(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 727 300**

(51) Int. CI.:

G10L 21/038 (2013.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **14.07.2011 E 17188331 (7)**

(97) Fecha y número de publicación de la concesión europea: **17.04.2019 EP 3291230**

(54) Título: **Procesamiento de señales de audio durante la reconstrucción de alta frecuencia**

(30) Prioridad:

19.07.2010 US 365518 P
27.09.2010 US 386725 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
15.10.2019

(73) Titular/es:

DOLBY INTERNATIONAL AB (100.0%)
Apollo Building, 3E, Herikerbergweg 1-35
1101 CN Amsterdam Zuidoost, NL

(72) Inventor/es:

KJOERLING, KRISTOFER

(74) Agente/Representante:

ELZABURU, S.L.P

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

Procesamiento de señales de audio durante la reconstrucción de alta frecuencia

Campo técnico

La solicitud se refiere a la HFR (regeneración/reconstrucción de alta frecuencia o «High Frequency Reconstruction/Regeneration») de señales de audio. En particular, la solicitud se refiere a un método y un sistema para realizar la HFR de señales de audio que tienen grandes variaciones en los niveles de energía a través del intervalo de baja frecuencia, que se utiliza para reconstruir las altas frecuencias de la señal de audio.

Antecedentes de la invención

Las tecnologías HFR, tales como la tecnología de replicación de la banda espectral (SBR, «Spectral Band Replication»), permiten mejorar significativamente la eficiencia de codificación de los códecs de audio perceptuales tradicionales. En combinación con la codificación de audio avanzada (AAC, «Advanced Audio Coding») MPEG-4, HFR forma un códec de audio muy eficaz, que ya está en uso en el sistema de radio por satélite XM y en Digital Radio Mondiale, y también estandarizado dentro del foro 3GPP, DVD y otros. La combinación de AAC y SBR se denomina aacPlus. Forma parte de la norma MPEG-4, en la que se denomina perfil AAC de alta eficiencia (HE-AAC, «High Efficiency AAC Profile»). En general, la tecnología de HFR se puede combinar con cualquier códec de audio perceptual de una manera compatible en ambos sentidos, ofreciendo de esta manera la posibilidad de mejorar los sistemas de difusión ya establecidos, como el MPEG capa 2 utilizado en el sistema Eureka DAB. Los métodos HFR también se pueden combinar con códecs de voz para permitir la transmisión de voz de banda ancha a tasas de bits ultrabajas.

La idea básica en la que se basa la HFR es la observación de que normalmente existe una fuerte correlación entre las características del intervalo de alta frecuencia de una señal y las características del intervalo de baja frecuencia de la misma señal. De esta manera, se puede lograr una buena aproximación para la representación del intervalo de alta frecuencia de entrada original de una señal mediante una transposición de la señal desde el intervalo de baja frecuencia hasta el intervalo de alta frecuencia.

Este concepto de transposición se estableció en el documento WO 98/57436 como método para recrear una banda de alta frecuencia a partir de una banda de menor frecuencia de una señal de audio. Al utilizar este concepto en la codificación de audio y/o en la codificación de voz se puede obtener un ahorro substancial en términos de la tasa de bits. A continuación, se hará referencia a la codificación de audio, pero cabe señalar que los métodos y sistemas descritos son igualmente aplicables a la codificación de voz y codificación unificada de voz y audio (USAC, «Unified Speech and Audio Coding»).El documento WO 02/41301 A1 describe un descodificador de audio con blanqueo espectral basado en una predicción lineal después de reconstrucción de alta frecuencia y antes de ajuste de envolvente.

La reconstrucción de alta frecuencia se puede realizar en el dominio del tiempo o en el dominio de la frecuencia, utilizando un banco de filtros o una transformación elegida. El proceso normalmente implica varias etapas, en donde las dos operaciones principales consisten en primer lugar en crear una señal de excitación de alta frecuencia y, posteriormente, conformar la señal de excitación de alta frecuencia para aproximarse a la envolvente espectral del espectro de alta frecuencia original. La etapa de creación de una señal de excitación de alta frecuencia se puede basar, por ejemplo, en la modulación de banda lateral única (SSB, «Single SideBand»), en donde una sinusoide con frecuencia ω se asigna a una sinusoide con frecuencia $\omega + \Delta\omega$, en donde $\Delta\omega$ es un desplazamiento de frecuencia fijo. En otras palabras, la señal de alta frecuencia puede ser generada a partir de la señal de baja frecuencia mediante una operación de «copia» de subbandas de baja frecuencia a subbandas de alta frecuencia. Un planteamiento adicional de cara a la creación de una señal de excitación de alta frecuencia puede involucrar la transposición armónica de subbandas de baja frecuencia. La transposición armónica de orden T se diseña típicamente para asignar una sinusoide con frecuencia ω de la señal de baja frecuencia a una sinusoide con frecuencia $T\omega$, con $T > 1$, de la señal de alta frecuencia.

La tecnología HFR se puede usar como parte de los sistemas de codificación de fuente, en los que una variedad de información de control destinada a guiar el proceso HFR desde un codificador hasta un descodificador junto con una representación de la señal de banda estrecha/baja frecuencia. En el caso de los sistemas en los que no se pueda transmitir ninguna señal de control adicional, el proceso se puede aplicar en el lado de descodificador con los datos de control adecuados estimados a partir de la información disponible en el lado de descodificador.

El ajuste de envolvente mencionado anteriormente de la señal de excitación de alta frecuencia tiene como objetivo lograr una forma espectral que se asemeje a la forma espectral de la banda alta original. Para ello, es necesario modificar la forma espectral de la señal de alta frecuencia. Dicho de manera diferente, el ajuste que se ha de aplicar a la banda alta es una función de la envolvente espectral existente y de la envolvente espectral objetivo deseada.

En el caso de los sistemas que operan en el dominio de la frecuencia, por ejemplo, sistemas HFR implementados en un banco de filtros seudo-QMF, los métodos de la técnica anterior son subóptimos a este respecto, ya que la creación de la señal de banda alta, mediante la combinación de varias contribuciones desde el intervalo de frecuencia de fuente,

introduce una envolvente espectral artificial en la banda alta que debe ser ajustada por envolvente. En otras palabras, la señal de alta frecuencia o de banda alta generada a partir de la señal de baja frecuencia durante el proceso HFR muestra típicamente una envolvente espectral artificial (que típicamente comprende discontinuidades espectrales). Esto plantea dificultades para el ajustador de envolvente espectral, ya que el ajustador no solo debe tener la capacidad de aplicar la envolvente espectral deseada con una resolución adecuada de tiempo y de frecuencia, sino que el ajustador también debe tener la capacidad de deshacer las características espectrales introducidas artificialmente por el generador de señal HFR. Esto plantea complejas restricciones de diseño del ajustador de envolvente. Como resultado, estas dificultades tienden a llevar a una pérdida percibida de la energía de alta frecuencia, y discontinuidades audibles en la forma espectral en la señal de banda alta, particularmente en el caso de señales de tipo voz. En otras palabras, los generadores de señales HFR convencionales tienden a introducir discontinuidades y variaciones de nivel en la señal de banda alta para señales que tienen grandes variaciones de nivel en el intervalo de banda baja, por ejemplo sibilancias. Cuando posteriormente el ajustador de envolvente se expone a esta señal de banda alta, el ajustador de envolvente no puede separar, de manera razonable y consistente, la discontinuidad recientemente introducida de cualquier característica espectral natural de la señal de banda baja.

El presente documento describe una solución al problema mencionado anteriormente, lo que tiene como resultado una mayor calidad de audio percibida. En particular, el presente documento describe una solución al problema de generar una señal de banda alta a partir de una señal de banda baja, en donde la envolvente espectral de la señal de banda alta se ajusta de manera efectiva con el fin de que se asemeje a la envolvente espectral original en la banda alta sin introducir artefactos no deseados.

Compendio de la invención

Según la invención, se proporciona un sistema como se presenta en la reivindicación 1, un método como se presenta en la reivindicación 2, un medio de almacenamiento como se presenta en la reivindicación 3, y un producto de programa informático como se presenta en la reivindicación 4.

El presente documento propone una etapa de corrección adicional como parte de la generación de señal de reconstrucción de alta frecuencia. Como resultado de la etapa de corrección adicional, se mejora la calidad de audio del componente de alta frecuencia o de la señal de banda alta. La etapa de corrección adicional se puede aplicar a todos los sistemas de codificación de fuente que utilicen técnicas de reconstrucción de alta frecuencia, así como a cualquier método o sistema de posprocesamiento de un solo extremo que tenga como objetivo recrear las altas frecuencias de una señal de audio.

De acuerdo con un aspecto, se describe un sistema configurado para generar una pluralidad de señales de subbanda de alta frecuencia que cubran un intervalo de alta frecuencia. El sistema puede estar configurado para generar la pluralidad de señales de subbanda de alta frecuencia a partir de una pluralidad de señales de subbanda de baja frecuencia. La pluralidad de señales de subbanda de baja frecuencia pueden ser señales de subbanda de una señal de audio de banda baja o banda estrecha, que se puede determinar utilizando una transformada o un banco de filtros de análisis. En particular, la pluralidad de señales de subbanda de baja frecuencia se puede determinar a partir de una señal en el dominio del tiempo de banda baja utilizando un banco de filtros de análisis QMF (filtro espejo en cuadratura, «Quadrature Mirror Filter») o una FFT (transformada rápida de Fourier, «Fast Fourier Transform»). La pluralidad de señales de subbanda de alta frecuencia generada puede corresponder a una aproximación de las señales de subbanda de alta frecuencia de una señal de audio original a partir de la cual se ha derivado la pluralidad de señales de subbanda de baja frecuencia. En particular, la pluralidad de señales de subbanda de baja frecuencia y la pluralidad de señales de subbanda de alta frecuencia (re)generadas puede corresponder a las subbandas de un banco de filtros QMF y/o de una transformada FFT.

El sistema puede comprender medios para recibir la pluralidad de señales de subbanda de baja frecuencia. Como tal, el sistema se puede colocar después de la transformada o del banco de filtros de análisis que genera la pluralidad de señales de subbanda de baja frecuencia a partir de una señal de banda baja. La señal de banda baja puede ser una señal de audio que se ha descodificado en un descodificador central a partir de un flujo de bits recibido. El flujo de bits se puede almacenar en un soporte de almacenamiento, por ejemplo, un disco compacto o un DVD, o el flujo de bits puede ser recibido en el descodificador a través de un medio de transmisión, por ejemplo, un medio de transmisión de radio u óptico.

El sistema puede comprender medios para recibir un conjunto de energías objetivo, que también se pueden denominar energías de factor de escala. Cada energía objetivo puede cubrir un intervalo objetivo diferente, que también se puede denominar una banda de factor de escala, dentro del intervalo de alta frecuencia. Típicamente, el conjunto de intervalos objetivo que corresponde al conjunto de energías objetivo cubre el intervalo de alta frecuencia completo. Una energía objetivo del conjunto de energías objetivo es normalmente indicativa de la energía deseada de una o más señales de subbanda de alta frecuencia dispuestas dentro del intervalo objetivo correspondiente. En particular, la energía objetivo puede corresponder a la energía deseada media de la una o más señales de subbanda de alta frecuencia que se encuentran dentro del intervalo objetivo correspondiente. La energía objetivo de un intervalo objetivo se deriva típicamente de la energía de la señal de banda alta de la señal de audio original dentro del intervalo objetivo. En otras palabras, el conjunto de energías objetivo describe típicamente la envolvente espectral de la zona de banda alta de la señal de audio original.

El sistema puede comprender medios para generar la pluralidad de señales de subbanda de alta frecuencia a partir de la pluralidad de señales de subbanda de baja frecuencia. Para este propósito, los medios para generar la pluralidad de señales de subbanda de alta frecuencia pueden estar configurados para realizar una transposición de copia de la pluralidad de señales de subbanda de baja frecuencia y/o realizar una transposición armónica de la pluralidad de señales de subbanda de baja frecuencia.

