación de la formación de matrices para cada sub-banda de frecuencias en la señal codificada, de manera que el receptor puede determinar cuándo debe utilizarse una matriz inversa complementaria. El receptor procesa y descodifica independientemente las señales de sub-banda para cada canal de la señal codificada, a menos que se reciba una indicación que indique que las señales de sub-banda fueron convertidas en matrices. El receptor puede invertir los efectos de la formación de matrices y recuperar los componentes espectrales de las señales de sub-banda de los canales izquierdo y derecho, aplicando una matriz inversa como sigue:

$$L'_i = M_i + D_i \quad (4a)$$

$$R'_i = M_i - D_i \quad (4b)$$

donde: L'_i = componente espectral i en la salida recuperada del canal izquierdo de la matriz; y R'_i = componente espectral i en la salida recuperada del canal derecho de la matriz.

En general, los componentes espectrales recuperados no son exactamente iguales a los componentes espectrales originales, debido a los efectos de la cuantificación.

Si la matriz inversa recibe los componentes espectrales con mantisas que están normalizadas, las operaciones de adición y substracción en la matriz inversa pueden dar como resultado componentes espectrales recuperados con mantisas que están infra-normalizadas o sobre-normalizadas, como se ha explicado anteriormente.

Esta situación es más complicada si el receptor sintetiza substitutos para uno o más componentes espectrales en señales de matrices de sub-bandas. El proceso de síntesis crea normalmente valores de componentes espectrales que son inciertos. Esta incertidumbre hace imposible determinar por adelantado qué componentes espectrales de la matriz inversa serán sobre-normalizados o infra-normalizados, a menos que los efectos totales del proceso de síntesis sean conocidos de antemano.

b) ACOPLAMIENTO

5 El acoplamiento puede ser utilizado para codificar componentes espectrales para múltiples canales. En modos de realización preferidos, el acoplamiento está restringido a componentes espectrales en sub-bandas de frecuencias más altas; sin embargo, puede utilizarse en principio el acoplamiento para cualquier parte del espectro.

10 El acoplamiento combina componentes espectrales de señales en múltiples canales como componentes espectrales de una sola señal de canal acoplado, y codifica la información que representa la señal de canal acoplado en lugar de codificar información que represente las múltiples señales originales. La señal codificada incluye también información lateral que representa la forma espectral de las señales originales. Esta información lateral permite al receptor sintetizar múltiples señales a partir de la señal del canal acoplado, que tienen sustancialmente la misma forma espectral que las múltiples señales de canal originales. En el Documento A/52 se describe una manera en la cual se puede realizar el acoplamiento.

15 La siguiente discusión describe una sola implementación en la cual puede realizarse al acoplamiento. De acuerdo con esta implementación, los componentes espectrales del canal acoplado se forman calculando el valor medio de los correspondientes componentes espectrales en múltiples canales. Esta información lateral que representa la forma espectral de las señales originales se denomina coordenadas de acoplamiento. Una coordenada de acoplamiento para un canal particular se calcula a partir de la relación de la energía de un componente espectral en ese canal particular, con respecto a la energía del componente espectral en la señal del canal acoplado.

20 En una implementación preferida, ambos componentes espectrales y las coordenadas de acoplamiento son transportadas en la señal codificada como números en coma flotante. El receptor sintetiza múltiples señales de canal a partir de la señal de canal acoplado, multiplicando cada componente espectral de la señal de canal acoplado por la coordenada de acoplamiento apropiada. El resultado es un conjunto de señales sintetizadas que tienen la misma o sustancialmente la misma forma espectral que las señales originales. Este proceso se puede representar como sigue:

$$s_{i,j} = C_i \cdot cc_{i,j} \tag{5}$$

30 donde: $s_{i,j}$ = componente espectral i sintetizado en el canal j ; C_i = componente espectral i en la señal de canal acoplado; y $cc_{i,j}$ = coordenada de acoplamiento para el componente espectral i en el canal j .

35 Si el componente espectral de canal acoplado y la coordenada de acoplamiento están representados por números en coma flotante que están normalizados, el producto de estos dos números dará como resultado un valor representado por una mantisa que puede ser infra-normalizada, pero que nunca puede estar sobre-normalizada por las razones que se han explicado anteriormente.

40 Esta situación es más complicada si el receptor sintetiza substitutos para uno o más componentes espectrales en la señal de canal acoplado. Como se ha mencionada anteriormente, el proceso de síntesis crea normalmente valores de componentes espectrales que son inciertos, y esta incertidumbre hace imposible determinar por adelantado qué componentes espectrales de la multiplicación serán infra-normalizados, a menos que se conozcan de antemano los efectos totales del proceso de síntesis.

