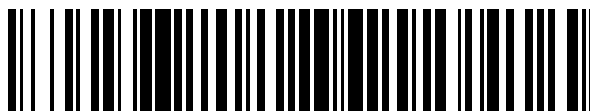


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 832 606**

51 Int. Cl.:

G10L 19/02 (2013.01)

G10L 21/038 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **30.04.2004** E 16169329 (6)

97 Fecha y número de publicación de la concesión europea: **21.10.2020** EP 3093844

54 Título: **Sistemas de codificación de audio mejorados y métodos que utilizan regeneración de componentes espectrales**

30 Prioridad:

08.05.2003 US 434449

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

10.06.2021

73 Titular/es:

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

72 Inventor/es:

**ANDERSEN, ROBERT L.;
TRUMAN, MICHAEL M.;
WILLIAMS, PHILLIP y
VERNON, STEPHEN D.**

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 832 606 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

Sistemas de codificación de audio mejorados y métodos que utilizan regeneración de componentes espectrales

Campo técnico

5 La presente invención se refiere a dispositivos de descodificación y codificación de audio y métodos para la transmisión, grabación y reproducción de señales de audio. Más en particular, la presente invención proporciona una reducción de información requerida para la transmisión o grabación de una señal de audio dada mientras que se mantiene un nivel dado de calidad percibida en la señal de salida de reproducción.

Antecedentes de la técnica

10 Numerosos sistemas de comunicación se enfrentan al problema de que la demanda para transmisión de información y capacidad de grabación a menudo supera la capacidad disponible. Como resultado, existe un considerable interés entre ellos en los campos de radiodifusión y grabación con el fin de reducir la cantidad de información requerida para transmitir o grabar una señal de audio destinada a la percepción humana sin degradar su calidad percibida. Existe también interés en la mejora de la calidad percibida de la señal de salida para un ancho de banda o capacidad de almacenamiento dados.

15 Los métodos tradicionales para reducir los requisitos de capacidad de información implican la transmisión o grabación de solamente partes seleccionadas de la señal de entrada. Las partes restantes se descartan. Las técnicas conocidas como codificación perceptual normalmente convierten una señal de audio original en componentes espectrales o señales de subbanda de frecuencia de modo que aquellas partes de la señal que son redundantes o irrelevantes se puedan identificar y descartar con mayor facilidad. Una parte de señal se considera redundante si puede recrear a partir de otras partes de la señal. Una parte de señal se considera irrelevante si es perceptualmente insignificante o inaudible. Un descodificador perceptual puede recrear las partes redundantes faltantes de una señal codificada, pero no puede crear ninguna información irrelevante faltante que no sea, además, redundante. La pérdida de información irrelevante es aceptable, sin embargo, puesto que su ausencia no tiene un efecto perceptible en la señal descodificada.

20 Una técnica de codificación de señal es perceptualmente transparente si descarta solo aquellas partes de una señal que son redundantes o perceptualmente irrelevantes. Si una técnica perceptualmente transparente no puede lograr una reducción suficiente en los requisitos de capacidad de información, entonces se necesita una técnica perceptualmente no transparente para descartar partes de señal adicionales que no son redundantes y son perceptualmente relevantes. El resultado inevitable es que se degrada la fidelidad percibida de la señal transmitida o grabada. Preferiblemente, una técnica perceptualmente no transparente descarta solamente aquellas partes de la señal que se considera que son menos significativas a nivel perceptual.

25 Una técnica de codificación denominada como "acoplamiento", que a menudo se considera una técnica perceptualmente no transparente, se puede utilizar para reducir los requisitos de capacidad de información. De acuerdo con esta técnica, las componentes espectrales en dos o más señales de audio de entrada se combinan para formar una señal de canal acoplado con una representación compuesta de estos componentes espectrales. Además, se genera información adicional que representa una envolvente espectral de las componentes espectrales en cada una de las señales de audio de entrada que se combinan para formar la representación compuesta. Una señal codificada que incluye la señal de canal acoplado y la información adicional es transmitida o grabada para una descodificación posterior por un receptor. El receptor genera señales desacopladas, que son réplicas inexactas de las señales de entrada originales, mediante la generación de copias de la señal de canal acoplado y utilizando la información adicional para escalar componentes espectrales, en las señales copiadas, de modo que las envolventes espectrales de las señales de entrada originales sean restauradas de forma sustancial. Una técnica de acoplamiento típica, para un sistema estéreo de dos canales, combina componentes de alta frecuencia de las señales de canal izquierdo y derecho para formar una sola señal de componentes de alta frecuencia compuestos y genera información adicional que representa las envolventes espectrales de las componentes de alta frecuencia en las señales originales de canal izquierdo y derecho. Un ejemplo de una técnica de acoplamiento se describe en el documento "Compresión de audio digital (AC-3)", documento de Normalización A/52 del Comité de Sistemas de Televisión Avanzados (ATSC).

30 Los requisitos de capacidad de información de la información adicional y de la señal de canal acoplado deberían elegirse para optimizar una compensación entre dos necesidades competitivas. Si el requisito de capacidad de información, para la información adicional se establece demasiado alto, el canal acoplado se verá forzada a transportar sus componentes espectrales con un nivel bajo de precisión. Niveles más bajos de precisión en las componentes espectrales de canal acoplado pueden provocar que se introduzcan niveles audibles de ruido de codificación o ruido de cuantización en las señales desacopladas. Por el contrario, si el requisito de capacidad de información de la señal de canal acoplado se establece demasiado alto, la información adicional se verá forzada a transportar las envolventes espectrales con un nivel bajo de detalle espectral. Niveles más bajos de detalle en las envolventes espectrales pueden provocar diferencias audibles en el nivel espectral y la forma de cada señal desacoplada.

55 En general, se puede conseguir una buena compensación si la información adicional transporte el nivel espectral de las subbandas de frecuencia que tienen anchos de banda proporcionales a las bandas críticas del sistema auditivo

humano. Puede observarse que las señales desacopladas pueden ser capaces de conservar los niveles espectrales de las componentes espectrales originales de las señales de entrada originales, pero generalmente no conservan la fase de las componentes espectrales originales. Esta pérdida de información de fase puede ser imperceptible si el acoplamiento está limitado a componentes espectrales de alta frecuencia puesto que el sistema auditivo humano es relativamente insensible a los cambios de fase, especialmente a altas frecuencias.

La información adicional que se genera mediante técnicas de acoplamiento tradicionales ha sido típicamente una medida de la amplitud espectral. Como resultado, el descodificador, en un sistema típico, calcula factores de escala basados en medidas de energía que se derivan de amplitudes espectrales. Estos cálculos generalmente requieren calcular la raíz cuadrada de la suma de los cuadrados de valores obtenidos a partir de la información adicional, que requiere recursos de cálculo sustanciales.

Una técnica de codificación a veces denominada "regeneración de alta frecuencia" (HFR) es una técnica perceptualmente no transparente que se puede utilizar para reducir los requisitos de capacidad de información. De acuerdo con esta técnica, se transmite o almacena una señal de banda base que contiene solamente componentes de baja frecuencia de una señal de audio de entrada. También se proporciona información adicional que representa una envolvente espectral de las componentes originales de alta frecuencia. Una señal codificada que incluye la señal de banda base y la información adicional es transmitida o grabada para una descodificación posterior por un receptor. El receptor regenera las componentes de alta frecuencia omitidas con niveles espectrales basados en la información adicional y combina la señal de banda base con las componentes de alta frecuencia regeneradas para producir una señal de salida. Una descripción de métodos conocidos para una HFR se puede encontrar en el documento de Makhoul y Berouti, "High-Frequency Regeneration in Speech Coding Systems", Proc. of the International Conf. on Acoust., Speech and Signal Proc., abril de 1979. Una técnica de HFR mejorada, que es adecuada para codificar música de alta calidad, se describe en la Solicitud de Patente de Estados Unidos número de serie 10/113.858, titulada "Broadband Frequency Translation for High Frequency Regeneration", presentada el 28 de marzo de 2002, a la que se hace referencia a continuación como la aplicación HFR.

El documento de Kjörling K et al.: Technical Description of Coding Technologies' Proposal for MPEG-4 v3 General Audio Bandwidth Extension: Spectral Band Replication (SBR) 59. MPEG meeting 11-15 marzo de 2002(ISO/IEC JTC1/SC29/WG11) también describe un método de mejora de codificación de audio a partir de codificar datos de audio de alta frecuencia al replicar la parte de alta frecuencia del espectro usando información de guiado de envolventes espectrales desde el codificador.

Los requisitos de capacidad de información de la información adicional y la señal de banda base deberían elegirse para optimizar una compensación entre dos necesidades competitivas. Si el requisito de capacidad de información para la información adicional se establece demasiado alto, la señal codificada será forzada a transportar las componentes espectrales en la señal de banda base con un nivel bajo de precisión. Niveles más bajos de precisión en las componentes espectrales de señal de banda base pueden provocar que se introduzcan niveles audibles de ruido de codificación o ruido de cuantificación en la señal de banda base y otras señales que se sintetizan a partir de ella. Por el contrario, si el requisito de capacidad de información de la señal de banda base se establece demasiado alto, la información adicional se verá forzada a transportar las envolventes espectrales con un nivel bajo de detalle espectral. Niveles más bajos de detalle, en las envolventes espectrales, pueden provocar diferencias audibles en el nivel espectral y la forma de cada señal sintetizada.

En general, se puede lograr una buena compensación si la información adicional transporte los niveles espectrales de subbandas de frecuencia que tienen anchos de banda proporcionales a las bandas críticas del sistema auditivo humano.

Al igual que para la técnica de acoplamiento anteriormente discutida, la información adicional que se genera mediante las técnicas tradicionales de HFR ha sido típicamente una medida de la amplitud espectral. Como resultado, el descodificador, en sistemas típicos, calcula factores de escala basados en medidas de energía que se derivan de amplitudes espectrales. Estos cálculos generalmente requieren calcular la raíz cuadrada de la suma de los cuadrados de valores obtenidos a partir de la información adicional, que requiere recursos de cálculo sustanciales.

Los sistemas tradicionales han utilizado técnicas de acoplamiento o técnicas de HFR, pero no ambas. En numerosas aplicaciones, las técnicas de acoplamiento pueden provocar menos degradación de señal que las técnicas de HFR, pero las técnicas de HFR pueden lograr mayores reducciones en los requisitos de capacidad de información. Las técnicas de HFR se pueden utilizar ventajosamente en aplicaciones multicanal y de un solo canal; sin embargo, las técnicas de acoplamiento no ofrecen ninguna ventaja en aplicaciones de un solo canal.