Además, los medios para generar la pluralidad de señales de subbanda de alta frecuencia pueden tener en cuenta una pluralidad de coeficientes de ganancia espectral durante el proceso de generación de la pluralidad de señales de subbanda de alta frecuencia. La pluralidad de coeficientes de ganancia espectral puede estar asociada con la pluralidad de señales de subbanda de baja frecuencia, respectivamente. En otras palabras, cada señal de subbanda de baja frecuencia de la pluralidad de señales de subbanda de baja frecuencia puede tener un coeficiente de ganancia espectral correspondiente a partir de la pluralidad de coeficientes de ganancia espectral. Un coeficiente de ganancia espectral de la pluralidad de coeficientes de ganancia espectral se puede aplicar a la señal de subbanda de baja frecuencia correspondiente.

La pluralidad de coeficientes de ganancia espectral se puede asociar con la energía de la respectiva pluralidad de señales de subbanda de baja frecuencia. En particular, cada coeficiente de ganancia espectral se puede asociar con la energía de su señal de subbanda de baja frecuencia correspondiente. En una realización, se determina un coeficiente de ganancia espectral sobre la base de la energía de la señal de subbanda de baja frecuencia correspondiente. Para este propósito, se puede determinar una curva dependiente de la frecuencia sobre la base de la pluralidad de valores de energía de la pluralidad de señales de subbanda de baja frecuencia. En este caso, un método para determinar la pluralidad de coeficientes de ganancia puede estar basado en la curva dependiente de la frecuencia que se determina a partir de una representación (por ejemplo, logarítmica) de las energías de la pluralidad de señales de subbanda de baja frecuencia.

En otras palabras, la pluralidad de coeficientes de ganancia espectral se puede derivar de una curva dependiente de la frecuencia ajustada a la energía de la pluralidad de señales de subbanda de baja frecuencia. En particular, la curva dependiente de la frecuencia puede ser un polinomio de un orden/grado predeterminado. De manera alternativa o además, la curva dependiente de la frecuencia puede comprender diferentes segmentos de curva, en donde los diferentes segmentos de curva están ajustados a la energía de la pluralidad de señales de subbanda de baja frecuencia en diferentes intervalos de frecuencia. Los diferentes segmentos de curva pueden ser diferentes polinomios de un orden predeterminado. En una realización, los diferentes segmentos de curva son polinomios de orden cero, de modo que los segmentos de curva representan los valores de energía media de la energía de la pluralidad de señales de subbanda de baja frecuencia dentro del intervalo de frecuencia correspondiente.

En una realización adicional, la curva dependiente de la frecuencia se ajusta a la energía de la pluralidad de señales de subbanda de baja frecuencia al realizar una operación de filtrado de media móvil sobre los diferentes intervalos de frecuencia.

En una realización, un coeficiente de ganancia de la pluralidad de coeficientes de ganancia se deriva de la diferencia de la energía media de la pluralidad de señales de subbanda de baja frecuencia y de un valor correspondiente de la curva dependiente de la frecuencia. El valor correspondiente de la curva dependiente de la frecuencia puede ser un valor de la curva a una frecuencia que se encuentra dentro del intervalo de frecuencia de la señal de subbanda de baja frecuencia a la cual corresponde el coeficiente de ganancia.

Típicamente, la energía de la pluralidad de señales de subbanda de baja frecuencia se determina en una determinada cuadrícula de tiempo, por ejemplo, trama a trama, es decir, la energía de una señal de subbanda de baja frecuencia dentro de un intervalo de tiempo definido por la cuadrícula de tiempo corresponde a la energía media de las muestras de la señal de subbanda de baja frecuencia dentro del intervalo de tiempo, por ejemplo, dentro de un trama. Como tal, se puede determinar una diferente pluralidad de coeficientes de ganancia espectral en la cuadrícula de tiempo elegida, por ejemplo, se puede determinar una diferente pluralidad de coeficientes de ganancia espectral para cada trama de la señal de audio. En una realización, la pluralidad de coeficientes de ganancia espectral se puede determinar muestra por muestra, por ejemplo, determinando la energía de la pluralidad de subbandas de baja frecuencia utilizando una ventana flotante a través de las muestras de cada señal de subbanda de baja frecuencia. Cabe señalar que el sistema puede comprender medios para determinar la pluralidad de coeficientes de ganancia espectral a partir de una pluralidad de señales de subbanda de baja frecuencia. Estos medios pueden estar configurados para realizar los métodos mencionados anteriormente con el fin de determinar la pluralidad de coeficientes de ganancia espectral.

Los medios para generar la pluralidad de señales subbanda de alta frecuencia pueden estar configurados para amplificar la pluralidad de señales de subbanda de baja frecuencia utilizando la respectiva pluralidad de coeficientes de ganancia espectral. Aunque en lo sucesivo se hace referencia a «amplificar» o «amplificación», la operación de «amplificación» se puede reemplazar por otras operaciones, tales como una operación de «multiplicación», una operación de «reajuste de escala» o una operación de «ajuste». La amplificación se puede realizar multiplicando una muestra de una señal de subbanda de baja frecuencia por su correspondiente coeficiente de ganancia espectral. En particular, los medios para generar la pluralidad de señales de subbanda de alta frecuencia se pueden configurar para determinar una muestra de una señal de subbanda de alta frecuencia en un instante de tiempo determinado a partir de muestras de una señal de subbanda de baja frecuencia en el instante de tiempo determinado y en al menos un

instante de tiempo anterior. Además, las muestras de una señal de subbanda de baja frecuencia pueden estar amplificadas por el respectivo coeficiente de ganancia espectral de la pluralidad de coeficientes de ganancia espectral. En una realización, los medios para generar la pluralidad de señales de subbanda de alta frecuencia están configurados para generar la pluralidad de señales de subbanda de alta frecuencia a partir de la pluralidad de señales de subbanda de baja frecuencia, de acuerdo con el algoritmo de «copia» especificado en MPEG-4 SBR. La pluralidad de señales de subbanda de baja frecuencia usada en este algoritmo de «copia» se puede haber amplificado utilizando la pluralidad de coeficientes de ganancia espectral, en donde la operación de «amplificación» se puede haber realizado según se describió anteriormente.

El sistema puede comprender medios para ajustar la energía de la pluralidad de señales de subbanda de alta frecuencia utilizando el conjunto de energías objetivo. Esta operación se denomina típicamente ajuste de envolvente espectral. El ajuste de envolvente espectral se puede realizar ajustando la energía de la pluralidad de señales de subbanda de alta frecuencia de tal manera que la energía media de la pluralidad de señales de subbanda de alta frecuencia que se encuentran dentro de un intervalo objetivo corresponda con la energía objetivo correspondiente. Esto se puede lograr determinando un valor de ajuste de envolvente a partir de los valores de energía de la pluralidad de señales de subbanda de alta frecuencia que se encuentran dentro de un intervalo objetivo y la energía objetivo correspondiente. En particular, el valor de ajuste de envolvente se puede determinar a partir de una proporción entre la energía objetivo y los valores de energía de la pluralidad de señales de subbanda de alta frecuencia que se encuentran dentro de un intervalo objetivo correspondiente. Este valor de ajuste de envolvente se puede usar para ajustar la energía de la pluralidad de señales de subbanda de alta frecuencia.

Según la invención, los medios para ajustar la energía comprenden medios para limitar el ajuste de la energía de las señales de subbanda de alta frecuencia que se encuentran dentro del intervalo limitador. Típicamente, el intervalo limitador cubre más de un intervalo objetivo. Los medios para limitar se usan para evitar una amplificación del ruido no deseable dentro de ciertas señales de subbanda de alta frecuencia. Por ejemplo, los medios para limitar pueden estar configurados para determinar un valor de ajuste de envolvente medio de los valores de ajuste de envolvente que corresponden a los intervalos objetivo cubiertos o que se encuentran dentro del intervalo limitador. Además, los medios para limitar se pueden configurar de manera que limiten el ajuste de la energía de las señales de subbanda de alta frecuencia que se encuentran dentro del intervalo limitador a un valor que es proporcional al valor de ajuste de envolvente medio.

Además, los medios para ajustar la energía de la pluralidad de señales de subbanda de alta frecuencia pueden comprender medios para asegurar que las señales de subbanda de alta frecuencia ajustadas que se encuentran dentro del intervalo objetivo particular tengan la misma energía. Estos últimos medios se denominan con frecuencia medios de «interpolación». En otras palabras, los medios de «interpolación» aseguran que la energía de cada una de las señales de subbanda de alta frecuencia que se encuentran dentro del intervalo objetivo particular corresponda a la energía objetivo. Los medios de «interpolación» se pueden implementar ajustando cada señal de subbanda de alta frecuencia dentro del intervalo objetivo particular por separado, de manera tal que la energía de la señal de subbanda de alta frecuencia ajustada corresponda con la energía objetivo asociada con el intervalo objetivo particular. Esto se puede lograr determinando un valor de ajuste de envolvente diferente para cada señal de subbanda de alta frecuencia dentro del intervalo objetivo particular. Se puede determinar un valor de ajuste de envolvente diferente sobre la base de la energía de la señal de subbanda de alta frecuencia particular y la energía objetivo correspondiente al intervalo objetivo particular. En una realización, el valor de ajuste de envolvente para una señal de subbanda de alta frecuencia particular se determina sobre la base de la proporción entre la energía objetivo y la energía de la señal de subbanda de alta frecuencia particular.

El sistema además puede comprender medios para recibir datos de control. Los datos de control pueden ser indicativos de si se aplica la pluralidad de coeficientes de ganancia espectral para generar la pluralidad de señales de subbanda de alta frecuencia. En otras palabras, los datos de control pueden ser indicativos de si se debe realizar o no el ajuste de ganancia adicional de las señales de subbanda de baja frecuencia. De manera alternativa o además, los datos de control pueden ser indicativos de un método que se utiliza para determinar la pluralidad de coeficientes de ganancia espectral. A modo de ejemplo, los datos de control pueden ser indicativos del orden predeterminado del polinomio que se ha de utilizar para determinar la curva dependiente de la frecuencia ajustada a las energías de la pluralidad de señales de subbanda de baja frecuencia. Los datos de control se reciben típicamente desde un codificador correspondiente que analiza la señal de audio original e informa al correspondiente descodificador o sistema HFR sobre cómo descodificar el flujo de bits.

De acuerdo con otro aspecto, se describe un método para generar una pluralidad de señales de subbanda de alta frecuencia que cubren un intervalo de alta frecuencia desde una pluralidad de señales de subbanda de baja frecuencia. El método puede comprender las etapas de recibir la pluralidad de señales de subbanda de baja frecuencia y/o recibir un conjunto de energías objetivo. Cada energía objetivo puede cubrir un intervalo objetivo diferente dentro del intervalo de alta frecuencia. Además, cada energía objetivo puede ser indicativa de la energía deseada de una o más señales de subbanda de alta frecuencia dispuestas dentro del intervalo objetivo. El método puede comprender la etapa de generar la pluralidad de señales de subbanda de alta frecuencia a partir de la pluralidad de señales de subbanda de baja frecuencia y a partir de una pluralidad de coeficientes de ganancia espectral asociados con la pluralidad de señales de subbanda de baja frecuencia, respectivamente. De manera alternativa o además, el método puede

comprender la etapa de ajustar la energía de la pluralidad de señales de subbanda de alta frecuencia utilizando el conjunto de energías objetivo. La etapa de ajustar la energía puede comprender la etapa de limitar el ajuste de la energía de las señales de subbanda de alta frecuencia que se encuentran dentro del intervalo limitador. Típicamente, el intervalo limitador cubre más de un intervalo objetivo.

De acuerdo con otro aspecto, se describe un soporte de almacenamiento. El soporte de almacenamiento puede comprender un programa de software adaptado para su ejecución en un procesador y para realizar las etapas del método descritas en el presente documento cuando se ejecuta en un dispositivo informático.