45 c) REGENERACIÓN ESPECTRAL

En sistemas de codificación que utilizan la regeneración espectral, un transmisor de codificación codifica solamente una parte de la banda base de una señal de audio de entrada y descarta el resto. El receptor de descodificación genera una señal sintetizada para sustituir la parte descartada. La señal codificada incluye información de estimación que es utilizada por el proceso de descodificación para controlar la síntesis de la señal, de manera que la señal sintetizada conserve en cierta medida los niveles espectrales de la parte de la señal de audio de entrada que se descarta.

55 Los componentes espectrales pueden ser regenerados de una diversidad de formas. Algunas formas utilizan un generador de números pseudo-aleatorios para generar o sintetizar componentes espectrales. Otras maneras convierten o copian componentes espectrales de la señal de banda base en partes del espectro que necesitan la regeneración.

60 No es importante ninguna forma particular para la presente invención; sin embargo, las descripciones de algunas implementaciones preferidas pueden ser obtenidas en las referencias citadas anteriormente.

La discusión siguiente describe una sola implementación de la regeneración de componentes espectrales. De acuerdo con esta implementación, se sintetiza un componente espectral copiando un componente espectral de la señal de banda base, combinando el componente copiado con un componente similar al ruido generado por un generador de números pseudo-aleatorios y estimando la combinación, de acuerdo con la información de estimación

transportada en la señal codificada. Los pesos relativos del componente copiado y el componente similar al ruido se ajustan también de acuerdo con un parámetro de mezcla transportado en la señal codificada. Este proceso puede ser representado por la expresión siguiente:

$$s_i = e_i \cdot [a_i \cdot T_i + b_i \cdot N_i] \quad (6)$$

5 donde: s_i = componente espectral i sintetizado; e_i = información de estimación de la envolvente para el componente espectral i ; T_i = componente espectral copiado para el componente espectral i ; N_i = componente similar al ruido generado para el componente espectral i ; a_i = parámetro de mezcla para el componente convertido T_i ; y b_i = parámetro de mezcla para el componente N_i similar al ruido.

10 Si el componente espectral copiado, la información de estimación de la envolvente, el componente similar al ruido y el parámetro de mezcla están representados por números en coma flotante que están normalizados, las operaciones de adición y multiplicación necesarias para generar el componente espectral sintetizado producirá un valor representado por una mantisa que puede ser infra-normalizada o sobre-normalizada por las razones que se han explicado anteriormente.

15 No es posible determinar por adelantado qué componentes espectrales sintetizados serán infra-normalizados o sobre-normalizados, a menos que se conozcan de antemano los efectos totales del proceso de síntesis.

20 B. TÉCNICAS MEJORADAS

La presente invención está dirigida a técnicas que permitan la transcodificación de señales codificadas perceptivamente, para ser realizadas más eficientemente y para proporcionar señales transcodificadas de mayor calidad. Esto se consigue eliminando algunas funciones del proceso de transcodificación, como el filtrado por análisis y por síntesis que son requeridos en los transmisores convencionales de codificación y en los receptores de descodificación. En su forma más simple, la transcodificación de acuerdo con la presente invención realiza un proceso de descodificación solamente en la medida necesaria para descuantificar información espectral, y realiza un proceso de codificación parcial solamente en la medida necesaria para recuantificar la información espectral descuantificada. Se puede realizar una descodificación y codificación adicionales si se desea. El proceso de transcodificación se simplifica aún más obteniendo los parámetros de control necesarios para controlar la descuantificación y recuantificación de la señal codificada. La discusión siguiente describe dos métodos que puede utilizar el transmisor de codificación para generar los parámetros de control necesarios para la transcodificación.

35 1. Suposiciones del caso peor

a) DESCRIPCIÓN GLOBAL

El primer método para generar parámetros de control supone condiciones del caso peor y modifica exponentes en coma flotante solamente en la medida necesaria para asegurar que nunca pueda tener lugar la sobre-normalización.

40 Es de esperar alguna infra-normalización innecesaria. Los exponentes modificados son utilizados por el controlador 14 de cuantificación para determinar los uno o más segundos parámetros de control. Los exponentes modificados no necesitan ser incluidos en la señal codificada, porque el proceso de transcodificación modifica también los exponentes bajo las mismas condiciones y modifica las mantisas que están asociadas con los exponentes modificados, de manera que la representación en coma flotante expresa el valor correcto.