El documento WO 98/57436 describe un sistema de codificación de fuente. El sistema emplea reducción de ancho de banda antes o en el codificador, seguida por replicación de banda espectral en el descodificador. Esto se consigue mediante el uso de métodos de transposición, en combinación con ajustes de envolventes espectrales. Se ofrece tasa de bits reducida para una calidad perceptual dada o una calidad perceptual mejorada para una tasa de bits dada. El sistema se implementa preferiblemente en un códec de hardware o software, pero también se puede implementar como procesador separado en combinación con un códec.

Descripción de la invención

Un objetivo de la presente invención es proporcionar mejoras en técnicas de procesamiento de señal similares a las que implementan acoplamiento y HFR en sistemas de codificación de audio.

5 Según un aspecto de la presente invención, un método para codificar una o más señales de audio de entrada se define en la reivindicación 1 adjunta.

Según otro aspecto de la presente invención, un método para decodificar una señal codificada que representa una o más señales de audio de entrada, se define en la reivindicación 7 adjunta.

Aspectos adicionales de la invención se definen en las reivindicaciones dependientes adjuntas.

10 Las diversas características de la presente invención y sus realizaciones preferidas se entenderán mejor con referencia a la discusión siguiente y los dibujos adjuntos, en los que números de referencia similares se refieren a elementos similares en las diversas figuras. Los contenidos de la discusión siguiente y de los dibujos se presentan solamente a modo de ejemplo y no deben entenderse como que representan limitaciones del alcance de la presente invención.

Breve descripción de los dibujos

15 La Figura 1 es un diagrama de bloques esquemático de un dispositivo que codifica una señal de audio para una posterior decodificación por un dispositivo utilizando una regeneración de alta frecuencia.

La Figura 2 es un diagrama de bloques esquemático de un dispositivo que decodifica una señal de audio codificada utilizando una regeneración de alta frecuencia.

20 La Figura 3 es un diagrama de bloques esquemático de un dispositivo que divide una señal de audio en señales de subbanda de frecuencia que tienen extensiones que se adaptan en respuesta a una o más características de la señal de audio.

La Figura 4 es un diagrama de bloques esquemático de un dispositivo que sintetiza una señal de audio procedente de señales de subbanda de frecuencia que tienen extensiones que están adaptadas.

25 Las Figuras 5 y 6 son diagramas de bloques esquemáticos de dispositivos que codifican una señal de audio, utilizando acoplamiento para una decodificación posterior, por un dispositivo, utilizando desacoplamiento y regeneración de alta frecuencia.

La Figura 7 es un diagrama de bloques esquemático de un dispositivo que decodifica una señal de audio codificada utilizando desacoplamiento y regeneración de alta frecuencia.

La Figura 8 es un diagrama de bloques esquemático de un dispositivo para codificar una señal de audio que utiliza un segundo banco de filtros con el fin de proporcionar componentes espectrales adicionales para cálculos de energía.

30 La Figura 9 es un diagrama de bloques esquemático de un aparato que puede implementar varios aspectos de la presente invención.

Modos para llevar a cabo la invención

A. Descripción general

35 La presente invención se refiere a sistemas de codificación de audio y métodos que reducen los requisitos de capacidad de información de una señal codificada al descargar una parte "residual" de una señal de audio de entrada original y codificar solamente una parte de banda base de la señal de audio de entrada original, y posteriormente decodificar la señal codificada al generar una señal sintetizada para sustituir la parte residual faltante. La señal codificada incluye información de escala que es utilizada por el proceso de decodificación para controlar síntesis de
40 señal, de modo que la señal sintetizada conserve algún grado de los niveles espectrales de la parte residual de la señal de audio de entrada original.

45 Esta técnica de codificación se denomina en este documento Regeneración de Alta Frecuencia (HFR) puesto que se prevé que, en numerosas implementaciones, la señal residual contendrá componentes espectrales de frecuencia más alta. En principio, sin embargo, esta técnica no está restringida a la síntesis de solamente componentes espectrales de alta frecuencia. La señal de banda base podría incluir algunas o la totalidad de las componentes espectrales de frecuencia más alta, o podría incluir componentes espectrales en subbandas de frecuencia dispersas a través de la totalidad del ancho de banda de una señal de entrada.

1. Codificador

La Figura 1 ilustra un codificador de audio que recibe una señal de audio de entrada y genera una señal codificada que representa la señal de audio de entrada. El banco de filtros de análisis 10 recibe la señal de audio de entrada

desde la ruta 9 y, en respuesta, proporciona información de subbanda de frecuencia que representa componentes espectrales de la señal de audio. La información que representa componentes espectrales de una señal de banda base se genera a lo largo de la ruta 12 y la información que representa componentes espectrales de una señal residual se genera a lo largo de la ruta 11. Las componentes espectrales de la señal de banda base representan el contenido 5
espectral de la señal de audio de entrada en una o más subbandas, en un primer conjunto de subbandas de frecuencia, que se representan mediante información de señal transportada en la señal codificada. En una implementación preferida, el primer conjunto de subbandas de frecuencia son las subbandas de menor frecuencia. Las componentes espectrales de la señal residual representan el contenido espectral de la señal de audio de entrada en una o más 10
subbandas en un segundo conjunto de subbandas de frecuencia, que no se representan en la señal de banda base y no son transportadas por la señal codificada. En una implementación, la unión de los conjuntos primero y segundo de subbandas de frecuencia constituye el ancho de banda completo de la señal de audio de entrada.

El calculador de energía 31 calcula una o más medidas de la energía espectral en una o más subbandas de frecuencia de la señal residual. En una implementación preferida, las componentes espectrales que se reciben desde la ruta 11 se disponen en subbandas de frecuencia que tienen anchos de banda proporcionales a las bandas críticas del sistema 15
auditivo humano, y el calculador de energía 31 proporciona una medida de energía para cada una de estas subbandas de frecuencia.

El modelo de síntesis 21 representa un proceso de síntesis de señal que tendrá lugar en un proceso de decodificación, que se utilizará para decodificar la señal codificada generada a lo largo de la ruta 51. El modelo de síntesis 21 puede llevar a cabo el propio proceso de síntesis o puede realizar algún otro proceso que pueda estimar la energía espectral 20
de la señal sintetizada sin realizar realmente el proceso de síntesis. El calculador de energía 32 recibe la salida del modelo de síntesis 21 y calcula una o más medidas de la energía espectral en la señal que ha de sintetizarse. En una implementación preferida, componentes espectrales de la señal sintetizada se disponen en subbandas de frecuencia que tienen anchos de banda proporcionales a las bandas críticas del sistema auditivo humano y el calculador de energía 32 proporciona una medida de energía para cada una de estas subbandas de frecuencia.

La ilustración de la Figura 1, así como las ilustraciones en las Figuras 5, 6 y 8, muestran conexiones entre el banco de filtros de energía y el modelo de síntesis que sugiere que el modelo de síntesis responde, al menos en parte, a la 25
señal de banda base; sin embargo, esta conexión es opcional. Algunas implementaciones del modelo de síntesis se discuten a continuación. Algunas de estas implementaciones funcionan independientemente de la señal de banda base.

El calculador de factor de escala 40 recibe una o más medidas de energía desde cada uno de los dos calculadores de energía y calcula factores de escala, según se explica con más detalle a continuación. La información de escala que 30
representa factores de escala calculados se pasa a lo largo de la ruta 41.

El formateador 50 recibe la información de escala de la ruta 41 y recibe de la ruta 12 información que representa las componentes espectrales de la señal de banda base. Esta información se ensambla en una señal codificada, que se 35
pasa a lo largo de la ruta 51 para transmisión o para grabar. La señal codificada puede transmitirse por rutas de comunicación moduladas o de banda base a través de todo el espectro, incluyendo frecuencias desde supersónicas a ultravioletas, o se puede grabar en un soporte utilizando, de forma esencial, cualquier tecnología de grabación, incluyendo cinta magnética, tarjetas o discos, tarjetas o discos ópticos, y marcas detectables en soportes tales como papel.

En las implementaciones preferidas, las componentes espectrales de la señal de banda base se codifican usando procesos de codificación perceptuales que reducen los requisitos de capacidad de información al descartar partes que 40
son redundantes o irrelevantes. Estos procesos de codificación no son esenciales para la presente invención.

2. Decodificador

La Figura 2 ilustra un decodificador de audio que recibe una señal codificada que representa una señal de audio y genera una representación descodificada de la señal de audio. El deformateador 60 recibe la señal codificada 45
procedente de la ruta 59 y obtiene información de escala e información de señal de la señal codificada. La información de escala representa factores de escala y la información de señal representa componentes espectrales de una señal de banda base que tiene componentes espectrales en una o más subbandas en un primer conjunto de subbandas de frecuencia. La componente de síntesis de señal 23 realiza un proceso de síntesis para generar una señal que tiene 50
componentes espectrales en una o más subbandas, en un segundo conjunto de subbandas de frecuencia, que representan componentes espectrales de una señal residual que no fue transportada por la señal codificada.

La ilustración en las Figuras 2 y 7 muestra una conexión entre el deformateador y la componente de síntesis de señal 23, que sugiere que la síntesis de señal responde, al menos en parte, a la señal de banda base; sin embargo, esta 55
conexión es opcional. Algunas implementaciones de síntesis de señal se discuten a continuación. Algunas de estas implementaciones funcionan independientemente de la señal de banda base.

La componente de escala de señal 70 obtiene factores de escala a partir de la información de escala recibida desde la ruta 61. Los factores de escala se utilizan para poner a escala las componentes espectrales de la señal sintetizada

generada por la componente de síntesis de señal 23. El banco de filtros de síntesis 80 recibe la señal sintetizada puesta a escala desde la ruta 71, recibe las componentes espectrales de la señal de banda base desde la ruta 62, y genera, en respuesta a lo largo de la ruta 89, una señal de audio de salida que es una representación descodificada de la señal de audio de entrada original. Aunque la señal de salida no es idéntica a la señal de audio de entrada original, se prevé que la señal de salida sea perceptualmente indistinguible de la señal de audio de entrada o sea, al menos, distinguible de una manera que sea perceptualmente agradable y aceptable para una aplicación dada.

En implementaciones preferidas, la información de señal representa las componentes espectrales de la señal de banda base en una forma codificada que ha de descodificarse utilizando un proceso de descodificación que es inverso al proceso de codificación utilizado en el codificador. Tal como se mencionó con anterioridad, estos procesos no son esenciales para la presente invención.

