

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 697 474**

51 Int. Cl.:

G10L 19/005 (2013.01)

G10L 21/038 (2013.01)

G10L 19/24 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **18.06.2014 PCT/EP2014/062902**

87 Fecha y número de publicación internacional: **24.12.2014 WO14202701**

96 Fecha de presentación y número de la solicitud europea: **18.06.2014 E 14733125 (0)**

97 Fecha y número de publicación de la concesión europea: **01.08.2018 EP 3011560**

54 Título: **Decodificador de audio que tiene un módulo de extensión de ancho de banda con un módulo de ajuste de energía**

30 Prioridad:

21.06.2013 EP 13173152

05.05.2014 EP 14167050

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

24.01.2019

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (100.0%)**

**Hansastraße 27c
80686 München, DE**

72 Inventor/es:

**LECOMTE, JÉRÉMIE;
BAUER, FABIAN;
SPERSCHNEIDER, RALPH y
TRITTHART, ARTHUR**

74 Agente/Representante:

SALVÀ FERRER, Joan

ES 2 697 474 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Decodificador de audio que tiene un módulo de extensión de ancho de banda con un módulo de ajuste de energía

5 **[0001]** El SBR (Spectral Band Replication, Replicación de Banda Espectral), como otras técnicas de extensión de banda, tiene como objeto codificar y decodificar partes de bandas altas espectrales de señales de audio en la parte superior de una etapa de codificador de núcleo. El SBR ha sido estandarizado en [ISO09] y se utiliza junto con AAC en el MPEG-4 Profile HE-AAC, que se emplea en diversos estándares de aplicación, por ejemplo 3GPP [3GP12a], DAB+ [EBU10] y DRM [EBU12].

10 **[0002]** El estado de la técnica para la decodificación de SBR en conjunto con AAC se describe en [ISO09, capítulo 4.6.18].

15 **[0003]** La Figura 1 ilustra el decodificador de SBR del estado de la técnica que comprende un análisis y un banco de filtros de síntesis, la decodificación de los datos de SBR que decodifica un generador de HF y un ajustador de HF.

• En la decodificación SBR del estado de la técnica, la salida del codificador de núcleo es una representación filtrada de bajo paso de la señal original. Es la entrada de x_{pcm_in} en el banco de filtros de análisis de QMF del decodificador de SBR.

20 • La salida de este banco de filtro x_{QMF_ana} se entrega al generador de HF, en el que tiene lugar el parcheado. Básicamente, el parcheado es una replicación del espectro de bajo paso en forma de bandas altas.

• El espectro parcheado $x_{HF_patched}$ se entrega ahora al ajustador de HF, junto con la información espectral de las bandas altas (envolturas), obtenidas a partir de la decodificación de los datos de SBR. La información de envoltura será decodificada por Huffman, decodificada diferencialmente y finalmente decuantificada a efectos de obtener los datos de envoltura (véase la Fig. 2). Los datos de envoltura obtenidos es un conjunto de factores de escala que cubre una determinada cantidad de tiempo, por ejemplo, un marco o partes de él. El ajustador de HF ajusta adecuadamente las energías de las bandas altas parcheadas a efectos de concordar lo mejor posible con las energías de las bandas altas originales en el lado codificador para cada banda k. Esto se aclara mediante la

30 Ecuación 1 y la Figura 2:

$$\begin{aligned} g_{sbr} [k] &= E_{Ref} [k] / E_{EstAvg} [l] \\ E_{Adj} [k] &= E_{Est} [k] \times g_{sbr} [k] \end{aligned} \tag{1}$$

35 Donde

$E_{Ref} [k]$ designa la energía para una banda k, que se transmite de forma codificada en la corriente de bits de SBR;

$E_{Est} [k]$ designa la energía de una banda alta k, parcheada por el generador de HF;

$E_{EstAvg} [l]$ designa la energía de banda alta media dentro de una banda de factor de escala l, que se define como un

40 intervalo de bandas entre una banda de partida k_{start}^l y una banda de stop k_{stop}^l :

$$E_{EstAvg} [l] = \frac{1}{N_l} \sum_{k=k_{start}^l}^{k_{stop}^l} E_{Est} (k) \tag{2}$$

$E_{Adj} [k]$ designa la energía de una banda alta k, ajustada por el ajustador de HF, mediante la ganancia $gain_{sbr}$;

45 $g_{sbr}[k]$ designa un factor de ganancia, resultante de la división mostrada en la ecuación (1).

