(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 526 333**

(51) Int. Cl.:

G10L 21/038 (2013.01)
G10L 19/028 (2013.01)
G10L 19/035 (2013.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **26.08.2008** **E 12196913 (3)**

(97) Fecha y número de publicación de la concesión europea: **22.10.2014** **EP 2571024**

(54) Título: **Frecuencia de transición adaptativa entre el rellenado con ruido y la extensión del ancho de banda**

(30) Prioridad:

27.08.2007 US 968134 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
09.01.2015

(73) Titular/es:

TELEFONAKTIEBOLAGET L M ERICSSON (PUBL) (100.0%)
164 83 Stockholm, SE

(72) Inventor/es:

TALEB, ANISSE y
BRIAND, MANUEL

(74) Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 526 333 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

Frecuencia de transición adaptativa entre el rellenado con ruido y la extensión del ancho de banda

Campo Técnico

La presente invención se refiere en general a métodos y dispositivos para la codificación y la descodificación de señales de audio, y en particular a métodos y dispositivos para el rellenado del espectro.

Antecedentes

Cuando las señales van a ser almacenadas y/o transmitidas, un planteamiento estándar actual es codificar las señales de audio en una representación digital de acuerdo con diferentes esquemas. Con el fin de ahorrar capacidad de almacenamiento y/o de transmisión, un deseo general es reducir el tamaño de la representación digital necesario para permitir la reconstrucción de las señales de audio con suficiente calidad. El compromiso entre el tamaño de la señal codificada y la calidad de la señal depende de la aplicación real.

Los codificadores de audio basados en transformada comprimen las señales de audio cuantificando los coeficientes de transformación. Para permitir bajas tasas de bits, los cuantificadores podrían concentrar los bits disponibles en los coeficientes más energéticos y perceptualmente relevantes y transmitir sólo esos, dejando “agujeros espectrales” de coeficientes no cuantificados en el espectro de frecuencia.

La llamada tecnología de SBR (Replicación de Banda Espectral – Spectral Band Replication, en inglés), véase por ejemplo el TS 26.404 V6.0.0 (2004-09) del 3GPP, “Enhanced aacPlus general audio codec – encoder SBR part (Versión 6)”, 2004 [1], cierra el hueco entre la señal de banda limitada de un codificador perceptual convencional y el ancho de banda audible de aproximadamente 15 kHz. La idea general tras la SBR es recrear el contenido de alta frecuencia faltante de una señal decodificada de una manera perceptualmente precisa. Las frecuencias por encima de 15 kHz son menos importantes desde un punto de vista psicoacústico, pero pueden ser también reconstruidas. No obstante, la SBR no puede ser utilizada como un códec autónomo. Siempre opera en conjunción con un códec de forma de onda convencional, un llamado códec de núcleo. El códec de núcleo es responsable de transmitir la parte inferior del espectro original mientras que el descodificador de SBR, que es principalmente un post-proceso al descodificador de forma de onda convencional, reconstruye el intervalo de frecuencia no transmitido. Los valores espectrales de la banda ancha no son transmitidos directamente como en los códecs convencionales. El sistema combinado ofrece una ganancia de codificación superior a la ganancia del códec de núcleo solo.

La metodología de SBR se basa en la definición de una frecuencia de transición fija entre una banda baja, bajas frecuencias perceptualmente relevantes codificadas, y una banda alta, altas frecuencias menos relevantes no codificadas. No obstante, en la práctica, esta frecuencia de transición se basa en el contenido de audio de la señal original. En otras palabras, de una señal a otra, la frecuencia de transición apropiada puede variar mucho. Éste es por ejemplo el caso cuando se comparan señales de conversación limpia y de música de banda completa.

Los “agujeros espectrales” del espectro descodificado pueden dividirse en dos tipos. El primero es pequeños agujeros a bajas frecuencias debido al efecto de enmascaramiento instantáneo, véase por ejemplo “Estimation of Perceptual Entropy Using Noise Masking Criteria”, Proc. ICASSP, pp. 2524-2527, Mayo de 1988, de J.D. Johnston [2]. El segundo es agujeros más grandes a partir de la saturación por el umbral absoluto de audición y la adición de enmascaramiento [2]. La SBR afecta principalmente al segundo tipo.

Además, un códec de audio típico basado en tal método que se dirige a rellenar el “agujero espectral”, es decir, coeficientes no codificados, para las altas frecuencias, es decir, el segundo tipo de “agujeros espectrales”, podría preferiblemente rellenar los agujeros espectrales sobre el espectro completo. Verdaderamente, incluso si un códec de SBR es capaz de proporcionar una señal de audio de ancho de banda completo, las altas frecuencias reconstruidas no enmascararán las molestas aberraciones introducidas por la descodificación, es decir, la cuantificación de la banda baja, es decir, las frecuencias bajas perceptualmente relevantes.

El documento WO 02/41302 A1 describe la adaptación de la frecuencia de transición de acuerdo con las características de la señal de entrada.

Compendio

Un objeto general de la presente invención es proporcionar métodos y dispositivos para permitir una supresión eficiente de las aberraciones perceptuales provocadas por agujeros espectrales sobre una señal de audio de banda completa.

Los objetos anteriores se logran mediante métodos y dispositivos de acuerdo con las reivindicaciones de patente adjuntas.

La presente invención tiene varias ventajas. Una ventaja es que un uso de la frecuencia de transición permite la utilización de un rellenado combinado del espectro utilizando tanto rellenado con ruido como extensión del ancho de banda. Además, la frecuencia de transición se define adaptativamente, por ejemplo, de acuerdo con el esquema de codificación utilizado, lo que hace al rellenado del espectro dependiente por ejemplo de la resolución de la frecuencia. Cualquier códec de conversación y/o de audio que utilice este método es capaz de proporcionar una alta calidad, es decir, con menos aberraciones molestas, y una señal de audio de ancho de banda completo. El método es flexible en el sentido de que puede ser combinado con cualquier tipo de representación de la frecuencia (DCT, MDCT, etc.) o bancos de filtros, es decir, con cualquier códec (perceptual, paramétrico, etc.).

Breve descripción de los dibujos

La invención, junto con otros objetos y ventajas de la misma, puede comprenderse mejor haciendo referencia a la siguiente descripción tomada junto con los dibujos que se acompañan, en los cuales:

la FIG. 1 es un esquema de bloques esquemático de un sistema de códec;

la FIG. 2 es un esquema de bloques esquemático de una realización de una realización de un codificador de señal de audio de acuerdo con la presente invención;

la FIG. 3 es una ilustración esquemática de coeficientes espectrales, grupos de los mismos y bandas de frecuencia;

la FIG. 4 es un esquema de bloques esquemático de una realización de una realización de un descodificador de señal de audio de acuerdo con la presente invención;

las FIGS. 5A-C son ilustraciones de principios para encontrar una frecuencia de transición;

la FIG. 6 es un diagrama de flujo de etapas de una realización de un método de acuerdo con la presente invención; y

la FIG. 7 es un diagrama de flujo de una etapa de una realización de un método de manejo de señal de acuerdo con la presente invención.

Descripción detallada

En todos los dibujos, los mismos números de referencia se utilizan para elementos similares o correspondientes.