45 Haciendo referencia a las figuras 2 y 4, el controlador 14 de cuantificación determina uno o más parámetros de control como se ha descrito anteriormente, y el estimador 43 analiza los componentes espectrales con respecto al proceso de síntesis del descodificador 24, para identificar qué exponentes deben ser modificados para asegurar que no tiene lugar la sobre-normalización en el proceso de síntesis. Estos exponentes son modificados y traspasados con otros exponentes no modificados al controlador 44 de cuantificación, que determina uno o más segundos parámetros de control para un proceso de re-codificación a realizar en el transcodificador 30. El estimador 43 necesita considerar solamente operaciones aritméticas en el proceso de síntesis, que pueden originar la sobre-normalización. Por esta razón, los procesos de síntesis para señales de canal acoplado como los descritos anteriormente no necesitan ser considerados porque, como se ha explicado anteriormente, este proceso particular no origina sobre-normalización.

Las operaciones aritméticas en otras implementaciones de acoplamiento pueden necesitar ser consideradas.

60 b) DETALLES DE PROCESO

(1) Formación de matrices

En la formación de matrices, el valor exacto de cada mantisa que será proporcionado a la matriz inversa no puede

ser conocido hasta después de realizar la cuantificación por el cuantificador 15, y de haber sintetizado cualquier componente similar al ruido generado por el proceso de descodificación. En esta implementación, puede suponerse el caso peor para cada operación con matrices porque los valores de la mantisa no son conocidos. Haciendo referencia a las ecuaciones 4a y 4b, la operación del caso peor en la matriz inversa es la suma de dos mantisas que tienen el mismo signo y magnitudes suficientemente grandes para sumarse a una magnitud mayor que uno, o bien la sustracción de dos mantisas que tienen signos diferentes y magnitudes suficientemente grandes para sumarse a una magnitud mayor que uno. Se puede impedir la sobre-normalización en el transcodificador para la situación del caso peor, desplazando cada mantisa un bit a la derecha y reduciendo sus exponentes en uno; por tanto, el estimador 43 disminuye los exponentes para cada componente espectral en el cálculo de la matriz inversa, y el controlador 44 de cuantificación utiliza estos exponentes modificados para determinar los uno o más segundos parámetros de control para el transcodificador. Se supone en este caso y en todo el resto de esta discusión que los valores de los exponentes antes de la modificación son mayores que cero.

Si las dos mantisas que se proporcionan realmente a la matriz inversa son conformes con la situación del caso peor, el resultado es una mantisa normalizada apropiadamente. Si la mantisa real no es conforme con la situación del caso peor, el resultado será una mantisa infra-normalizada.

(2) Regeneración espectral (HFR)

En la regeneración espectral, el valor exacto de cada mantisa que será proporcionado al proceso de regeneración, no puede ser conocido hasta después de que el cuantificador 15 haya realizado la cuantificación y de haber sintetizado cualquier componente similar al ruido generado por el proceso de descodificación. En esta implementación, se puede suponer el caso peor para cada operación aritmética, porque los valores de la mantisa no son conocidos. Haciendo referencia a la ecuación 6, la operación del caso peor es la suma de mantisas para un componente espectral convertido y un componente similar al ruido, que tengan el mismo signo y magnitudes suficientemente grandes para sumarse a una magnitud mayor que uno. Las operaciones de multiplicación no pueden originar la sobre-normalización pero tampoco pueden asegurar que no ocurra la sobre-normalización; por tanto, debe suponerse que el componente espectral sintetizado está sobre-normalizado. Puede impedirse la sobre-normalización en el transcodificador desplazando la mantisa del componente espectral y la mantisa del componente similar al ruido en un bit a la derecha y reducir los exponentes en uno; por tanto, el estimador 43 disminuye el exponente del componente convertido y el controlador 44 de cuantificación utiliza este exponente modificado para determinar los uno o más segundos parámetros de control para el transcodificador.

Si las dos mantisas que se proporcionan realmente al proceso de regeneración son conformes con la situación del caso peor, el resultado es una mantisa normalizada apropiadamente. Si las mantisas reales no son conformes con la situación del caso peor, el resultado será una mantisa infra-normalizada.

c) VENTAJAS Y DESVENTAJAS

El primer método que hace las suposiciones del caso peor, puede ser implementado económicamente. Sin embargo, requiere que el transcodificador fuerce a algunos componentes espectrales a que sean infra-normalizados y transportados con menos precisión en su señal codificada, a menos que se asignen más bits para representarlos. Además, debido a que el valor de algunos exponentes ha disminuido, las curvas de enmascaramiento basadas en estos exponentes modificados son menos precisas.