• El banco de filtros QMF de síntesis decodifica las muestras de QMF procesadas x_{HF_adj} a PCM audio x_{pcm_out} .

50 **[0004]** Si el espectro reconstruido tiene una falta de ruido, que estaba presente en las bandas altas originales pero no parcheado por el Generador de HF, existe la posibilidad de añadir algún ruido adicional con un determinado ruido de fondo Q para cada banda k.

$$Q [k] = \frac{\text{Energía}_{\text{adicional_ruido}} [k]}{\text{Energía}_{\text{HF_generada}} [k]} \quad (3)$$

5

[0005] Por otra parte, el SBR del estado de la técnica permite mover márgenes de marco de SBR dentro de determinados límites y múltiples envolturas por marco.

[0006] La decodificación por SBR en conjunto con CELP/HVXC se describe en [EBU12, capítulo 5.6.2.2]. El decodificador CELP/HVXC+SBR en DRM está estrechamente relacionado con la decodificación por SBR del estado de la técnica en HEAAC, descrito en el capítulo 1.1.1. Básicamente, es válida la Figura 1.

[0007] La decodificación de la información de envoltura está adaptada a las propiedades espectrales de las señales similares al habla, como se describe en [EBU12, capítulo 5.6.2.2.4].

15

[0008] En la decodificación AMR-WB normal, la excitación de banda alta se obtiene mediante la generación de ruido blanco $u_{HB1}(n)$. La potencia de la excitación de banda alta se ajusta con un valor igual a la potencia de la excitación de banda inferior $u_2(n)$, lo que significa que

$$u_{HB2}(n) = u_{HB1}(n) \sqrt{\frac{\sum_{i=0}^{63} u_2^2(k)}{\sum_{i=0}^{63} u_{HB1}^2(k)}} \quad (4)$$

20

[0009] Finalmente la excitación de banda alta se obtiene mediante

$$u_{HB}(n) = \hat{g}_{HB} \cdot u_{HB2}(n) \quad (5)$$

25

en el que \hat{g}_{HB} es un factor de ganancia.

[0010] En el modo 23,85 kbit/s, se decodifica \hat{g}_{HB} a partir del índice de ganancia recibido (información secundaria).

30

[0011] En los modos de 6,60, 8,85, 12,65, 14,25, 15,85, 18,25, 19,85 y 23,05 kbit/s, se estima \hat{g}_{HB} mediante información de voz delimitada por [0.1, 1.0]. Primero se calcula el sesgo de síntesis e_{tilt}

$$e_{\text{tilt}} = \frac{\sum_{i=0}^{63} \hat{s}_{hp}(n) \cdot \hat{s}_{hp}(n-1)}{\sum_{i=0}^{63} \hat{s}_{hp}^2(n)} \quad (6)$$

35

en el que \hat{s}_{hp} es la síntesis de habla de banda inferior filtrada de elevado paso $\hat{s}_{hp12.8}(n)$ con una frecuencia de corte de 400 Hz. Seguidamente se encuentra g_{HB} del siguiente modo

$$g_{HB} = w_{SP} \cdot g_{SP} + (1 - w_{SP}) \cdot g_{BG} \quad (7)$$

40

en el que $g_{SP} = 1 - e_{\text{tilt}}$ es la ganancia para la señal de habla, $g_{BG} = 1.25 g_{SP}$ es la ganancia para la señal de ruido de segundo plano, y w_{SP} es una función de ponderación ajustada en 1, cuando la detección de la actividad de voz (VAD, voice activity detection) está en ON, y 0 cuando VAD se halla en OFF. g_{HB} está delimitado entre [0.1, 1.0]. En

el caso de segmentos vocalizados en los que hay menos energía presente a elevadas frecuencias, e_{lit} se aproxima 1 con lo que resulta una menor ganancia g_{HB} . Esto reduce la energía del ruido generado en el caso de los segmentos vocalizados.

- 5 **[0012]** A continuación se deriva el filtro de síntesis LP de banda alta $A_{\text{HB}}(z)$ a partir del filtro de síntesis de banda baja ponderado:

$$A_{\text{HB}}(z) = \hat{A}\left(\frac{z}{0.8}\right) \quad (8)$$

- 10 en el que $\hat{A}(z)$ es el filtro de síntesis LP interpolado. $\hat{A}(z)$ ha sido calculado mediante el análisis de la señal con la velocidad de muestreo de 12,8 kHz pero ahora se utiliza para una señal de 16 kHz. Esto significa que la banda entre 5,1–5,6 kHz en el dominio 12,8 kHz será mapeada a 6,4–7,0 kHz en el dominio de 16 kHz.