De acuerdo con un aspecto adicional, se describe un producto de programa informático. El programa informático puede comprender instrucciones ejecutables para realizar las etapas del método descritas en el presente documento cuando se ejecutan en un ordenador.

Cabe señalar que los métodos y los sistemas, incluidas sus realizaciones preferentes como se describe en la presente solicitud de patente, se pueden usar de manera independiente o en combinación con los demás métodos y sistemas descritos en este documento. Además, todos los aspectos de los métodos y sistemas descritos en la presente solicitud de patente se pueden combinar de manera arbitraria. En particular, las características de las reivindicaciones se pueden combinar entre sí de manera arbitraria.

Breve descripción de los dibujos

La invención se explica a continuación por medio de ejemplos ilustrativos, haciendo referencia a los dibujos que se acompañan, en donde

la figura 1a ilustra el espectro absoluto de un ejemplo de señal de banda alta antes del ajuste de envolvente espectral;

la figura 1b ilustra un ejemplo de relación entre tramas de tiempo de los datos de audio y bordes de tiempo de envolvente en las envolventes espectrales;

la figura 1c ilustra el espectro absoluto de un ejemplo de señal de banda alta antes del ajuste de envolvente espectral, y las bandas de factor de escala, las bandas limitadoras y los parches HF (alta frecuencia, «High Frequency») correspondientes;

la figura 2 ilustra una realización de un sistema HFR en el que el proceso de copia se complementa con una etapa de ajuste de ganancia adicional;

la figura 3 ilustra una aproximación de la envolvente espectral bruta de un ejemplo de señal de banda baja;

la figura 4 ilustra una realización de un ajustador de ganancia adicional que opera con datos de control opcionales, las muestras de subbandas QMF, y envía como salida una curva de ganancia;

la figura 5 ilustra una realización más detallada del ajustador de ganancia adicional de la figura 4;

la figura 6 ilustra una realización de un sistema HFR con una señal de banda estrecha como entrada y una señal de banda ancha como salida;

la figura 7 ilustra una realización de un sistema HFR incorporado en el módulo SBR de un descodificador de audio;

la figura 8 ilustra una realización del módulo de reconstrucción de alta frecuencia de un ejemplo de descodificador de audio;

la figura 9 ilustra una realización de un ejemplo de codificador;

la figura 10a ilustra el espectrograma de un ejemplo de segmento vocal que se ha descodificado utilizando un descodificador convencional;

la figura 10b ilustra el espectrograma del segmento vocal de la figura 10a, que se ha descodificado utilizando un descodificador que aplica el procesamiento de ajuste de ganancia adicional; y

la figura 10c ilustra el espectrograma del segmento vocal de la figura 10a para la señal no codificada original.

Descripción de realizaciones preferidas

Las realizaciones descritas a continuación son meramente ilustrativas de los principios de la presente invención, PROCESAMIENTO DE SEÑALES DE AUDIO DURANTE LA RECONSTRUCCIÓN DE ALTA FRECUENCIA. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles descritos en esta invención resultarán evidentes para otros expertos en la materia. Por consiguiente, la intención es limitar la invención únicamente por el alcance de las reivindicaciones de patente pendientes y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones de esta invención.

Como se describió anteriormente, los descodificadores de audio que utilizan técnicas HFR comprenden típicamente una unidad HFR para generar una señal de audio de alta frecuencia y una posterior unidad de ajuste de envolvente espectral para ajustar la envolvente espectral de la señal de audio de alta frecuencia. Cuando se ajusta la envolvente espectral de la señal de audio, esto se realiza típicamente mediante la implementación de un banco de filtros, o mediante un filtrado en el dominio del tiempo. El ajuste puede tratar de realizar una corrección de la envolvente espectral absoluta o se puede realizar mediante un filtrado que también corrija las características de fase. De cualquier manera, el ajuste es típicamente una combinación de dos etapas, la eliminación de la envolvente espectral actual y la aplicación de la envolvente espectral objetivo.

Es importante señalar que los métodos y los sistemas descritos en el presente documento no están meramente dirigidos a la eliminación de la envolvente espectral de la señal de audio. Los métodos y sistemas tratan de realizar una corrección espectral adecuada de la envolvente espectral de la señal de banda baja como parte de la etapa de regeneración de alta frecuencia, con el fin de no introducir discontinuidades de la envolvente espectral del espectro de alta frecuencia creadas al combinar diferentes segmentos de la banda baja, es decir, la señal de baja frecuencia, desplazados o transpuestos a intervalos de frecuencia diferentes de la banda alta, es decir, la señal de alta frecuencia.

En la figura 1a se muestra un espectro de dibujo estilizado 100, 110 de la salida de una unidad HFR antes de pasar al ajustador de envolvente. En el panel superior, se usa un método de copia (con dos parches) para generar una señal de banda alta 105 a partir de la señal de banda baja 101, por ejemplo, el método de copia utilizado en la MPEG-4 SBR (replicación de banda espectral, «Spectral Band Replication») que se describe en la norma «ISO/IEC 14496-3 Tecnología de la información - Codificación de objetos audiovisuales - Parte 3: Audio». El método de copia traduce partes de las frecuencias inferiores 101 a frecuencias superiores 105. En el panel inferior, se usa un método de transposición armónica (con dos parches) para generar la señal de banda alta 115 a partir de la señal de banda baja 111, por ejemplo, el método de transposición armónica de MPEG-D USAC, que se describe en la norma «MPEG-D USAC: ISO/IEC 23003-3 - Codificación unificada de voz y audio».

En la posterior fase de ajuste de envolvente, se aplica una envolvente espectral objetivo sobre los componentes de alta frecuencia 105, 115. Como se puede ver en el espectro 105, 115 que se dirige al ajustador de envolvente, se pueden observar discontinuidades (de manera notable en los bordes del parche) en la forma espectral de la señal de excitación de banda alta 105, 115, es decir, de la señal de banda alta que entra al ajustador de envolvente. Estas discontinuidades tienen su origen en el hecho de que se usan varias contribuciones de las bajas frecuencias 101, 111 con el fin de generar la banda alta 105, 115. Como se puede observar, la forma espectral de la señal de banda alta 105, 115 está relacionada con la forma espectral de la señal de banda baja 101, 111. En consecuencia, las formas espectrales particulares de la señal de banda baja 101, 111, por ejemplo, una forma de gradiente ilustrada en la figura 1a, pueden llevar a discontinuidades en el espectro total 100, 110.

Además del espectro 100, 110, la figura 1a ilustra ejemplos de bandas de frecuencia 130 de los datos de envolvente espectral que representan la envolvente espectral objetivo. Estas bandas de frecuencia 130 se denominan bandas de factor de escala o intervalos objetivo. Típicamente, se especifica un valor de energía objetivo, es decir, una energía de factor de escala, para cada intervalo objetivo, es decir, banda de factor escala. En otras palabras, las bandas de factor de escala definen la resolución de frecuencia efectiva de la envolvente espectral objetivo, ya que típicamente solo hay un único valor de energía objetivo por intervalo objetivo. Utilizando los factores de escala o energías objetivo especificados para las bandas de factor de escala, el posterior ajustador de envolvente trata de ajustar la señal de banda alta de manera tal que la energía de la señal de banda alta dentro de las bandas de factor de escala sea igual a la energía de los datos de envolvente espectral recibidos, es decir, la energía objetivo, para las bandas de factor de escala respectivas.

En la figura 1c se proporciona una descripción más detallada utilizando un ejemplo de señal de audio. En el gráfico se ilustra el espectro de una señal de audio del mundo real 121 que se dirige al ajustador de envolvente, así como la señal original correspondiente 120. En este ejemplo particular, el intervalo SBR, es decir, el intervalo de la señal de alta frecuencia, empieza en 6,4 kHz y consiste en tres replicaciones diferentes del intervalo de frecuencia de banda baja. Los intervalos de frecuencia de las diferentes replicaciones vienen indicados por «parche 1», «parche 2» y «parche 3». Queda claro a partir del espectrograma que los parches introducen discontinuidades en la envolvente espectral alrededor de 6,4 kHz, 7,4 kHz, y 10,8 kHz. En el presente ejemplo, estas frecuencias corresponden a bordes de parche.

La figura 1c ilustra además las bandas de factor de escala 130 así como las bandas limitadoras 135, cuya función se describirá con más detalle a continuación. En la realización ilustrada, se utiliza el ajustador de envolvente de MPEG-4 SBR. Este ajustador de envolvente opera utilizando un banco de filtros QMF. Los aspectos principales de la operación de un ajustador de envolvente de este tipo son:

- calcular la energía media a través de una banda de factor de escala 130 de la señal de entrada al ajustador de envolvente, es decir, la señal que sale de la unidad HFR; en otras palabras, la energía media de la señal de banda alta regenerada se calcula dentro de cada intervalo objetivo/banda de factor de escala 130;

- determinar un valor de ganancia, también denominado valor de ajuste de envolvente, para cada banda de factor de escala 130, en donde el valor de ajuste de envolvente es la raíz cuadrada de la proporción de

energía entre la energía objetivo (es decir, el objetivo de energía recibido de un codificador) y la energía media de la señal de banda alta regenerada 121 dentro de la banda de factor de escala respectiva 130;

- aplicar el respectivo valor de ajuste de envolvente a la banda de frecuencia de la señal de banda alta regenerada 121, en donde la banda de frecuencia corresponde a la respectiva banda de factor de escala 130.

Además, el ajustador de envolvente puede comprender etapas y variaciones adicionales, en particular:

- una función limitadora, que limita el valor de ajuste de envolvente máximo permitido que se aplicará sobre una cierta banda de frecuencia, es decir, sobre una banda limitadora 135. El valor de ajuste de envolvente máximo permitido es una función de los valores de ajuste de envolvente determinados para las diferentes bandas de factor de escala 130 que caen dentro de una banda limitadora 135. En particular, el valor de ajuste de envolvente máximo permitido es una función de la media de los valores de ajuste de envolvente determinados para las diferentes bandas de factor de escala 130 que caen dentro de una banda limitadora 135. A modo de ejemplo, el valor de ajuste de envolvente máximo permitido puede ser el valor medio de los valores de ajuste de la envolvente relevantes multiplicado por un factor limitador (tal como 1,5). La funcionalidad limitadora se aplica típicamente con el fin de limitar la introducción de ruido en la señal de banda alta regenerada 121. Esto es particularmente relevante en el caso de señales de audio que comprendan sinusoides prominentes, es decir, señales de audio que tengan un espectro con picos inequívocos a ciertas frecuencias. Sin el uso de la funcionalidad limitadora, se determinarían valores de ajuste de envolvente significativos para las bandas de factor de escala 130 para las cuales la señal de audio original comprende estos picos inequívocos. Como resultado, se ajustará el espectro de la banda de factor de escala completa 130 (y no solo el pico inequívoco), introduciendo de esta manera ruido.

- una funcionalidad de interpolación, que permite calcular los valores de ajuste de envolvente para cada subbanda QMF individual dentro de una banda de factor de escala, en lugar de calcular un único valor de ajuste de envolvente para toda la banda de factor de escala. Ya que las bandas de factor de escala comprenden típicamente más de una subbanda QMF, un valor de ajuste de envolvente se puede calcular como la proporción de la energía de una subbanda QMF particular dentro de la banda de factor de escala y la energía objetivo que se recibe del codificador, en lugar de calcular la proporción de la energía media de todas las subbandas QMF dentro de la banda de factor de escala y la energía objetivo que se recibe del codificador. Como tal, se puede determinar un valor de ajuste de envolvente diferente para cada subbanda QMF dentro de una banda de factor de escala. Cabe señalar que el valor de energía objetivo recibido para una banda de factor de escala corresponde típicamente a la energía media de este intervalo de frecuencia dentro de la señal original. La manera de aplicar la energía objetivo media recibida a la banda de frecuencia correspondiente de la señal de banda alta regenerada depende del funcionamiento del descodificador. Esto se puede realizar aplicando un valor de ajuste de envolvente total a las subbandas QMF dentro de una banda de factor de escala de la señal de banda alta regenerada o aplicando un valor de ajuste de envolvente individual a cada subbanda QMF. Se puede considerar que este último planteamiento es como si la información de envolvente recibida (es decir, una energía objetivo por banda de factor escala) se «interpolase» a lo largo de las subbandas QMF dentro de una banda de factor de escala, con el fin de obtener una resolución de frecuencia superior. Por consiguiente, este planteamiento se denomina «interpolación» en MPEG-4 SBR.