3. Bancos de filtros

Los bancos de filtros de análisis y síntesis se pueden implementar, esencialmente de cualquier manera que se desee, incluyendo una amplia gama de tecnologías de filtros digitales, transformadas de bloques y transformadas de ondículas. En un sistema de codificación de audio que tiene un codificador y un descodificador, similares a los mostrados en las Figuras 1 y 2, respectivamente, el banco de filtros de análisis 10 se implementa mediante una Transformada de Coseno Discreta Modificada (MDCT) y el banco de filtros de síntesis 80 se implementa mediante una Transformada de Coseno Discreta Inversa modificada, que se describe en el documento de Princen et al., "Subband/Transform Coding Using Filter Bank Designs 6ase don Time Domain Aliasing Cancellation", Proc. Of the International Conf. On Acoust., Speech and Signal Proc., mayo de 1987, páginas 2161-64. Ninguna implementación de banco de filtros particular es importante, en principio.

Los bancos de filtros de análisis que se implementan mediante transformadas de bloques dividen 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 señal. Un grupo de uno o más coeficientes de transformada adyacentes representan el contenido espectral dentro de una subbanda de frecuencia particular que tiene un ancho de banda proporcional al número de coeficientes en el grupo.

Los bancos de filtros de análisis que se implementan por algún tipo de filtro digital, tal como un filtro polifásico, en lugar de una transformada de bloque, dividen una señal de entrada en un conjunto de señales de subbanda. Cada señal de subbanda es una representación basada en el tiempo del contenido espectral de la señal de entrada dentro de una subbanda de frecuencia particular. Preferentemente, la señal de subbanda es objeto de decimación de modo que cada señal de subbanda tenga un ancho de banda que es proporcional al número de muestras en la señal de subbanda durante un intervalo de tiempo unitario.

La siguiente discusión se refiere, más particularmente, a implementaciones que utilizan transformadas de bloques similares a la Transformada de Cancelación de Solapamiento de Dominio de Tiempo (TDAC) mencionada anteriormente. En esta discusión, el término "componentes espectrales" se refiere a los coeficientes de transformada y los términos "subbanda de frecuencia" y "señal de subbanda" se refieren a grupos de uno o más coeficientes de transformada adyacentes. Sin embargo, los principios de la presente invención se pueden aplicar a otros tipos de implementaciones, de modo que los términos "subbanda de frecuencia" y "señal de subbanda" se refieren también a una señal que representa el contenido espectral de una parte del ancho de banda completo de una señal, y el término "componentes espectrales" generalmente se puede entender como que se refiere a muestras o elementos de la señal de subbanda.

B. Factores de escala

En sistemas de codificación que utilizan una transformada similar a la transformada TDAC, a modo de ejemplo, los coeficientes de transformada $X(k)$ representan componentes espectrales de una señal de audio de entrada original $x(t)$. Los coeficientes de transformada se dividen en diferentes conjuntos que representan una señal de banda base y una señal residual. Los coeficientes de transformada $Y(k)$ de una señal sintetizada se generan durante el proceso de descodificación mediante el uso de un proceso de síntesis tal como uno de los descritos a continuación.

1. Cálculo

En una implementación preferida, el proceso de codificación proporciona información de escala que transporte factores de escala calculados a partir de la raíz cuadrada de una relación de una medida de energía espectral de la señal residual a una medida de energía espectral de la señal sintetizada. Las medidas de energía espectral para la señal residual y la señal sintetizada se pueden calcular a partir de las expresiones

$$E(k) = X^2(k) \quad (1a)$$

$$ES(k) = Y^2(k) \quad (1b)$$

donde $X(k)$ = coeficiente de transformada k en la señal residual;

$E(k)$ = medida de energía de la componente espectral $X(k)$;

$Y(k)$ = coeficiente de transformada k en la señal sintetizada; y

$ES(k)$ = medida de energía de la componente espectral $Y(k)$.

5 Los requisitos de capacidad de información, para información adicional, que se basan en medidas de energía para cada componente espectral son demasiado altos para la mayoría de las aplicaciones; por lo tanto, se calculan factores de escala a partir de medidas de energía de grupos o subbandas de frecuencia de componentes espectrales según las expresiones

$$E(m) = \sum_{k=m_1}^{m_2} X^2(k) \quad (2a)$$

$$ES(m) = \sum_{k=m_1}^{m_2} Y^2(k) \quad (2b)$$

10 donde $E(m)$ = medida de energía para subbanda de frecuencia m de la señal residual; y

$ES(m)$ = medida de energía para subbanda de frecuencia m de la señal sintetizada.

Los límites de la suma m_1 y m_2 especifican las componentes espectrales de frecuencia más baja y más alta en la subbanda m . En implementaciones preferidas, las subbandas de frecuencia tienen anchos de banda proporcionales a las bandas críticas del sistema auditivo humano.

15 Los límites de suma también pueden representarse utilizando una notación establecida tal como $k \in \{M\}$, donde $\{M\}$ representa el conjunto de todas las componentes espectrales que se incluyen en el cálculo de energía. Esta notación se utiliza durante el resto de esta descripción por las razones que se explican a continuación. Mediante el uso de esta notación, las expresiones 2a y 2b se pueden escribir tal como se muestra en las expresiones 2c y 2d, respectivamente,

$$E(m) = \sum_{k \in \{M\}} X^2(k) \quad (2c)$$

$$ES(m) = \sum_{k \in \{M\}} Y^2(k) \quad (2d)$$

20

donde $\{M\}$ = conjunto de todas las componentes espectrales en la subbanda m .

El factor de escala $SF(m)$, para la subbanda m , se puede calcular a partir de cualquiera de las siguientes expresiones

$$SF(m) = \sqrt{\frac{E(m)}{ES(m)}} \quad (3a)$$

$$SF(m) = \frac{\sqrt{E(m)}}{\sqrt{ES(m)}} \quad (3b)$$

25 pero un cálculo basado en la primera expresión suele ser más eficiente.

2. Representación de los factores de escala

30 Preferentemente, el proceso de codificación proporciona información de escala en la señal codificada que transporte los factores de escala calculados en una forma que requiere una capacidad de información menor que estos propios factores de escala. Se puede utilizar una diversidad de métodos para reducir los requisitos de capacidad de información de la información de escala.

35 Un método representa cada factor de escala, en sí mismo, como un número puesto a escala con un valor de escala asociado. Una forma en que se puede realizar lo anterior es representar cada factor de escala como un número de punto flotante en el que una mantisa es el número puesto a escala y un exponente asociado representa el valor de escala. La precisión de las mantisas o números puestos a escala se puede elegir para transportar los factores de escala con suficiente precisión. El intervalo permitido de los exponentes o valores de escala se puede elegir para proporcionar un intervalo dinámico suficiente para los factores de escala. El proceso que genera la información de escala también puede permitir que dos o más mantisas de punto flotante o números escalados compartan un exponente o un valor de escala comunes.

Otro método reduce los requisitos de capacidad de información mediante la normalización de los factores de escala con respecto a algún valor de base o valor de normalización. El valor de base se puede especificar por anticipado para los procesos de codificación y decodificación de la información de escala, o se puede determinar de forma adaptativa. A modo de ejemplo, los factores de escala para todas las subbandas de frecuencia de una señal de audio se pueden normalizar con respecto al mayor de los factores de escala para un intervalo de la señal de audio, o se pueden normalizar con respecto a un valor que se selecciona a partir de un conjunto de valores. Con la información de escala se incluye alguna indicación del valor base de modo que el proceso de decodificación pueda revertir los efectos de la normalización.

El procesamiento necesario para codificar y decodificar la información de escala se puede facilitar en numerosas implementaciones si los factores de escala se pueden representar mediante valores que están dentro de un intervalo de cero a uno. Este intervalo se puede asegurar si los factores de escala se normalizan con respecto a algún valor base que sea igual o mayor que todos los factores de escala posibles. De forma alternativa, los factores de escala se pueden normalizar con respecto a algún valor base mayor que cualquier factor de escala que pueda esperarse razonablemente, y establecerse igual a uno si algún evento no previsto o raro provoca que un factor de escala supere este valor.

Si se restringe el valor base para que sea una potencia de dos, los procesos que normalizan los factores de escala e invierten la normalización se pueden implementar de manera eficiente mediante funciones aritméticas de enteros binarios u operaciones de desplazamiento binario.

Se pueden usar más de uno de estos métodos juntos. A modo de ejemplo, la información de escala puede incluir representaciones de punto flotante de factores de escala normalizados.

C. Síntesis de señal

La señal sintetizada se puede generar de varias formas.

1. Desplazamiento de frecuencia

Una técnica genera componentes espectrales $Y(k)$ de la señal sintetizada mediante el desplazamiento lineal de las componentes espectrales $X(k)$ de una señal de banda base. Este desplazamiento puede expresarse como

$$Y(j) = X(k) \tag{4}$$

donde la diferencia $(j-k)$ es la cantidad de desplazamiento de frecuencia para la componente espectral k .

Cuando las componentes espectrales, en la subbanda m , se desplazan a la subbanda de frecuencia p , el proceso de codificación puede calcular un factor de escala para la subbanda de frecuencia p a partir de una medida de energía de las componentes espectrales en la subbanda de frecuencia m , según la expresión

$$SF(p) = \sqrt{\frac{E(p)}{ES(p)}} = \sqrt{\frac{\sum_{j \in \{P\}} X^2(j)}{\sum_{j \in \{P\}} Y^2(j)}} = \sqrt{\frac{\sum_{j \in \{P\}} X^2(j)}{\sum_{k \in \{M\}} X^2(k)}} \tag{5}$$

donde $\{P\}$ = conjunto de todas las componentes espectrales en la subbanda de frecuencia p ; y

$\{M\}$ = conjunto de las componentes espectrales en la subbanda de frecuencia m que son objeto de desplazamiento.

No se requiere que el conjunto $\{M\}$ contenga todas las componentes espectrales en la subbanda de frecuencia m , y algunas de las componentes espectrales en la subbanda de frecuencia m se pueden representar en el conjunto más de una vez. Esto se debe a que el proceso de desplazamiento de frecuencia puede no desplazar algunas componentes espectrales en la subbanda de frecuencia m , y puede desplazar otros componentes espectrales en subbanda de frecuencia m más de una vez en diferentes cantidades cada vez. Cualquiera o ambas de estas situaciones se producirán cuando la subbanda de frecuencia p no tenga el mismo número de componentes espectrales que la subbanda de frecuencia m .

El siguiente ejemplo ilustra una situación en la que algunas componentes espectrales, en una subbanda m , se omiten y otras se representan más de una vez. La extensión de frecuencia de la subbanda de frecuencia m es desde 200 Hz a 3.5 kHz y la extensión de frecuencia de la subbanda de frecuencia p es desde 10 kHz a 14 kHz. Se sintetiza una señal en la subbanda de frecuencia p desplazando componentes espectrales desde 500 Hz a 3.5 kHz al intervalo de 10 kHz a 13 kHz, donde la cantidad de desplazamiento para cada componente espectral es de 9.5 kHz y desplazando las componentes espectrales desde 500 Hz a 1.5 kHz, al intervalo de 13 kHz a 14 kHz, donde la cantidad de desplazamiento para cada componente espectral es 12.5 kHz. El conjunto $\{M\}$, en este ejemplo, no incluiría ninguna componente espectral desde 200 Hz a 500 Hz, pero podría incluir las componentes espectrales desde 1.5 kHz a 3.5 kHz y podría incluir dos presencias de cada componente espectral desde 500 Hz a 1.5 kHz.