- 15 **[0013]** A continuación se filtra $u_{\text{HB}}(n)$ a través de $A_{\text{HB}}(z)$. La salida de esta síntesis de banda alta $s_{\text{HB}}(n)$ se filtra a través de un filtro de paso de banda FIR $H_{\text{HB}}(z)$, que tiene un paso de banda de 6 a 7 kHz. Finalmente, se añade s_{HB} al habla sintetizada de manera que se produzca la señal de habla emitida sintetizada.

- 20 **[0014]** En AMR–WB+ la señal de HF está compuesta de los componentes de frecuencia anteriores ($f_s/4$) de la señal introducida. Para representar la señal de HF con una baja velocidad, se emplea una estrategia BWE (bandwidth extension, extensión del ancho de banda). En el BWE, la información sobre la energía se envía al decodificador en la forma de envoltura espectral y de energía de marco, pero la estructura fina de la señal se extrapola en el decodificador a partir de la señal de excitación recibida (decodificada) en la señal de LF.

- 25 **[0015]** El espectro de la señal muestreada en baja s_{HF} puede considerarse como una versión plegada de la banda de alta frecuencia antes del muestreo descendente. Se lleva a cabo un análisis de LP en $s_{\text{HF}}(n)$ de manera que se obtenga un conjunto de coeficientes, que modelan la envoltura espectral de esta señal. Típicamente, es necesaria una cantidad de parámetros menor que en la señal de LF. En este caso, se utiliza un filtro de orden 8. Los coeficientes de LP son transformados a continuación en representación de ISP y cuantificados para su transmisión.

- 30 **[0016]** La síntesis de la señal de HF implementa una especie de mecanismo de extensión del ancho de banda (BWE) y utiliza algunos datos del decodificador de LF. Es una evolución del mecanismo de BWE utilizado en el decodificador de habla de AMR–WB (véase más arriba). El decodificador de HF se detalla en la Figura 3.

- 35 **[0017]** La señal de HF se sintetiza en 2 etapas:

1. Cálculo de la excitación de HF;
2. Cálculo de la señal de HF procedente de la excitación de HF.

- 40 **[0018]** La excitación de HF se obtiene configurando la señal de excitación de LF en el dominio del tiempo con factores escalares (o ganancias) en una base de submarco de 64 muestras. Esta excitación de HF se postprocesa para reducir el “grado de ocupación” de la salida, y seguidamente se filtra mediante un filtro de síntesis de predicción lineal de HF $1/A_{\text{HF}}(z)$. El resultado es sometido a un postprocesamiento adicional para compensar las variaciones de energía. Para mayor información, consulte el documento [3GP09].

- 45 **[0019]** La ocultación de pérdida de paquete en SBR en conjunto con AAC se especifica en el documento 3GPP TS 26.402 [3GP12a, capítulo 5.2] y fue subsiguientemente reutilizada en DRM [EBU12, capítulo 5.6.3.1] y DAB [EBU10, capítulo A2].

- 50 **[0020]** En caso de una pérdida de marco, el número de envolturas por marco se establece en uno y los últimos datos de envoltura válidos recibidos se reutiliza y disminuyen en energía en una relación constante para cada marco oculto.

- 55 **[0021]** Los datos de envoltura resultantes se introducen seguidamente en el proceso de decodificación normal en el que el ajustador de HF los utiliza para calcular las ganancias, que se utilizan para ajustar las bandas altas parcheadas fuera del generador de HF. El resto de la decodificación de SBR se lleva a cabo como de costumbre.

[0022] Por otra parte, los valores delta del ruido de fondo codificados se están ajustando a cero, lo que permite que el ruido de fondo decodificado por delta permanezca estático. Al final del proceso de la decodificación, esto significa que la energía del ruido de fondo sigue la energía de la señal de HF.

5

[0023] Por otra parte, se despejan las banderas para añadir senos.

[0024] La ocultación de SBR del estado de la técnica también se ocupa de la recuperación. Proporciona una transición suave desde la señal oculta a la señal correctamente decodificada en términos de brechas de energía que pueden resultar de límites de marco no concordantes.

