(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 484 794**

(51) Int. CI.:

G10L 19/26 (2013.01)
G10L 19/02 (2013.01)
G10L 19/107 (2013.01)
G10L 19/20 (2013.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **23.06.2011 E 11743032 (2)**

(97) Fecha y número de publicación de la concesión europea: **28.05.2014 EP 2589046**

(54) Título: **Pos-filtro selectivo**

(30) Prioridad:

02.07.2010 US 361237 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
12.08.2014

(73) Titular/es:

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

(72) Inventor/es:

RESCH, BARBARA;
KJÖRLING, KRISTOFER y
VILLEMOES, LARS

(74) Agente/Representante:

PÉREZ BARQUÍN, Eliana

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

Pos-filtro selectivo

Campo técnico

La presente invención se refiere, en general, a la codificación de audio digital y, más precisamente, a técnicas de codificación para señales de audio que contienen componentes de distintos caracteres.

Antecedentes

Una extendida clase de métodos de codificación para señales de audio que contienen habla o canto incluye la predicción lineal excitada por código (CELP) aplicada en la alternación en el tiempo con distintos métodos de codificación, que incluyen métodos de codificación del dominio de la frecuencia, especialmente adaptados para la música, o métodos de naturaleza general, para admitir variaciones en el carácter entre sucesivos periodos de tiempo de la señal de audio. Por ejemplo, un descodificador simplificado de Codificación Unificada de Habla y Audio (USAC; véase la norma ISO / IEC 23003-3) del Grupo de Expertos en Imágenes en Movimiento (MPEG) es operable en al menos tres modalidades de descodificación, la Codificación Avanzada de Audio (AAC; véase la norma ISO / IEC 13818-7), la CELP algebraica (ACELP) y la excitación codificada por transformación (TCX), según se muestra en la parte superior de la figura 2 adjunta.

Las diversas realizaciones de la CELP están adaptadas a las propiedades de los órganos humanos del habla y, posiblemente, al sentido humano del oído. Según se usa en esta solicitud, la CELP se referirá a todas las posibles realizaciones y variantes, incluyendo, pero sin limitarse a, la ACELP, la CELP de banda ancha y estrecha, la SB-CELP (CELP de sub-banda), la CELP de baja y alta velocidad, la RCELP (CELP relajada), la LD-CELP (CELP de bajo retardo), la CS-CELP (CELP de estructura conjugada), la CS-ACELP (ACELP de estructura conjugada), la PSI-CELP (CELP de innovación sincrónica en tono) y la VSELP (predicción lineal excitada por suma vectorial). Los principios de la CELP son expuestos por R. Schroeder y S. Atal en los Anales de la Conferencia Internacional del IEEE sobre Acústica, Habla y Procesamiento de Señales (ICASSP), vol. 10, págs. 937 a 940, 1985, y algunas de sus aplicaciones están descritas en las referencias 25 a 29 citadas en la publicación de Chen y Gersho en las Transacciones del IEEE sobre Habla y Procesamiento de Audio, vol. 3, nº 1, de 1995. Según lo adicionalmente detallado en el artículo anterior, un descodificador de CELP (o, análogamente, un sintetizador de habla de CELP) puede incluir un predictor de tono, que restaura el componente periódico de una señal de habla codificada, y un libro de códigos de pulsos, desde el cual se añade una secuencia de innovación. El predictor de tono puede incluir, a su vez, un predictor de retardo largo para restaurar el tono y un predictor de retardo corto para restaurar las regiones formantes por medio de la modelación del envolvente espectral. En este contexto, el tono se concibe, en general, como la frecuencia fundamental del componente sonoro tonal producido por las cuerdas vocales, y adicionalmente realzado por las partes resonantes del conducto vocal. Esta frecuencia, junto con sus armónicos, dominará el habla o el canto. En términos generales, los métodos de CELP son los más adecuados para el procesamiento del canto solo o unipersonal, para el cual la frecuencia del tono está bien definida y es relativamente fácil de determinar.

Para mejorar la calidad percibida del habla codificada por CELP, es práctica común combinarla con el pos-filtrado (o realce tonal, en otras palabras). La Patente Estadounidense Nº 4.969.192 y la sección II del artículo de Chen y Gersho revelan propiedades deseables de tales pos-filtros, esto es, su capacidad de suprimir los componentes de ruido situados entre los armónicos del tono vocal detectado (parte a largo plazo, véase la sección IV). Se cree que una parte importante de este ruido emana del modelado del envolvente espectral. La parte a largo plazo de un pos-filtro sencillo puede ser diseñada para que tenga la siguiente función de transferencia:

$$H_E(z) = 1 + \alpha\left(\frac{z^T + z^{-T}}{2} - 1\right),$$

donde T es un periodo tonal estimado en términos del número de muestras y α es una ganancia del pos-filtro, según se muestra en las figuras 1 y 2. De manera similar a un filtro peine, un filtro de ese tipo atenúa las frecuencias 1/(2T), 3/(2T), 5/(2T), ..., que están situadas a mitad de camino entre los armónicos de la frecuencia tonal y las frecuencias adyacentes. La atenuación depende del valor de la ganancia α. Los pos-filtros levemente más sofisticados aplican esta atenuación solamente a las frecuencias bajas – de aquí, el término habitualmente usado de pos-filtro bajo – donde el ruido es más perceptible. Esto puede ser expresado aplicando en cascada la función de transferencia H_E descrita anteriormente y un filtro H_{LP} de paso bajo. Así, el valor descodificado pos-procesado de S_E proporcionado por el pos-filtro estará dado, en el dominio de transformación, por

$$S_E(z) = S(z) - \alpha S(z) P_{LT}(z) H_{LP}(z),$$

donde

$$P_{LT}(z) = 1 - \frac{z^{T} + z^{-T}}{2}$$

y S es la señal descodificada que es suministrada como entrada al pos-filtro. La Figura 3 muestra una realización de un pos-filtro con estas características, que es adicionalmente expuesto en la sección 6.1.3 de la Especificación Técnica ETSI TS 126 290, versión 6.3.0, edición 6. Como sugiere esta figura, la información tonal está codificada como un parámetro en la señal del flujo de bits y es extraída por un módulo de rastreo tonal, conectado comunicativamente con el filtro de predicción a largo plazo que lleva a cabo las operaciones expresadas por P_{LT}.

La parte a largo plazo descrita en el párrafo anterior puede ser usada sola. Alternativamente, se dispone en serie con un filtro modelador del ruido que preserva los componentes en los intervalos de frecuencia correspondientes a las regiones formantes y que atenúa el ruido en otras regiones espectrales (parte a corto plazo; véase la sección III), es decir, en los 'valles espectrales' de la envolvente de región formante. Como otra posible variación, este grupo de filtros está adicionalmente suplementado por un filtro de tipo de paso alto, para reducir un deterioro percibido, debido al declive espectral de la parte a corto plazo.

Los miembros de un comité de estandarización del MPEG pudieron acceder al documento XP00002658193 [M. Neuendorf (ed.): WD7 de USAC, 92ésima Reunión del MPEG, Dresde, N11299] en línea a partir del 26 de abril de 2010, y la Autoridad Examinadora Preliminar Internacional para esta solicitud ha expresado la opinión de que este documento estuvo públicamente disponible desde esa fecha. Tanto XP00002658193 como la Solicitud Internacional publicada como WO 99/38144 A1 revelan sistemas de procesamiento de audio con la aplicación selectiva del pos-filtrado.

Las señales de audio que contienen una mezcla de componentes de distintos orígenes – por ejemplo, tonal, no tonal, vocal, instrumental, no musical – no siempre son reproducidas por las tecnologías disponibles de codificación digital de manera satisfactoria. Más precisamente, ha sido observado que las tecnologías disponibles son deficientes en la manipulación de tal material de audio no homogéneo, favoreciendo generalmente a uno de los componentes en perjuicio del otro. En particular, la música que contiene canto acompañado por uno o más instrumentos o partes corales, que ha sido codificada por métodos de la naturaleza descrita anteriormente, será a menudo descodificada con distorsiones perceptibles que arruinan parte de la experiencia auditiva.

Resumen de la invención

A fin de mitigar al menos algunos de los inconvenientes esbozados en la sección anterior, es un objeto de la presente invención proporcionar métodos y dispositivos adaptados para la codificación y descodificación de audio de las señales que contienen una mezcla de componentes de distintos orígenes. Como objetos específicos, la invención busca proporcionar aquellos métodos y dispositivos que sean adecuados desde el punto de vista de la eficacia de codificación o la fidelidad (percibida) de reproducción, o ambas.

La invención logra al menos uno de estos objetos según lo definido en las reivindicaciones independientes. Las reivindicaciones dependientes definen realizaciones de la invención.

Los inventores han percibido que algunas distorsiones percibidas en las señales de audio descodificadas de origen no homogéneo provienen de una conmutación inadecuada entre varias modalidades de codificación, de las cuales al menos una incluye el pos-filtrado en el descodificador y al menos una no lo hace. Más precisamente, los pos-filtros disponibles eliminan no solamente el ruido inter-armónico (y, donde corresponda, el ruido en los valles espectrales), sino también los componentes de señales que representan acompañamiento instrumental o vocal, y otro material de naturaleza 'deseable'. El hecho de que la diferencia apenas perceptible en los valles espectrales puede ser tan grande como de 10 dB (según lo observado por Ghitza y Goldstein, Trans. del IEEE sobre Acústica, Habla y Procesamiento de Señales, vol. ASSP-4, págs. 697 a 708, de 1986) puede haber sido tomado como justificación por muchos diseñadores para filtrar con severidad estas bandas de frecuencia. La degradación de la calidad por la misma atenuación inter-armónica (y de valle espectral) puede, sin embargo, ser menos importante que la de las ocasiones de conmutación. Cuando el pos-filtro está activado, el fondo de una voz cantante suena repentinamente atenuado y, cuando el filtro es desactivado, el fondo se hace instantáneamente más sonoro. Si la conmutación tiene lugar frecuentemente, debido a la naturaleza de la señal de audio o a la configuración del dispositivo de codificación, habrá una distorsión de conmutación. Como ejemplo, un descodificador de USAC puede ser operable tanto en una modalidad de ACELP combinada con el pos-filtrado como en una modalidad de TCX sin pos-filtrado. La modalidad de ACELP se usa en episodios donde está presente un componente vocal dominante. De tal modo, la conmutación a la modalidad de ACELP puede ser activada por el inicio del canto, tal como al principio de una nueva frase musical, al comienzo de un nuevo verso, o simplemente después de un episodio donde se considera que el acompañamiento sofoca la voz cantante, en el sentido de que el componente vocal ya no es prominente. Los experimentos han confirmado que una solución alternativa, o más bien una elusión del problema, por la cual se usa la codificación TCX en toda su extensión (y la modalidad ACELP es inhabilitada), no remedia el problema, ya que aparecen distorsiones como de reverberación.