La aplicación de HFR mencionada con anterioridad describe otras consideraciones que pueden incorporarse en un sistema de codificación con el fin de mejorar la calidad percibida de la señal sintetizada. Una consideración es una característica que modifica las componentes espectrales desplazadas, según sea necesario, para garantizar que se mantenga una fase coherente en la señal desplazada. En implementaciones preferidas de la presente invención, la cantidad de desplazamiento de frecuencia está restringida de modo que las componentes desplazadas mantengan una fase coherente sin ninguna modificación adicional. Para implementaciones que utilizan la transformada TDAC, a modo de ejemplo, lo anterior se puede conseguir asegurando que la cantidad de desplazamiento sea un número par.

Otra consideración es el carácter similar a ruido, o similar al tono, de una señal de audio. En muchas situaciones, la parte de frecuencia más alta de una señal de audio es más ruidosa que la parte de frecuencia más baja. Si una señal de banda base de baja frecuencia es más similar al tono, y una señal residual de alta frecuencia es más similar al ruido, el desplazamiento de frecuencia generará una señal sintetizada de alta frecuencia que es más parecida al tono que la señal residual original. El cambio en el carácter de la parte de alta frecuencia de la señal puede provocar una degradación audible, pero la audibilidad de la degradación se puede reducir o evitar mediante una técnica de síntesis descrita a continuación que utiliza el desplazamiento de frecuencia y la generación de ruido para conservar el carácter similar a ruido de la parte de alta frecuencia.

En otras situaciones en las que las partes de frecuencia más baja y frecuencia más alta de una señal son ambas semejantes a tonos, el desplazamiento de frecuencia aún puede provocar una degradación audible puesto que las componentes espectrales desplazadas no conservan la estructura armónica de la señal residual original. Los efectos audibles de esta degradación se pueden reducir o evitar restringiendo la frecuencia más baja de la señal residual que ha de sintetizarse mediante desplazamiento de frecuencias. La aplicación de HFR sugiere que la frecuencia más baja para el desplazamiento no debe ser inferior a aproximadamente 5 kHz.

2. Generación de ruido

Una segunda técnica que se puede utilizar para generar la señal sintetizada, es sintetizar una señal similar a ruido, tal como generando una secuencia de números pseudo-aleatorios para representar las muestras de una señal en el dominio del tiempo. Esta técnica particular tiene el inconveniente de que se debe utilizar un banco de filtros de análisis para obtener las componentes espectrales de la señal generada para la subsiguiente síntesis de señal. Como alternativa, se puede generar una señal similar a ruido utilizando un generador de números pseudo-aleatorios para generar, de forma directa, las componentes espectrales. Cualquiera de los métodos puede ser representado, esquemáticamente, por la expresión

$$Y(j) = N(j) \tag{6}$$

donde $N(j)$ = componente espectral j de la señal similar a ruido.

Con cualquiera de los métodos, sin embargo, el proceso de codificación sintetiza la señal similar a ruido. Los recursos de cálculo adicionales necesarios para generar esta señal aumentan la complejidad y los costes de implementación del proceso de codificación.

3. Desplazamiento y ruido

Una tercera técnica para la síntesis de señal es combinar un desplazamiento de frecuencia de la señal de banda base con las componentes espectrales de una señal sintetizada similar a ruido. En una implementación preferida, las partes relativas de la señal desplazada y la señal similar a ruido se adaptan tal como se describe en la aplicación de HFR en función de la información de control de mezcla de ruido que se trasporte en la señal codificada. Esta técnica se puede expresar como

$$Y(j) = a \cdot X(k) + b \cdot N(j) \tag{7}$$

donde a = parámetro de mezcla para la componente espectral desplazada; y

b = parámetro de mezcla para la componente espectral similar a ruido.

En una implementación, el parámetro de mezcla b se calcula tomando la raíz cuadrada de una Medida de Planeidad Espectral (SFM), que es igual a un logaritmo de la relación de la media geométrica a la media aritmética de los valores de componente espectral, que es objeto de escala y delimitado para variar dentro de un intervalo de cero a uno. Para esta implementación particular, $b=1$ indica una señal similar a ruido. Preferentemente, el parámetro de mezcla a se deriva de b , tal como se muestra en la siguiente expresión

$$a = \sqrt{c - b^2} \tag{8}$$

donde c es una constante.

En una implementación preferida, la constante c , en la expresión 8, es igual a uno y la señal similar a ruido se genera de modo que sus componentes espectrales $N(j)$ tengan un valor medio de cero y medidas de energía que sean estadísticamente equivalentes a las medidas de energía de las componentes espectrales desplazadas con las que se combinan. El proceso de síntesis puede mezclar las componentes espectrales de la señal similar a ruido con las componentes espectrales desplazadas, como se muestra arriba en la expresión 7. La energía de la subbanda de frecuencia p , en esta señal sintetizada, se puede calcular a partir de la expresión

$$ES(p) = \sum_{(j \in \{P\})} Y^2(j) = \sum_{k \in \{M\}, (j \in \{P\})} [a \cdot X(k) + b \cdot N(j)]^2 \quad (9)$$

En una implementación alternativa, los parámetros de mezcla representan funciones especificadas de frecuencia, o transporten expresamente funciones de frecuencia $a(j)$ y $b(j)$ que indican cómo el carácter similar a ruido de la señal de audio de entrada original varía con la frecuencia. En otra forma de realización alternativa, se proporcionan parámetros de mezcla para subbandas de frecuencia individuales, que se basan en medidas de ruido que se pueden calcular para cada subbanda.

El cálculo de las medidas de energía para la señal sintetizada se realiza mediante los procesos de codificación y decodificación. Los cálculos que incluyen componentes espectrales de la señal similar a ruido no son deseables puesto que el proceso de codificación debe utilizar recursos de cálculo adicionales para sintetizar la señal similar a ruido solamente para la finalidad de realizar estos cálculos de energía. La propia señal sintetizada no es necesaria para ninguna otra finalidad por el proceso de codificación.

La implementación preferida, descrita con anterioridad, permite que el proceso de codificación obtenga una medida de energía de las componentes espectrales de la señal sintetizada mostrada en la expresión 7 sin sintetizar la señal similar a ruido puesto que la energía de una subbanda de frecuencia de las componentes espectrales, en la señal sintetizada, es estadísticamente independiente de la energía espectral de la señal similar a ruido. El proceso de codificación puede calcular una medida de energía sobre la base, solamente, de las componentes espectrales desplazadas. Una medida de energía que se calcula de esta manera será, en promedio, una medida precisa de la energía real. Como resultado, el proceso de codificación puede calcular un factor de escala para la subbanda de frecuencia p a partir de solamente una medida de energía de la subbanda de frecuencia m de la señal de banda base según la expresión 5.

En una implementación alternativa, las medidas de energía espectral son transportadas por la señal codificada en lugar de los factores de escala. En esta implementación alternativa, la señal similar a ruido se genera de modo que sus componentes espectrales tengan una media igual a cero y una varianza igual a uno, y las componentes espectrales desplazadas son objeto de escala de modo que su varianza sea uno. La energía espectral de la señal sintetizada, que se obtiene combinando componentes, tal como se muestra en la expresión 7 es, en promedio, igual a la constante c . El proceso de decodificación puede poner a escala esta señal sintetizada para tener las mismas medidas de energía que la señal residual original. Si la constante c no es igual a uno, el proceso de escalado también debe tenerse en cuenta para esta constante.

D. Acoplamiento

Se pueden conseguir reducciones en los requisitos de información de una señal codificada, para un nivel dado de calidad de señal percibida en la señal decodificada, utilizando un acoplamiento en sistemas de codificación que genera una señal codificada que representa dos o más canales de señales de audio.

1. Codificador

Las Figuras 5 y 6 ilustran codificadores de audio que reciben dos canales de señales de audio de entrada desde las rutas 9a y 9b, y generan, a lo largo de la ruta 51, una señal codificada que representa los dos canales de señales de audio de entrada. Los detalles y características de los bancos de filtros de análisis 10a y 10b, los calculadores de energía 31a, 32a, 31b y 32b, los modelos de síntesis 21a y 21b, los calculadores de factor de escala 40a y 40b, y el formateador 50 son esencialmente los mismos que los descritos anteriormente para las componentes del codificador de único canal, que se ilustra en la Figura 1.

a) Características comunes

Los codificadores ilustrados en la Figura 5 y 6 son similares. Las características que son comunes a las dos implementaciones se discuten antes de que se discutan las diferencias.

Con referencia a las Figuras 5 y 6, los bancos de filtros de análisis 10a y 10b generan componentes espectrales a lo largo de las rutas 13a y 13b, respectivamente, que representan componentes espectrales de una señal de audio de entrada respectiva, en una o más subbandas en un tercer conjunto de subbandas de frecuencia. En una implementación preferida, el tercer conjunto de subbandas de frecuencia son una o más subbandas de frecuencia media que están por encima de subbandas de baja frecuencia en el primer conjunto de subbandas de frecuencia, y están por debajo de subbandas de alta frecuencia en el segundo conjunto de subbandas de frecuencia. Los

calculadores de energía 35a y 35b calculan, cada uno, una o más medidas de energía espectral en una o más subbandas de frecuencia. Preferentemente, estas subbandas de frecuencia tienen anchos de banda que son proporcionales a las bandas críticas del sistema auditivo humano y los calculadores de energía 35a y 35b proporcionan una medida de energía para cada una de estas subbandas de frecuencia.

5 El acoplador 26 genera, a lo largo de la ruta 27, una señal de canal acoplado que tiene componentes espectrales que representan un compuesto de las componentes espectrales recibidas desde las rutas 13a y 13b. Esta representación compuesta se puede formar en una diversidad de formas. A modo de ejemplo, cada componente espectral en la representación compuesta se puede calcular a partir de la suma o el promedio de los valores de componente espectral correspondiente recibidos de las rutas 13a y 13b. El calculador de energía 37 calcula una o más medidas de energía espectral en una o más subbandas de frecuencia de la señal de canal acoplado. En una implementación preferida, estas subbandas de frecuencia tienen anchos de banda que son proporcionales a las bandas críticas del sistema auditivo humano y el calculador de energía 37 proporciona una medida de energía para cada una de estas subbandas de frecuencia.