Una realización de un sistema de códec general para señales de audio se ilustra esquemáticamente en la Fig. 1. Una fuente de audio 10 produce una señal de audio 15. La señal de audio 15 es manejada en un codificador 20, el cual produce un flujo binario 25 que comprende datos que representan a la señal de audio 15. El flujo binario 25 puede ser transmitido, como por ejemplo, en el caso de la comunicación multimedia, mediante una disposición de transmisión y/o de almacenamiento 30. La disposición de transmisión y/o de almacenamiento 30 opcionalmente también puede comprender alguna capacidad de almacenamiento. El flujo binario 25 puede también ser almacenado sólo en la disposición de transmisión y/o de almacenamiento 30, sólo introduciendo un retardo de tiempo en la utilización del flujo binario. La disposición de transmisión y/o de almacenamiento 30 es así una disposición que introduce al menos uno de un reposicionamiento en el espacio o retardo en el tiempo del flujo binario 25. Cuando se está utilizando, el flujo binario 25 es manejado en un descodificador 40, que produce una salida de audio 35 de los datos comprendidos en el flujo binario. Típicamente, la salida de audio 35 debería recoger la señal de audio 15 original lo mejor posible bajo ciertas restricciones.

En muchas aplicaciones en tiempo real, el retardo en el tiempo entre la producción de la señal de audio 15 original y la salida de audio 35 producida típicamente no está permitido que exceda un cierto tiempo. Si los recursos de transmisión en el mismo tiempo son limitados, la tasa de bits disponible es también típicamente baja. Con el fin de utilizar la tasa de bits disponible de una mejor manera posible, se ha desarrollado la codificación de audio perceptual. La codificación de audio perceptual, por lo tanto, se ha convertido actualmente en una parte importante para muchos servicios multimedia. El principio básico es convertir la señal de audio en coeficientes espectrales en un dominio de la frecuencia y utilizar un modelo perceptual para determinar un enmascaramiento dependiente de la frecuencia y del tiempo de los coeficientes espectrales.

La Fig. 2 ilustra una realización de un codificador de audio 20 de acuerdo con la presente invención. En esta realización particular, el codificador de audio 20 perceptual es un codificador espectral basado en un transformador perceptual o en un banco de filtros perceptual. Se recibe una fuente de audio 15, que comprende tramas de señales de audio $x[n]$.

En un codificador espectral típico, un convertidor 21 está dispuesto para convertir la señal de audio 15 en el dominio del tiempo en un conjunto 24 de coeficientes espectrales $X_b[n]$ de un dominio de frecuencia. En un codificador de transformada típico, la conversión puede, por ejemplo, ser llevada a cabo mediante una Transformada de Fourier

Discreta (DFT – Discrete Fourier Transform, en inglés), una Transformada del Coseno Discreta (DCT – Discrete Cosine Transform, en inglés) o una Transformada del Coseno Discreta Modificada (MDCT – Modified Discrete Cosine Transform, en inglés). El convertidor 21 puede por ello típicamente estar constituido por un transformador espectral. Los detalles de la actual transformación no son de particular importancia para las ideas básicas de la presente invención y por lo tanto no se explicarán más.

El conjunto 24 de coeficientes espectrales, es decir, se proporciona una representación de frecuencia de la señal de audio de entrada a una sección de cuantificación y codificación 28, en la que los coeficientes espectrales son cuantificados y codificados. Típicamente, la cuantificación está operando para concentrar los bits disponibles en los coeficientes más energéticos y perceptualmente más relevantes. Esto puede ser llevado a cabo utilizando, por ejemplo, diferentes tipos de umbrales de enmascaramiento o reducciones del ancho de banda. El resultado será típicamente “agujeros espectrales” de coeficientes no cuantificados en el espectro de la frecuencia. En otras palabras, algunos de los coeficientes se dejan fuera a propósito, puesto que son perceptualmente menos importantes, para no ocupar recursos de transmisión que se necesitan más para otros propósitos. Tales agujeros espectrales pueden entonces mediante diferentes estrategias de reconstrucción ser corregidos o reconstruidos en el lado del descodificador. Típicamente, aparecen agujeros espectrales de dos clases. La primera clase comprende agujeros espectrales, solos o unos pocos vecinos que ocurren en diferentes lugares en la región de frecuencia más baja. El segundo tipo es un grupo más o menos continuo de agujeros espectrales en el extremo de la alta frecuencia del espectro.

De acuerdo con la presente invención, resulta favorable tratar estas dos clases distintas de agujeros espectrales de diferentes maneras, con el fin de lograr un rellenado del espectro lo más eficiente posible. Un parámetro que se debe determinar es entonces una frecuencia, a la cual se encuentran los diferentes planteamientos de rellenado, una llamada frecuencia de transición. Puesto que la distribución de agujeros espectrales difiere entre las diferentes clases de señales de audio, la elección óptima de frecuencia de transición también difiere. De acuerdo con la presente invención, la frecuencia de transición está adaptada a un contenido espectral de la señal de audio. Típicamente, la frecuencia de transición está adaptada a un contenido espectral de una trama presente de la señal de audio, no obstante, la frecuencia de transición puede también depender del contenido espectral de las tramas previas de la señal de audio, y si no hay ningún requisito de retardo serio, la frecuencia de transición puede también depender del contenido espectral de tramas futuras de la señal de audio. Esta adaptación puede ser llevada a cabo en el lado del codificador mediante unos circuitos de determinación de transición 60, típicamente integrados con la sección de cuantificación y codificación 28. No obstante, en realizaciones alternativas, los circuitos de determinación de transición 60 pueden ser proporcionados como una sección de operación separadamente, por lo que sólo un parámetro que representa la frecuencia de transición es proporcionado a las diferentes funcionalidades del codificador 20. La frecuencia de transición puede ser utilizada en el lado del codificador por ejemplo proporcionando una codificación de envoltura para los intervalos de frecuencia en los diferentes lados de la frecuencia de transición.

La sección de cuantificación y de codificación 28 está también dispuesta para empaquetar los coeficientes espectrales codificados junto con la información lateral adicional en una corriente de bits de acuerdo con el estándar de transmisión o de almacenamiento que va a ser utilizado. Un flujo binario 25 que tiene datos que representan al conjunto de coeficientes espectrales es por ello extraído de la sección de cuantificación y codificación 28. Puesto que la frecuencia de transición puede obtenerse directamente a partir del contenido espectral de la señal de audio, la misma derivación puede ser llevada a cabo en ambos lados de la interfaz de transmisión, es decir, tanto en el codificador como en el descodificador. Esto significa que el propio valor de la frecuencia de transición no necesariamente tiene que ser transmitido entre la información lateral adicional. No obstante, por supuesto es también posible hacer eso si hay una capacidad de tasa de bits disponible.

En una realización particular, se utiliza una transformada MDCT. Después de que se ha realizado la ponderación mediante un modelo psico-acústico, los coeficientes de la MDCT son cuantificados utilizando un vector de cuantificación. En el vector de cuantificación, VQ (Quantization Vector, en inglés), los coeficientes espectrales son divididos en pequeños grupos. Cada grupo de coeficientes puede verse como un único vector, y cada vector es cuantificado individualmente.

Por ejemplo, debido a las elevadas restricciones en la tasa de bits, el cuantificador puede centrar los bits disponibles en los grupos más energéticos y perceptualmente más relevantes, resultando en que algunos grupos son puestos a cero. Estos grupos forman agujeros espectrales en el espectro cuantificado. Esto se ilustra en la FIGURA 3. En la presente realización, los grupos 70 comprenden el mismo número de coeficientes espectrales 71, en este caso cuatro. No obstante, en realizaciones alternativas pueden ser también posibles grupos que tienen diferente número de coeficientes espectrales. En una realización particular, todos los grupos comprenden sólo un coeficiente espectral cada uno, es decir, el grupo es el mismo que el propio coeficiente espectral. Los grupos cuantificados 72 se ilustran en la figura mediante rectángulos sin relleno, mientras que los grupos puestos a cero 73 se ilustran como rectángulos negros. Son típicamente sólo los grupos cuantificados 72 los que son transmitidos a cualquier usuario final.