10

[0025] La ocultación de SBR del estado de la técnica en conjunto con CELP/HVXC se describe in [EBU12, capítulo 5.6.3.2] y se describe en términos generales en lo que sigue:

Cuando quiera que se haya detectado un marco corrupto, se aplica un conjunto predeterminado de valores de datos al decodificador de SBR. Esto permite obtener una "envoltura espectral estática de banda alta con un nivel de reproducción relativamente bajo, que presenta una tendencia hacia las frecuencias más elevadas" [EBU12, capítulo 5.6.3.2]. En este caso, la ocultación de SBR inserta algún tipo de ruido de confort, que no tiene un desvanecimiento dedicado en el dominio del SBR. Esto protege los oídos del escucha contra ráfagas de audio relativamente fuertes y mantiene la impresión de un ancho de banda constante.

20

[0026] La ocultación del BWE del estado de la técnica según G.718 se describe en [ITU08, 7.11.1.7.1] y se señala a grandes rasgos como sigue:

En el modo de bajo retardo, que está disponible exclusivamente para la capa 1 y 2, la ocultación de la banda de alta frecuencia de 6000 a 7000 Hz se realiza exactamente de la misma manera como cuando no se producen borrados de marco. La operación del decodificador de canal limpio para las capas 1, 2 y 3 es como sigue: se aplica una extensión de ancho de banda ciego. El espectro en el rango de 6400–7000 Hz se rellena con una señal de ruido blanco, adecuadamente escalada en el dominio de excitación (la energía de la banda alta debe coincidir con la energía de banda baja). Se sintetiza a continuación, con un filtro derivado mediante la ponderación del mismo filtro de síntesis de LP que el utilizado en el dominio de 12,8 kHz. Para las capas 4 y 5 no se realiza ninguna extensión de ancho de banda, dado que estas capas cubren la banda completa hasta 8 kHz.

30

[0027] En la operación por defecto se lleva a cabo un procesamiento de baja complejidad para reconstruir la banda de alta frecuencia de la señal sintetizada con una frecuencia de muestreo de 16 kHz. En primer lugar, la excitación de banda de alta frecuencia escalada, $u''_{HB}(n)$, es linealmente atenuada a lo largo del marco como

35

$$u'''_{HB}(n) = u''_{HB}(n) \cdot g_{att}(n), \quad \text{para } n = 0, \dots, 319 \quad (9)$$

en el que la longitud del marco es de 320 muestras y $g_{att}(n)$ es un factor de atenuación que está dada por

$$g_{att}(n) = 1,0 - n \frac{1,0 - \bar{g}_p}{320}, \quad \text{para } n = 0, \dots, 319 \quad (10)$$

40

[0028] En la ecuación anterior, \bar{g}_p es la ganancia promedio de pitch (tono). Es la misma ganancia utilizada durante la ocultación del libro de código adaptativo. A continuación, se atenúa la memoria del filtro de paso de banda en el intervalo de frecuencias 6000 – 7000 Hz para lo cual se utiliza $g_{att}(n)$, como se deriva en la ecuación 10, para impedir cualquier discontinuidad. Finalmente, la señal de excitación de elevada frecuencia, $u'''(n)$, se filtra a través del filtro de síntesis. La señal sintetizada es añadida a continuación a la síntesis oculta con una frecuencia de muestreo de 16 kHz.

45

[0029] La ocultación en el estado de la técnica del ancho de banda a ciegas en AMR–WB se señala a grandes rasgos en [3GP12b, 6.2.4] y se indica abreviadamente como sigue:

50

Cuando un marco se pierde o se pierde parcialmente, el parámetro de ganancia de banda alta no se recibe y se utiliza una estimación para la ganancia de banda alta en su lugar. Esto significa que en el caso de marcos de habla malos/perdidos, la reconstrucción de banda alta funciona de la misma manera para todos los modos diferentes.

[0030] En el caso de que se pierda un marco, el filtro de síntesis de LP de banda alta se deriva como es habitual a partir de los coeficientes de LP procedentes de la banda de núcleo. La única excepción es que los coeficientes de LPC no han sido decodificados a partir de la corriente de bits, sino que han sido extrapolados mediante la estrategia normal de ocultación de AMR–WB.