2. Procesos deterministas

a) DESCRIPCIÓN GLOBAL

El segundo método para generar parámetros de control lleva a cabo un proceso que permite determinar ejemplos específicos de sobre-normalización e infra-normalización. Los exponentes en coma flotante son modificados para impedir la sobre-normalización y para hacer mínimas las ocurrencias de infra-normalización. Los exponentes modificados son utilizados por el controlador 14 de cuantificación para determinar los uno o más segundos parámetros de control. Los exponentes modificados no necesitan estar incluidos en la señal codificada porque el proceso de transcodificación modifica también los exponentes bajo las mismas condiciones y modifica las mantisas que están asociadas con los exponentes modificados, de manera que la representación en coma flotante expresa el valor correcto.

Haciendo referencia a las figuras 2 y 5, el controlador 14 de cuantificación determina uno o más primeros parámetros de control como se ha descrito anteriormente, y el modelo 53 de síntesis analiza los componentes espectrales con respecto al proceso de síntesis del descodificador 24, para identificar qué exponentes deben ser modificados para asegurar que no tiene lugar la sobre-normalización en el proceso de síntesis y para hacer mínimas las ocurrencias de infra-normalización que tiene lugar en el proceso de síntesis. Estos exponentes son modificados y traspasados con otros exponentes sin modificar al controlador 54 de cuantificación, que determina uno o más segundos parámetros de control para un proceso de re-codificación a realizar en el transcodificador 30. El modelo 53 de síntesis realiza todo o partes del proceso de síntesis, o simula sus efectos para permitir que los efectos sobre

la normalización de todas las operaciones aritméticas del proceso de síntesis sean determinados por adelantado.

El valor de cada mantisa cuantificada y cualquier componente sintetizado debe estar disponible para el proceso de análisis que se realiza en el modelo 53 de síntesis. Si los procesos de síntesis utilizan un generador de números pseudos-aleatorios u otro proceso cuasi-aleatorio, los valores de inicialización o de origen deben estar sincronizados entre el proceso de análisis del transmisor y el proceso de síntesis del receptor. Esto puede conseguirse haciendo que el codificador 10 del transmisor determine todos los valores de inicialización e incluya alguna indicación de estos valores en la señal codificada. Si la señal codificada está organizada en intervalos o tramas independientes, puede ser deseable incluir esta información en cada trama para hacer mínimos los retardos de arranque en la decodificación y para facilitar una diversidad de actividades en la producción de programas, como la edición.

b) DETALLES DE PROCESO

(1) Formación de matrices

En la formación de matrices, es posible que el proceso de decodificación utilizado por el decodificador 24 sintetice uno o ambos componentes espectrales que se introducen en la matriz inversa. Si se sintetiza cualquiera de esos componentes, es posible que los componentes espectrales calculados por la matriz inversa estén sobre-normalizados o infra-normalizados. Los componentes espectrales calculados por la matriz inversa pueden estar también sobre-normalizados o infra-normalizados debido a los errores de cuantificación en las mantisas. El modelo 53 de síntesis puede comprobar estas condiciones no normalizadas, porque puede determinar el valor exacto de las mantisas y exponentes que son introducidos en la matriz inversa.

Si el modelo 53 de síntesis determina que se perderá la normalización, el exponente de uno o ambos componentes que se introducen en la matriz inversa pueden ser reducidos para impedir la sobre-normalización y pueden ser aumentados para impedir la infra-normalización. Los exponentes modificados no están incluidos en la señal codificada pero son utilizados por el controlador 54 de cuantificación para determinar los uno o más segundos parámetros de control. Cuando el transcodificador 30 hace las mismas modificaciones a los exponentes, las mantisas asociadas se ajustarán también, de manera que los números de coma flotante resultantes expresen los valores correctos de los componentes.

(2) Regeneración espectral (HFR)

En la regeneración espectral, es posible que el proceso de decodificación utilizado por el decodificador 24 sintetice el componente espectral convertido y que pueda sintetizar también un componente similar al ruido para sumarlo al componente convertido. Como resultado, es posible que el componente espectral calculado por el proceso de regeneración espectral sea sobre-normalizado o infra-normalizado. El componente regenerado puede estar también sobrenormalizado o infra-normalizado debido a los errores de cuantificación en la mantisa del componente convertido. El modelo 53 de síntesis puede comprobar estas condiciones no normalizadas porque puede determinar el valor exacto de las mantisas y exponentes que son introducidos en el proceso de regeneración.