Los grupos 70 de coeficientes se dividen a su vez en diferentes bandas de frecuencia 74. Esta división es preferiblemente llevada a cabo de acuerdo con algún criterio psico-acústico. Los grupos que tienen esencialmente

similares propiedades psico-acústicas pueden por ello ser tratados colectivamente. El número de miembros de cada banda de frecuencia 74, es decir, el número de grupos 70 asociados con las bandas de frecuencia 74 puede por lo tanto diferir. Si grandes porciones de frecuencia tienen similares propiedades, una banda de frecuencia que cubre estas frecuencias puede tener un gran intervalo de frecuencia. Si las propiedades psico-acústicas cambian rápidamente con las frecuencias, esto por el contrario requiere bandas de frecuencia de un pequeño intervalo de frecuencia. Las rutinas para el rellenado del espectro pueden preferiblemente depender de la banda de frecuencia que va a ser rellenada como se explica con detalle a continuación.

En la etapa de descodificación, la operación inversa es básicamente lograda. En la Fig. 4, se ilustra una realización de un descodificador 40 de audio de acuerdo con la presente invención. Se recibe un flujo binario 25, que tiene propiedades causadas por el codificador descrito anteriormente en esta memoria. La descuantificación y la descodificación del flujo binario 25 recibido, por ejemplo, una corriente de bits es llevada a cabo en un descodificador 41 de coeficiente espectral. El descodificador 41 de coeficiente espectral está dispuesto para descodificar coeficientes espectrales recuperados del flujo binario en los coeficientes espectrales descodificados $X^Q[n]$ de un conjunto inicial de coeficientes espectrales 42, posiblemente agrupados en grupos de frecuencias $X_b^Q[n]$. El conjunto inicial de coeficientes espectrales 42 preferiblemente agrupa al conjunto de coeficientes espectrales proporcionados por el convertidor del lado del codificador, posiblemente tras un post-procesamiento tal como, por ejemplo, enmascaramiento de umbrales o reducciones del ancho de banda.

Como se explica también a continuación, la aplicación de enmascarar umbrales o de reducciones del ancho de banda en el codificador típicamente resulta en que el conjunto de coeficientes espectrales 42 está incompleto en el sentido de que típicamente comprende los llamados “agujeros espectrales”. Los “agujeros espectrales” corresponden a coeficientes espectrales que no son recibidos en el flujo binario. En otras palabras, los agujeros espectrales son coeficientes espectrales no definidos o no codificados $X^Q[n]$ o coeficientes espectrales automáticamente puestos a un valor predeterminado, típicamente cero, por el descodificador espectral 41. Para evitar aberraciones audibles, estos coeficientes tienen que ser reemplazados por estimaciones (rellenadas) en el descodificador.

Los agujeros espectrales a menudo vienen en dos tipos. Los agujeros espectrales pequeños están típicamente en las bajas frecuencias, y uno o unos pocos agujeros espectrales grandes típicamente ocurren en las altas frecuencias.

Para minimizar las aberraciones en la señal de audio descodificada, el descodificador “rellena” el espectro reemplazando los agujeros espectrales en el espectro con estimaciones de los coeficientes. Estas estimaciones pueden basarse en información lateral transmitida por el descodificador y/o pueden depender de la propia señal. Ejemplos de tal información lateral útil podrían ser la envoltura de potencia del espectro y la tonalidad, es decir, medida de planicidad espectral, de los coeficientes faltantes.

Pueden utilizarse dos métodos diferentes para rellenar distintas clases de agujeros espectrales. El “rellenado con ruido” funciona bien para agujeros espectrales en las frecuencias bajas, mientras que la “extensión del ancho de banda” es más adecuada a frecuencias altas. La presente invención describe un método para decidir dónde deben utilizarse el rellenado con ruido y la extensión del ancho de banda, respectivamente.

La presente invención se basa en la definición de una frecuencia de transición entre las partes relevantes baja y alta del espectro. Basándose en esta información, un algoritmo de codificación típico que se basa en un procedimiento de “rellenado con ruido” de alta calidad será capaz de reducir las aberraciones de codificación que ocurren para bajas tasas y también de regenerar una señal de audio de ancho de banda completo incluso a bajas tasas y con un esquema de baja complejidad basado en la “extensión del ancho de banda”. Esto se explicará con más detalle a continuación.

El conjunto inicial de coeficientes espectrales 42 del descodificador de coeficiente espectral 41, que comprende típicamente una cierta cantidad de agujeros espectrales, se proporciona a unos circuitos de determinación de transición 60. Los circuitos de determinación de transición 60 están dispuestos para determinar una frecuencia de transición f_t.

El conjunto inicial de coeficientes espectrales 42 del descodificador de coeficiente espectral 41 es proporcionado también a un rellenador de espectro 43. El rellenador de espectro 43 está dispuesto para llevar a cabo un rellenado del espectro en el conjunto inicial de coeficientes espectrales 42, creando un conjunto 44 completo de coeficientes espectrales reconstruidos $X_b'[n]$. El conjunto 44 de coeficientes espectrales reconstruidos tiene típicamente todos los coeficientes espectrales dentro de un cierto intervalo de frecuencia definido.

El rellenador de espectro 43 a su vez comprende un filtro de ruido 50. El filtro de ruido 50 está dispuesto para proporcionar un proceso para el rellenado con ruido de los agujeros espectrales, preferiblemente en la región de la frecuencia baja, es decir, por debajo de la frecuencia de transición f_t. Un valor es por ello asignado a los coeficientes espectrales en el conjunto inicial de coeficientes espectrales por debajo de la frecuencia de transición que “falta”, como resultado de no estar incluida en la corriente de bits codificados recibida. En este sentido, una salida 65 de los

circuitos de determinación de transición 60 está conectada al filtro de ruido 50, proporcionando información asociada con la frecuencia de transición f_t.

El rellenador de espectro 43 comprende también un extensor del ancho de banda 55, dispuesto para extender el ancho de banda del conjunto inicial de coeficientes espectrales por encima de la frecuencia de transición con el fin de producir el conjunto 44 de coeficientes espectrales reconstruidos. Por lo tanto, la salida 65 de los circuitos de determinación de transición 60 está también conectada al extensor del ancho de banda 55.

Como se ha mencionado anteriormente, el resultado del rellenador de espectro 43 es un conjunto 44 completo de coeficientes espectrales reconstruidos $X_b'[n]$, que tiene todos los coeficientes espectrales dentro de un cierto intervalo de frecuencia definido.

El conjunto 44 de coeficientes espectrales reconstruidos es proporcionado a un convertidor 45 conectado al rellenador de espectro 43. El convertidor 45 está dispuesto para convertir el conjunto 44 de coeficientes espectrales de un dominio de la frecuencia en una señal de audio 46 de un dominio del tiempo. El convertidor 45 está en la presente realización basado en un transformador perceptual, correspondiente a la técnica de transformada utilizada en el codificador 20 (Fig. 2). En una realización particular, la señal es proporcionada de nuevo al dominio del tiempo con una transformada inversa, por ejemplo, MDCT Inversa – IMDCT o DFT inversa – IDFT, etc. En otras realizaciones puede utilizarse un banco de filtros inverso. Como en el lado del codificador, la técnica de tal convertidor 45 es conocida de la técnica anterior, y no se explicará de nuevo. Una señal de audio reconstruida perceptualmente final 34 $x'[n]$ es proporcionada en una salida 35 para la señal de audio, posiblemente con otras etapas de tratamiento.