5

[0031] La ocultación en el estado de la técnica de la extensión del ancho de banda en AMR–WB+ se señala a grandes rasgos en [3GP09, 6.2] y se resume sumariamente en lo que sigue:

En el caso de una pérdida de paquetes, los datos de control que son internos para el decodificador de HF se

generan a partir del vector indicador de marcos malos BFI = (bfi0, bfi1, bfi2, bfi3). Estos datos son $bfi_{isf_{n_f}}$, BFI_{GAIN},
 10 y el número de submarcos para la interpolación de ISF. La naturaleza de estos datos se define con mayor detalle en lo que sigue:

$bfi_{isf_{n_f}}$ es una bandera binaria que indica la pérdida de los parámetros ISF. Dado que los parámetros de ISF para la señal de HF son siempre transmitidos en el primer paquete (que contiene el primer submarco) que es sea de HF20, 40 o 80, la bandera de pérdida siempre se coloca al indicador bfi del primer submarco (bfi0). Lo mismo es
 15 válido para la indicación de ganancias de HF pérdidas. Si el primer paquete/submarco del modo vigente se pierde (HF20, 40 o 80) se pierde la ganancia y es necesario ocultarla.

[0032] La ocultación de los vectores HF ISF es muy similar a la ocultación ISF para los ISF de núcleo. La idea principal es la de reutilizar el último vector ISF bueno, pero desplazarlo hacia el vector ISF medio (en el que el vector
 20 de ISF medio es entrenado offline):

$$isf_q[i] = 0.9 \cdot isf_q[i] + 0.1 \cdot mean_isf_hf[i] \quad (11)$$

[0033] Las ganancias de BWE ($\bar{g}_0, \dots, \bar{g}_{nb-1}$) se estiman según el siguiente código de fuente (en el código:
 25 $\bar{g}_i = gain_q[i]$; 2.807458 es una constante del decodificador).

/* utilizar las ganancias pasadas ligeramente desplazadas hacia la media */

past_q = (0.9f(*past_q* + 20.0f)) - 20.0f;

para (i=0; i<4; i++) {

30 *gain_q*[i] = *past_q* + 2.807458f;

}

tmp = 0.0;

para (i=0; i<4; i++) {

} *tmp* += *gain_q*[i];

35 *past_q = 0.25f**tmp* - 2.807458f;

[0034] Con el fin de derivar las “ganancias para hacer concordar la magnitud en fs/4” se utiliza el mismo algoritmo que en la decodificación de canal limpio, pero con la excepción de que los ISF para la parte de HF y/o de LF ya pueden haberse ocultado. Todas las etapas siguientes tales como la interpolación lineal, suma y aplicación de
 40 ganancias son las mismas que en el caso del canal limpio.

[0035] Para derivar la excitación, se utiliza el mismo procedimiento que en un marco correctamente recibido, donde la excitación de banda inferior se utiliza después:

45 • de que haya sido aleatorizado

• de que haya sido ampliado en el dominio del tiempo con ganancias de submarco

• de que haya sido configurado en el dominio de las frecuencias con un filtro de LP

• la energía haya sido alisada a lo largo del tiempo

50 **[0036]** A continuación, la síntesis se lleva a cabo según la Figura 3.

[0037] El documento: AES convention paper 6789: Schneider, Krauss y Ehret [SKE06] describe una técnica de ocultación que reutiliza los últimos datos de envoltura de SBR válidos. Si se pierde más de un marco de SBR, se aplica un desvanecimiento. “El principio básico es el de bloquear simplemente los últimos valores de SBR de
 55 envoltura válidos conocidos hasta que sea posible continuar el procesamiento de SBR con datos transmitidos de

nuevo. Además, se lleva a cabo un desvanecimiento si más de un marco de SBR no es decodificable”.

5 **[0038]** En el documento AES convention paper 6962: Sang-Uk Ryu y Kenneth Rose [RR06] se describe una técnica de ocultación que estima la información paramétrica, para lo cual utiliza datos de SBR del marco anterior y del marco siguiente. Las envolturas de banda alta son estimadas de manera adaptativa a partir de la evolución de la energía en los marcos circundantes.

10 **[0039]** Los conceptos de la ocultación de pérdidas de paquetes pueden producir una señal de audio perceptualmente degradada durante la pérdida de paquetes.

15 **[0040]** El documento WO201/127617 A1 describe un procedimiento de ocultación de error por el que los coeficientes de dominio de frecuencia son copiados desde un marco anterior. La señal de banda alta para el marco actual se escala de forma adaptativa con el fin de mantener el índice de energía entre la señal de banda alta y la señal de banda baja.