15 El calculador de factor de escala 44 recibe una o más medidas de energía de cada uno de los calculadores de energía 35a, 35b y 37, y calcula los factores de escala según se explicó con anterioridad. La información de escala, que representa los factores de escala para cada señal de audio de entrada que se representa en la señal de canal acoplado, se pasa a lo largo de las rutas 45a y 45b, respectivamente. Esta información de escala se puede codificar tal como se explicó anteriormente. En una implementación preferida, se calcula un factor de escala para cada señal de canal de entrada, en cada subbanda de frecuencia, según se representa por cualquiera de las siguientes expresiones

$$SF_i(m) = \sqrt{\frac{E_i(m)}{EC(m)}} \quad (10a)$$

$$SF_i(m) = \frac{\sqrt{E_i(m)}}{\sqrt{EC(m)}} \quad (10b)$$

donde $SF_i(m)$ = factor de escala para la subbanda de frecuencia m del canal de señal i ;

$E_i(m)$ = medida de energía para la subbanda de frecuencia m del canal de señal de entrada i ; y

25 $EC(m)$ = medida de energía para la subbanda de frecuencia m del canal acoplado.

El formateador 50 recibe información de escala de las rutas 41a, 41b, 45a y 45b, recibe información que representa componentes espectrales de señales de banda base de las rutas 12a y 12b, y recibe información que representa componentes espectrales de la señal de canal acoplado de la ruta 27. Esta información se une en una señal codificada, según se explicó con anterioridad, para transmisión o grabación.

30 Los codificadores mostrados en las Figuras 5 y 6, así como el descodificador mostrado en la Figura 7, son dispositivos de dos canales; sin embargo, varios aspectos de la presente invención se pueden aplicar en sistemas de codificación para un mayor número de canales. Las descripciones y los dibujos se refieren a implementaciones de dos canales simplemente por conveniencia de explicación e ilustración.

b) Características diferentes

35 Las componentes espectrales, en la señal de canal acoplado, se pueden utilizar en el proceso de descodificación para HFR. En tales implementaciones, el codificador debe proporcionar información de control en la señal codificada para que el proceso de descodificación la utilice en la generación de señales sintetizadas a partir de la señal de canal acoplado. Esta información de control se puede generar de varias maneras.

40 Una forma es según se ilustra en la Figura 5. Según esta implementación, el modelo de síntesis 21a responde a componentes espectrales de banda base recibidos de la ruta 12a y responde a componentes espectrales recibidos de la ruta 13a que han de ser acoplados por el acoplador 26. El modelo de síntesis 21a, los calculadores de energía asociados 31a y 32a, y el calculador de factor de escala 40a realizan cálculos de una manera análoga a los cálculos discutidos con anterioridad. La información de escala que representa estos factores de escala se pasa a lo largo de la ruta 41a al formateador 50. El formateador también recibe información de escala procedente de la ruta 41b, que representa los factores de escala calculados de manera similar para las componentes espectrales procedentes de las rutas 12b y 13b.

45 En una implementación alternativa del codificador mostrado en la Figura 5, el modelo de síntesis 21a funciona independientemente de las componentes espectrales procedentes de una o ambas de las rutas 12a y 13a, y el modelo de síntesis 21b funciona independientemente de las componentes espectrales desde una o ambas de las rutas 12b y 13b, tal como se describió con anterioridad.

5 En otra implementación, los factores de escala para HFR no se calculan para la señal de canal acoplado y/o las señales de banda base. En cambio, se pasa una representación de las medidas de energía espectral al formateador 50 y se incluye en la señal codificada en lugar de una representación de los factores de escala correspondientes. Esta implementación aumenta la complejidad del cálculo del proceso de descodificación puesto que el proceso de descodificación debe calcular al menos alguno de los factores de escala; sin embargo, reduce la complejidad de cálculo del proceso de codificación.

10 Otra manera de generar la información de control se ilustra en la Figura 6. Según esta implementación, las componentes de escala 91a y 91b reciben la señal de canal acoplado procedente de la ruta 27, y los factores de escala del calculador de factor de escala 44, y realizan un procesamiento equivalente al realizado en el proceso de descodificación, discutido a continuación, para generar señales desacopladas de la señal de canal acoplado. Las señales desacopladas se pasan a los modelos de síntesis 21a y 21b, y se calculan los factores de escala de una manera análoga a la discutida anteriormente en relación con la Figura 5.

15 En una implementación alternativa del codificador mostrado en la Figura 6, los modelos de síntesis 21a y 21b pueden funcionar independientemente de las componentes espectrales para las señales de banda base y/o la señal de canal acoplado, si estos componentes espectrales no se requieren para el cálculo de las medidas de energía espectral y factores de escala. Además, los modelos de síntesis pueden funcionar independientemente de la señal de canal acoplado, si las componentes espectrales, en la señal de canal acoplado, no se utilizan para la regeneración HFR.

2. Descodificador

20 La Figura 7 ilustra un descodificador de audio que recibe una señal codificada que representa dos canales de señales de audio de entrada desde la ruta 59, y genera a lo largo de las rutas 89a y 89b representantes descodificadas de las señales. Los detalles y características del deformateador 60, las componentes de síntesis de señal 23a y 23b, las componentes de escala de señal 70a y 70b, y los bancos de filtros de síntesis 80a y 80b, son esencialmente los mismos que los descritos anteriormente para las componentes del descodificador de único canal ilustrado en Figura 2.

25 El deformateador 60 obtiene, a partir de la señal codificada, una señal de canal acoplado y un conjunto de factores de escala de acoplamiento. La señal de canal acoplado, que tiene componentes espectrales que representan un compuesto de componentes espectrales en las dos señales de audio de entrada, se pasa a lo largo de la ruta 64.

Los factores de escala de acoplamiento, para cada una de las dos señales de audio de entrada, se pasan a lo largo de las rutas 63a y 63b, respectivamente.

30 La componente de escala de señal 92a genera, a lo largo de la ruta 93a, las componentes espectrales de una señal desacoplada que se aproxima a los niveles de energía espectral de las correspondientes componentes espectrales en una de las señales de audio de entrada originales. Estas componentes espectrales desacopladas se pueden generar multiplicando cada componente espectral, en la señal de canal acoplado, por un factor de escala de acoplamiento apropiado. En implementaciones que disponen componentes espectrales de la señal de canal acoplado en subbandas de frecuencia y proporcionan un factor de escala para cada subbanda, las componentes espectrales de una señal desacoplada se pueden generar según la expresión

$$XD_i(k) = SF_i(m) \cdot XC(k) \quad (11)$$

donde $XC(k)$ = componente espectral k en subbanda m de la señal de canal acoplado;

$SF_i(m)$ = factor de escala para la subbanda de frecuencia m del canal de señal i ; y

40 $XD_i(k)$ = componente espectral desacoplado k para el canal de señal i .

Cada señal desacoplada se pasa a un banco de filtros de síntesis respectivo. En la implementación preferida descrita con anterioridad, las componentes espectrales de cada señal desacoplada están en una o más subbandas en un tercer conjunto de subbandas de frecuencia, que son intermedias a las subbandas de frecuencia de los conjuntos primero y segundo de subbandas de frecuencia.

45 Las componentes espectrales desacopladas también se pasan a un componente de síntesis de señal respectivo 23a o 23b si se necesitan para la síntesis de señal.

E. Creación adaptativa de bandas

50 Los sistemas de codificación que disponen componentes espectrales en dos o tres conjuntos de subbandas de frecuencia, tal como se discutió anteriormente, pueden adaptar las extensiones o intervalos de frecuencias de las subbandas que se incluyen en cada conjunto. Puede ser ventajoso, a modo de ejemplo, disminuir el extremo inferior del intervalo de frecuencias, del segundo conjunto de subbandas de frecuencia, para la señal residual durante los intervalos de una señal de audio de entrada que tiene componentes espectrales de alta frecuencia que se consideran similares al ruido. Las extensiones de frecuencia se pueden adaptar, además, para eliminar todas las subbandas en

un conjunto de subbandas de frecuencia. A modo de ejemplo, el proceso de HFR se puede inhibir para las señales de audio de entrada que tienen grandes cambios bruscos en la amplitud al eliminar todas las subbandas procedentes del segundo conjunto de subbandas de frecuencia.

5 Las Figuras 3 y 4 ilustran una manera en la que las extensiones de frecuencia de la banda base, las señales de canal residual y/o acoplado se pueden adaptar por cualquier motivo, incluyendo una respuesta a una o más características de una señal de audio de entrada. Para la implementación de esta característica, cada uno de los bancos de filtros de análisis mostrados en las Figuras 1, 5, 6 y 8, se pueden sustituir por el dispositivo mostrado en la Figura 3 y cada uno de los bancos de filtros de síntesis mostrados en las Figuras 2 y 7 se pueden sustituir por el dispositivo mostrado en la Figura 4. Estas Figuras muestran cómo se pueden adaptar subbandas de frecuencia para tres conjuntos de subbandas de frecuencia; sin embargo, los mismos principios de implementación se pueden utilizar para adaptar un número diferente de conjuntos de subbandas.

10 Haciendo referencia a la Figura 3, el banco de filtros de análisis 14 recibe una señal de audio de entrada desde la ruta 9, y genera en respuesta un conjunto de señales de subbanda de frecuencia que se pasa al componente de creación adaptativa de bandas 15. La componente de análisis de señal 17 analiza la información derivada directamente de la señal de audio de entrada y/o derivada de las señales de subbanda, y genera información de control de banda en respuesta a este análisis. La información de control de banda se pasa al componente de creación adaptativa de bandas 15, y pasa la información de control de banda a lo largo de la ruta 18 hacia el formateador 50. El formateador 50 incluye una representación de esta información de control de banda en la señal codificada.

15 La componente de creación adaptativa de bandas 15 responde a la información de control de banda asignando las componentes espectrales de señal de subbanda a conjuntos de subbandas de frecuencia. Las componentes espectrales asignadas al primer conjunto de subbandas se pasan a lo largo de la ruta 12. Las componentes espectrales asignadas al segundo conjunto de subbandas se pasan a lo largo de la ruta 11. Las componentes espectrales asignadas al tercer conjunto de subbandas se pasan a lo largo de la ruta 13. Si existe un espacio o intervalo de frecuencias que no está incluido en ninguno de los conjuntos, esto se puede conseguir al no asignar componentes espectrales en este espacio o intervalo para ninguno de los conjuntos.

20 La componente de análisis de señal 17 puede generar, además, información de control de banda para adaptar las extensiones de frecuencia en respuesta a condiciones no relacionadas con la señal de audio de entrada. A modo de ejemplo, se pueden adaptar extensiones en respuesta a una señal que representa un nivel deseado de calidad de señal o la capacidad disponible para transmitir o grabar la señal codificada.