En consecuencia, en un primer y un segundo aspecto, la invención proporciona un método de codificación de audio (y un sistema de codificación de audio con las características correspondientes) caracterizado por la toma de una decisión en cuanto a si el dispositivo que descodificará el flujo de bits, que es emitido por el método de codificación, debería o no aplicar el pos-filtrado, incluyendo la atenuación del ruido inter-armónico. El resultado de la decisión es codificado en el flujo de bits y es accesible para el dispositivo de descodificación.

Por parte de la invención, la decisión en cuanto a usar o no el pos-filtro se toma por separado de la decisión en cuanto a la modalidad de codificación más adecuada. Esto posibilita mantener un estado de pos-filtrado en toda la extensión de un periodo de longitud tal que la conmutación no moleste al oyente. De tal modo, el método de codificación puede prescribir que el pos-filtro se mantendrá inactivo incluso aunque conmute a una modalidad de codificación donde el filtro esté convencionalmente activo.

Se hace notar que la decisión en cuanto a aplicar o no el pos-filtrado se toma normalmente trama a trama. De tal modo, en primer lugar, el pos-filtrado no es aplicado para menos de una trama a la vez. En segundo lugar, la decisión en cuanto a inhabilitar o no el pos-filtrado es solamente válida durante una trama actual, y puede ser mantenida, o bien re-evaluada, para la trama subsiguiente. En un formato de codificación que habilite un formato de trama principal y un formato reducido, que es una fracción del formato normal, por ejemplo, 1/8 de su longitud, puede no ser necesario tomar decisiones de pos-filtrado para tramas reducidas individuales. En cambio, puede ser considerado un cierto número de tramas reducidas que se acumulan hasta formar una trama normal, y los parámetros relevantes para la decisión de filtrado pueden ser obtenidos calculando la media o la mediana de las tramas reducidas comprendidas en la misma.

En un tercer y un cuarto aspecto de la invención, se proporciona un método de descodificación de audio (y un sistema de descodificación de audio con las características correspondientes) con una etapa de descodificación seguida por una etapa de pos-filtrado, que incluye la atenuación del ruido inter-armónico, y que está caracterizado por una etapa de inhabilitación del pos-filtro, de acuerdo a información de pos-filtrado codificada en la señal del flujo de bits.

Un método de descodificación con estas características está bien adaptado para la codificación de señales de audio de origen mixto, en virtud de su capacidad de desactivar el pos-filtro según la información de pos-filtrado solamente y, por tanto, independientemente de factores tales como la modalidad actual de codificación. Cuando se aplica a técnicas de codificación en las cuales la actividad de pos-filtro está convencionalmente asociada a modalidades específicas de codificación, la capacidad de inhabilitación del pos-filtrado habilita una nueva modalidad operativa, a saber, la aplicación no filtrada de una modalidad de descodificación convencionalmente filtrada.

En un aspecto adicional, la invención también proporciona un producto de programa de ordenador para realizar uno de los métodos anteriores. Y aún más, la invención proporciona un pos-filtro para atenuar el ruido inter-armónico, que es operable ya sea en una modalidad activa o en una modalidad de traspasamiento, según lo indicado por una señal de pos-filtrado proporcionada al pos-filtro. El pos-filtro puede incluir una sección de decisión para controlar autónomamente la actividad de pos-filtrado.

Como apreciará el experto, un codificador adaptado para cooperar con un descodificador está equipado con módulos funcionalmente equivalentes, a fin de habilitar la reproducción fiel de la señal codificada. Tales módulos equivalentes pueden ser módulos idénticos o similares, o módulos con idénticas o similares características de transferencia. En particular, los módulos en el codificador y el descodificador, respectivamente, pueden ser unidades de procesamiento similares o disímiles, que ejecutan respectivos programas de ordenador que realizan conjuntos equivalentes de operaciones matemáticas.

En una realización, la codificación del presente método incluye la toma de una decisión en cuanto a un pos-filtro que incluye adicionalmente la atenuación de valles espectrales (con respecto a la envolvente de la región formante, véase lo precedente). Esto corresponde a la parte a corto plazo del pos-filtro. Es entonces ventajoso adaptar el criterio, sobre el cual se basa la decisión, a la naturaleza del pos-filtro.

Una realización está orientada a un codificador específicamente adaptado para la codificación del habla. Como algunos de los problemas que motivan la invención han sido observados cuando se codifica una mezcla de componentes vocales y otros, la combinación de codificación del habla y la toma de decisión independiente con respecto al pos-filtrado, prestada por la invención, es particularmente ventajosa. En particular, un descodificador de ese tipo puede incluir un módulo de codificación de predicción lineal excitada por código.

En una realización, el codificador basa su decisión en una presencia simultánea detectada de un componente de señal con frecuencia fundamental dominante (tono) y de otro componente de señal situado por debajo de la frecuencia fundamental. La detección también puede estar orientada a hallar la co-ocurrencia de un componente con frecuencia fundamental dominante y de otro componente con energía entre los armónicos de esta frecuencia fundamental. Esta es una situación en la cual se encuentran con frecuencia distorsiones del tipo en consideración. Por tanto, si se establece tal presencia simultánea, el codificador decidirá que el pos-filtrado no es adecuado, lo que

será indicado, en consecuencia, por la información de pos-filtrado contenida en el flujo de bits.

Una realización usa como su criterio de detección el contenido total de potencia de señal en la señal de tiempo de audio por debajo de una frecuencia tonal, posiblemente una frecuencia tonal estimada por una predicción a largo plazo en el codificador. Si esto es mayor que un umbral predeterminado, se considera que hay otros componentes relevantes además del componente tonal (incluyendo los armónicos), que provocarán que el pos-filtro sea inhabilitado.

En un codificador que comprende un módulo de CELP, puede hacerse uso del hecho de que un módulo de ese tipo estima la frecuencia tonal de la señal de tiempo de audio. Entonces, un criterio adicional de detección es comprobar el contenido de energía entre, o por debajo de, los armónicos de esta frecuencia, según lo descrito en más detalle en lo que antecede.

Como un desarrollo adicional de la realización precedente, que incluye un módulo de CELP, la decisión puede incluir una comparación entre una potencia estimada de la señal de audio cuando es codificada con CELP (es decir, codificada y descodificada) y una potencia estimada de la señal de audio cuando es codificada con CELP y pos-filtrada. Si la diferencia de potencia es mayor que un umbral, lo que puede indicar que se perderá un componente relevante, no de ruido, de la señal, el codificador decidirá inhabilitar el pos-filtro.

En una realización ventajosa, el codificador comprende un módulo de CELP y un módulo de TCX. Como es conocido en la tecnología, la codificación TCX es ventajosa con respecto a ciertas clases de señales, notablemente, las señales no vocales. No es práctica común aplicar el pos-filtrado a una señal codificada por TCX. Por tanto, el codificador puede seleccionar la codificación TCX, la codificación CELP con pos-filtrado o bien la codificación CELP sin pos-filtrado, abarcando por ello una gama considerable de tipos de señal.

Como un desarrollo adicional de la realización precedente, la decisión entre las tres modalidades de codificación se toma sobre la base de un criterio de velocidad-distorsión, es decir, aplicando un procedimiento de optimización conocido per se en la técnica.

En otro desarrollo adicional de la realización precedente, el codificador comprende adicionalmente un codificador de Codificación Avanzada de Audio (AAC), del que también se sabe que es particularmente adecuado para ciertos tipos de señales. Preferiblemente, la decisión en cuanto a aplicar o no la codificación AAC (del dominio de frecuencia) se toma por separado de la decisión en cuanto a cuál usar de las otras modalidades (de predicción lineal). De tal modo, el codificador puede ser percibido como operable en dos super-modalidades, AAC o TCX/CELP, en el segundo caso de las cuales el codificador seleccionará entre TCX, CELP pos-filtrada y CELP no filtrada. Esta realización habilita el procesamiento de una gama incluso más amplia de tipos de señales de audio.

En una realización, el codificador puede decidir que un pos-filtrado en la descodificación ha de aplicarse gradualmente, es decir, con ganancia gradualmente creciente. Análogamente, puede decidir que el pos-filtrado ha de ser eliminado gradualmente. Tal aplicación y eliminación gradual hace que la conmutación entre regímenes con y sin pos-filtrado sea menos perceptible. Como ejemplo, un episodio de canto, para el cual se halla que la codificación de CELP pos-filtrada es adecuada, puede ser precedido por un episodio instrumental, en el cual la codificación TCX es óptima; un descodificador según la invención puede entonces aplicar el pos-filtrado gradualmente en, o cerca de, el principio del episodio de canto, de modo que las ventajas del pos-filtrado sean preservadas incluso aunque sean evitadas las molestas distorsiones de conmutación.