Volviendo a la figura 1c, se puede ver que el ajustador de envolvente deberá aplicar valores de ajuste de envolvente elevados con el fin de hacer coincidir el espectro 121 de la señal que se dirige al ajustador de envolvente con el espectro 120 de la señal original. También se puede ver que, debido a las discontinuidades, se producen grandes variaciones de los valores de ajuste de envolvente dentro de las bandas limitadoras 135. Como resultado de estas grandes variaciones, los valores de ajuste de envolvente que corresponden a los mínimos locales del espectro regenerado 121 estarán limitados por la función limitadora del ajustador de envolvente. En consecuencia, las discontinuidades dentro del espectro regenerado 121 permanecerán, incluso después de realizar la operación de ajuste de envolvente. Por otra parte, si no se usa una funcionalidad limitadora, se podría introducir ruido no deseado, como se describió anteriormente.

Por lo tanto, ocurre un problema para la regeneración de una señal de banda alta para cualquier señal que tenga grandes variaciones de nivel a lo largo del intervalo de banda baja. Este problema se debe a las discontinuidades que se introducen durante la regeneración de alta frecuencia de la banda alta. Cuando posteriormente el ajustador de envolvente se expone a esta señal regenerada, no se puede separar, de manera razonable y consistente, la discontinuidad recientemente introducida de cualquier característica espectral del «mundo real» de la señal de banda baja. Los efectos de este problema son dobles. En primer lugar, las formas espectrales se introducen en la señal de banda alta que el ajustador de envolvente no puede compensar. En consecuencia, la salida tiene una forma espectral errónea. En segundo lugar, se percibe un efecto de inestabilidad, debido al hecho de que este efecto va y viene como función de las características espectrales de banda baja.

El presente documento aborda el problema mencionado anteriormente al describir un método y un sistema que proporcionan una señal de banda alta HFR en la entrada del ajustador de envolvente que no muestra discontinuidades espectrales. Para este propósito, se propone el retirar o reducir la envolvente espectral de la señal de banda baja al

realizar la regeneración de alta frecuencia. Al hacer esto, se evita introducir cualesquiera discontinuidades espectrales en la señal de banda alta antes de realizar el ajuste de envolvente. Como resultado, el ajustador de envolvente no tendrá que gestionar estas discontinuidades espectrales. En particular, se puede usar un ajustador de envolvente convencional, en donde la función limitadora del ajustador de envolvente se utiliza para evitar la introducción de ruido en la señal de banda alta regenerada. En otras palabras, el método y el sistema descritos se pueden usar para regenerar una señal de banda alta HFR que tenga pocas o ningunas discontinuidades espectrales y un bajo nivel de ruido.

Cabe señalar que la resolución en el tiempo del ajustador de envolvente puede ser diferente de la resolución en el tiempo del procesamiento propuesto de la envolvente espectral durante la generación de señal de banda alta. Como se indicó anteriormente, se pretende que el procesamiento de la envolvente espectral durante la regeneración de la señal de banda alta modifique la envolvente espectral de la señal de banda baja, con el fin de aliviar el procesamiento dentro del posterior ajustador de envolvente. Este procesamiento, es decir, la modificación de la envolvente espectral de la señal de banda baja, se puede realizar, por ejemplo, una vez por trama de audio, en donde el ajustador de envolvente puede ajustar la envolvente espectral a lo largo de varios intervalos de tiempo, es decir, utilizando varias envolventes espectrales recibidas. Esto se describe en la figura 1b, en la que la cuadrícula de tiempo 150 de los datos de envolvente espectral se representa en el panel superior, y la cuadrícula de tiempo 155 para el procesamiento de la envolvente espectral de la señal de banda baja durante la regeneración de la señal de banda alta se representa en el panel inferior. Como se puede ver en el ejemplo de la figura 1b, los bordes de tiempo de los datos de envolvente espectral varían con el tiempo, mientras que el procesamiento de la envolvente espectral de la señal de banda baja opera en una cuadrícula de tiempo fija. También se puede ver que se podrían realizar varios ciclos de ajuste de envolvente (representados por bordes de tiempo 150) durante un ciclo de procesamiento de la envolvente espectral de la señal de banda baja. En el ejemplo ilustrado, el procesamiento de la envolvente espectral de la señal de banda baja opera trama por trama, lo que significa que se determina una pluralidad diferente de coeficientes de ganancia espectral para cada trama de la señal. Cabe señalar que el procesamiento de la señal de banda baja puede operar en cualquier cuadrícula de tiempo y que la cuadrícula de tiempo de este procesamiento no tiene que coincidir con la cuadrícula de tiempo de los datos de la envolvente espectral.

En la figura 2, se representa un sistema HFR basado en banco de filtros 200. El sistema HFR 200 opera utilizando un banco de filtros seudo-QMF y el sistema 200 se puede usar para producir la señal de banda alta y banda baja 100 ilustrada en el panel superior de la figura 1a. Sin embargo, se ha añadido una etapa adicional de ajuste de ganancia como parte del proceso de generación de alta frecuencia, que en el ejemplo ilustrado es un proceso de copia. La señal de entrada de baja frecuencia es analizada por un QMF de 32 subbandas 201, con el fin de generar una pluralidad de señales de subbanda de baja frecuencia. A algunas o todas las señales de subbanda de baja frecuencia se les aplican parches en ubicaciones de frecuencia superior de acuerdo con un algoritmo de generación de HF (alta frecuencia, «High Frequency»). Adicionalmente, la pluralidad de subbandas de baja frecuencia se envía directamente al banco de filtros de síntesis 202. El banco de filtros de síntesis mencionado anteriormente 202 es un QMF inverso de 64 subbandas 202. En el caso de la implementación particular ilustrada en la figura 2, el uso de un banco de filtros de análisis QMF de 32 subbandas 201 y el uso de un banco de filtros de síntesis QMF de 64 subbandas 202 dará lugar a una tasa de muestreo de salida de la señal de salida del doble de la tasa de muestreo de entrada de la señal de entrada. Sin embargo, cabe señalar que los sistemas descritos en el presente documento no se limitan a sistemas con diferentes tasas de muestreo de entrada y salida. Los expertos en la técnica podrán contemplar una multitud de diferentes relaciones de tasas de muestreo.

Como se describió en la figura 2, las subbandas de las frecuencias inferiores son asignadas a subbandas de frecuencias superiores. Como parte de este proceso de copia se introduce una fase de ajuste de ganancia 204. La señal de alta frecuencia creada, es decir, la pluralidad generada de señales de subbanda de alta frecuencia, se envía al ajustador de envolvente 203 (que posiblemente comprende una funcionalidad de interpolación y/o limitadora), antes de su combinación con la pluralidad de señales de subbanda de baja frecuencia en el banco de filtros de síntesis 202. Al utilizar tal sistema HFR 200, y en particular al utilizar una fase de ajuste de ganancia 204, es posible evitar la introducción de las discontinuidades de envolvente espectral, como se ilustra en la figura 1. Para este propósito, la fase de ajuste de ganancia 204 modifica la envolvente espectral de la señal de banda baja, es decir, la envolvente espectral de la pluralidad de señales de subbanda de baja frecuencia, de tal manera que la señal de banda baja modificada se puede utilizar para generar una señal de banda alta, es decir, una pluralidad de señales de subbanda de alta frecuencia, que no muestran discontinuidades, en particular discontinuidades en los bordes de parche. Haciendo referencia a la figura 1c, la fase de ajuste de ganancia adicional 204 asegura que la envolvente espectral 101, 111 de la señal de banda baja se modifique de tal manera que no haya discontinuidades en la señal de banda alta generada 105, 115, o que estas sean limitadas.

La modificación de la envolvente espectral de la señal de banda baja se puede lograr aplicando una curva de ganancia a la envolvente espectral de la señal de banda baja. Esta curva de ganancia puede ser determinada por una unidad de determinación de curva de ganancia 400 ilustrada en la figura 4. El módulo 400 toma como entrada los datos QMF 402 correspondientes al intervalo de frecuencia de la señal de banda baja utilizados para recrear la señal de banda alta. En otras palabras, la pluralidad de señales de subbanda de baja frecuencia se envía a la unidad de determinación de curva de ganancia 400. Como ya se indicó, solo se puede utilizar un subconjunto de las subbandas QMF disponibles de la señal de banda baja para generar la señal de banda alta, es decir, solo se puede enviar un subconjunto de las

subbandas QMF disponibles a la unidad de determinación de curva de ganancia 400. Además, el módulo 400 puede recibir datos de control opcionales 404, por ejemplo, datos de control enviados desde un codificador correspondiente. El módulo 400 enviará como salida una curva de ganancia 403 que se aplicará durante el proceso de regeneración de alta frecuencia. En una realización, la curva de ganancia 403 se aplica a las subbandas QMF de la señal de banda baja, que se usan para generar la señal de banda alta. Es decir, la curva de ganancia 403 se puede usar dentro del proceso de copia del proceso HFR.

Los datos de control opcional 404 pueden comprender información de la resolución de la envolvente espectral bruta, que se estimará en el módulo 400, y/o información sobre la conveniencia de aplicar el proceso de ajuste de ganancia. Así pues, los datos de control 404 pueden controlar la cantidad de procesamiento adicional involucrado durante el proceso de ajuste de ganancia. Los datos de control 404 también pueden activar una derivación del procesamiento de ajuste de ganancia adicional, si se producen señales que no permitan fácilmente realizar la estimación de envolvente espectral bruta, por ejemplo, señales que comprendan sinusoides sencillas.