Si el modelo 53 de síntesis determina que se perderá la normalización, el exponente de uno o ambos componentes que se introducen en el proceso de regeneración pueden ser reducidos para impedir la sobre-normalización y pueden ser aumentados para impedir la infra-normalización. Los exponentes modificados no están incluidos en la señal codificada pero son utilizados por el controlador 54 de cuantificación para determinar los uno o más segundos parámetros de control. Cuando el transcodificador 30 hace las mismas modificaciones a los exponentes, las mantisas asociadas se ajustarán también, de manera que los números de coma flotante resultantes expresen los valores correctos de los componentes.

(3) Acoplamiento

En los procesos de síntesis para señales de canales acoplados, es posible que el proceso de decodificación utilizado por el decodificador 24 sintetice los componentes similares al ruido para uno o más componentes espectrales de la señal de canal acoplado. Como resultado, es posible que el componente espectral calculado por el proceso de síntesis sea infra-normalizado. Los componentes sintetizados pueden estar también infra-normalizados debido a los errores de cuantificación en la mantisa de los componentes espectrales de la señal de canal acoplado. El modelo 53 de síntesis puede comprobar estas condiciones no normalizadas porque puede determinar el valor exacto de las mantisas y exponentes que son introducidos en el proceso de síntesis.

Si el modelo 53 de síntesis determina que se perderá la normalización, el exponente de uno o ambos componentes que se introducen en el proceso de síntesis pueden ser aumentados para impedir la infra-normalización. Los exponentes modificados no están incluidos en la señal codificada pero son utilizados por el controlador 54 de cuantificación para determinar los uno o más segundos parámetros de control. Cuando el transcodificador 30 hace las mismas modificaciones a los exponentes, las mantisas asociadas se ajustarán también, de manera que los números de coma flotante resultantes expresen los valores correctos de los componentes.

c) VENTAJAS Y DESVENTAJAS

Los procesos que realizan el método determinista son más costosos de implementar que los que realizan el método de estimación del caso peor; sin embargo, estos costes adicionales de implementación pertenecen a los transmisores de codificación y permiten a los transcodificadores ser implementados de una manera mucho menos costosa. Además, se pueden evitar o hacer mínimas las imprecisiones originadas por las mantisas no normalizadas y las curvas de enmascaramiento basadas en exponentes, que han sido modificados de acuerdo con el método determinista, son más precisas que las curvas calculadas por el método de la estimación del caso peor.

10 C. IMPLEMENTACIÓN

Se pueden implementar diversos aspectos de la presente invención en una diversidad de maneras, incluyendo el software para la ejecución por un ordenador o algún otro aparato que incluya componentes más especializados, tales como circuitos procesadores de señales digitales (DSP) acoplados a componentes similares a los encontrados en un ordenador de propósito general. La figura 6 es un diagrama de bloques del dispositivo 70, que puede ser utilizado para implementar aspectos de la presente invención. El DSP 72 proporciona recursos informáticos. La RAM 73 es una memoria de acceso aleatorio (RAM) del sistema utilizada por el DSP 72 para el proceso de señales. La ROM 74 representa alguna forma de almacenamiento persistente, tal como una memoria de sólo lectura (ROM), para almacenar programas necesarios para hacer funcionar el dispositivo 70 y para llevar a cabo diversos aspectos de la presente invención. El control 75 de E/S representa los circuitos de interfaz para recibir y transmitir señales por medio de canales 76, 77 de comunicaciones. Los convertidores de analógico a digital y los convertidores de digital a analógico pueden incluir el control 75 de E/S según se desee, para recibir y/o transmitir señales analógicas de audio. En el modo de realización ilustrado, todos los componentes principales del sistema se conectan al bus 71, que puede representar más de un bus físico; sin embargo, no se requiere una arquitectura de bus para implementar la presente invención.