En una realización, la decisión en cuanto a si ha de aplicarse o no el pos-filtrado se basa en una señal de diferencia aproximada, que se aproxima a aquel componente de señal que ha de ser eliminado de una futura señal descodificada por el pos-filtrado. Como una opción, la señal de diferencia aproximada es calculada como la diferencia entre la señal de tiempo de audio y la señal de tiempo de audio cuando se somete al pos-filtrado (simulado). Como otra opción, una sección de codificación extrae una señal descodificada intermedia, por lo cual la señal de diferencia aproximada puede ser calculada como la diferencia entre la señal de tiempo de audio y la señal descodificada intermedia cuando se somete al pos-filtrado. La señal descodificada intermedia puede ser almacenada en un almacén temporal de predicción a largo plazo del codificador. Puede representar además la excitación de la señal, lo que implica que se necesitaría aplicar un filtrado adicional de síntesis (conducto vocal, resonancias) para obtener la señal descodificada final. El motivo para usar una señal descodificada intermedia es que captura algunas de las particularidades, notablemente, las debilidades, del método de codificación, permitiendo por ello una estimación más realista del efecto del pos-filtro. Como una tercera opción, una sección de descodificación extrae una señal descodificada intermedia, por lo cual la señal de diferencia aproximada puede ser calculada como la diferencia entre la señal descodificada intermedia y la señal descodificada intermedia cuando se somete al pos-filtrado. Este procedimiento da, probablemente, una estimación menos fiable que las dos primeras opciones, pero, por otra parte, puede ser llevado a cabo por el descodificador de forma autónoma.

La señal de diferencia aproximada así obtenida es luego evaluada con respecto a uno de los siguientes criterios, que, cuando se decantan por la afirmativa, llevarán a una decisión para inhabilitar el pos-filtro.

a) si la potencia de la señal de diferencia aproximada supera o no un umbral predeterminado, lo que indica que una parte significativa de la señal sería eliminada por el pos-filtro;

b) si el carácter de la señal de diferencia aproximada es más bien tonal, antes que similar al ruido;

c) si una diferencia entre los espectros de frecuencia de magnitud de la señal de diferencia aproximada y los de la señal de tiempo de audio está o no desigualmente distribuida con respecto a la frecuencia, lo que sugiere que no es ruido sino, más bien, una señal que tendría sentido para un oyente humano;

d) si un espectro de frecuencia de magnitud de la señal de diferencia aproximada está o no localizado en intervalos de frecuencia dentro de un envolvente de relevancia predeterminada, en base a lo que puede esperarse habitualmente de una señal del tipo a procesar; y

e) si un espectro de frecuencia de magnitud de la señal de diferencia aproximada está o no localizado en intervalos de frecuencia dentro de un envolvente de relevancia obtenido nivelando un espectro de frecuencia de magnitud de la señal de tiempo de audio con una magnitud del mayor componente de señal en el mismo, reducido a escala en un factor de escala predeterminado.

Al evaluar el criterio e), es ventajoso aplicar el rastreo de máximos en el espectro de magnitud, es decir, distinguir partes con formas similares a picos, normalmente asociadas a componentes tonales, antes que al ruido. Los componentes identificados por el rastreo de máximos, que puede tener lugar por medio de algún algoritmo conocido per se en la técnica, pueden ser clasificados adicionalmente aplicando un umbral a la altura de los picos, por lo cual los componentes restantes son material tonal de una cierta magnitud. Tales componentes representan habitualmente contenido de señal relevante, antes que ruido, lo que motiva una decisión para inhabilitar el pos-filtro.

En una realización de la invención como un descodificador, la decisión de inhabilitar el pos-filtro es ejecutada por un conmutador controlable por la sección de control, y capaz de omitir el pos-filtro en el circuito. En otra realización, el pos-filtro tiene ganancia variable, controlable por la sección de control, o un controlador de ganancia en el mismo, en el cual la decisión de inhabilitar es llevada a cabo fijando la ganancia del pos-filtro (véase la sección anterior) en cero o fijando su valor absoluto por debajo de un umbral predeterminado.

En una realización, la descodificación de acuerdo a la presente invención incluye extraer información de pos-filtrado de la señal del flujo de bits que está siendo descodificada. Más precisamente, la información de pos-filtrado puede ser codificada en un campo de datos que comprende al menos un bit en un formato adecuado para su transmisión. Ventajosamente, el campo de datos es un campo existente definido por un estándar aplicable pero no en uso, de modo que la información de pos-filtrado no aumente la carga útil a transmitir.

Se hace notar que los métodos y aparatos revelados en esta sección pueden ser aplicados, después de modificaciones adecuadas dentro de las capacidades del experto, que incluyen la experimentación rutinaria, a la codificación de señales con varios componentes, posiblemente correspondientes a distintos canales, tales como canales estéreo. A lo largo de la presente solicitud, el realce tonal y el pos-filtrado son usados como sinónimos. Se hace notar además que la AAC está expuesta como un ejemplo representativo de métodos de codificación del dominio de la frecuencia. En efecto, la aplicación de la invención a un descodificador o codificador operable en una modalidad de codificación del dominio de frecuencia, distinta a la AAC, solamente requerirá pequeñas modificaciones, si acaso, dentro de las capacidades del experto. De manera similar, TCX se menciona como un ejemplo de codificación por transformación de predicción lineal ponderada y de codificación por transformación en general.

Las características de dos o más realizaciones descritas anteriormente en la presente memoria pueden ser combinadas, a menos que sean claramente complementarias, en realizaciones adicionales. El hecho de que dos características sean mencionadas en distintas reivindicaciones no excluye que puedan ser combinadas con ventaja. Análogamente, también pueden ser proporcionadas realizaciones adicionales por la omisión de ciertas características que no sean necesarias o no sean esenciales para el propósito deseado.

Breve descripción de los dibujos

Se describirán ahora realizaciones de la presente invención con referencia a los dibujos adjuntos, en los cuales:

la figura 1 es un diagrama de bloques que muestra un descodificador convencional con pos-filtro;

la figura 2 es un diagrama de bloques esquemático de un descodificador convencional operable en modalidad de AAC, ACELP y TCX, y que incluye un pos-filtro permanentemente conectado aguas abajo del módulo de ACELP;

la figura 3 es un diagrama de bloques que ilustra la estructura de un pos-filtro;

las figuras 4 y 5 son diagramas de bloques de dos descodificadores de acuerdo a la invención;

las figuras 6 y 7 son diagramas de bloques que ilustran diferencias entre un descodificador convencional (figura 6) y un descodificador (figura 7) de acuerdo a la invención;

la figura 8 es un diagrama de bloques de un codificador de acuerdo a la invención;

las figuras 9 y 10 son diagramas de bloques que ilustran diferencias entre un descodificador convencional (figura 9) y un descodificador (figura 10) de acuerdo a la invención; y

la figura 11 es un diagrama de bloques de un pos-filtro autónomo que puede ser activado y desactivado selectivamente.

Descripción detallada de realizaciones

La figura 4 es un dibujo esquemático de un sistema descodificador 400 de acuerdo a una realización de la invención, que tiene como su entrada una señal de flujo de bits y como su salida una señal de audio. Como en los descodificadores convencionales mostrados en la figura 1, un pos-filtro 440 está dispuesto aguas abajo de un módulo descodificador 410, pero puede ser incorporado a, o apartado de, el trayecto de descodificación, operando un conmutador 442. El pos-filtro está habilitado en la posición de conmutación mostrada en la figura. Sería inhabilitado si el conmutador se fijara en la posición opuesta, por lo cual la señal del módulo descodificador 410, en cambio, sería conducida por la línea 444 de omisión. Como una contribución inventiva, el conmutador 442 es controlable por la información de pos-filtrado contenida en la señal del flujo de bits, de modo que el pos-filtrado pueda ser aplicado y eliminado, independientemente del estado actual del módulo descodificador 410. Debido a que un pos-filtro 440 funciona con algún retardo – por ejemplo, el pos-filtro mostrado en la figura 3 introducirá un retardo equivalente al menos al periodo tonal T – un módulo 443 de retardo de compensación está dispuesto en la línea 444 de omisión para mantener los módulos en una condición sincronizada en la conmutación. El módulo 443 de retardo retarda la señal en el mismo periodo que lo haría el pos-filtro 440, pero no procesa de otro modo la señal. Para minimizar el tiempo de traspaso, el módulo 443 de retardo de compensación recibe la misma señal que el pos-filtro 440 en todo momento. En una realización alternativa, donde el pos-filtro 440 es reemplazado por un pos-filtro de retardo cero (por ejemplo, un filtro causal, tal como un filtro con dos tomas, independiente de futuros valores de señal), el módulo 443 de retardo de compensación puede ser omitido.

La figura 5 ilustra un desarrollo adicional de acuerdo a las revelaciones de la invención del sistema descodificador 500 de triple modalidad de la figura 2. Un módulo descodificador 511 de ACELP está dispuesto en paralelo con un módulo descodificador 512 de TCX y un módulo descodificador 513 de AAC. En serie con el módulo descodificador 511 de ACELP está dispuesto un pos-filtro 540 para atenuar el ruido, en particular, el ruido situado entre armónicos de una frecuencia tonal, directamente o indirectamente obtenible de la señal del flujo de bits para el cual está adaptado el sistema descodificador 500. La señal del flujo de bits también codifica información de pos-filtrado que gobierna las posiciones del conmutador superior 541 operable para conmutar el pos-filtro 540 fuera del trayecto de procesamiento y reemplazarlo por un retardo 543 de compensación como en la figura 4. Un conmutador inferior 542 se usa para conmutar entre distintas modalidades de descodificación. Con esta estructura, la posición del conmutador superior 541 es irrelevante cuando se usa uno de los módulos 512, 513 de TCX o AAC; por tanto, la información de pos-filtrado no necesariamente indica esta posición, excepto en la modalidad de ACELP. Cualquiera que sea la modalidad de descodificación que se está usando actualmente, la señal es suministrada desde el punto de conexión aguas abajo del conmutador inferior 542 hasta un módulo 550 de copia de banda espectral (SBR), que emite una señal de audio. El experto se dará cuenta de que el dibujo es de naturaleza conceptual, como queda notablemente claro a partir de los conmutadores, que se muestran esquemáticamente como entidades físicas distintas con medios móviles de contacto. En una posible implementación realista del sistema descodificador, los conmutadores, así como los otros módulos, estarán realizados por instrucciones legibles por ordenador.