25 La información de control de banda se puede generar de muchas formas. En una implementación, la información de control de banda especifica la frecuencia más baja y/o más alta para cada conjunto en el que se deben asignar componentes espectrales. En otra implementación, la información de control de banda especifica una de entre una pluralidad de disposiciones predefinidas de extensiones de frecuencia.

30 Con referencia a la Figura 4, la componente de creación adaptativa de bandas 81 recibe conjuntos de componentes espectrales a partir de las rutas 71, 93 y 62, y recibe información de control de banda de la ruta 68. La información de control de banda se obtiene de la señal codificada por el deformateador 60. La componente de creación adaptativa de bandas 81 responde a la información de control de banda distribuyendo las componentes espectrales en los conjuntos recibidos de componentes espectrales en un conjunto de señales de subbanda de frecuencia, que se pasan al banco de filtros de síntesis 82. El banco de filtros de síntesis 82 genera, a lo largo la ruta 89, una señal de audio de salida en respuesta a las señales de subbanda de frecuencia.

F. Segundo banco de filtros de análisis

35 Las medidas de energía espectral que se calculan a partir de la expresión 1a, en codificadores de audio que implementan el banco de filtros de análisis 10, con una transformada tal como la transformada de TDAC mencionada con anterioridad, a modo de ejemplo, tienden a ser menores que la energía espectral verdadera de la señal de audio de entrada puesto que el banco de filtros de análisis proporciona solamente coeficientes de transformada de valor real. Las implementaciones que utilizan transformadas como la Transformada de Fourier Discreta (DFT) son capaces de proporcionar cálculos de energía más precisos puesto que cada coeficiente de transformada está representado por un valor complejo que trasporte con mayor precisión la verdadera magnitud de cada componente espectral.

40 La imprecisión inherente de los cálculos de energía basados en coeficientes de transformada, con solamente valores reales de transformadas tales como la transformada de TDAC, puede superarse utilizando un segundo banco de filtros de análisis con funciones de base que son ortogonales a las funciones de base del banco de filtros de análisis 10. La Figura 8 ilustra un codificador de audio que es similar al codificador mostrado en la Figura 1, pero incluye un segundo banco de filtros de análisis 19. Si el codificador utiliza la TCMD de la transformada de TDAC para implementar el banco de filtros de análisis 10, se puede utilizar una Transformada Sinusoidal Discreta Modificada (MDST) correspondiente para implementar el segundo banco de filtros de análisis 19.

El calculador de energía 39 calcula medidas más precisas de la energía espectral $E'(k)$ a partir de la expresión

$$E'(k) = X_1^2(k) + X_2^2(k) \quad (12)$$

donde $X_1(k)$ = coeficiente de transformada k del primer banco de filtros de análisis; y

$X_2(k)$ = coeficiente de transformada k del segundo banco de filtros de análisis.

- 5 En implementaciones que calculan medidas de energía para subbandas de frecuencia, el calculador de energía 39 realiza el cálculo de las medidas para una subbanda de frecuencia m a partir de la expresión

$$E'(m) = \sum_{k \in \{M\}} X_1^2(k) + X_2^2(k) \quad (13)$$

- 10 El calculador de factor de escala 49 calcula factores de escala $SF'(m)$ a partir de estas medidas más precisas de energía de una manera que es análoga a las expresiones 3a o 3b. Un cálculo análogo a la expresión 3a se muestra en la expresión 14.

$$SF'(m) = \sqrt{\frac{E'(m)}{ES(m)}} = \sqrt{\frac{\sum_{k \in \{M\}} X_1^2(k) + X_2^2(k)}{\sum_{k \in \{M\}} Y^2(k)}} \quad (14)$$

- 15 Se debe tener algún cuidado cuando se utilizan los factores de escala $SF'(m)$ que se calculan a partir de estas medidas de energía más precisas. Las componentes espectrales de la señal sintetizada que son objeto de escala de acuerdo con los factores de escala más precisos $SF'(m)$ casi, con toda seguridad, distorsionarán el balance espectral relativo de la parte de banda base de una señal y la parte sintetizada regenerada puesto que las medidas de energía más precisas siempre serán mayores o iguales a las medidas de energía calculadas solamente a partir de los coeficientes de transformada de valor real. Una manera con la que se puede compensar esta diferencia es reducir a la mitad la medición de energía más precisa porque, en promedio, la medida más precisa será el doble que la medida menos precisa. Esta reducción proporcionará un nivel estadísticamente coherente de energía en las partes de banda base y sintetizadas de una señal mientras se retiene el beneficio de una medida más precisa de la energía espectral.

- 20 Puede ser útil señalar que el denominador de la relación en la expresión 14 se debe calcular solamente a partir de los coeficientes de transformada de valor real procedentes del banco de filtros de análisis 10, incluso si están disponibles coeficientes adicionales del segundo banco de filtros de análisis 19. El cálculo de los factores de escala se debe hacer de esta manera debido a que la escala realizada durante el proceso de descodificación se basará en componentes espectrales sintetizados que son análogos solamente los coeficientes de transformada obtenidos del banco de filtros de análisis 10. El proceso de descodificación no tendrá acceso a ningunos coeficientes que correspondan o puedan derivarse de componentes espectrales obtenidos del segundo banco de filtros de análisis 19.

G. Implementación

- 30 Varios aspectos de la presente invención se pueden implementar en una amplia variedad de maneras, que incluyen software en un sistema informático de uso general, o en algún otro aparato que incluya componentes más especializados tales como circuitos de procesador de señal digital (DSP) acoplados a componentes similares a los que se encuentran en un sistema informático de uso general. La Figura 9 es un diagrama de bloques del dispositivo 70, que se puede utilizar para implementar diversos aspectos de la presente invención en un codificador de audio o descodificador de audio. El procesador DSP 72 proporciona recursos informáticos. La memoria RAM 73 es una memoria de acceso aleatorio (RAM) utilizado por el procesador DSP 72 para el procesamiento de la señal. La memoria ROM 74 representa alguna forma de almacenamiento persistente tal como memoria de solo lectura (ROM) para almacenar programas necesarios para operar el dispositivo 70 y para llevar a cabo diversos aspectos de la presente invención. El control de E/S 75 representa un circuito de interfaz para recibir y transmitir señales por medio de los canales de comunicación 76, 77. En el control I/O 75 se pueden incluir convertidores analógico a digital y los convertidores digital a analógico según se desee para recibir y/o transmitir señales de audio analógicas. En la realización mostrada, todas las componentes principales de 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.

- 45 En realizaciones implementadas en un sistema informático de uso general, se pueden incluir componentes adicionales para interconectarse a dispositivos tales como un teclado o un ratón y una pantalla, y para controlar un dispositivo de almacenamiento que tiene un soporte de almacenamiento tal como cinta o disco magnéticos, o un soporte óptico. El soporte de almacenamiento se puede utilizar para grabar programas de instrucciones para sistemas operativos, utilidades y aplicaciones, y puede incluir realizaciones de programas que implementan diversos aspectos de la presente invención.

Las funciones requeridas para practicar diversos aspectos de la presente invención se pueden realizar mediante componentes que se implementan en una amplia diversidad de maneras que incluyen componentes lógicos discretos, circuitos integrados, uno o más circuitos integrados dependientes de la aplicación (ASIC) y/o procesadores controlados por programa. La manera con la que se implementan estos componentes no es importante para la presente invención.

- 5 Las implementaciones de software de la presente invención se pueden transportar mediante una diversidad de soportes legibles por máquina, tales como rutas de comunicación de banda base o moduladas en todo el espectro, incluidas frecuencias de supersónicas a ultravioletas, o soportes de almacenamiento que transporten información utilizando, esencialmente, cualquier tecnología de grabación, incluida la cinta magnética, tarjetas o discos, tarjetas o discos ópticos, y marcas detectables en soportes similares al papel.

10

REIVINDICACIONES

1. Un método para codificar señales de audio de entrada (9), el método comprende:

5 derivar (10), de una o más señales de entrada (9), una señal de banda base (12) y una señal residual (11), la señal de banda base (12) comprende componentes espectrales en un conjunto de primeras subbandas de frecuencia, la señal residual (11) comprende componentes espectrales en un conjunto de segundas subbandas de frecuencia;

10 calcular, para cada subbanda de frecuencia en el conjunto de segundas subbandas de frecuencia de la señal residual (11), un valor de raíz cuadrada de una suma de una pluralidad de valores de potencia espectral para una pluralidad de componentes espectrales en esa subbanda de frecuencia en el conjunto de segundas subbandas de frecuencia de la señal residual (11);

generar (40), para cada subbanda de frecuencia en el conjunto de segundas subbandas de frecuencia de la señal residual (11), un factor de escala basado al menos en parte en un valor de raíz cuadrada respectivo de una suma de una pluralidad de valores de potencia espectral como se calcula para esa subbanda de frecuencia en el conjunto de segundas subbandas de frecuencia de la señal residual (11);

15 codificar la señal de banda base (12), junto con información de control (41), hasta una señal codificada (51), la información de control (41) incluye un conjunto de factores de escala, cada factor de escala en el conjunto de factores de escala es generado respectivamente para una subbanda de frecuencia en el conjunto de segundas subbandas de frecuencia de la señal residual (11) basado al menos en parte en un valor de raíz cuadrada de una suma de una pluralidad de valores de potencia espectral como se calcula para esa subbanda de frecuencia en el conjunto de segundas subbandas de frecuencia de la señal residual (11).

2. El método de la reivindicación 1, en donde el conjunto de factores de escala va a ser usado, por un descodificador de audio que recibe la señal codificada, para escalar componentes espectrales de una señal sintetizada que va a ser generada por el descodificador de audio.

3. El método de la reivindicación 2, en donde la señal sintetizada se va a generar basándose al menos en parte en las componentes espectrales en el conjunto de primeras subbandas de frecuencia de la señal de banda base (12), en donde opcionalmente la señal sintetizada comprende al menos en parte una componente de señal similar a ruido,

30 en donde además opcionalmente: componentes espectrales de al menos una segunda subbanda de frecuencia en el conjunto de segundas subbandas de frecuencia de la señal residual (11) se obtiene en parte al desplazar componentes espectrales de al menos una primera subbanda de frecuencia en el conjunto de primeras subbandas de frecuencia de la señal de banda base (12).

4. El método de cualquier reivindicación anterior, en donde la señal sintetizada se genera independiente de la señal de base banda (12).

5. El método de cualquier reivindicación anterior, en donde las componentes espectrales en el conjunto de primeras subbandas de frecuencia de la señal de banda base (12) se disponen en subbandas de frecuencia que tienen anchos de banda proporcionales a bandas críticas del sistema auditivo humano.