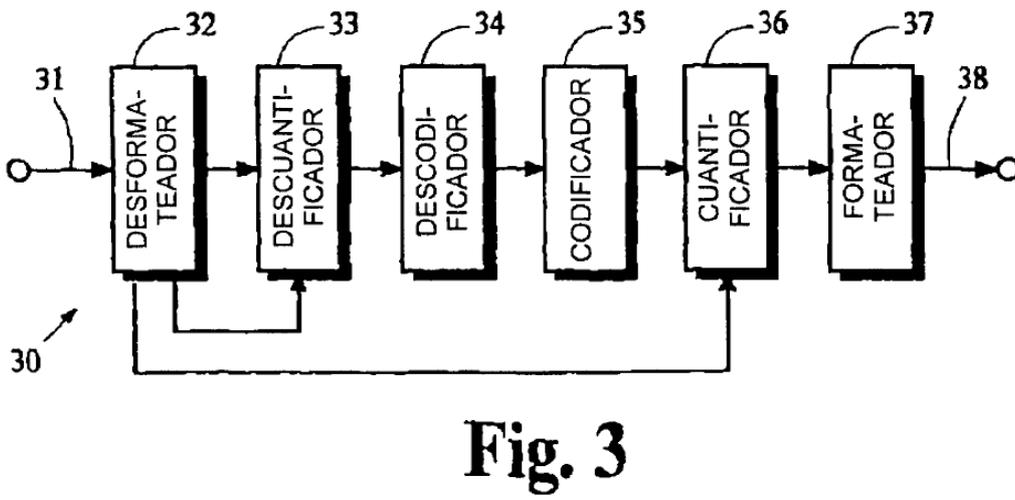
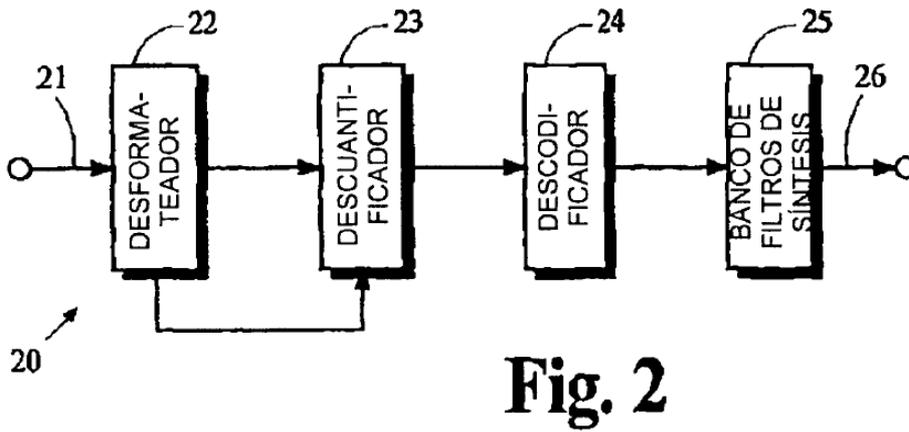
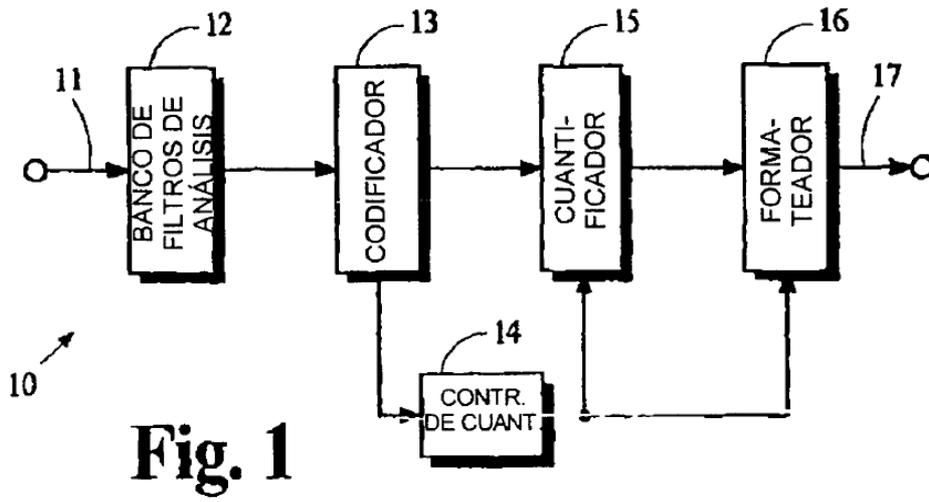
En modos de realización implementados en un sistema de ordenadores de propósito general, pueden incluirse componentes adicionales para hacer el interfaz con dispositivos tales como un teclado o un ratón y una pantalla, y para controlar un dispositivo de almacenamiento que tenga un medio de almacenamiento tal como una cinta o un disco magnéticos, o un medio óptico. El medio de almacenamiento puede ser utilizado para grabar programas de instrucciones para los sistemas operativos, utilidades y aplicaciones, y puede incluir modos de realización de programas que implementen diversos aspectos de la presente invención.

Las funciones requeridas para poner en práctica diversos aspectos de la presente invención, pueden ser realizadas por componentes que están implementados en una amplia diversidad de formas que incluyen componentes lógicos discretos, circuitos integrados, uno o más ASIC y/o procesadores controlados por programa. La manera en la cual se implementan estos componentes no es importante para la presente invención.