Las figuras 6 y 7 también son diagramas de bloques de dos sistemas descodificadores de triple modalidad, operables en una modalidad de descodificación ACELP, TCX o del dominio de frecuencia. Con referencia a la última figura, que muestra una realización de la invención, una señal del flujo de bits es suministrada a un punto 701 de entrada, que está a su vez permanentemente conectado, mediante los ramales respectivos, con los tres módulos descodificadores 711, 712, 713. El punto 701 de entrada también tiene un ramal conector 702 (no presente en el sistema descodificador convencional de la figura 6) con un módulo 740 de realce tonal, que actúa como un pos-filtro del tipo general descrito anteriormente. Como es práctica común en la técnica, un primer módulo 703 de ventanas de transición está dispuesto aguas abajo de los módulos 711, 712 de ACELP y TCX, para llevar a cabo las transiciones entre los módulos descodificadores. Un segundo módulo 704 de transición está dispuesto aguas abajo del módulo descodificador 713 del dominio de frecuencia y del primer módulo 703 de ventanas de transición, para llevar a cabo la transición entre las dos super-modalidades. Además, un módulo 750 de SBR está proporcionado inmediatamente aguas arriba del punto 705 de salida. Es claro que la señal de flujo de bits es suministrada directamente (o después del demultiplexado, según corresponda) a los tres módulos descodificadores 711, 712, 713 y al módulo 740 de realce tonal. La información contenida en el flujo de bits controla qué módulo de descodificación ha de estar activo. Por parte de la invención, sin embargo, el módulo 740 de realce tonal realiza una análoga auto-activación que, en respuesta a información de pos-filtrado en el flujo de bits, puede actuar como un pos-filtro o,

simplemente, como un dispositivo de traspasamiento. Esto, por ejemplo, puede ser realizado mediante la provisión de una sección de control (no mostrada) en el módulo 740 de realce tonal, por medio del cual la acción de pos-filtrado puede ser activada o desactivada. El módulo 740 de realce tonal está siempre en su modalidad de traspasamiento cuando el sistema descodificador funciona en el dominio de frecuencia o la modalidad de descodificación TCX, en donde, en rigor, no es necesaria ninguna información de pos-filtrado. Se entiende que los módulos que no forman parte de la contribución inventiva, y cuya presencia es obvia para el experto, por ejemplo, un demultiplexador, han sido omitidos de la figura 7 y de otros dibujos similares, para aumentar la claridad.

Como variación, el sistema descodificador de la figura 7 puede estar equipado con un módulo de control (no mostrado) para decidir si ha de aplicarse o no el pos-filtrado, usando un enfoque de análisis por síntesis. Tal módulo de control está comunicativamente conectado con el módulo 740 de realce tonal y con el módulo 711 de ACELP, desde el cual extrae una señal descodificada intermedia $s_{i_DEC}(n)$ que representa una etapa intermedia en el proceso de descodificación, preferiblemente, una correspondiente a la excitación de la señal. El módulo de detección tiene la información necesaria para simular la acción del módulo 740 de realce tonal, según lo definido por las funciones de transferencia $P_{LT}(z)$ y $H_{LP}(z)$ (véase la sección Antecedentes y la figura 3) o, equivalentemente, sus respuestas de impulso de filtro $p_{LT}(z)$ y $h_{LP}(n)$. Como se desprende de la exposición en la sección Antecedentes, el componente a restar en el pos-filtrado puede ser estimado por una señal de diferencia aproximada $s_{AD}(n)$ que es proporcional a $[(s_{i_DEC} * P_{LT}) * h_{LP}](n)$, donde * indica la convolución discreta. Esto es una aproximación de la verdadera diferencia entre la señal de audio original y la señal descodificada pos-filtrada, a saber,

$$s_{ORIG}(n) - s_E(n) = s_{ORIG}(n) - (s_{DEC}(n) - \alpha[s_{DEC} * p_{LT} * h_{LP}](n)),$$

donde α es la ganancia del pos-filtro. Estudiando la energía total, la energía de banda baja, la tonalidad, el espectro efectivo de magnitudes o los espectros de magnitudes pasadas de la señal, según lo revelado en la sección Resumen y las reivindicaciones, la sección de control puede hallar una base para la decisión de activar o desactivar el módulo 740 de realce tonal.

La figura 8 muestra un sistema codificador 800 de acuerdo a una realización de la invención. El sistema codificador 800 está adaptado para procesar señales de audio digital, que son generalmente obtenidas capturando una onda sonora por parte de un micrófono, y transduciendo la onda hacia una señal eléctrica analógica. La señal eléctrica es luego muestreada en una señal digital susceptible de ser suministrada, en un formato adecuado, al sistema codificador 800. El sistema consiste, en general, en un módulo codificador 810, un módulo 820 de decisión y un multiplexador 830. En virtud de los conmutadores 814, 815 (simbólicamente representados), el módulo codificador 810 es operable en una modalidad CELP, TCX o bien AAC, activando selectivamente los módulos 811, 812, 813. El módulo 820 de decisión aplica uno o más criterios predefinidos para decidir si se inhabilita o no el pos-filtrado durante la descodificación de una señal de flujo de bits producida por el sistema codificador 800 para codificar una señal de audio. Con este fin, el módulo 820 de decisión puede examinar la señal de audio directamente o puede recibir datos desde el módulo codificador 810 mediante una línea 816 de conexión. Una señal indicativa de la decisión adoptada por el módulo 820 de decisión es suministrada, junto con la señal de audio codificada proveniente del módulo codificador 810, a un multiplexador 830, que concatena las señales en un flujo de bits que constituye la salida del sistema codificador 800.

Preferiblemente, el módulo 820 de decisión basa su decisión sobre una señal de diferencia aproximada calculada a partir de una señal descodificada intermedia s_{i_DEC}, que puede ser restada del módulo codificador 810. La señal descodificada intermedia representa una etapa intermedia en el proceso de descodificación, según lo expuesto en los párrafos precedentes, pero puede ser extraída desde una etapa correspondiente del proceso de codificación. Sin embargo, en el sistema codificador 800 está disponible la señal de audio original s_{ORIG}, de modo que, ventajosamente, la señal de diferencia aproximada se forme como:

$$s_{ORIG}(n) - (s_{i_DEC}(n) - \alpha\,[(s_{i_DEC} * p_{LT}) * h_{LP}](n)).$$

La aproximación reside en el hecho de que la señal descodificada intermedia se usa en lugar de la señal descodificada final. Esto permite una valoración de la naturaleza del componente que un pos-filtro eliminaría en la descodificación y, aplicando uno de los criterios expuestos en la sección Resumen, el módulo 820 de decisión podrá adoptar una decisión en cuanto a inhabilitar o no el pos-filtrado.

Como una variación de esto, el módulo 820 de decisión puede usar la señal original en lugar de una señal descodificada intermedia, de modo que la señal de diferencia aproximada sea $[(s_{i_DEC} * p_{LT}) * h_{LP}](n)$. Es probable que esto sea una aproximación menos fiel pero, por otra parte, hace optativa la presencia de una línea 816 de conexión entre el módulo 820 de decisión y el módulo codificador 810.

En otras variaciones de ese tipo de esta realización, donde el módulo 820 de decisión estudia directamente la señal de audio, pueden aplicarse uno o más de los siguientes criterios:

- ¿Contiene la señal de audio tanto un componente con frecuencia fundamental dominante como un componente

situado por debajo de la frecuencia fundamental? (La frecuencia fundamental puede ser suministrada como un sub-producto del módulo codificador 810).

• ¿Contiene la señal de audio tanto un componente con frecuencia fundamental dominante como un componente situado entre los armónicos de la frecuencia fundamental?

• ¿Contiene la señal de audio energía significativa de señal por debajo de la frecuencia fundamental?

• ¿Es (probable que sea) preferible la descodificación pos-filtrada a la descodificación no filtrada con respecto a la optimalidad de velocidad-distorsión?

En todas las variaciones descritas de la estructura codificadora mostrada en la figura 8 – es decir, independientemente de la base del criterio de detección – la sección 820 de decisión puede ser habilitada para decidir sobre un avance gradual o una eliminación gradual del pos-filtrado, a fin de lograr transiciones suaves. El avance y la eliminación graduales pueden ser controlados ajustando la ganancia de pos-filtro.

La figura 9 muestra un descodificador convencional operable en una modalidad de descodificación de frecuencia y en una modalidad de descodificación de CELP, según la señal de flujo de bits suministrada al descodificador. El pos-filtrado se aplica toda vez que se selecciona la modalidad de descodificación de CELP. Una mejora de este descodificador se ilustra en la figura 10, que muestra un descodificador 1000 de acuerdo a una realización de la invención. Este descodificador es operable no solamente en una modalidad de descodificación basada en el dominio de frecuencia, en la cual está activo el módulo descodificador 1013 del dominio de frecuencia, y en una modalidad de descodificación de CELP filtrada, en la cual están activos el módulo descodificador 1011 de CELP y el pos-filtro 1040, sino también en una modalidad de CELP no filtrada, en la cual el módulo 1011 de CELP suministra su señal a un módulo 1043 de retardo de compensación, mediante una línea 1044 de omisión. Un conmutador 1042 controla qué modalidad de descodificación se usa actualmente, en respuesta a la información de pos-filtrado contenida en la señal del flujo de bits proporcionada al descodificador 1000. En este descodificador y en el de la figura 9, la última etapa de procesamiento es efectuada por un módulo 1050 de SBR, desde el cual es emitida la señal de audio final.