El códec debe decidir en qué bandas de frecuencia utilizar el rellenado con ruido y en qué bandas de frecuencia utilizar la extensión del ancho de banda. El rellenado con ruido proporciona el mejor resultado cuando la mayoría de los grupos de la banda de frecuencia para ser rellenados están cuantificados, y hay sólo agujeros espectrales menores en la banda. La extensión del ancho de banda es preferible cuando una gran parte de la señal en las frecuencias altas se deja sin cuantificar.

Un método básico sería establecer una frecuencia de transición fija entre el rellenado con ruido y la extensión del ancho de banda. Los agujeros espectrales en las bandas o grupos de frecuencia por debajo de tal frecuencia son rellenados mediante rellenado con ruido y los agujeros espectrales en grupos o bandas de frecuencia por encima de esa frecuencia son rellenados mediante extensión del ancho de banda.

Un problema con este planteamiento es, no obstante, que la frecuencia de transición óptima no es la misma para todas las señales de audio. Algunas señales tienen la mayoría de la energía concentrada en las frecuencias bajas y una gran parte de la señal podría ser sometida a extensión del ancho de banda. Otras señales tienen su energía más uniformemente distribuida sobre el espectro y estas señales pueden beneficiarse de utilizar sólo rellenado con ruido.

De acuerdo con la presente invención la frecuencia de transición es adaptativamente dependiente de una distribución de agujeros espectrales en el citado conjunto de coeficientes espectrales. Una rutina para encontrar una frecuencia de transición adecuada podría ser pasar por todas las bandas de frecuencia, empezando en la más alta (BN) hasta 1. Si no hay coeficientes cuantificados en la banda actual, será rellenada mediante extensión del ancho de banda. Si hay coeficientes cuantificados en la banda, los agujeros de esta banda así como las siguientes bandas son rellenados utilizando rellenado con ruido. Así, una frecuencia de transición es puesta en el límite superior de la primera banda de frecuencia vista desde el lado de la alta frecuencia que tiene un coeficiente cuantificado en ella. Esto se ilustra en la Fig. 5A. Los agujeros espectrales 77 en la banda N, es decir, por encima de la frecuencia de transición f_t son así rellenados con planteamientos de extensión del ancho de banda. Los agujeros espectrales 76 por debajo de la frecuencia de transición f_t son, por el contrario, rellenados mediante rellenado con ruido.

Un ejemplo alternativo se ilustra en la Fig. 5B. Aquí la definición de la frecuencia de transición se basa directamente en los grupos 70, despreciando la división de banda de frecuencia. Aquí, la extensión del ancho de banda se utiliza para todos los grupos desde las frecuencias más altas hasta el grupo inmediatamente por encima del primer grupo cuantificado 78. Los agujeros espectrales 76 por debajo de la frecuencia de transición f_t son, por el contrario, rellenados mediante rellenado con ruido.

Estos métodos son más adaptativos a la señal de audio y al cuantificador, es decir, al esquema de codificación, pero pueden experimentar problemas menores cuando la señal es cuantificada, por ejemplo de acuerdo con la Fig. 5C. Aquí, una gran parte de las altas frecuencias de la señal es puesta a cero, y la extensión del ancho de banda preferiblemente se utilizaría desde la banda B9 a la B12. No obstante, puesto que existe un solo grupo cuantificado 79 codificado en la banda de frecuencia B11, la extensión del ancho de banda estará completamente deshabilitada por debajo de este grupo cuantificado 79 y el rellenado con ruido será utilizado en todas las bandas hasta este grupo 79.

Para evitar también este problema, se propone también otro ejemplo alternativo, en el que la frecuencia de transición f_t es seleccionada dependiendo de una proporción de agujeros espectrales en las bandas de frecuencia. Como en las realizaciones previas, el códec pasa por todas las bandas de frecuencia, empezando en la más alta hacia el 1. Para cada banda de frecuencia, el número de coeficientes espectrales o grupos codificados se cuenta. Si el número de coeficientes cuantificados o grupos dividido por el número total de coeficientes espectrales o grupos, es decir, la proporción de coeficientes espectrales codificados, de la banda de frecuencia excede un cierto umbral, los agujeros espectrales de esa banda de frecuencia y las siguientes bandas de frecuencia son rellenados con rellenado con ruido. Si no, se utiliza la extensión del ancho de banda. Análogamente, se puede monitorizar la proporción de agujeros espectrales en las bandas de frecuencia. En otras palabras, debe encontrarse una banda de frecuencia de transición, la cual es una banda de la frecuencia más alta en la cual una proporción de agujeros espectrales es menor que un primer umbral.

Existen también criterios alternativos para seleccionar la banda de frecuencia de transición. Una posibilidad es dejar que el propio umbral dependa de la frecuencia. De tal manera, una cierta proporción de agujeros espectrales puede ser aceptada en las partes de alta frecuencia para utilizar todavía técnicas de expansión del ancho de banda, pero no en las partes de baja frecuencia. Cualquier experto en la materia se da cuenta de que los detalles en la selección de criterios apropiados pueden ser variados de muchas maneras, por ejemplo, siendo dependientes de otras propiedades relativas a la señal o de otra información lateral.

En un ejemplo, la frecuencia de transición se hace dependiente de, y preferiblemente igual a, un límite de frecuencia superior de la banda de frecuencia de transición. No obstante, existen también varias alternativas. Una alternativa es buscar el coeficiente espectral o grupo codificado en frecuencia más alto y establecer la frecuencia de transición en el lado de alta frecuencia de ese grupo.

El algoritmo del ejemplo descrito anteriormente puede ser también descrito con el siguiente pseudo código:

```
For currentBand = N to 1
   ratio = numCodedCoeffInBand(currentBand) /
numCoeffInBand(currentBand)
   If ratio > threshold
         Transition is between currentBand and
currentBand + 1
         Return
   End if
Next
Transition is at the start of band 1
```

Se prefiere que la frecuencia de transición no varíe demasiado entre tramas consecutivas. Cambios demasiado grandes pueden ser percibidos como molestos. Por lo tanto, en una realización de ejemplo, la frecuencia de transición es también dependiente de una frecuencia de transición usada previamente. Sería posible, por ejemplo, prohibir que la frecuencia de transición cambie más de una cantidad predeterminada absoluta o relativa entre dos tramas consecutivas. Alternativamente, una frecuencia de transición provisional podría ser introducida como un valor en un filtro junto con frecuencias de transición previas, proporcionando una frecuencia de transición modificada que tiene un comportamiento de cambio más suave. La frecuencia de transición dependerá entonces de más de una frecuencia de transición previa.

Estas rutinas son típicamente llevadas a cabo en los circuitos de determinación de transición, es decir, preferiblemente en la sección de cuantificación y de codificación del codificador y del descodificador, respectivamente.