20 **[0041]** Uno de los objetivos de la presente invención es el de proporcionar un decodificador de audio y un procedimiento que tenga un concepto mejorado en cuanto a la ocultación de pérdidas de paquetes.

25 **[0042]** Este objeto puede alcanzarse mediante un decodificador de audio según la reivindicación 1. El decodificador de audio según la invención une el módulo de extensión de ancho de banda para el módulo de decodificación de banda de núcleo en términos de energía o, en otras palabras, asegura que el módulo de extensión de ancho de banda siga al módulo de decodificación de banda de núcleo en cuanto a energía durante la ocultación, independientemente de lo que haga el módulo de decodificación de la banda de núcleo.

30 **[0043]** La innovación con esta estrategia es que –en caso de ocultación– la generación de banda alta ya no se adapta estrictamente a las energías de envoltura. Con la técnica del bloqueo de la ganancia, las energías de banda alta se adaptan a las energías de banda baja durante la ocultación y, por tanto, ya no se basan únicamente en los datos transmitidos en el último marco bueno. Este procedimiento utiliza la idea de utilizar la información de banda baja para la reconstrucción de la banda alta.

35 **[0044]** Con esta estrategia, no es necesario transferir datos adicionales (por ejemplo, el factor de desvanecimiento) desde el codificador de núcleo al codificador de extensión de ancho de banda. Esto hace que la técnica sea de fácil aplicación a cualquier codificador con extensión de ancho de banda, especialmente a SBR, donde el cálculo de la ganancia ya se realiza inherentemente (Ecuación 1).

40 **[0045]** La ocultación del decodificador de audio de la invención tiene en cuenta la pendiente de desvanecimiento del módulo de decodificación de banda de núcleo. Esto conduce a un comportamiento previsto del desvanecimiento en su conjunto. Se evitan las situaciones en las que las energías de las bandas de frecuencia del módulo de decodificación de banda de núcleo se desvanecen más lentamente que las energías de las bandas de frecuencia del módulo de ampliación de ancho de banda, que se harían perceptibles y causarían la impresión desagradable de una banda de la señal limitada.

45 **[0046]** Por otra parte, también se evitan las situaciones en las que las energías en las bandas de frecuencia del módulo de decodificación de banda de núcleo se desvanecen más rápidamente que las energías de las bandas de frecuencia del módulo de ampliación de ancho de banda, lo que introduciría artefactos porque las bandas de frecuencia de los módulos de extensión de ancho de banda se amplifican excesivamente, en comparación con las bandas de frecuencia del módulo de decodificación de banda de núcleo.

50 **[0047]** En contraste con un decodificador sin desvanecimiento que tiene una extensión de ancho de banda con niveles de energía predefinidos (como por ejemplo un decodificador CELP/HVXC + SBR decodificador), que conserva solamente la inclinación espectral de un cierto tipo de señal, funciona el decodificador de audio de la invención de forma independiente con respecto a las características espectrales de las señales, de manera que se evita una degradación perceptualmente decodificada de la señal de audio.

55 **[0048]** La técnica propuesta podría ser utilizada con cualquier procedimiento de extensión de ancho de banda (BWE) en la parte superior de un módulo de decodificación de banda de núcleo (en lo que sigue: codificador de núcleo). La mayor parte de la técnica de extensión de ancho de banda se basa en la ganancia por banda entre los niveles originales de energía y los niveles de energía que obtienen después de copiar el espectro de núcleo. La técnica propuesta no actúa sobre las energías del marco de audio anterior, como lo hace el estado de la técnica,

sino sobre las ganancias del marco de audio anterior.

[0049] Cuando un marco de audio se pierde o es ilegible (o en otras palabras, si se produce una pérdida de marco de audio) las ganancias derivadas del último marco bueno se introducen en el proceso de decodificación normal del módulo de decodificación de banda de núcleo, que ajusta las energías de las bandas de frecuencia del módulo de extensión de ancho de banda (véase la ecuación 1). Esto forma la ocultación. Cualquier desvanecimiento, que se aplique en el módulo de decodificación de banda de núcleo por una ocultación del módulo de decodificación de banda de núcleo, se aplicará de forma automática a las energías de las bandas de frecuencia del módulo de ampliación de ancho de banda mediante el bloqueo de las energías entre las bandas alta y baja.