Las implementaciones por software de la presente invención pueden ser transportadas en una diversidad de medios legibles por máquina, tal como caminos de comunicación en banda base o modulados, a través del espectro que incluye desde frecuencias supersónicas a ultravioleta, o medios de almacenamiento que transportan información que utiliza esencialmente cualquier tecnología de grabación, incluyendo cinta, tarjetas o discos magnéticos, tarjetas o discos ópticos y marcas detectables en medios como el papel.

REIVINDICACIONES

1. Un método para transcodificar información codificada de audio en una segunda señal codificada, comprendiendo el método:
- 5 recibir una primera señal codificada transporta primeras mantisas cuantificadas, primeros exponentes, uno o más primeros parámetros de control, y uno o más segundos parámetros de control, donde las primeras mantisas cuantificadas y los primeros exponentes representan componentes espectrales de una señal de audio en una representación de coma flotante, donde cada primer exponente se asocia con las una o más primeras mantisas
- 10 cuantificadas, cada primera mantisa cuantificada se estima de acuerdo con su primer exponente asociado que se representa mediante un número fijo de bits o, alternativamente, se restringe a tener un valor dentro de una gama prescrita, y cada primera mantisa cuantificada y cada primer exponente asociado representan un respectivo componente espectral;
- 15 obtener mantisas descuantificadas a partir de las primeras mantisas cuantificadas, descuantificándolas de acuerdo con resoluciones de cuantificación adaptadas en respuesta a los uno o más primeros parámetros de control;
- 20 derivar segundos exponentes, donde dicha derivación comprende reducir primeros exponentes en uno y desplazar hacia la derecha las una o más primeras mantisas descuantificadas asociadas;
- 25 obtener segundas mantisas cuantificadas cuantificando las mantisas descuantificadas usando resoluciones de cuantificación adaptadas en respuesta a los uno o más segundos parámetros de control, donde cada segundo exponente se asocia a una o más segundas mantisas cuantificadas, cada segunda mantisa cuantificada se estima de acuerdo con su segundo exponente asociado, y cada segunda mantisa cuantificada y cada segundo exponente asociado representan un respectivo componente de control; y
- ensamblar las segundas mantisas cuantificadas, los segundos exponentes y uno o más segundos parámetros de control en la segunda señal codificada.
- 30 2. Un transcodificador para transcodificar información codificada de audio, en el que el transcodificador comprende medios adaptados para realizar todos los pasos del método de acuerdo con la reivindicación 1.
3. Un medio que transporta un programa de instrucciones ejecutable por un dispositivo, en el que la ejecución del programa de instrucciones hace que el dispositivo realice todos los pasos del método de acuerdo con la
- 35 reivindicación 1.