La Fig. 6 es un diagrama de flujo que ilustra etapas de una realización de un método de acuerdo con la presente invención. Un método para la recuperación del espectro en la descodificación espectral de una señal de audio se inicia en la etapa 200. En la etapa 210, se obtiene un conjunto inicial de coeficientes espectrales que representan a la señal de audio. En la etapa 212, se determina una frecuencia de transición. La frecuencia de transición es adaptada al contenido espectral de la señal de audio. El rellenado con ruido de agujeros espectrales en el conjunto inicial de coeficientes espectrales por debajo de la frecuencia de transición se lleva a cabo en la etapa 214 y la extensión del ancho de banda del conjunto inicial de coeficientes espectrales por encima de la frecuencia de transición es llevado a cabo en la etapa 216. El proceso finaliza en la etapa 249.

Análogamente, la Fig. 7 es un diagrama de flujo que ilustra una etapa de una realización de otro método de acuerdo con la presente invención. Un método para su uso en la codificación espectral de una señal de audio se inicia en la etapa 200. En la etapa 212, se determina una frecuencia de transición. La frecuencia de transición para un conjunto inicial de coeficientes espectrales que representan a la señal de audio es adaptada a un contenido espectral de la señal de audio. La frecuencia de transición que define una frontera entre un intervalo de frecuencia, prevista para ser un sujeto para el rellenado con ruido de agujeros espectrales, y un intervalo de frecuencia, previsto para ser un sujeto para la extensión del ancho de banda.

La presente invención adquiere un número de ventajas mediante la definición adaptativa de la frecuencia de transición de acuerdo con el esquema de codificación utilizado. La frecuencia de transición adaptada permite el uso eficiente de un rellenado de espectro combinado utilizando tanto rellenado con ruido como extensión del ancho de banda. Cualquier códec de conversación y/o de audio utilizando este método es capaz de proporcionar una señal de audio de alta calidad y ancho de banda completo con menos aberraciones molestas. El método es flexible en el sentido de que puede ser combinado con cualquier clase de representación de la frecuencia (DCT, MDCT, etc.) o bancos de filtros, es decir, con cualquier códec (perceptual, paramétrico, etc.).

Las realizaciones descritas anteriormente deben ser entendidas como unos pocos ejemplos ilustrativos de la presente invención. Los expertos en la materia deben entender que pueden realizarse varias modificaciones, combinaciones y cambios sin separarse del alcance de la presente invención. En particular, distintas soluciones parciales en las diferentes realizaciones pueden ser combinadas en otras configuraciones, donde sea técnicamente posible. El alcance de la presente invención está, no obstante, definido por las reivindicaciones adjuntas.

REFERENCIAS

[1] 3GPP TS 26.404 V6.0.0 (2004-09), “Enhanced aacPlus general audio codec – encoder SBR part (Versión 6)”, 2004.

[2] J. D. Johnston, “Estimation of Perceptual Entropy Using Noise Masking Criteria”, Proc. ICASSP, pp. 2524-2527, Mayo de 1988.

REIVINDICACIONES

1. Método para la recuperación de espectro para la descodificación espectral de una señal de audio, que comprende las etapas de:

obtener (210) un conjunto (42) inicial de coeficientes espectrales que representan a la citada señal de audio;

dividir los citados coeficientes espectrales del citado conjunto (42) inicial de coeficientes espectrales en una pluralidad de bandas de frecuencia (74);

determinar (212) una frecuencia de transición (f_t);

rellenado con ruido (214) de agujeros espectrales en el citado conjunto (42) inicial de coeficientes espectrales por debajo de la citada frecuencia de transición (f_t); y

extender el ancho de banda (216) del citado conjunto (42) inicial de coeficientes espectrales por encima de la citada frecuencia de transición (f_t);

donde

la citada frecuencia de transición (f_t) es ajustada al límite superior de la primera banda de frecuencia visto desde el lado de la frecuencia alta que tiene en él un coeficiente cuantificado.

2. Método de acuerdo con la reivindicación 1, en el que las citadas bandas de frecuencia (74) tienen un ancho de frecuencia constante.

3. El método de acuerdo con la reivindicación 1, en el que al menos dos de las citadas bandas de frecuencia (74) tienen diferentes anchos de frecuencia.

4. Método de acuerdo con cualquier reivindicación precedente, en el que la citada etapa de ajustar la citada frecuencia de transición (f_t) depende también de una frecuencia de transición previamente utilizada

5. Método de acuerdo con la reivindicación 4, en el que la citada etapa de ajustar la citada frecuencia de transición (f_t) depende también de más de una frecuencia de transición previamente utilizada.

6. Método de acuerdo con la reivindicación 4 ó 5, en el que la citada frecuencia de transición (f_t) está prohibido que cambie más de una cantidad absoluta o relativa predeterminada entre dos tramas consecutivas.

7. Método para su uso para la codificación espectral de una señal de audio, que comprende:

dividir los coeficientes espectrales de un conjunto inicial (24; 42) de coeficientes espectrales que representan a la citada señal de audio en una pluralidad de bandas de frecuencia;

determinar (212) una frecuencia de transición (f_t) para el conjunto (24; 42) inicial de coeficientes espectrales que representan a la citada señal de audio;

definiendo la citada frecuencia de transición (f_t) una frontera entre un intervalo de frecuencia, prevista para ser un sujeto para el rellenado con ruido de agujeros espectrales, y un intervalo de frecuencia, previsto para ser un sujeto para la extensión del ancho de banda;

estando la citada frecuencia de transición (f_t) ajustada al límite superior de la primera banda de frecuencia vista desde el lado de la frecuencia superior que tiene un coeficiente cuantificado en ella.

8. Descodificador (40) para la descodificación espectral de una señal de audio, que comprende:

una entrada para obtener un conjunto (42) inicial de coeficientes espectrales que representan a la citada señal de audio;

unos circuitos de determinación de transición (60) dispuestos para determinar una frecuencia de transición (f_t);

un rellenador con ruido (50) para el rellenado con ruido de agujeros espectrales en el citado conjunto (42) inicial de coeficientes espectrales por debajo de la citada frecuencia de transición (f_t); y

un extensor de ancho de banda (55) dispuesto para extender el ancho de banda del citado conjunto (42) inicial de coeficientes espectrales por encima de la citada frecuencia de transición (f_t);

estando los citados circuitos de determinación de transición (60) dispuestos

para dividir los citados coeficientes espectrales del citado conjunto inicial de coeficientes espectrales en una pluralidad de bandas de frecuencia (74), y

para determinar la citada frecuencia de transición (f_t) como el límite superior de la primera banda de frecuencia vista desde el lado de la frecuencia superior que tiene un coeficiente cuantificado en ella.

9. Codificador (20) para la codificación espectral de una señal de audio, que comprende:

circuitos de determinación de transición (60) dispuestos para determinar una frecuencia de transición (f_t) para un conjunto (24) inicial de coeficientes espectrales que representan a la citada señal de audio;

definiendo la citada frecuencia de transición (f_t) una frontera entre un intervalo de frecuencia, previsto para ser un sujeto para el rellenado con ruido de agujeros espectrales, y un intervalo de frecuencia, previsto para ser un sujeto para la extensión del ancho de banda;

estando los citados circuitos de determinación de transición (60) dispuestos para

dividir los citados coeficientes espectrales del citado conjunto inicial de coeficientes espectrales en una pluralidad de bandas de frecuencia (74), y para

determinar la citada frecuencia de transición (f_t) como el límite superior de la primera banda de frecuencia vista desde el lado de la frecuencia superior que tiene un coeficiente cuantificado en ella.