En la figura 5 se muestra una vista más detallada del módulo 400 de la figura 4. Los datos QMF 402 de la señal de banda baja se envían a la unidad de estimación de envolvente 501 que estima la envolvente espectral, por ejemplo, en una escala de energía logarítmica. La envolvente espectral posteriormente se envía al módulo 502 que estima la envolvente espectral bruta de la envolvente espectral de alta resolución (frecuencia) que se recibe de la unidad de estimación de envolvente 501. En una realización, esto se realiza ajustando un polinomio de orden bajo a los datos de envolvente espectral, es decir, un polinomio de un orden en el intervalo, por ejemplo, de 1, 2, 3 o 4. La envolvente espectral bruta también se puede determinar realizando una operación de media móvil de la envolvente espectral de alta resolución a lo largo del eje de frecuencia. En la figura 3 se ilustra la determinación de una envolvente espectral bruta 301 de una señal de banda baja. Se puede ver que el espectro absoluto 302 de la señal de banda baja, es decir, la energía de las bandas QMF 302, se aproxima mediante una envolvente espectral bruta 301, es decir, mediante una curva dependiente de la frecuencia ajustada a la envolvente espectral de la pluralidad de señales de subbanda de baja frecuencia. Además, se muestra que solo se usan 20 señales de subbanda QMF para generar las señales de banda alta, es decir, solo una parte de las 32 señales de subbanda QMF se usan dentro del proceso HFR.

```
function GainVec = calculateGainVec(LowEnv)
%%    function GainVec = calculateGainVec(LowEnv)
%   Entrada: Energía de envolvente de banda baja en dB
%   Salida: vector de ganancia a aplicar a la banda baja antes de la generación de HF
%
%   La función hace un ajuste polinómico de orden bajo de la envolvente espectral de banda
%   baja, como representación de la pendiente espectral global de banda baja. La pendiente
%   global de acuerdo con esto se traduce seguidamente en un vector de ganancia que se
%   puede aplicar antes de una generación de HF para quitar la pendiente global (o forma
%   espectral bruta).
%
%   Esto impide que la generación de HF introduzca discontinuidades en la forma espectral,
%   que serían “confusas” para el siguiente proceso de limitador y ajuste de envolvente. La
%   “confusión” se produce cuando el limitador y el ajustador de envolvente necesitan tener
%   cuidado de una discontinuidad grande, y de este modo un valor de ganancia elevado. Es
%   muy difícil sintonizar y tener un funcionamiento apropiado de estos módulos si se tiene
%   que tener cuidado tanto variaciones “naturales” en la banda alta como las variaciones
%   “artificiales” introducidas por el proceso de generación de HF.
%
%


  polyOrderWhite = 3;
  x_lowBand = 1:length(LowEnv);

  p=polyfit(x_lowBand,LowEnv,polyOrderWhite);
  lowBandEnvSlope = zeros(size(x_lowBand));
  for k=polyOrderWhite:-1:0
    tmp = (x_lowBand.^k).*p(polyOrderWhite - k + 1);
    lowBandEnvSlope = lowBandEnvSlope + tmp;
  end

  GainVec = 10.^((mean(LowEnv) - lowBandEnvSlope)./20);
```

El método usado para determinar la envolvente espectral bruta de la envolvente espectral de alta resolución, y en particular el orden del polinomio que se adapta o se ajusta a la envolvente espectral de alta resolución, puede estar controlado por los datos de control opcional 404. El orden del polinomio puede ser una función del tamaño del intervalo

de frecuencia 302 de la señal de banda baja, para la cual se ha de determinar una envolvente espectral bruta 301 y/o puede ser una función de otros parámetros relevantes para la forma espectral bruta total del intervalo de frecuencia relevante 302 de la señal de banda baja. El ajuste polinómico calcula un polinomio que aproxima los datos en un sentido de error de mínimos cuadrados. A continuación se establece una realización preferente, mediante código Matlab:

En el código anterior, la entrada es la envolvente espectral (LowEnv) de la señal de banda baja que se obtiene al promediar muestras de subbanda QMF sobre la base de subbandas a lo largo de un intervalo de tiempo correspondiente a la trama de tiempo actual de datos operado por el ajustador de envolvente subsecuente. Como se indicó anteriormente, el procesamiento de ajuste de ganancia de la señal de banda baja se puede realizar en diversas otras cuadrículas de tiempo. En el ejemplo anterior, la envolvente espectral absoluta estimada se expresa en un dominio logarítmico. Un polinomio de bajo orden, en el ejemplo anterior un polinomio de orden 3, se ajusta a los datos. Una vez hallado el polinomio, se calcula una curva de ganancia (GainVec) a partir de la diferencia en energía media de la señal de banda baja y la curva (lowBandEnvSlope) que se obtiene del polinomio ajustado a los datos.

En el ejemplo anterior, la operación de determinar la curva de ganancia se realiza en el dominio logarítmico.

El cálculo de curva de ganancia es realizado por la unidad de cálculo de la curva de ganancia 503. Como se indicó anteriormente, la curva de ganancia se puede determinar a partir de la energía media de la parte de la señal de banda baja usada para regenerar la señal de banda alta, y de la envolvente espectral de la parte de la señal de banda baja usada para regenerar la señal de banda alta. En particular, la curva de ganancia se puede determinar a partir de la diferencia entre la energía media y la envolvente espectral bruta, representada, por ejemplo, por un polinomio. Es decir, el polinomio calculado se puede usar para determinar una curva de ganancia que comprende un valor de ganancia separado, también denominado coeficiente de ganancia espectral, para cada subbanda QMF relevante de la señal de banda baja. Esta curva de ganancia que comprende los valores de ganancia se usa posteriormente en el proceso HFR.

Como ejemplo, a continuación se describirá un proceso de generación de HFR de acuerdo con MPEG-4 SBR. La señal generada HF se puede derivar de la siguiente fórmula (véase el documento MPEG-4 Parte 3 (ISO/IEC 14496-3), subparte 4, sección 4.6.18.6.2):

$$\mathbf{X}_{High}\left(k,l+t_{HFAdj}\right) = \mathbf{X}_{Low}\left(p,l+t_{HFAdj}\right) + \mathbf{bwArray}\left(g\left(k\right)\right)\cdot\alpha_0\left(p\right)\cdot\mathbf{X}_{Low}\left(p,l-1+t_{HFAdj}\right) + \left[\mathbf{bwArray}\left(g\left(k\right)\right)\right]^2\cdot\alpha_1\left(p\right)\cdot\mathbf{X}_{Low}\left(p,l-2+t_{HFAdj}\right),$$

en donde p es el índice de subbanda de la señal de banda baja, es decir, p identifica una de la pluralidad de señales de subbanda de baja frecuencia. La fórmula de generación de HF anterior se puede reemplazar por la siguiente fórmula, que realiza de manera combinada un ajuste de ganancia y una generación de HF:

$$\mathbf{X}_{High}\left(k,l+t_{HFAdj}\right) = preGain\left(p\right)\cdot\left(\mathbf{X}_{Low}\left(p,l+t_{HFAdj}\right) + \mathbf{bwArray}\left(g\left(k\right)\right)\cdot\alpha_0\left(p\right)\cdot\mathbf{X}_{Low}\left(p,l-1+t_{HFAdj}\right) + \left[\mathbf{bwArray}\left(g\left(k\right)\right)\right]^2\cdot\alpha_1\left(p\right)\cdot\mathbf{X}_{Low}\left(p,l-2+t_{HFAdj}\right)\right),$$

en donde la curva de ganancia se denomina *preGain(p).*

En el documento mencionado anteriormente MPEG-4, Parte 3, se ofrecen más detalles del proceso de copia, por ejemplo, respecto a la relación entre p y k. En la fórmula anterior, $X_{Low}(p,l)$ indica una muestra en la instancia de tiempo l de la señal de subbanda de baja frecuencia que tiene un índice de subbanda p. Esta muestra en combinación con muestras anteriores se usa para generar una muestra de la señal de subbanda de alta frecuencia $X_{High}(k,l)$ que tiene un índice de subbanda k.

Cabe señalar que el aspecto del ajuste de ganancia se puede usar en cualquier sistema de reconstrucción de alta frecuencia basado en banco de filtros. Esto se ilustra en la figura 6, en donde la presente invención es parte de una unidad HFR 601 independiente, que opera en una señal de banda estrecha o banda baja 602 y envía como salida una señal de banda ancha o banda alta 604. El módulo 601 puede recibir datos de control adicional 603 como entrada, en donde los datos de control 603 pueden especificar, entre otras cosas, la cantidad de procesamiento usado para el ajuste de ganancia descrito, así como, por ejemplo, información de la envolvente espectral objetiva de la señal de banda alta. Sin embargo, estos parámetros son solo ejemplos de datos de control opcionales 603. En una realización, también se puede derivar información relevante de la señal de banda estrecha 602 que se introduce en el módulo 601, o por otros medios. Es decir, los datos de control 603 se pueden determinar dentro del módulo 601 sobre la base de la información disponible en el módulo 601. Cabe señalar que la unidad HFR 601 independiente podría recibir la pluralidad de señales de subbanda de baja frecuencia y podría enviar como salida la pluralidad de señales de

subbanda de alta frecuencia, es decir, las transformadas o los bancos de filtros para síntesis/análisis podrían colocarse fuera de la unidad HFR 601.

Como ya se indicó anteriormente, puede resultar beneficioso señalar la activación del procesamiento de ajuste de ganancia en el flujo de bits desde un codificador hasta un descodificador. En el caso de ciertos tipos de señal, por ejemplo, una sinusoide sencilla, el procesamiento de ajuste de ganancia puede no ser relevante y, por consiguiente, resultar beneficiosa de cara a permitir que el sistema de codificador/descodificador desactive el procesamiento adicional con el fin de no introducir un comportamiento no deseado para estas señales de caso límite. Para este propósito, el codificador puede estar configurado para analizar las señales de audio y generar datos de control que activan y desactivan el procesamiento de ajuste de ganancia en el descodificador.

En la figura 7, la fase de ajuste de ganancia propuesta se incluye en una unidad de reconstrucción de alta frecuencia 703 que es parte de un códec de audio. Un ejemplo de esta unidad HFR 703 es la herramienta de replicación de la banda espectral MPEG-4 usada como parte del códec AAC de alta eficiencia o el MPEG-D USAC (códec de audio y voz unificado, «Unified Speech and Audio Codec»). En esta realización, se recibe un flujo de bits 704 en un descodificador de audio 700. El flujo de bits 704 se desmultiplexa en el desmultiplexor 701. La parte relevante SBR del flujo de bits 708 se envía al módulo SBR o la unidad HFR 703, y el flujo de bits relevante de descodificador central 707, por ejemplo, datos AAC o datos de descodificador central USAC, se envían al módulo codificador central 702. Además, la señal de banda baja o banda estrecha 706 se pasa desde el descodificador central 702 a la unidad HFR 703. La presente invención se incorpora como parte del proceso SBR en la unidad HFR 703, por ejemplo, de acuerdo con el sistema descrito en la figura 2. La unidad HFR 703 envía como salida una señal de banda ancha o banda alta 705 utilizando el procesamiento descrito en el presente documento.

En la figura 8, se describe con más detalle una realización del módulo de reconstrucción de alta frecuencia 703. La figura 8 ilustra que la generación de señal HF (alta frecuencia, «High Frequency») se puede derivar de diferentes módulos de generación de HF en diferentes instantes en el tiempo. La generación de HF se puede basar en un transpondedor de copia basado en QMF 803, o bien la generación de HF se puede basar en un transpondedor armónico basado en FFT 804. En el caso de ambos módulos de generación de señales HF, la señal de banda baja se procesa 801, 802 como parte de la generación de HF con el fin de determinar una curva de ganancia que se utiliza en el proceso de copia 803 o de transposición armónica 804. Las salidas de los dos transpondedores se envían selectivamente al ajustador de la envolvente 805. La decisión de qué señal de transpondedor se va a usar está controlada por el flujo de bits 704 o 708. Cabe señalar que, debido a la naturaleza de copia del transpondedor basado en QMF, la forma de la envolvente espectral de la señal de banda baja se mantiene más claramente que cuando se utiliza un transpondedor armónico. Esto típicamente dará como resultado discontinuidades más claras de la envolvente espectral de la señal de banda alta cuando se usen transpondedores de copia. Esto se ilustra en los paneles superior e inferior de la figura 1a. En consecuencia, puede ser suficiente con solo incorporar el ajuste de ganancia para el método de copia basado en QMF realizado en el módulo 803. No obstante, la aplicación del ajuste de ganancia para la transposición armónica realizada en el módulo 804 también puede resultar beneficiosa.

En la figura 9 se describe un módulo descodificador correspondiente. El codificador 901 puede estar configurado para analizar la señal de entrada particular 903 y determinar la cantidad de procesamiento de ajuste de ganancia que resulta adecuada para el tipo particular de señal de entrada 903. En particular, el codificador 901 puede determinar el grado de discontinuidad en la señal de subbanda de alta frecuencia que será provocada por la unidad HFR 703 en el descodificador. Para este propósito, el codificador 901 puede comprender una unidad HFR 703, o al menos partes relevantes de la unidad HFR 703. Sobre la base del análisis de la señal de entrada 903, se pueden generar los datos de control 905 para el descodificador correspondiente. La información 905, que se refiere al ajuste de ganancia que se ha de realizar en el descodificador, se combina en el multiplexor 902 con un flujo de bits de audio 906, formando de esta manera el flujo de bits completo 904 que se transmite al descodificador correspondiente.