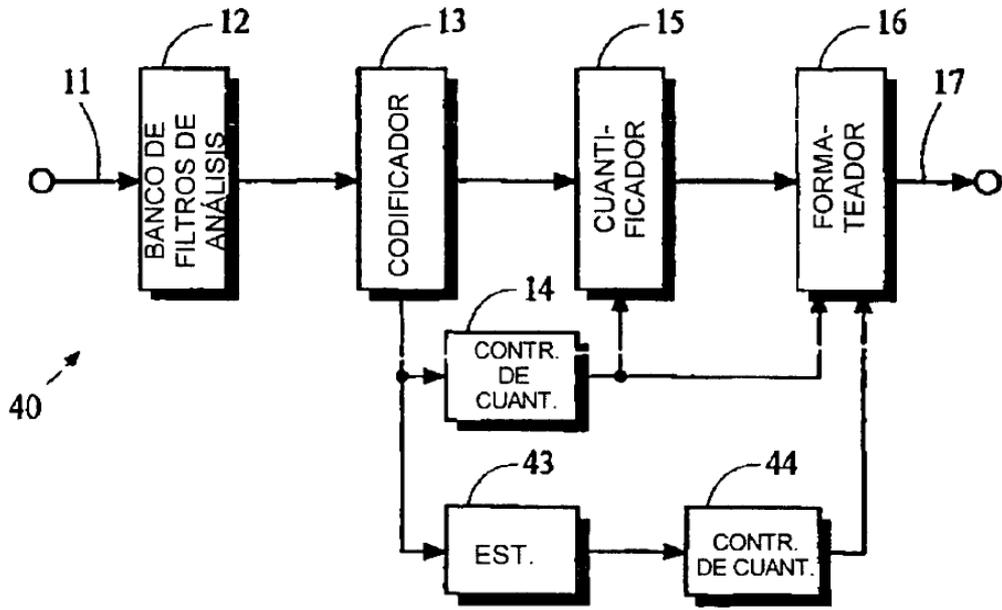


Fig. 4

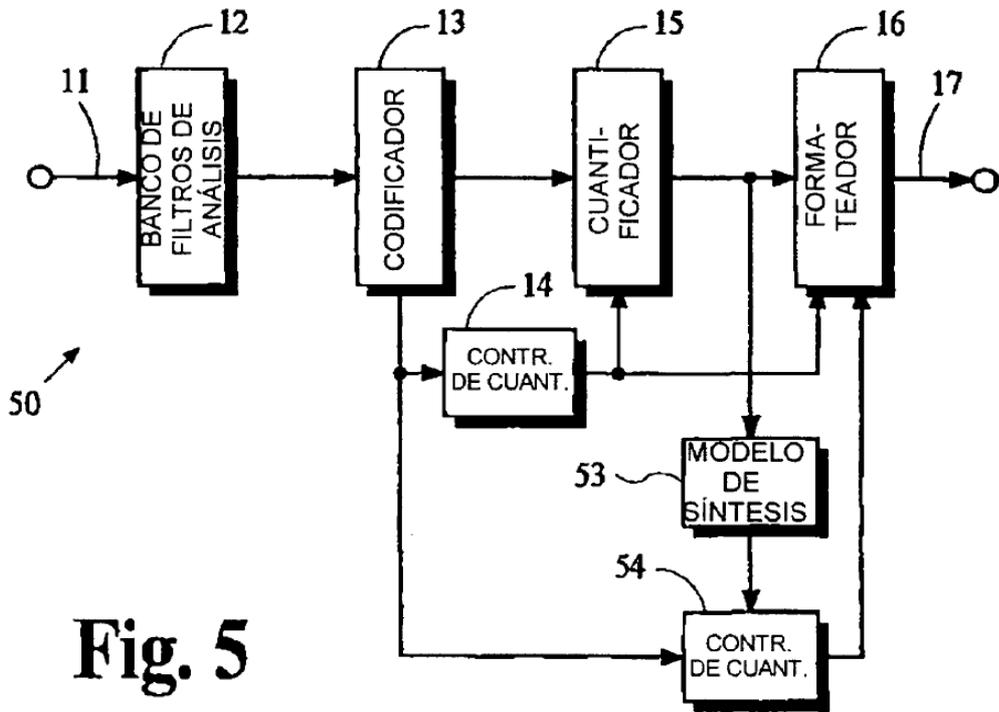


Fig. 5

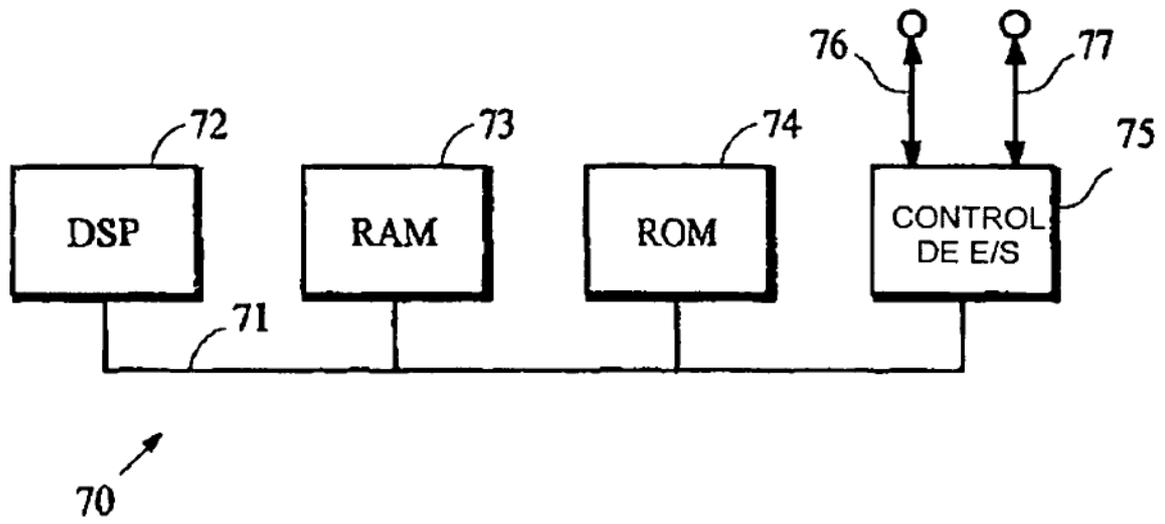


Fig. 6