La figura 11 muestra un pos-filtro 1100 adecuado para ser dispuesto aguas abajo de un descodificador 1199. El filtro 1100 incluye un módulo 1140 de pos-filtrado, que es habilitado o inhabilitado por un módulo de control (no mostrado), notablemente, un controlador de ganancia binario o no binario, en respuesta a una señal de pos-filtrado recibida desde un módulo 1120 de decisión dentro del pos-filtro 1100. El módulo de decisión realiza una o más pruebas sobre la señal obtenida desde el descodificador, para llegar a una decisión en cuanto a si el módulo 1140 de pos-filtrado ha de estar activo o inactivo. La decisión puede ser tomada según la funcionalidad del módulo 820 de decisión en la figura 8, que usa la señal original y/o una señal descodificada intermedia para predecir la acción del pos-filtro. La decisión del módulo 1120 de decisión también puede basarse en información similar a la que los módulos de decisión usan en esas realizaciones donde se forma una señal descodificada intermedia. Como ejemplo, el módulo 1120 de decisión puede estimar una frecuencia de tono (a menos que esto sea inmediatamente extraíble desde la señal del flujo de bits) y calcular el contenido de energía en la señal por debajo de la frecuencia de tono y entre sus armónicos. Si este contenido de energía es significativo, probablemente representa un componente de señal relevante, en lugar de ruido, lo que motiva una decisión para inhabilitar el módulo 1140 de pos-filtrado.

Una prueba de escucha de 6 personas ha sido llevada a cabo, durante la cual muestras musicales codificadas y descodificadas de acuerdo a la invención fueron comparadas con muestras de referencia que contenían la misma música codificada, aplicando el pos-filtrado de la manera convencional, pero manteniendo sin cambios todos los otros parámetros. Los resultados confirman una mejora percibida de la calidad.

Realizaciones adicionales de la presente invención devendrán evidentes para una persona experta en la técnica, después de leer la descripción anterior. Incluso aunque la presente descripción y los dibujos revelan realizaciones y ejemplos, la invención no está restringida a estos ejemplos específicos. Pueden hacerse numerosas modificaciones y variaciones sin apartarse del alcance de la presente invención, que está definido por las reivindicaciones adjuntas.

Los sistemas y métodos revelados anteriormente en la presente memoria pueden ser implementados como software, firmware, hardware o una combinación de los mismos. Ciertos componentes, o todos los componentes, pueden ser implementados como software ejecutado por un procesador de señales digitales o un micro-procesador, o ser implementados como hardware, o como un circuito integrado específico de la aplicación. Tal software puede ser distribuido en medios legibles por ordenador, que pueden comprender medios de almacenamiento de ordenadores (o medios no transitorios) y medios de comunicación (o medios transitorios). Como es bien sabido para una persona experta en la técnica, los medios de almacenamiento de ordenadores incluyen medios tanto volátiles como no volátiles, extraíbles como no extraíbles, implementados en cualquier método o tecnología para el almacenamiento de información, tal como instrucciones legibles por ordenador, estructuras de datos, módulos de programa u otros datos. Los medios de almacenamiento de ordenadores incluyen, pero no se limitan a, las memorias RAM, ROM, EE-PROM, la memoria flash u otra tecnología de memoria, CD-ROM, discos versátiles digitales (DVD) u otro almacenamiento de disco óptico, casetes magnéticos, cinta magnética, almacenamiento de disco magnético u otros dispositivos de almacenamiento magnético, o cualquier otro medio que pueda ser usado

para almacenar la información deseada y que pueda ser objeto de acceso por parte de un ordenador. Además, es bien sabido para el experto que los medios de comunicación realizan habitualmente instrucciones legibles por ordenador, estructuras de datos, módulos de programa u otros datos en una señal de datos modulados, tal como una onda portadora u otro mecanismo de transporte, e incluyen cualquier medio de suministro de información.

REIVINDICACIONES

1. Un pos-filtro (440; 550; 740; 1040; 1140) de atenuación de ruido inter-armónico, adaptado para recibir una señal de entrada, que comprende una señal preliminar de audio descodificada de acuerdo a una entre una pluralidad de modalidades de descodificación, en donde la actividad de pos-filtro está convencionalmente asociada a modalidades específicas de descodificación, y para suministrar una señal de audio de salida, comprendiendo adicionalmente el pos-filtro de atenuación de ruido inter-armónico una sección de control para operar selectivamente el pos-filtro en una de las siguientes modalidades:

i) una modalidad de filtrado, en la cual filtra la señal preliminar de audio para obtener una señal filtrada y suministra esto como la señal de audio de salida; y

ii) una modalidad de traspasamiento, en la cual suministra la señal preliminar de audio como la señal de audio de salida,

estando dicha sección de control configurada para ingresar a la modalidad de traspasamiento en respuesta al valor de una señal de pos-filtrado, por lo cual una modalidad de descodificación convencionalmente filtrada se aplica sin filtrar.

2. El pos-filtro de la reivindicación 1, en el cual la señal de pos-filtrado está incluida en la señal de entrada.

3. El pos-filtro de la reivindicación 1, que comprende adicionalmente un módulo (1120) de decisión adaptado para estimar una frecuencia tonal de la señal preliminar de audio y para evaluar al menos uno de los siguientes criterios:

a) si la potencia de los componentes espectrales por debajo de la frecuencia de tono supera o no un umbral predeterminado;

b) si los componentes espectrales por debajo de la frecuencia de tono son o no tonales;

c) si la potencia de los componentes espectrales entre armónicos de la frecuencia de tono supera o no un umbral predeterminado; y

d) si los componentes espectrales entre armónicos de la frecuencia de tono son o no tonales;

y, en respuesta a una determinación positiva, tomar una decisión para generar una señal de pos-filtrado negativa que inhabilita el pos-filtro.

4. Un sistema descodificador (400; 500; 700; 1000) para descodificar una señal de flujo de bits como una señal de tiempo de audio, que incluye:

una sección (410; 511, 512, 513; 711, 712, 713; 1011, 1013) de descodificación para descodificar una señal de flujo de bits como una señal preliminar de tiempo de audio; y

el pos-filtro de atenuación de ruido inter-armónico de cualquiera de las reivindicaciones 1 a 3.

5. El sistema descodificador de la reivindicación 4,

en el cual el pos-filtro tiene ganancia variable que determina la atenuación inter-armónica y la sección de control incluye un controlador de ganancia operable para fijar el valor absoluto de la ganancia por debajo de un umbral predeterminado, por lo cual el pos-filtro es inhabilitado.

6. El sistema descodificador de la reivindicación 4 o 5,

en el cual el pos-filtro está adaptado para atenuar solamente aquellos componentes espectrales que estén situados por debajo de una frecuencia predeterminada de corte.

7. El sistema descodificador de la reivindicación 4,

comprendiendo adicionalmente la sección de descodificación un módulo (511; 711; 1011) descodificador de predicción lineal excitada por código, CELP, y un módulo (512; 712) descodificador de excitación codificada por transformación, TCX, para descodificar una señal de flujo de bits como una señal de tiempo de audio,

estando la sección de control adaptada para operar el sistema descodificador al menos en las siguientes modalidades:

a) el módulo de TCX está habilitado y el pos-filtro está inhabilitado;

b) el módulo de CELP y el pos-filtro están habilitados; y

c) el módulo de CELP está habilitado y el pos-filtro está inhabilitado, en donde la señal preliminar de tiempo de audio y la señal de tiempo de audio coinciden.

8. El sistema descodificador de la reivindicación 6,

comprendiendo adicionalmente la sección descodificadora un módulo (513; 713) descodificador de Codificación Avanzada de Audio, AAC, para descodificar una señal de flujo de bits como una señal de tiempo de audio,

estando la sección de control adaptada para operar el descodificador también en la siguiente modalidad:

d) el módulo de AAC está habilitado y el pos-filtro está inhabilitado.

9. El sistema descodificador de la reivindicación 4,

en el cual la señal del flujo de bits es un flujo de bits del Grupo de Expertos de Imágenes en Movimiento, MPEG, y está segmentada en tramas de tiempo, y la sección de control está adaptada para inhabilitar una trama de tiempo entera o una secuencia de tramas de tiempo enteras; y la sección de control está adicionalmente adaptada para recibir, para cada trama de tiempo, un campo de datos asociado a esta trama de tiempo, y es operable, en respuesta al valor del campo de datos, para inhabilitar el pos-filtro, por lo cual la señal preliminar de tiempo de audio es emitida como la señal de tiempo de audio.

10. Un método de descodificación de una señal de flujo de bits como una señal de tiempo de audio, que incluye las etapas de:

descodificar una señal de flujo de bits como una señal preliminar de tiempo de audio, en una entre una pluralidad de modalidades de descodificación, en donde la actividad de pos-filtro está convencionalmente asociada a modalidades específicas de descodificación; y

pos-filtrar la señal preliminar de tiempo de audio atenuando el ruido inter-armónico, obteniendo por ello una señal de tiempo de audio,

caracterizado porque la etapa de pos-filtrado es omitida selectivamente, en respuesta a información de pos-filtrado codificada en la señal del flujo de bits, por lo cual una modalidad de descodificación convencionalmente filtrada se aplica sin filtrar.

11. Un sistema codificador (800) para codificar una señal de tiempo de audio como una señal de flujo de bits, que incluye una sección codificadora (810) operable en varias modalidades de codificación, para codificar una señal de tiempo de audio como una señal de flujo de bits,

caracterizado por una sección (820) de decisión adaptada para decidir si el pos-filtrado, que incluye la atenuación del ruido inter-armónico, ha de ser inhabilitado o no en la descodificación de la señal de flujo de bits, por separado de la decisión sobre la modalidad de codificación, y para codificar esta decisión en la señal de flujo de bits como información de pos-filtrado.

12. El sistema codificador de la reivindicación 11,

estando la sección de decisión adaptada para:

detectar una co-presencia de un componente de señal con frecuencia fundamental dominante, y un componente de señal situado por debajo de la frecuencia fundamental y, optativamente, entre sus armónicos; y

en respuesta a una determinación positiva, tomar una decisión para inhabilitar.

13. El sistema codificador de la reivindicación 11,

que comprende adicionalmente un módulo de codificación de predicción lineal excitada por código, CELP, incluyendo adicionalmente dicha sección de codificación un módulo de codificación de excitación codificada por transformación, TCX,

en el cual la sección de decisión está adaptada para seleccionar una de las siguientes modalidades de codificación, preferiblemente sobre la base de una optimización de velocidad-distorsión.

a) codificación de TCX;

b) codificación de CELP con pos-filtrado; y

c) codificación de CELP sin pos-filtrado,

comprendiendo adicionalmente el sistema codificador un selector (814) de codificación, adaptado para seleccionar una de las siguientes super-modalidades:

i) codificación de Codificación Avanzada de Audio, AAC, en donde la sección de decisión está inhabilitada; y

ii) codificación de TCX/CELP, en donde la sección de decisión está habilitada para seleccionar una de las modalidades de codificación a), b) and c).