En la figura 10, se muestran los espectros de salida de una señal del mundo real. En la figura 10a, se representa la salida de un descodificador MPEG USAC que descodifica un flujo de bits mono de 12 kbps. La sección de la señal de mundo real es una parte vocal de una grabación a capela. La abscisa corresponde al eje del tiempo, mientras que la ordenada corresponde al eje de la frecuencia. Comparando el espectrograma de la figura 10a con la figura 10c, que muestra el espectrograma correspondiente de la señal original, queda claro que aparecen espacios vacíos (véanse los números de referencia 1001, 1002) en el espectro correspondientes a las partes fricativas del segmento vocal. En la figura 10b se representa el espectrograma de la salida del descodificador MPEG USAC que incluye la presente invención. Se puede ver en el espectrograma que los espacios vacíos en el espectro han desaparecido (véanse los números de referencia 1003, 1004 correspondientes a los números de referencia 1001, 1002).

La complejidad del algoritmo de ajuste de ganancia propuesto se calculó como MOPS ponderados, en donde las funciones como POW/DIV/TRIG se ponderan como 25 operaciones, y todas las demás operaciones se ponderan como una operación. Dadas estas suposiciones, la complejidad calculada representa aproximadamente 0,1 WMOPS y un uso de RAM/ROM insignificante.

En otras palabras, el procesamiento de ajuste de ganancia propuesto requiere una baja capacidad de procesamiento y de memoria.

En el presente documento se han descrito un método y sistema para generar una señal de banda alta a partir de una señal de banda baja. El método y el sistema se adaptan para generar una señal de banda alta con pocas o ninguna discontinuidades espectrales, mejorando de esta manera el rendimiento perceptual de los métodos y sistemas de reconstrucción de alta frecuencia. El método y el sistema se pueden incorporar fácilmente en sistemas de codificación/descodificación de audio existentes. En particular, el método y el sistema se pueden incorporar sin necesidad de modificar el procesamiento de ajuste de envolvente de los sistemas de codificación/descodificación de audio existentes. De manera notable, esto se aplica a la funcionalidad de interpolación y limitación del procesamiento de ajuste de envolvente que puede realizar sus tareas previstas. Como tal, el método y el sistema descritos se pueden usar para regenerar señales de banda alta que tengan pocas o ninguna discontinuidades espectrales y un bajo nivel de ruido. Además, se ha descrito el uso de datos de control, en donde los datos de control se pueden usar para adaptar los parámetros del método y el sistema descritos (y la complejidad computacional) al tipo de señal de audio.

Los métodos y sistemas descritos en el presente documento pueden estar implementados como software, firmware y/o hardware. Ciertos componentes pueden, por ejemplo, estar implementados como software que se ejecuta en un procesador o microprocesador de señales digitales. Otros componentes pueden, por ejemplo, estar implementados como hardware y/o como circuitos integrados específicos de aplicación. Las señales que se encuentran en los métodos y sistemas descritos se pueden almacenar en soportes tales como memoria de acceso aleatorio o soportes de almacenamiento óptico. Estos pueden ser transferidos a través de redes, tales como redes de radio, redes de satélite, redes inalámbricas o redes por cable, por ejemplo, internet. Los dispositivos típicos que utilizan los métodos y sistemas descritos en el presente documento son dispositivos electrónicos portátiles u otros equipos de consumo que se utilizan para almacenar y/o procesar señales de audio. Los métodos y sistemas también se pueden usar en sistemas informáticos, por ejemplo, servidores web de internet, que almacenan y proporcionan señales de audio, por ejemplo, señales musicales, para su descarga.

REIVINDICACIONES

1. Un sistema (601, 703) configurado para generar una pluralidad de señales de subbanda de audio de alta frecuencia (604) que cubren un intervalo de alta frecuencia desde una pluralidad de señales de subbanda de baja frecuencia (602), el sistema (601, 703) comprende:

- medios para recibir la pluralidad de señales de subbanda de baja frecuencia (602);

- menos para recibir un conjunto de energías objetivo, cada energía objetivo cubre un intervalo objetivo diferente (130) dentro del intervalo de alta frecuencia y que es indicativo de la energía deseada de una o más señales de subbanda de alta frecuencia que se encuentran dentro del intervalo objetivo (130);

- medios para generar la pluralidad de señales de subbanda de alta frecuencia (604) a partir de la pluralidad de señales de subbanda de baja frecuencia (602) y a partir de una pluralidad de coeficientes de ganancia espectral asociados con la pluralidad de señales de subbanda de baja frecuencia (602), respectivamente; y

- medios para ajustar la energía (203) de la pluralidad de señales de subbanda de alta frecuencia (604) utilizando el conjunto de energías objetivo, en donde los medios para ajustar comprenden medios para limitar el ajuste de la energía de las señales de subbanda de alta frecuencia (604) que se encuentran en el intervalo limitador (135).

2. Un método para generar una pluralidad de señales de subbanda de alta frecuencia (604) que cubren un intervalo de alta frecuencia desde una pluralidad de señales de subbanda de audio de baja frecuencia (602), el método comprende:

- recibir la pluralidad de señales de subbanda de baja frecuencia (602);

- recibir un conjunto de energías objetivo, cada energía objetivo cubre un intervalo objetivo diferente (130) dentro del intervalo de alta frecuencia y que es indicativo de la energía deseada de una o más señales de subbanda de alta frecuencia (604) que se encuentran dentro del intervalo objetivo (130);

- generar la pluralidad de señales de subbanda de alta frecuencia (604) a partir de la pluralidad de señales de subbanda de baja frecuencia (602) y a partir de una pluralidad de coeficientes de ganancia espectral asociados con la pluralidad de señales de subbanda de baja frecuencia (602), respectivamente; y

- ajustar la energía de la pluralidad de señales de subbanda de alta frecuencia (604) utilizando el conjunto de energías objetivo, en donde ajustar la energía de la pluralidad de señales de subbanda de alta frecuencia comprende limitar el ajuste de la energía de las señales de subbanda de alta frecuencia (604) que se encuentran en el intervalo limitador (135).

3. Un soporte de almacenamiento que comprende un programa de software adaptado para su ejecución en un procesador y para realizar las etapas del método de la reivindicación 2 cuando se ejecuta en un dispositivo informático.

4. Un producto de programa informático que comprende instrucciones ejecutables para realizar el método de la reivindicación 2 cuando se ejecuta en un ordenador.