6. El método de cualquier reivindicación anterior, en donde un valor de potencia espectral para una componente espectral en una subbanda de frecuencia en el conjunto de segundas subbandas de frecuencia de la señal residual se calcula a partir de un coeficiente de transformada de la componente espectral.

7. Un método para decodificar señales de audio, el método comprende:

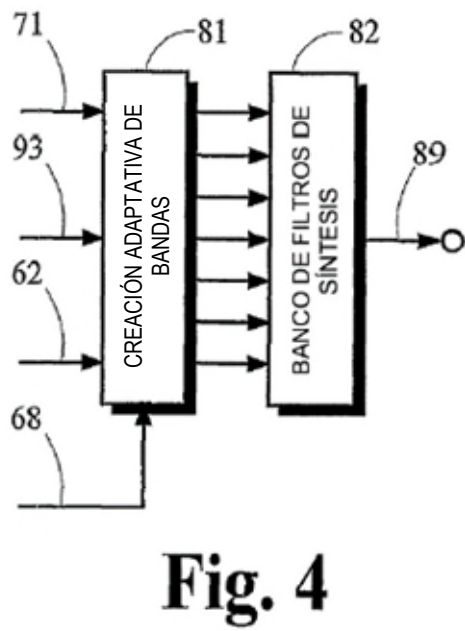
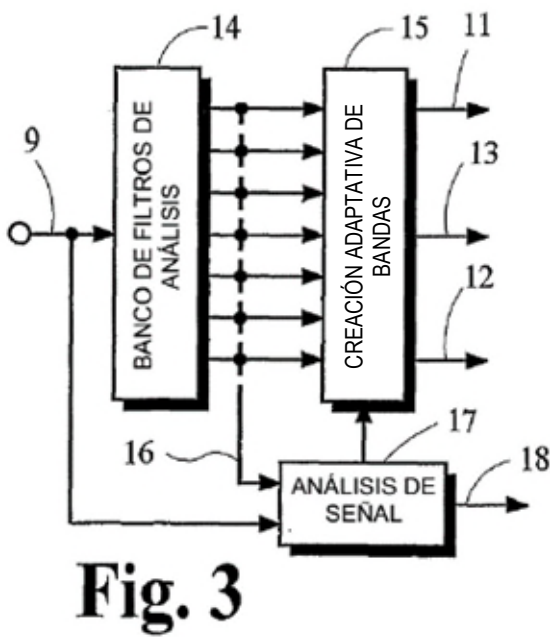
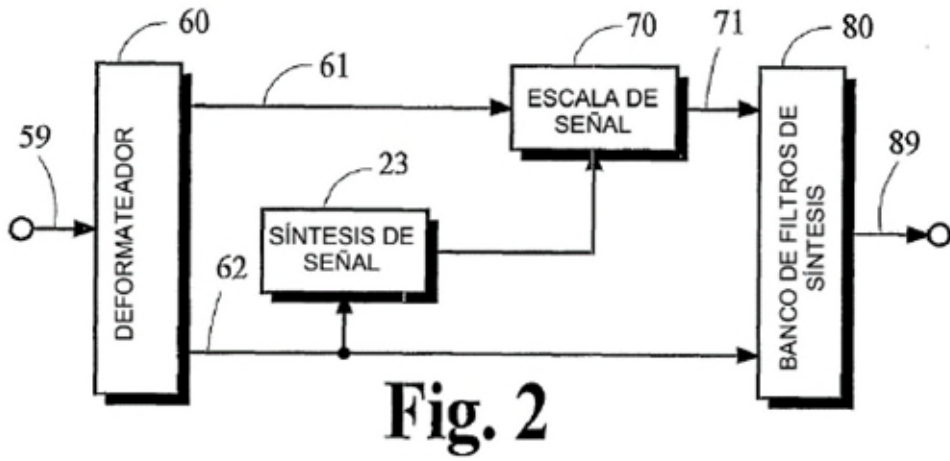
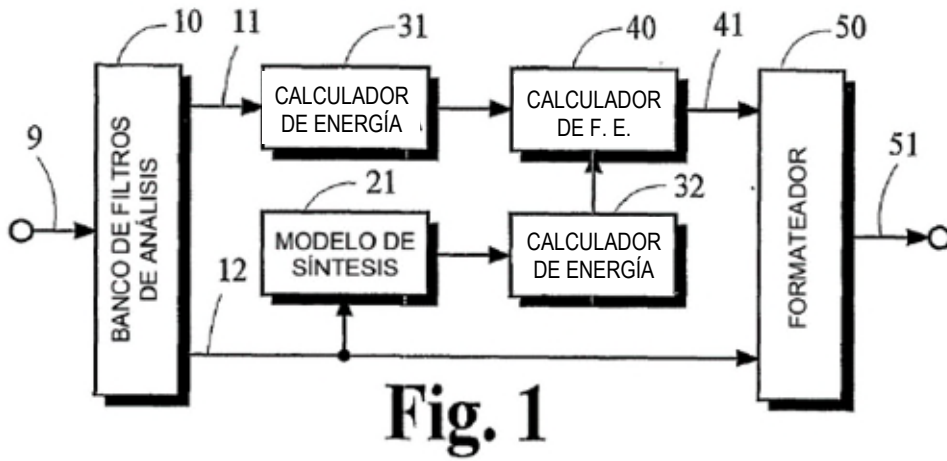
40 descodificar una señal codificada (59) hasta una señal de banda base (62) e información de control (61), la señal de banda base (62) que tiene componentes espectrales en un conjunto de primeras subbandas de frecuencia, la información de control (61) incluye un conjunto de factores de escala, cada factor de escala en el conjunto de factores de escala es generado respectivamente para una subbanda de frecuencia en un conjunto de segundas subbandas de frecuencia de una señal residual (11) basado al menos en parte en un valor de raíz cuadrada de una suma de una pluralidad de valores de potencia espectral como se calcula para esa subbanda de frecuencia en el conjunto de segundas subbandas de frecuencia de la señal residual (11);

45 generar componentes espectrales en el conjunto de segundas subbandas de frecuencia, las componentes espectrales en una segunda subbanda de frecuencia en el conjunto de segundas subbandas de frecuencia se escalan sobre la base de un factor de escala respectivo en el conjunto de factores de escala como se decodifica de la señal codificada (59);

50 generar una señal sintetizada, sobre la base de las componentes espectrales en el conjunto de primeras subbandas de frecuencia de la señal de banda base (62) y las componentes espectrales en el conjunto de segundas subbandas de frecuencia de la señal sintetizada.

8. Un medio que transporte un programa de instrucciones ejecutables por un dispositivo, en donde la ejecución del programa de instrucciones provoca que el dispositivo realice el método de cualquiera de las reivindicaciones 1-7.

9. Un aparato que comprende uno o más procesadores configurados para realizar el método de una cualquiera de las reivindicaciones 1-7.