14. Un método de codificación de una señal de tiempo de audio como una señal de flujo de bits, incluyendo el método la etapa de codificación de una señal de tiempo de audio como una señal de flujo de bits en una de varias modalidades de codificación,

caracterizado por la etapa adicional de decidir si el pos-filtrado, que incluye la atenuación del ruido inter-armónico, ha de ser inhabilitado o no en la descodificación de la señal del flujo de bits, por separado de la decisión sobre la modalidad de codificación, y la codificación de esta decisión en la señal de flujo de bits como información de pos-filtrado.

15. Un producto de programa de ordenador que incluye un portador de datos que almacena instrucciones para realizar el método de la reivindicación 10 o 14.

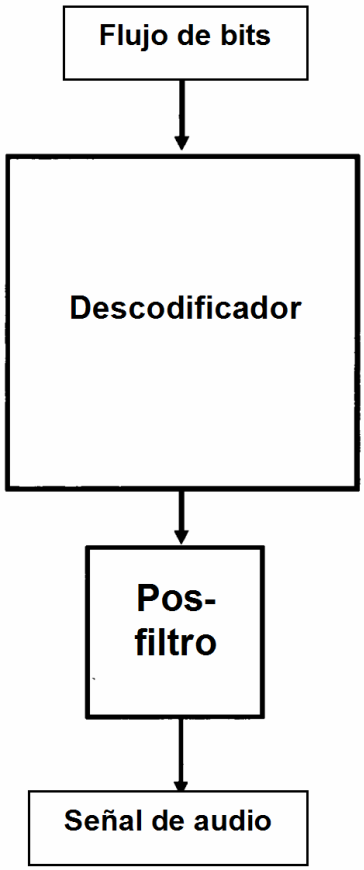

Fig. 1
(técnica anterior)

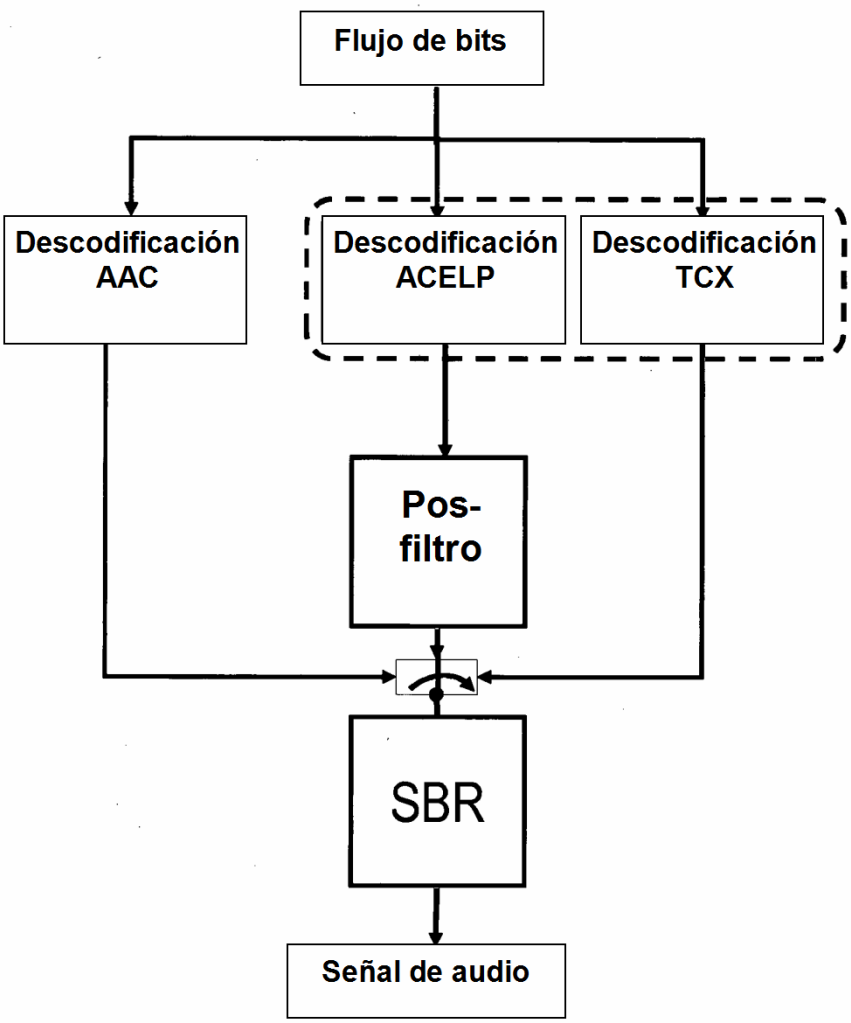

Fig. 2
(técnica anterior)

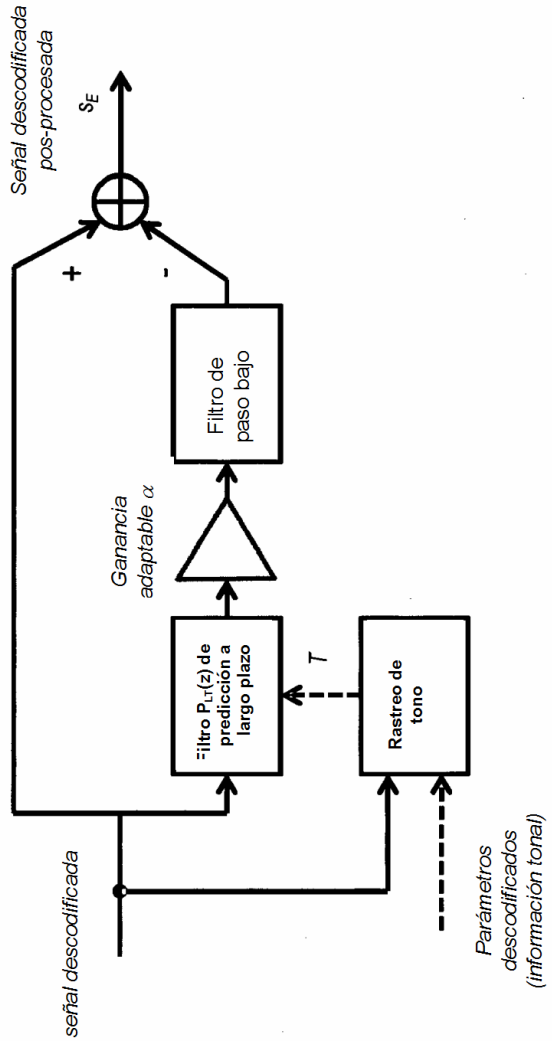
Señal descodificada pos-procesada
s_E
+
-
Filtro de paso bajo
Ganancia adaptable α
Filtro $P_{LT}(z)$ de predicción a largo plazo
T
Rastreo de tono
señal descodificada
Parámetros descodificados (información tonal)

Fig. 3

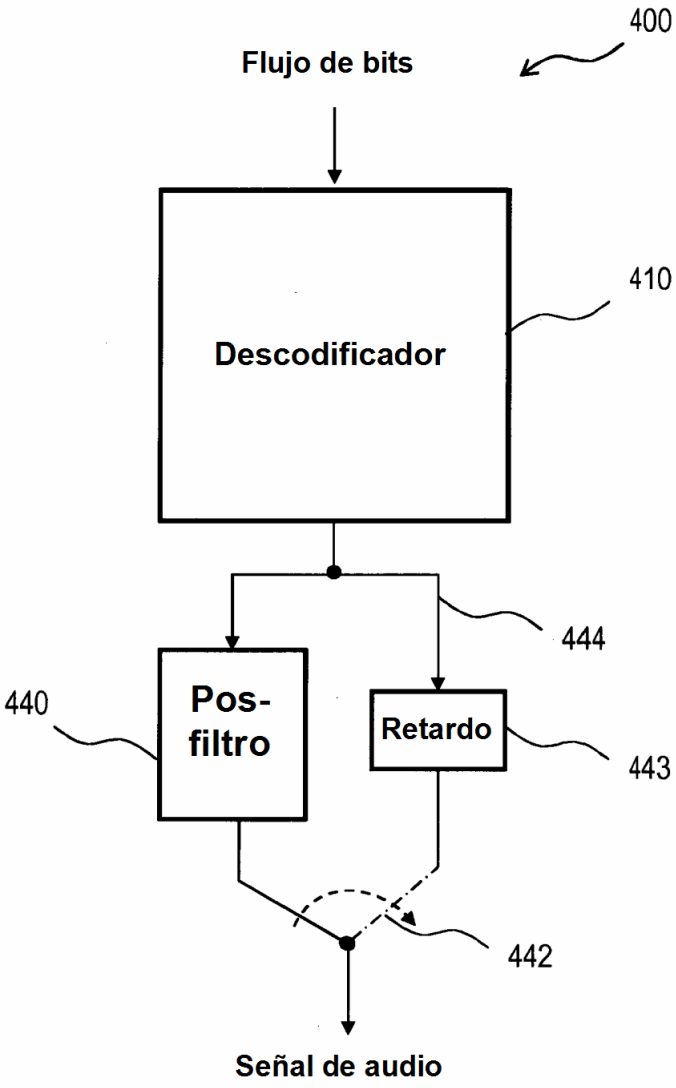
400
Flujo de bits
410
Descodificador
444
440
Pos-
filtro
Retardo
443
442
Señal de audio