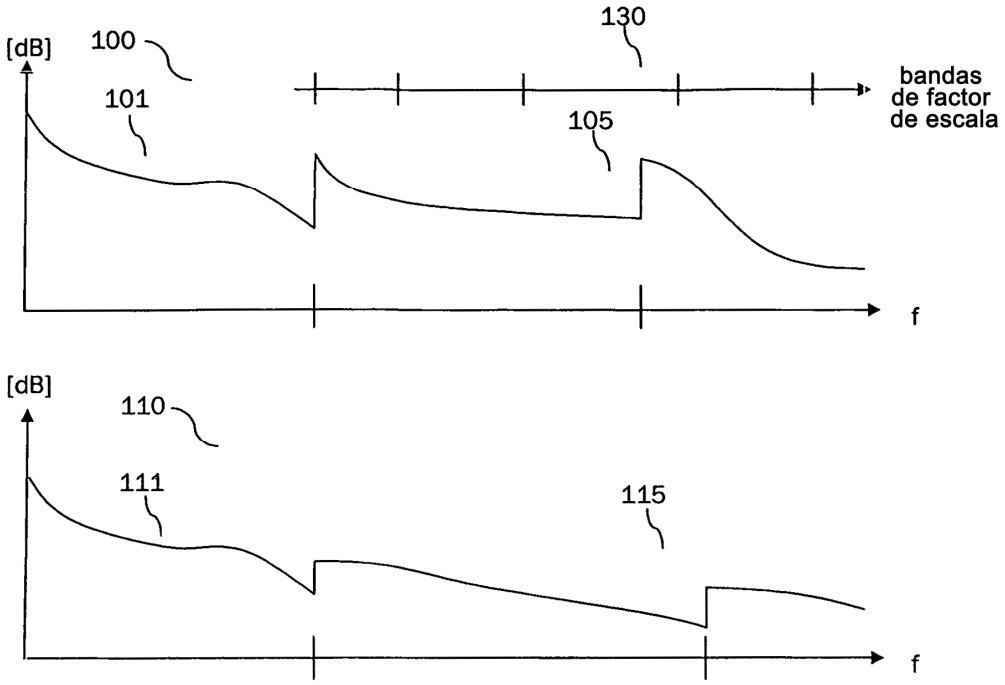

[dB]
100
130
bandas
de factor
de escala
101
105
f
[dB]
110
111
115
f

Fig. 1a

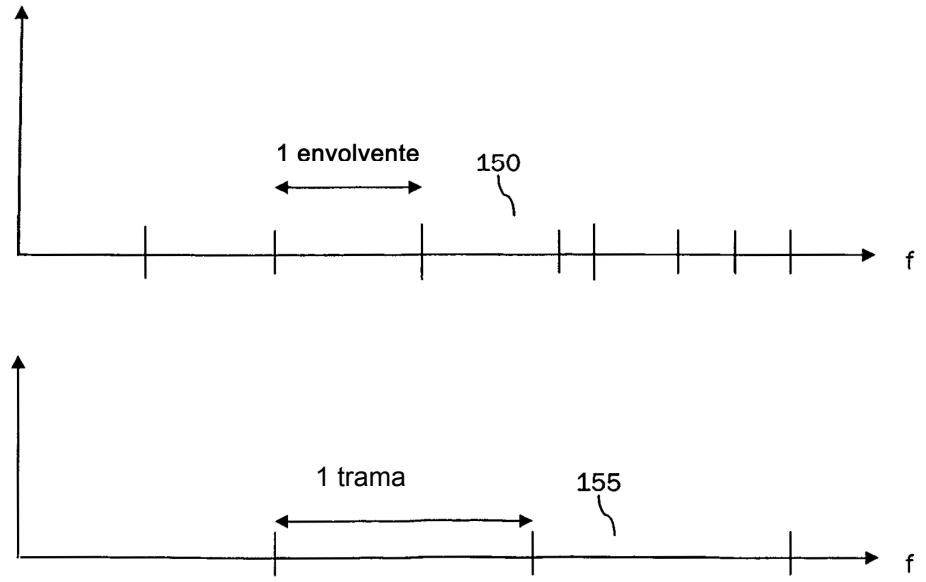

1 envolvente
150
f
1 trama
155
f

Fig. 1b

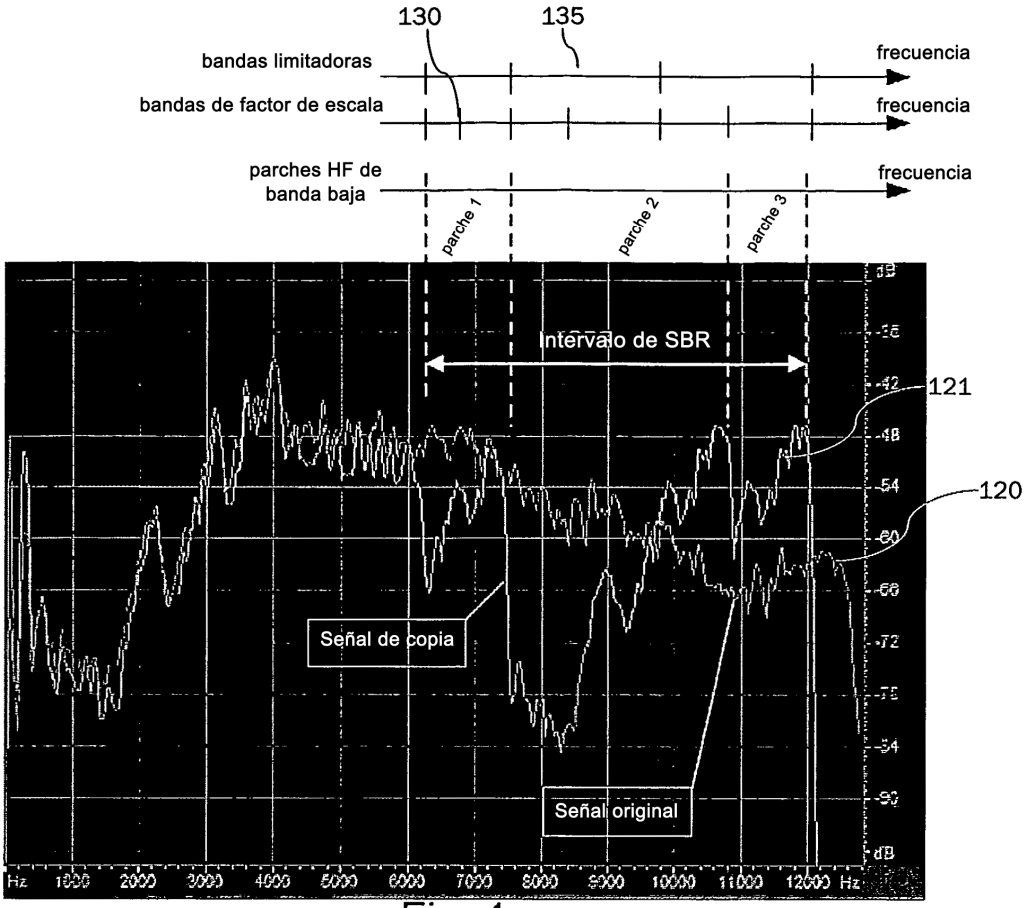

130
135
bandas limitadoras
frecuencia
bandas de factor de escala
frecuencia
parches HF de banda baja
frecuencia
parche 1
parche 2
parche 3
Intervalo de SBR
121
120
Señal de copia
Señal original
dB
Hz 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000 11000 12000 Hz

Fig. 1c

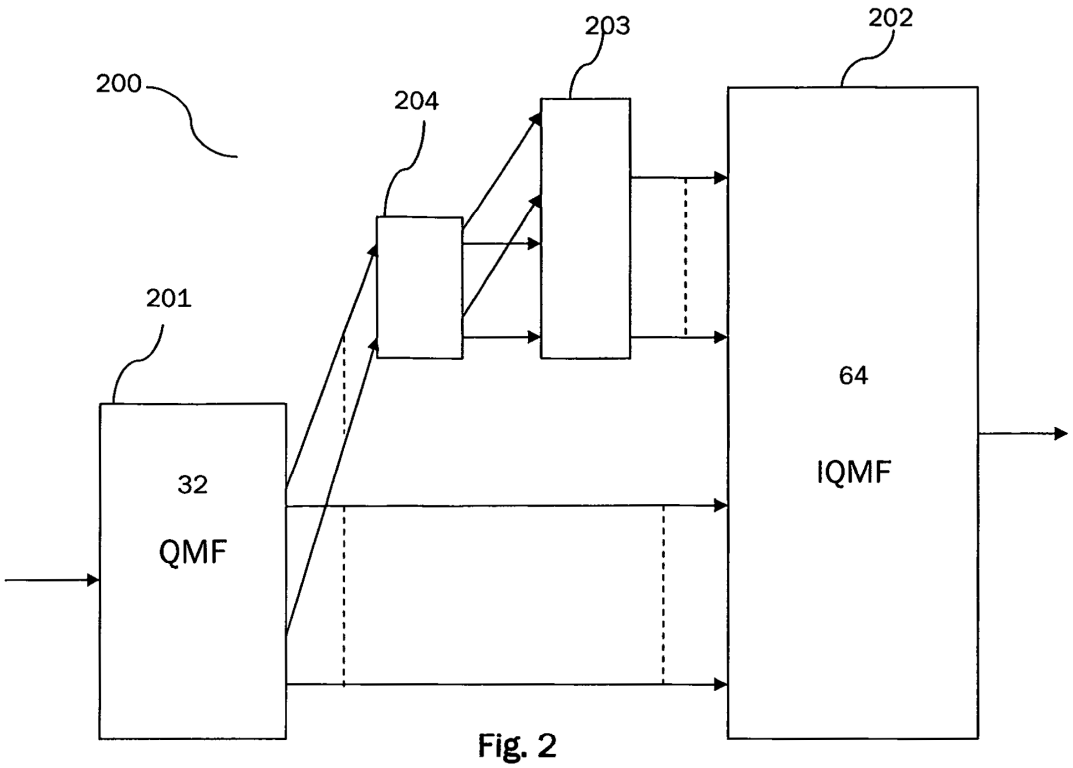

200
201
32
QMF
204
203
202
64
IQMF

Fig. 2

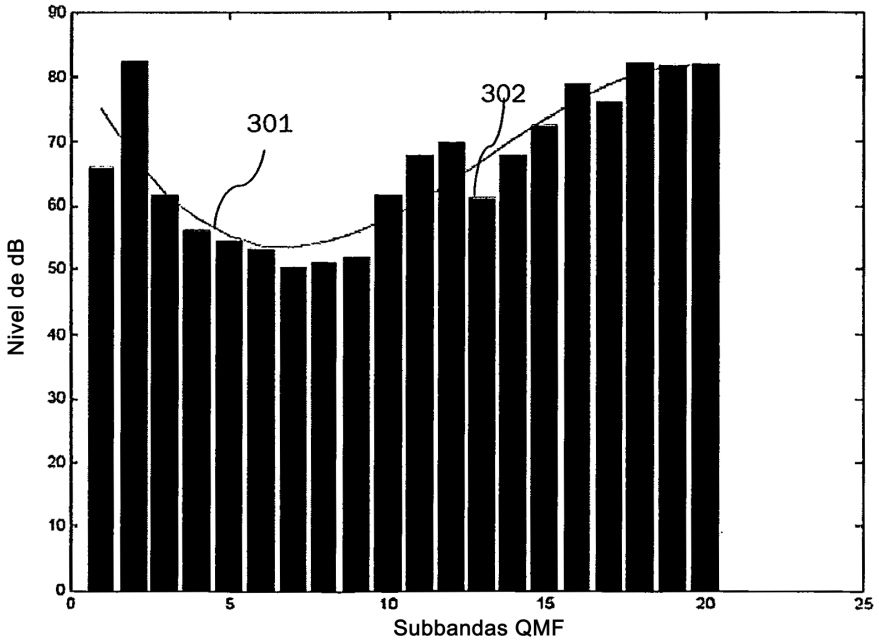

301
302
Nivel de dB
Subbandas QMF
0
10
20
30
40
50
60
70
80
90
0
5
10
15
20
25

Fig. 3

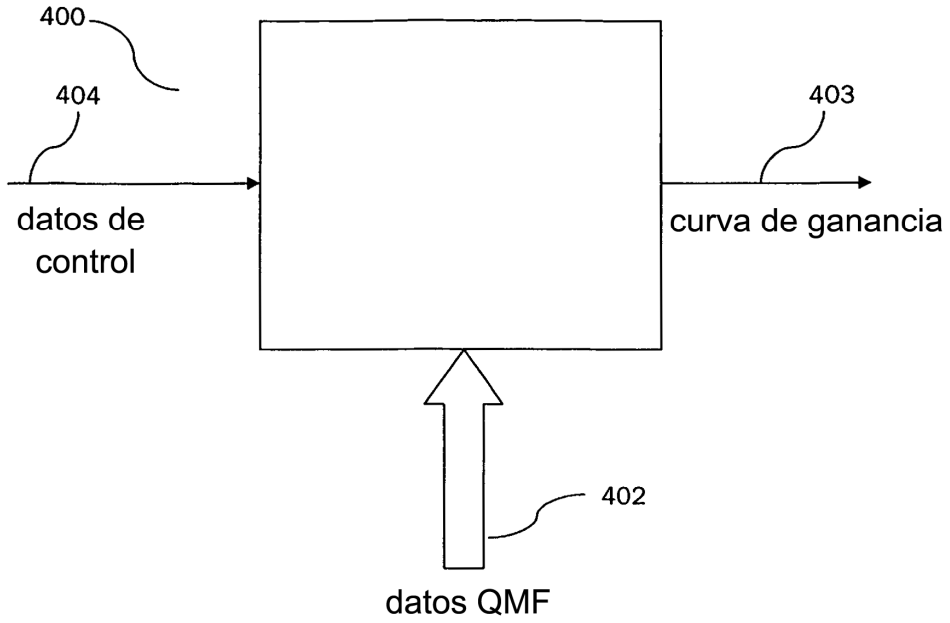

400
404
403
datos de control
curva de ganancia
402
datos QMF

Fig. 4

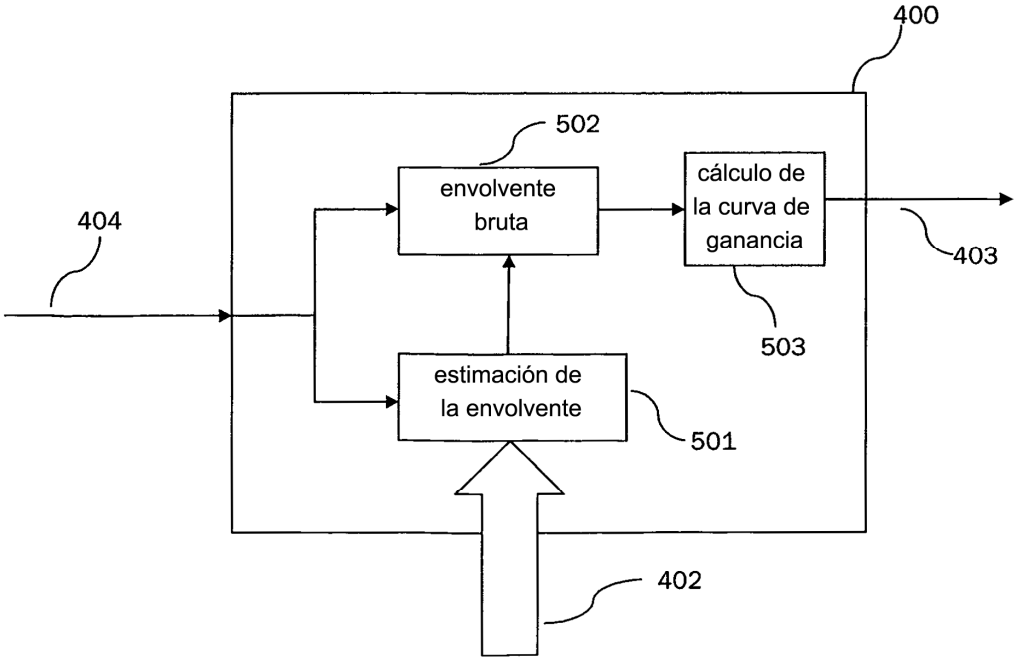

400
502
envolvente bruta
cálculo de la curva de ganancia
404
403
503
estimación de la envolvente
501
402

Fig. 5

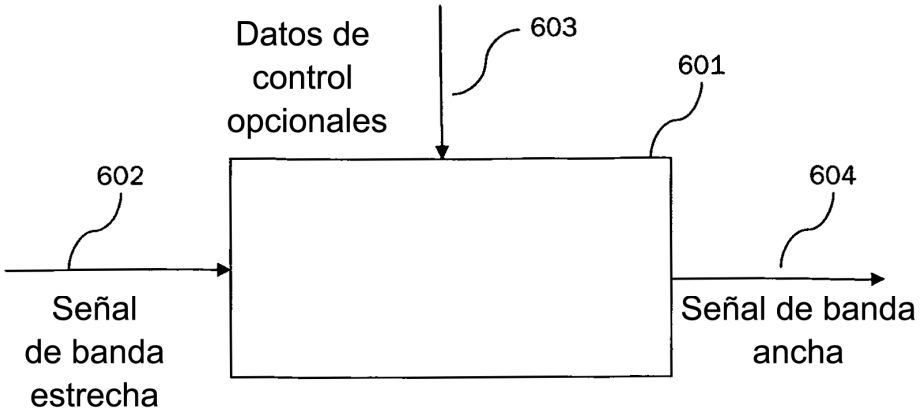

Datos de
control
opcionales
603
601
602
604
Señal
de banda
estrecha
Señal de banda
ancha