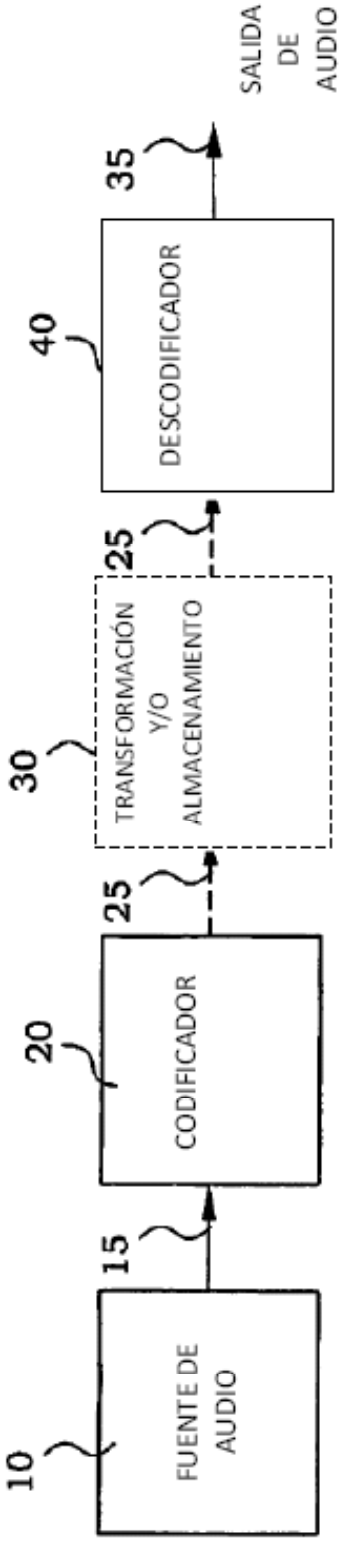
10
FUENTE DE AUDIO
15
20
CODIFICADOR
25
30
TRANSFORMACIÓN Y/O ALMACENAMIENTO
25
40
DESCODIFICADOR
35
SALIDA DE AUDIO

Fig. 1

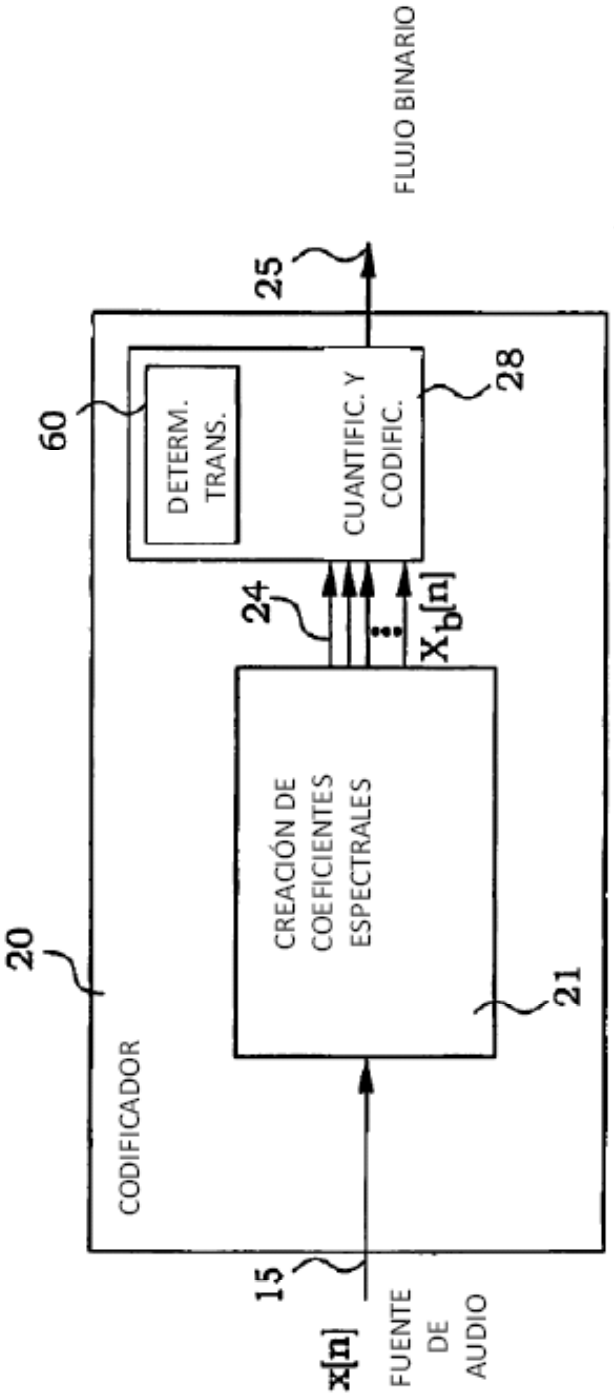
20
CODIFICADOR
15
x[n]
FUENTE DE AUDIO
CREACIÓN DE COEFICIENTES ESPECTRALES
21
24
$X_b[n]$
60
DETERM. TRANS.
CUANTIFIC. Y CODIFIC.
28
25
FLUJO BINARIO