Fig. 4

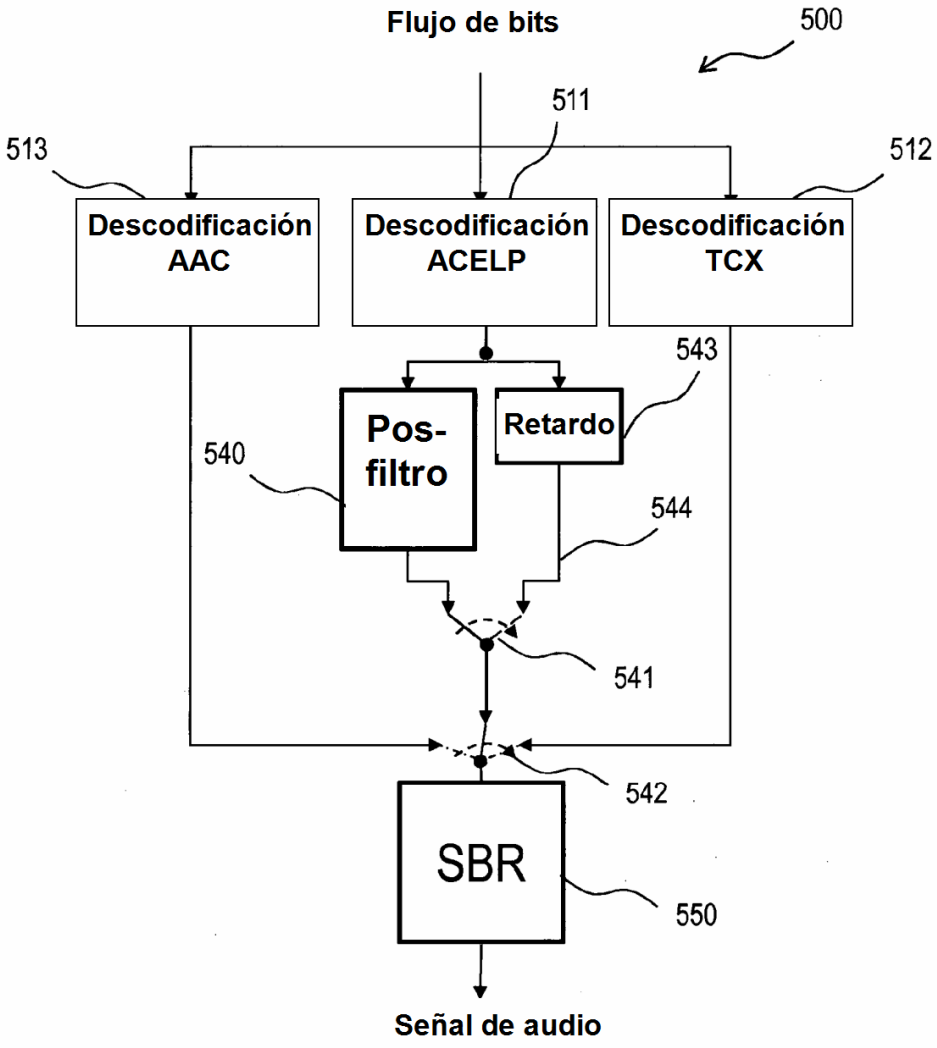
Flujo de bits
500
511
513
512
Descodificación AAC
Descodificación ACELP
Descodificación TCX
543
Pos-filtro
Retardo
540
544
541
542
SBR
550
Señal de audio

Fig. 5

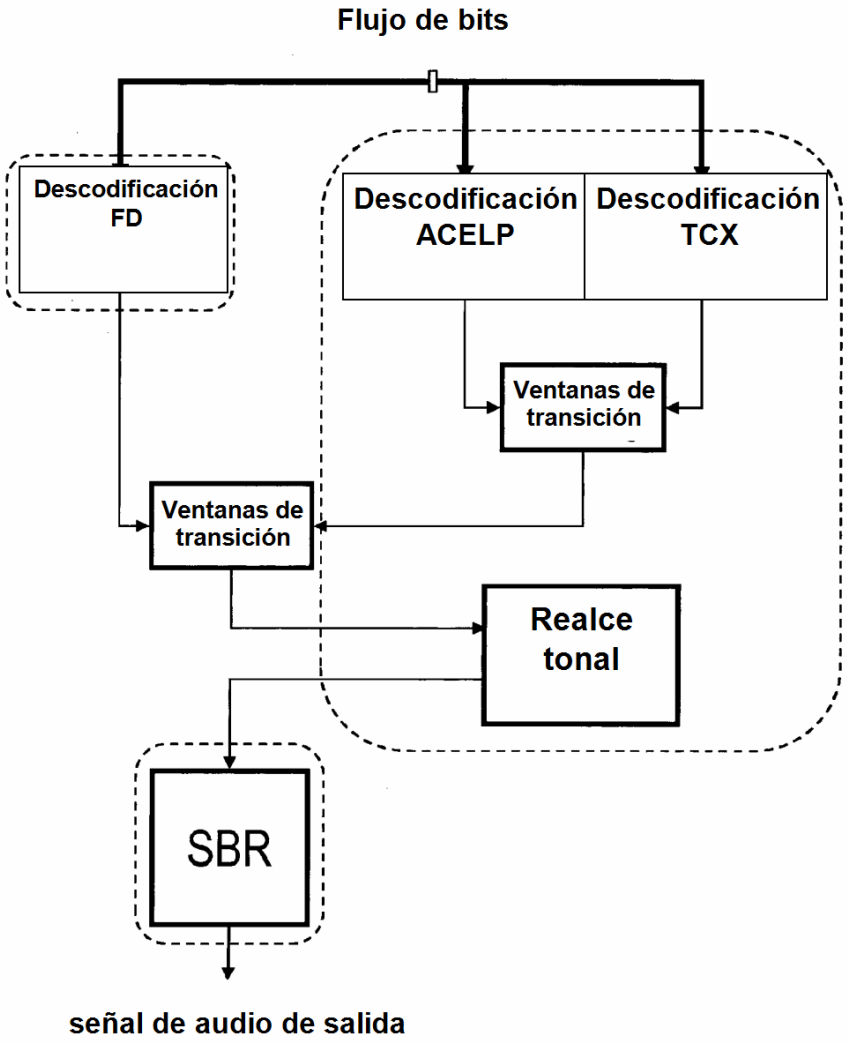

Fig. 6
(técnica anterior)

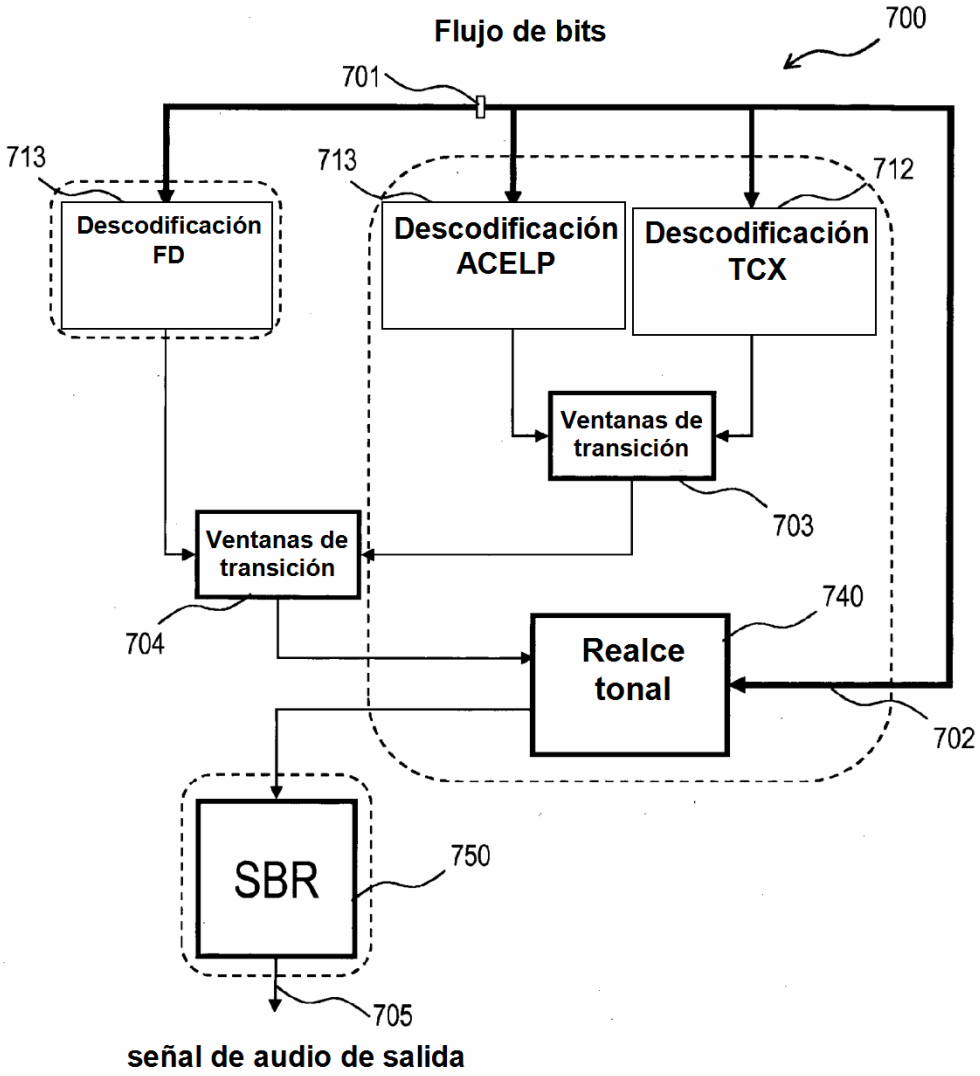
Flujo de bits
700
701
713
713
712
Descodificación FD
Descodificación ACELP
Descodificación TCX
Ventanas de transición
703
Ventanas de transición
704
740
Realce tonal
702
SBR
750
705
señal de audio de salida