Fig. 6

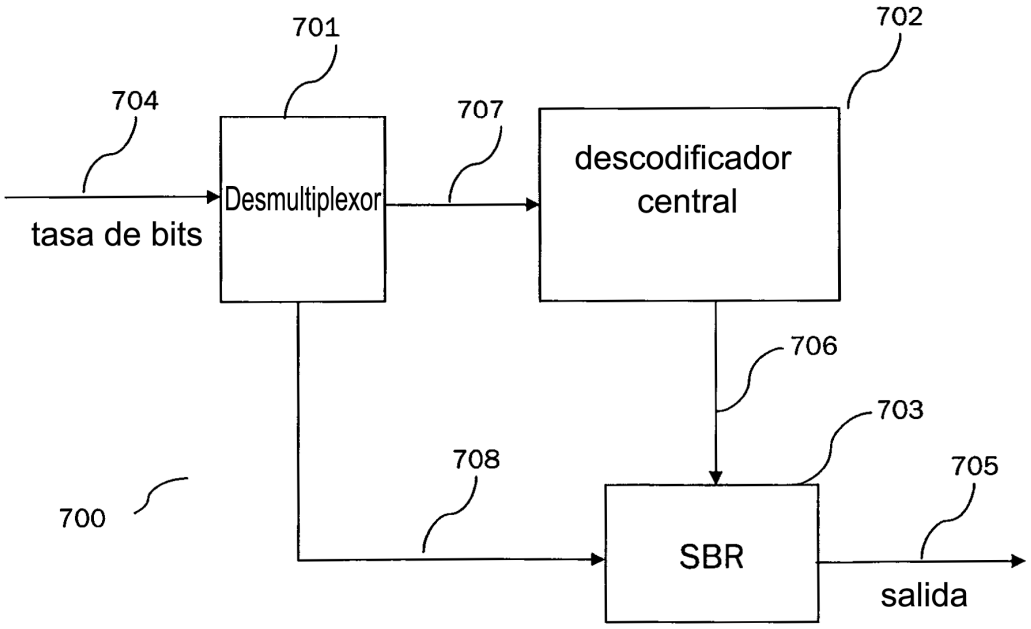

701
702
704
707
Desmultiplexor
descodificador
central
tasa de bits
706
703
708
705
700
SBR
salida

Fig. 7

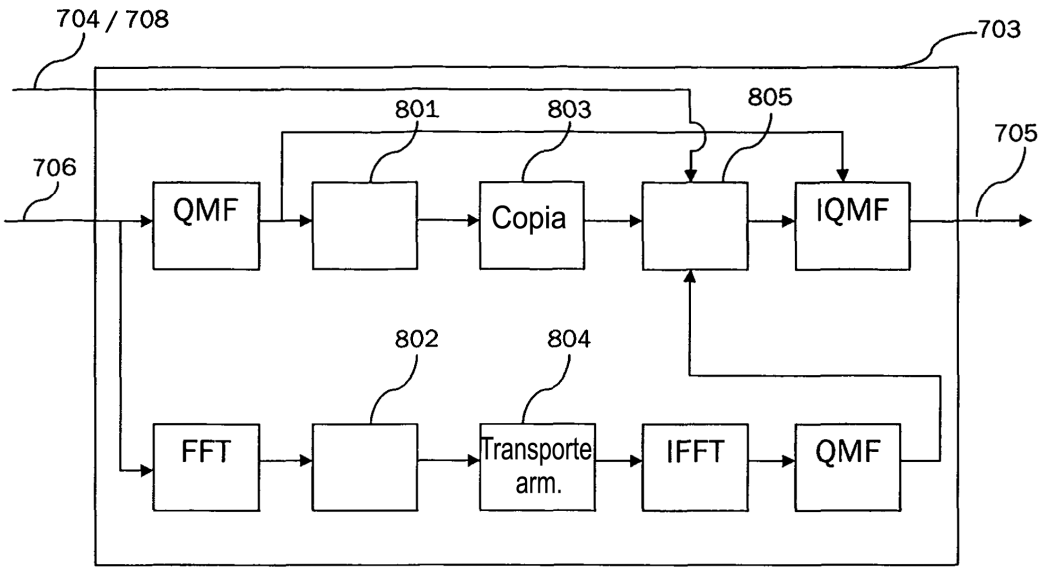

704 / 708
703
801
803
805
705
706
QMF
Copia
IQMF
802
804
FFT
Transporte arm.
IFFT
QMF

Fig. 8

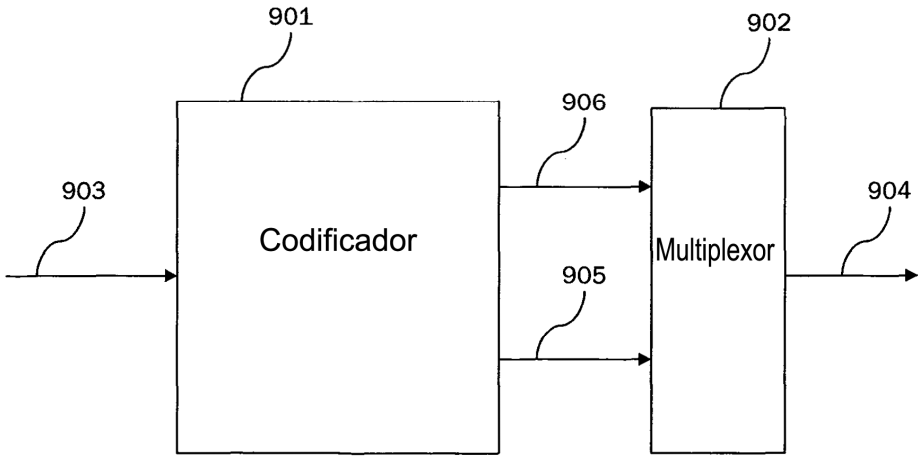

901
902
906
903
Codificador
Multiplexor
904
905

Fig. 9

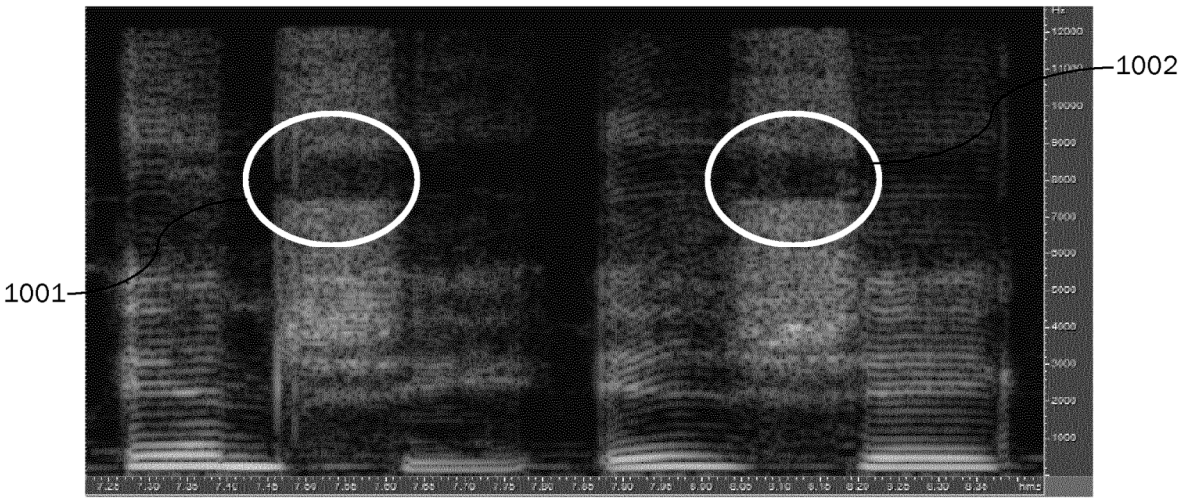


Fig. 10a

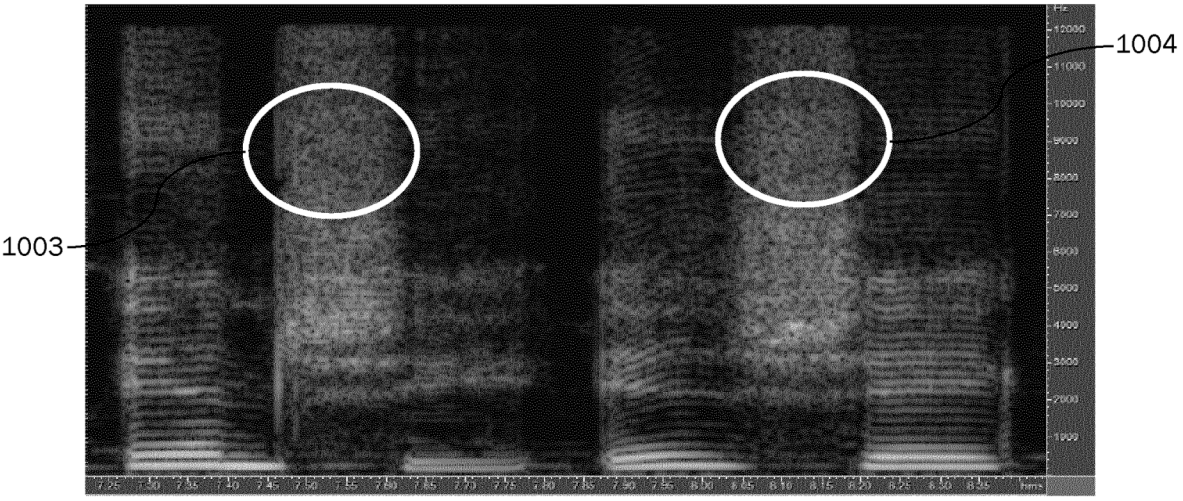


Fig. 10b

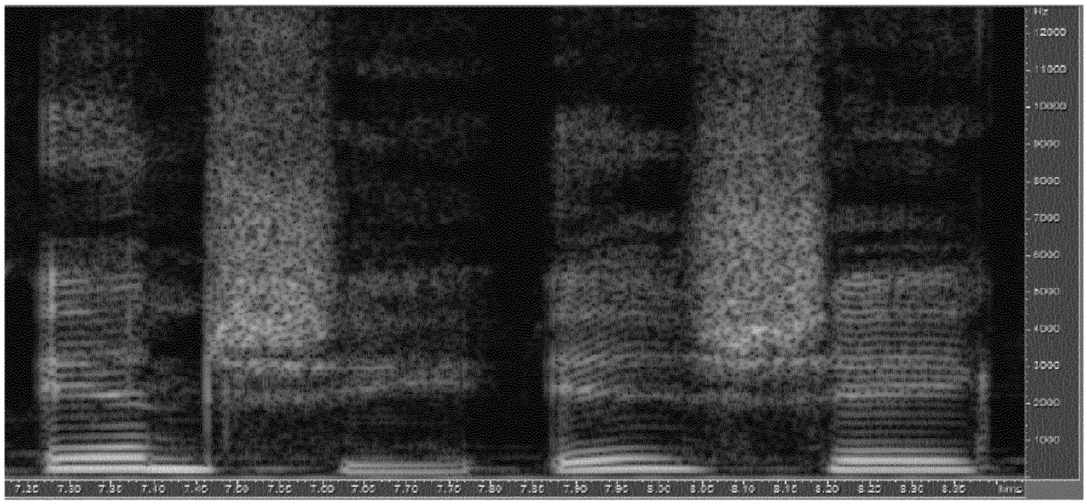

Hz
12000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000

Fig. 10c