Fig. 2

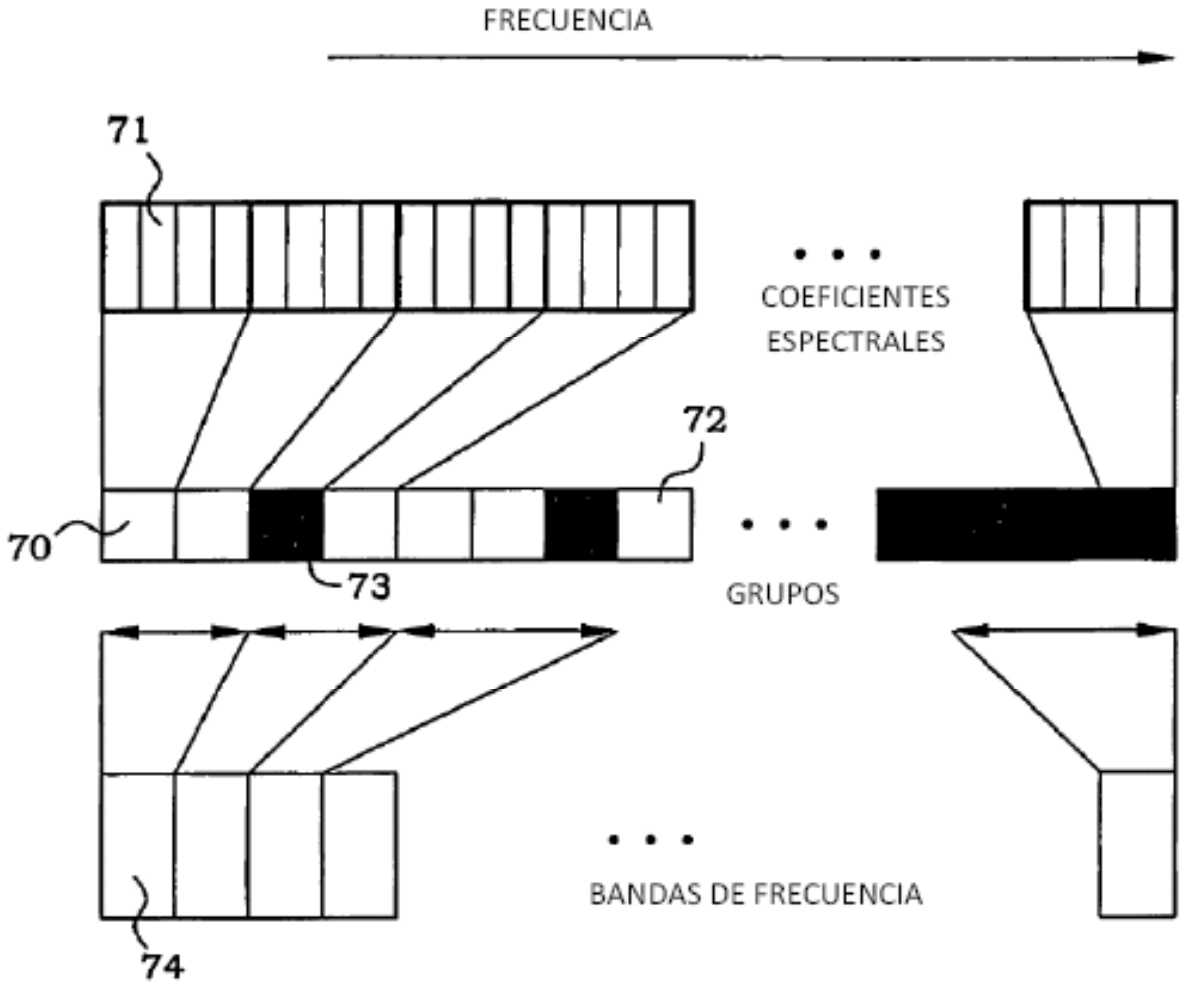
FRECUENCIA
71
COEFICIENTES
ESPECTRALES
72
70
73
GRUPOS
BANDAS DE FRECUENCIA
74

Fig. 3

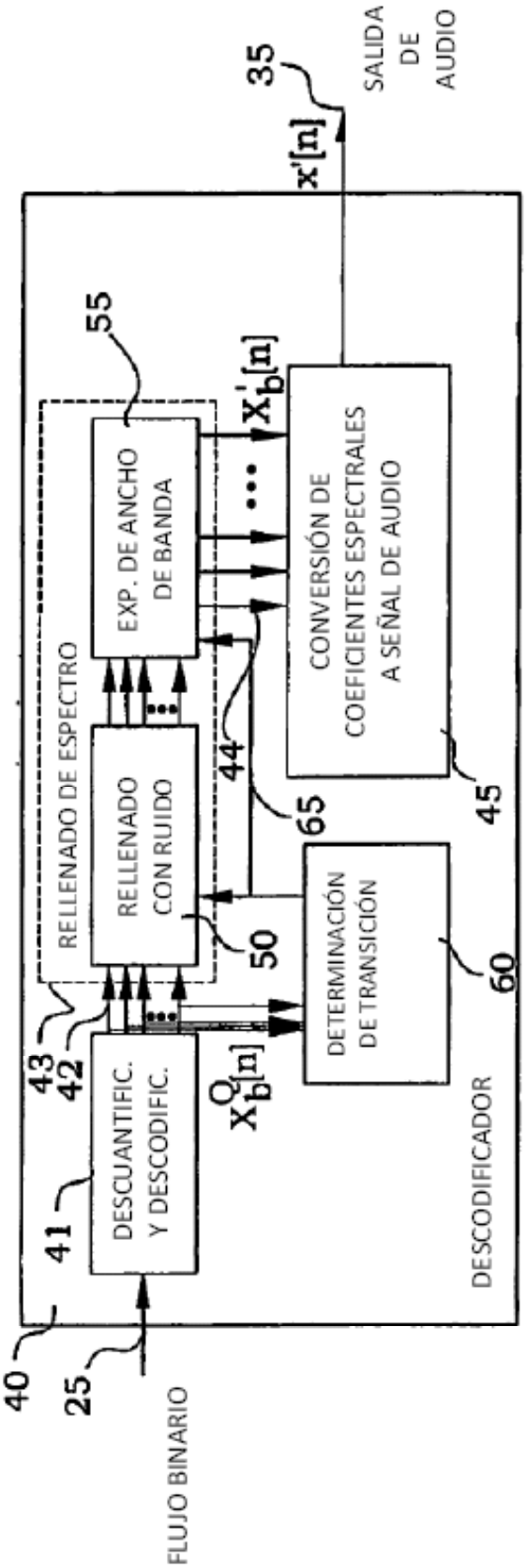
35
SALIDA DE AUDIO
x'[n]
55
EXP. DE ANCHO DE BANDA
RELLENADO DE ESPECTRO
RELLENADO CON RUIDO
x'b[n]
CONVERSIÓN DE COEFICIENTES ESPECTRALES A SEÑAL DE AUDIO
44
65
45
50
60
DETERMINACIÓN DE TRANSICIÓN
43
42
41
DESCUANTIFIC. Y DESCODIFIC.
Q x b[n]
DESCODIFICADOR
40
25
FLUJO BINARIO

Fig. 4

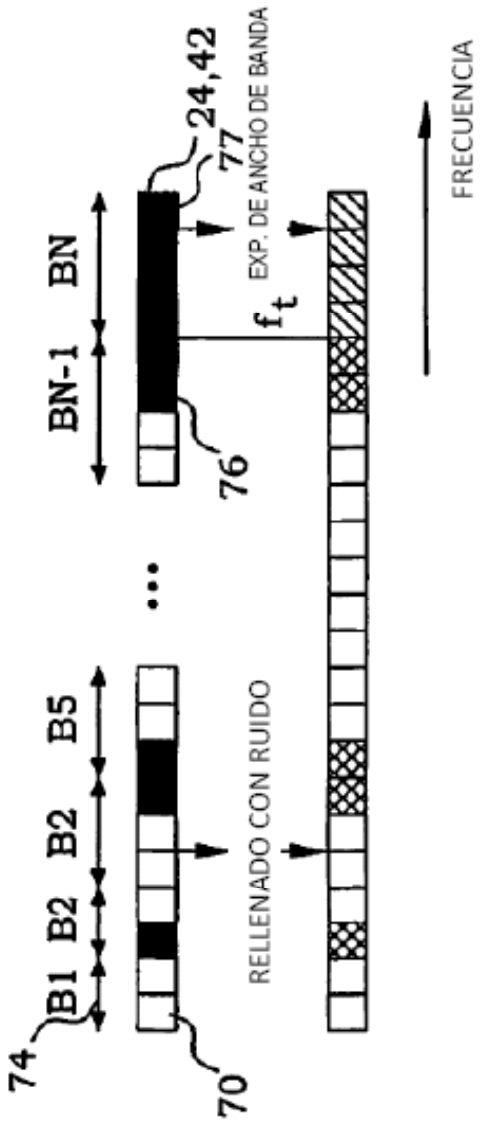
24,42
77
EXP. DE ANCHO DE BANDA
BN
BN-1
ft
76
B5
B2
B2
B1
74
70
RELLENADO CON RUIDO
FRECUENCIA

Fig. 5A

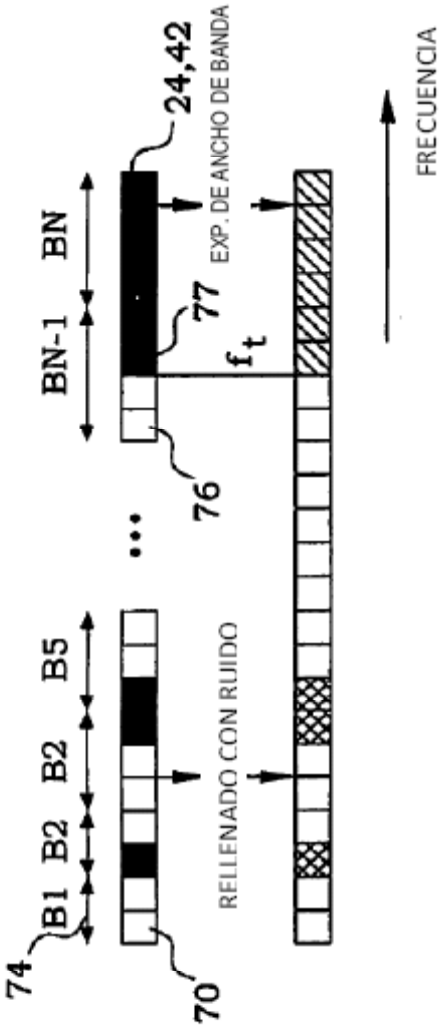
74
B1
B2
B2
B5
70
RELLENADO CON RUIDO
...
76
BN-1
BN
77
f_t
24,42
EXP. DE ANCHO DE BANDA
FRECUENCIA