Fig. 7

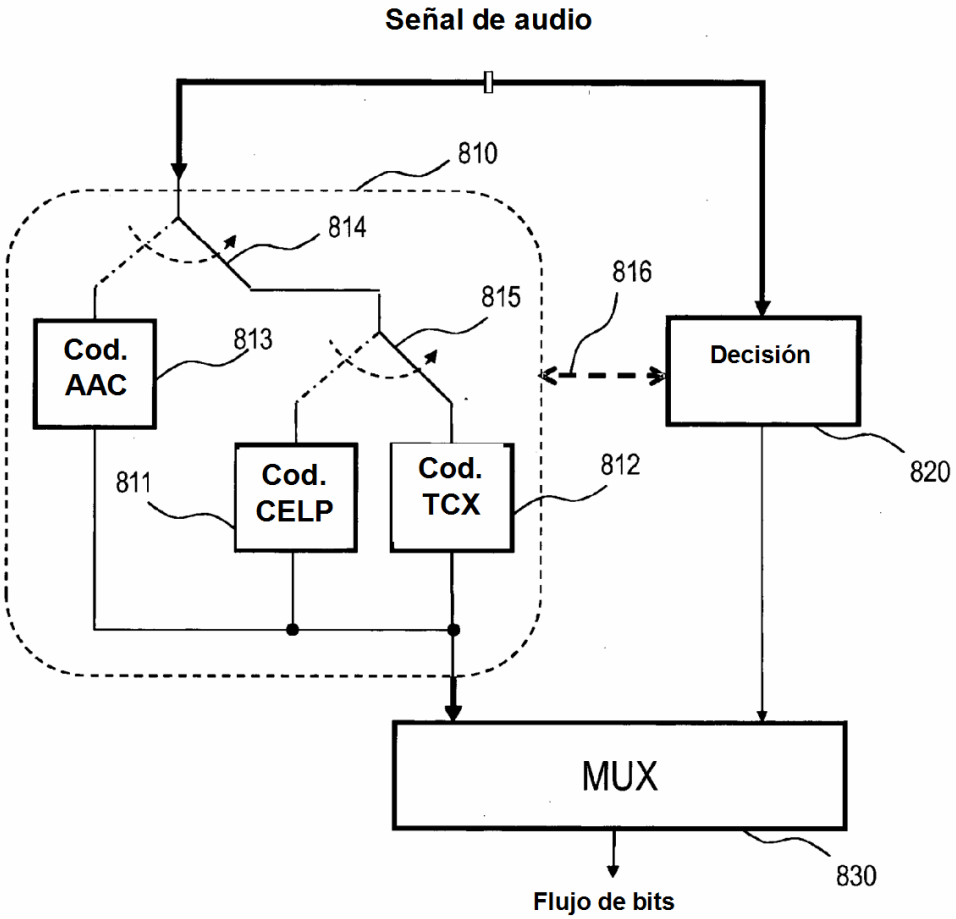
Señal de audio
810
814
815
816
813
Cod. AAC
Decisión
820
811
Cod. CELP
Cod. TCX
812
MUX
830
Flujo de bits

Fig. 8

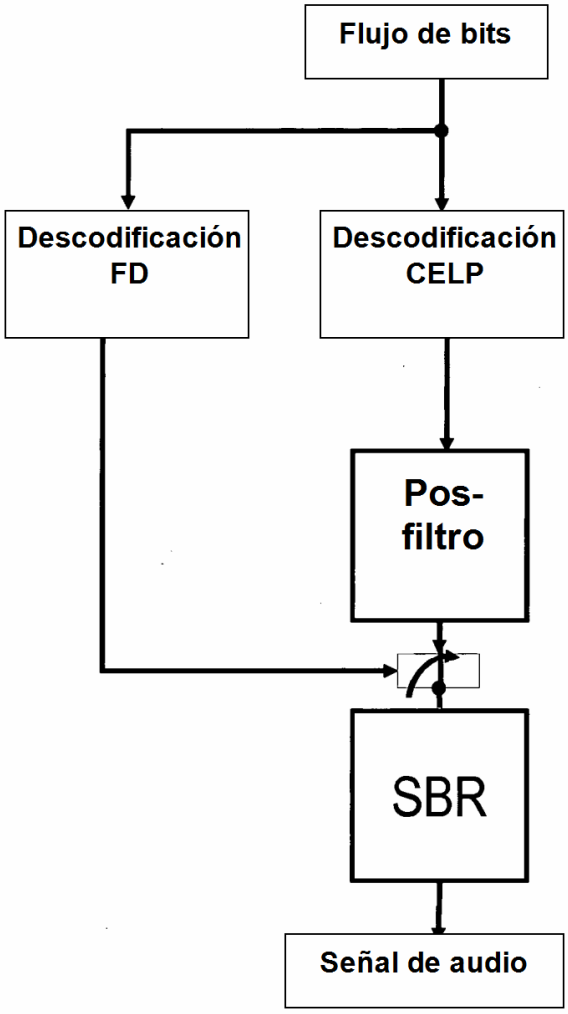

Fig. 9
(técnica anterior)

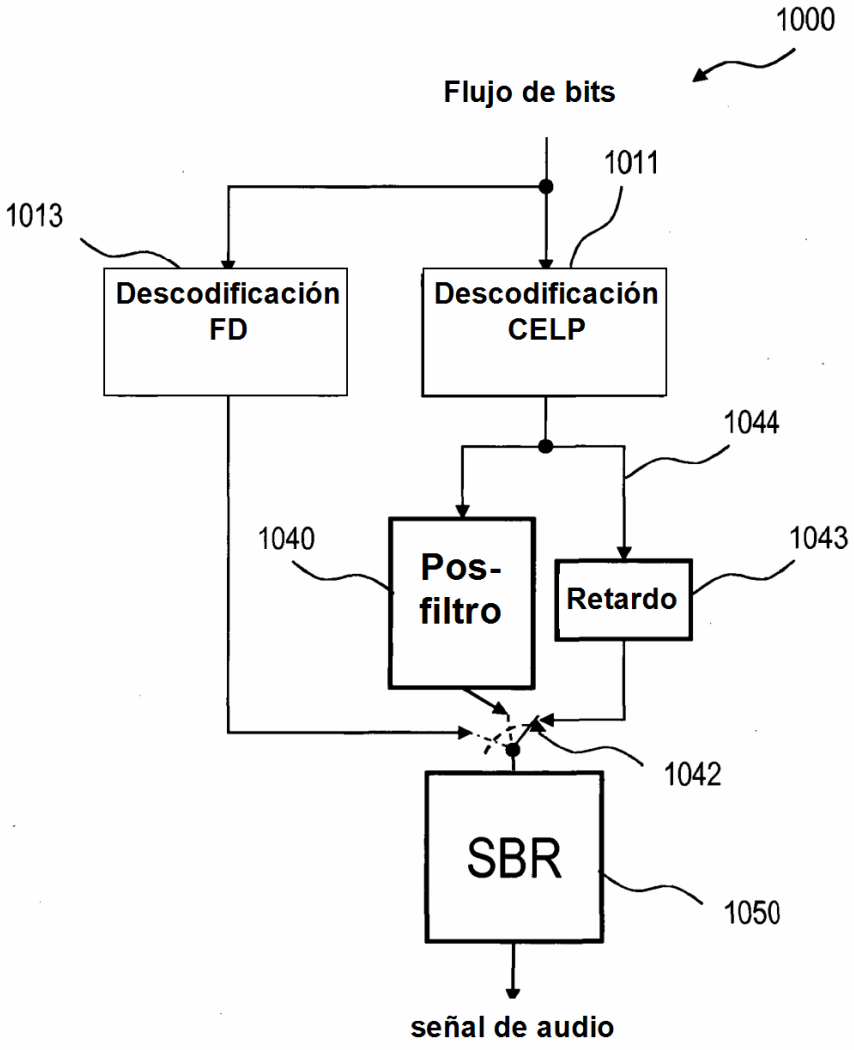
1000
Flujo de bits
1011
1013
Descodificación FD
Descodificación CELP
1044
1040
Pos-filtro
Retardo
1043
1042
SBR
1050
señal de audio

Fig. 10

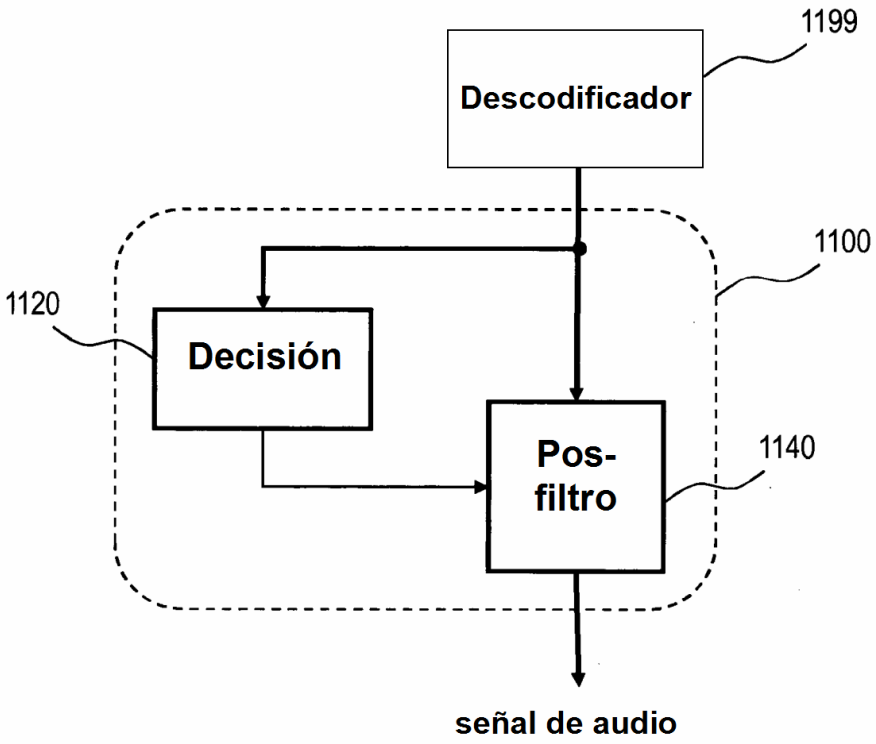
1199
Descodificador
1100
1120
Decisión
Pos-
filtro
1140
señal de audio

Fig. 11