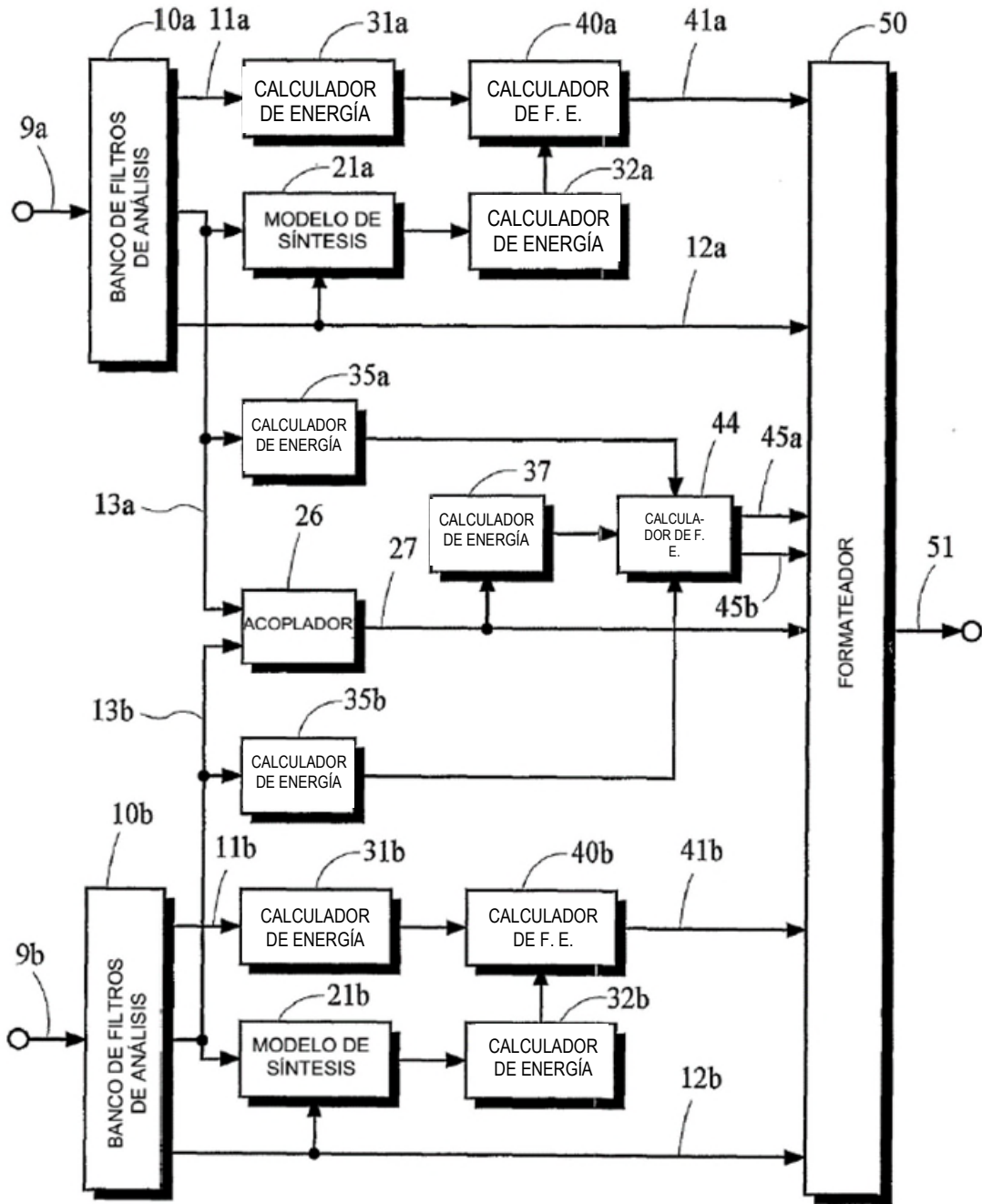


Fig. 5

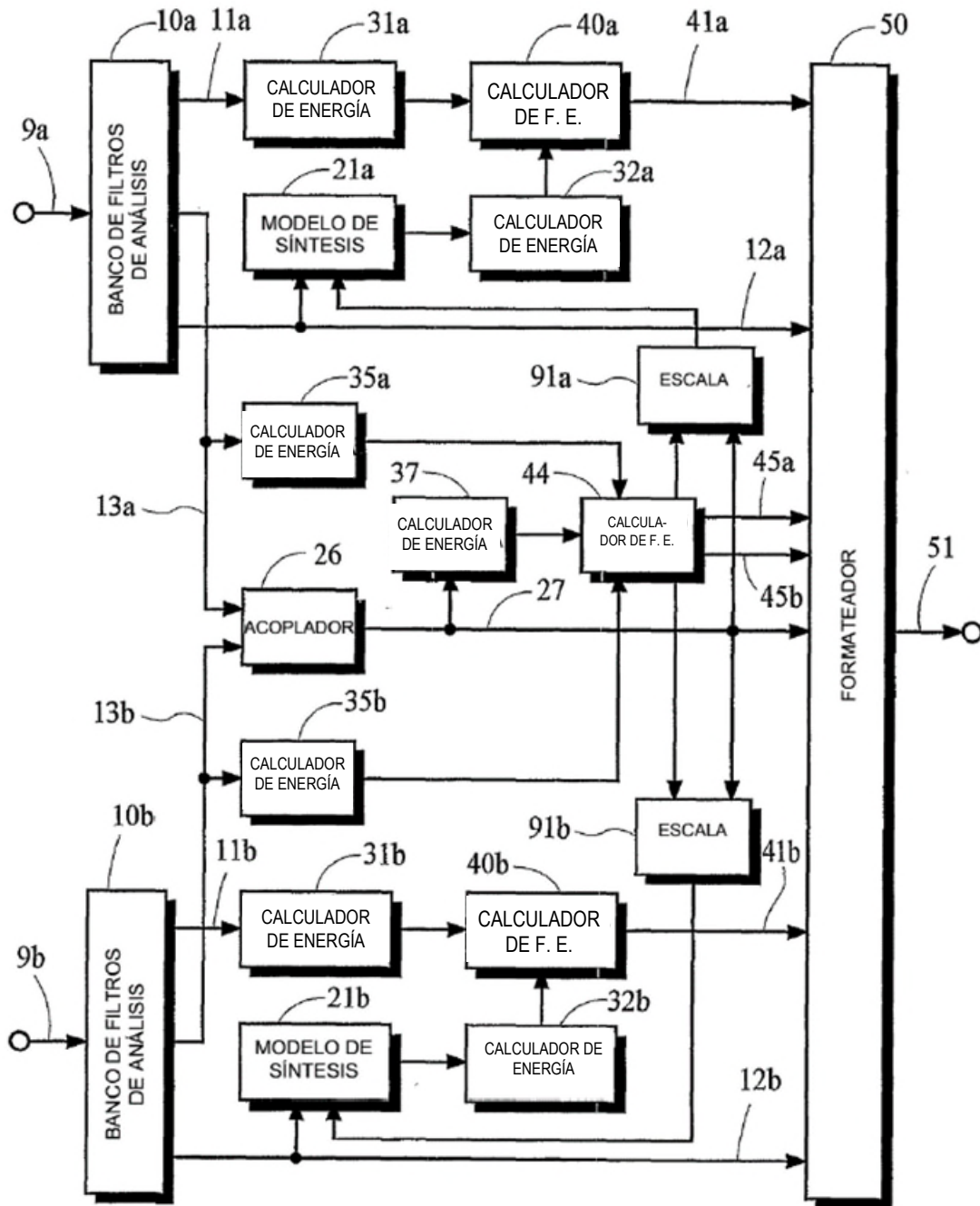


Fig. 6

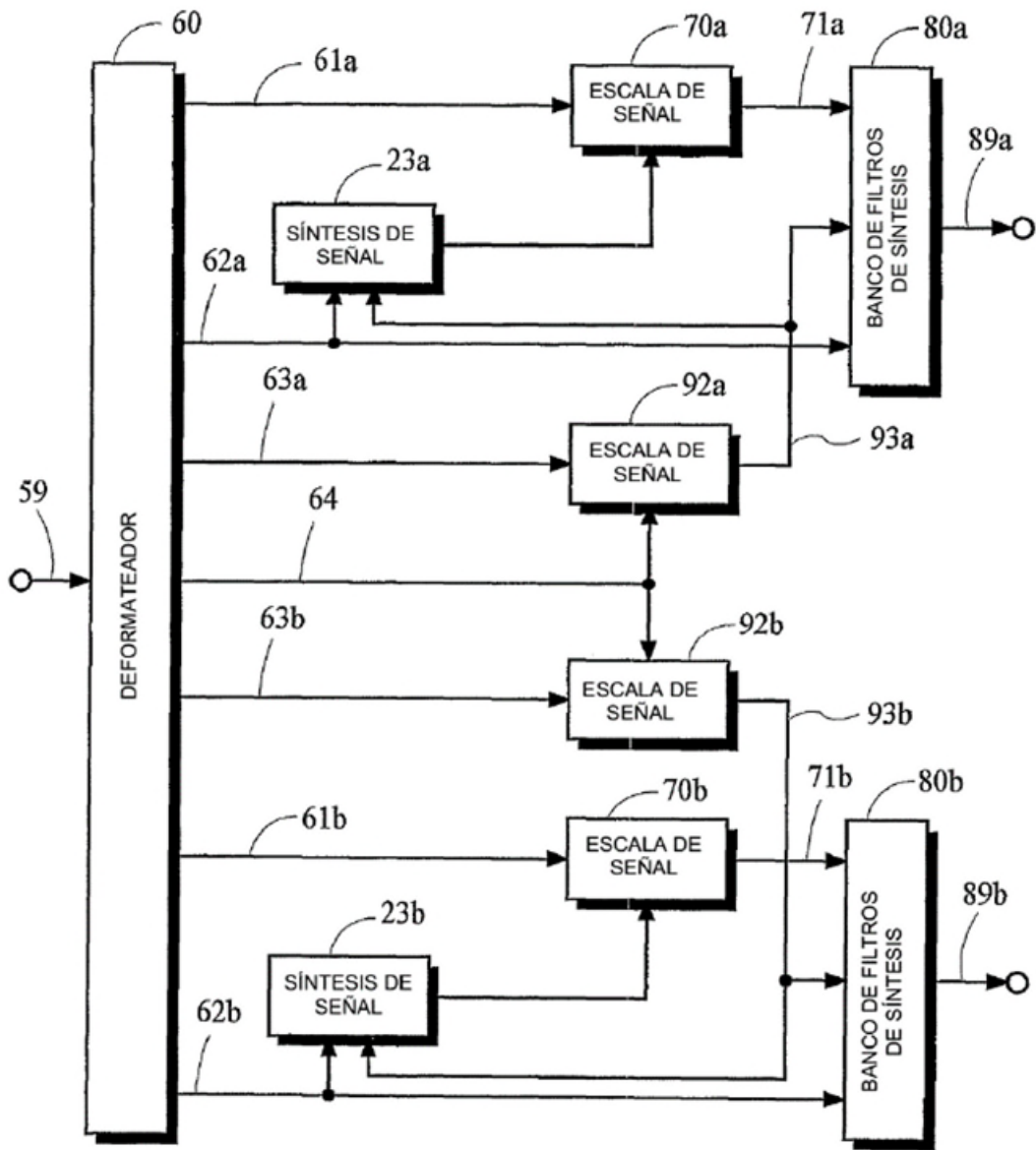


Fig. 7

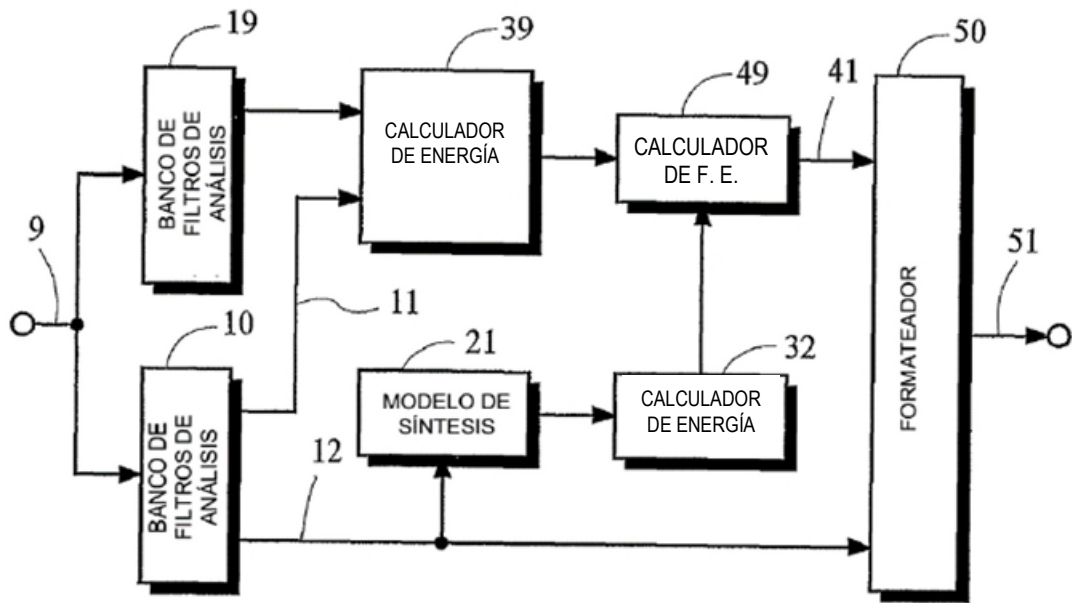
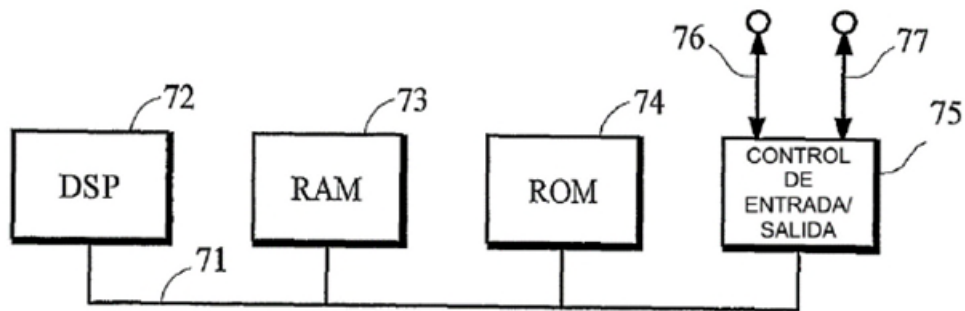


Fig. 8



70 ↗

Fig. 9