Fig. 5B

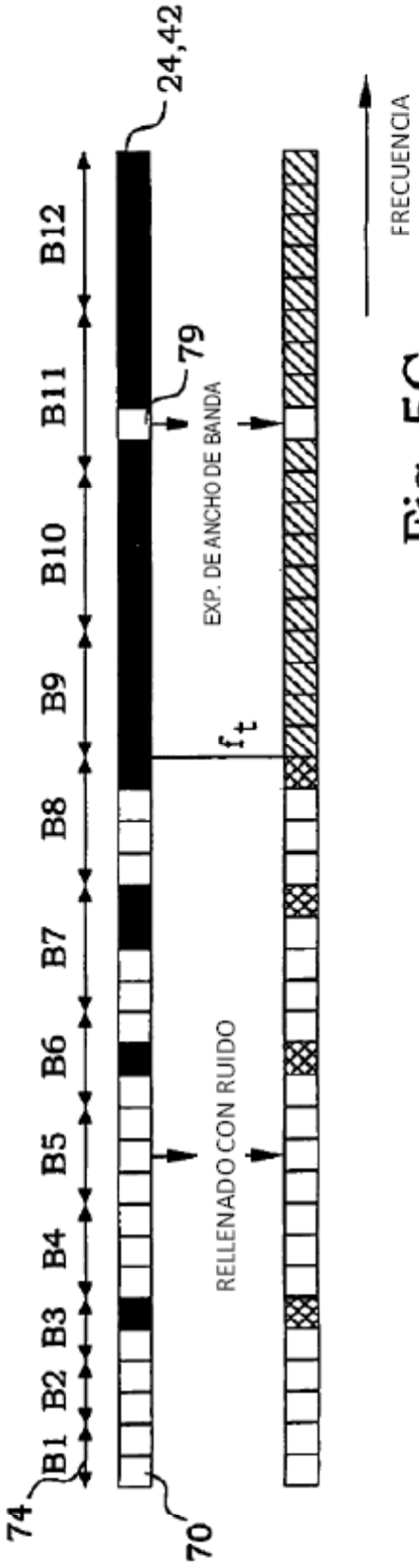
74
B1
B2
B3
B4
B5
B6
B7
B8
B9
B10
B11
B12
70
RELLENADO CON RUIDO
f_t
79
EXP. DE ANCHO DE BANDA
24,42
FRECUENCIA

Fig. 5C

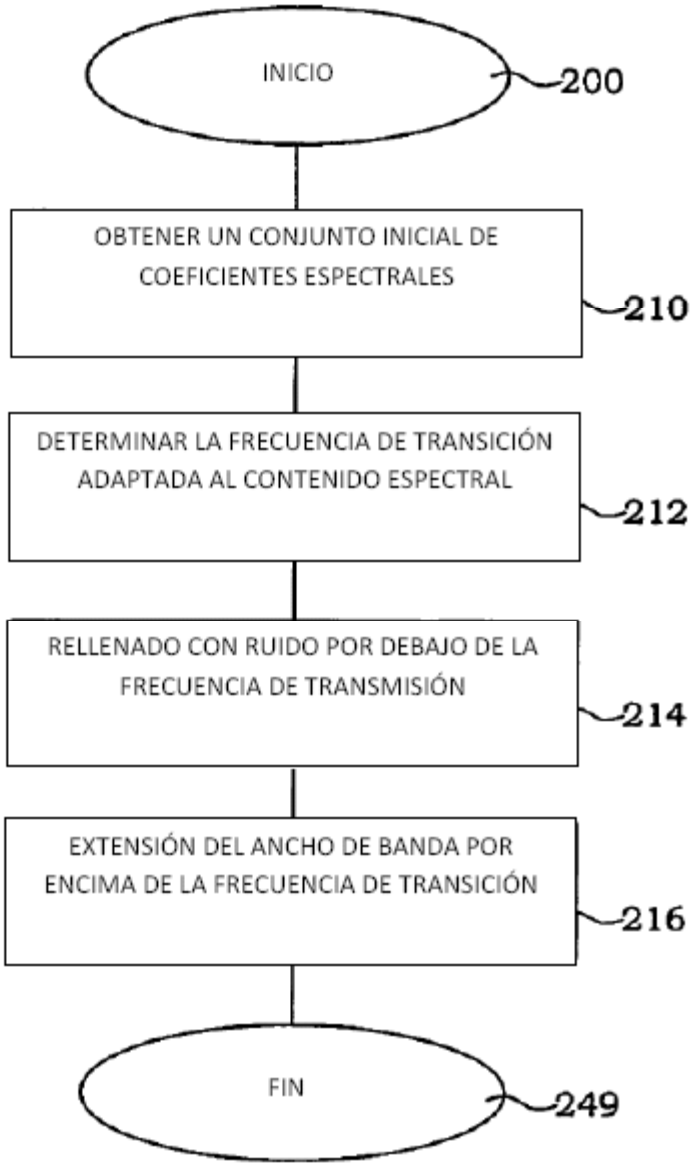
INICIO
200
OBTENER UN CONJUNTO INICIAL DE COEFICIENTES ESPECTRALES
210
DETERMINAR LA FRECUENCIA DE TRANSICIÓN ADAPTADA AL CONTENIDO ESPECTRAL
212
RELLENADO CON RUIDO POR DEBAJO DE LA FRECUENCIA DE TRANSMISIÓN
214
EXTENSIÓN DEL ANCHO DE BANDA POR ENCIMA DE LA FRECUENCIA DE TRANSICIÓN
216
FIN
249

Fig. 6

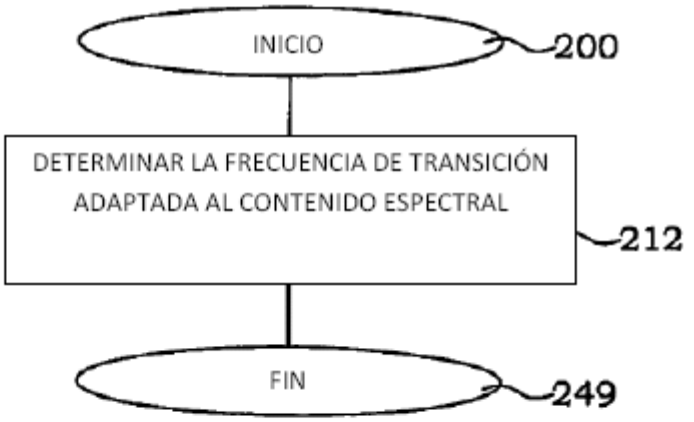
INICIO
200
DETERMINAR LA FRECUENCIA DE TRANSICIÓN ADAPTADA AL CONTENIDO ESPECTRAL
212
FIN
249

Fig. 7