[0050] La señal de dominio de frecuencia que tiene al menos una banda de frecuencia puede ser, por ejemplo, una señal de excitación de predicción lineal excitada por código algebraico (señal de excitación ACELP).

[0051] En algunas formas de realización, el módulo de extensión de ancho de banda comprende un módulo que proporciona un factor de ganancia para enviar el factor de ganancia actual, al menos en el marco de audio real en el que la pérdida de marco de audio se produce en el módulo de ajuste de energía.

[0052] En una forma de realización preferida, el módulo que proporciona el factor de ganancia está configurado de tal manera que en el marco de audio actual en el que tiene lugar la pérdida de marco de audio el factor de ganancia actual es el factor de ganancia del marco de audio anterior. Esta forma de realización desactiva por completo el desvanecimiento contenido en el módulo que decodifica la extensión del ancho de banda mediante el solo bloqueo de las ganancias derivadas para la última envoltura en el último tramo bueno:

$$\begin{aligned} g_{bwe}^{[n]}[k] &= g_{bwe}^{[n-1]}[k] \\ E_{Adj}[k] &= E_{Est}[k] * g_{bwe}[k] \end{aligned} \quad (12)$$

en la que $E_{Adj}[k]$ designa la energía de una banda de frecuencia k del módulo de extensión de ancho de banda, ajustada para expresar lo mejor posible la distribución original de la energía; $g_{bwe}^{[n]}[k]$ designa el factor de ganancia del marco actual; y $g_{bwe}^{[n-1]}[k]$ designa el factor de ganancia del marco anterior.

[0053] En otra forma de realización preferida, el módulo que proporciona el factor de ganancia está configurado de tal manera que en el marco de audio actual en el que tiene lugar la pérdida de marco el factor de ganancia actual se calcula a partir del factor de ganancia del marco de audio anterior y a partir de una clase de señal del marco de audio anterior.

[0054] Esta forma de realización utiliza un clasificador de señales para calcular las ganancias sobre la base de las ganancias pasadas y también adaptativamente sobre la clase de señal del marco anteriormente recibido:

$$\begin{aligned} g_{bwe}^{[n]}[k] &= f(g_{bwe}^{[n-1]}[k], c_{sig}^{[n-1]}) \\ E_{Adj}[k] &= E_{Est}[k] * g_{bwe}[k] \end{aligned} \quad (13)$$

en el que $f(g_{bwe}^{[n-1]}, c_{sig}^{[n-1]})$ designa una función que depende del factor de ganancia $g_{bwe}^{[n-1]}$ del marco de audio anterior y de la clase de señal $c_{sig}^{[n-1]}$ del marco de audio anterior. Las clases de señal pueden referirse a clases de sonidos del habla tales como: obstruyente (con subclases: stop, africtivo, fricativo), sonoras (estas subclases: nasal, flap approximant, vocal), lateral, trill.

[0055] En una forma de realización preferida el módulo que proporciona el factor de ganancia está configurado para calcular un número de marcos de audio subsiguientes en los que tiene lugar una pérdida de marco de audio y configurado para llevar a cabo un procedimiento reductor del factor de ganancia en el que el número de marcos de audio subsiguientes en los que tiene lugar una pérdida de marco de audio que supera un número predefinido.

[0056] Si tiene lugar un fricativo inmediatamente antes de una pérdida de marco en forma de ráfaga (múltiples pérdidas de marco en marcos de audio subsiguientes), el desvanecimiento inherente por defecto del módulo de decodificación de banda de núcleo puede ser demasiado lento para asegurar un sonido agradable y natural en combinación con el bloqueo de ganancia. El resultado percibido de esto puede ser un fricativo prolongado con mucha energía en las bandas de frecuencia del módulo de extensión de ancho de banda. Por esta razón es posible que se lleve a cabo una verificación de múltiples pérdidas de marco. Si esta verificación es positiva, puede llevarse a cabo un procedimiento para disminuir el factor de ganancia.

10 **[0057]** En una forma de realización preferida, el procedimiento para reducir el factor de ganancia comprende la etapa de reducción del factor de ganancia actual dividiendo el factor de ganancia actual por un primer número en el caso en que el factor de ganancia actual supere un primer umbral. Mediante estos rasgos se reducen las ganancias que superen un primer umbral (que puede determinarse empíricamente).

15 **[0058]** En una forma de realización preferida, el procedimiento para reducir el factor de ganancia comprende la etapa de reducción del factor de ganancia actual para lo cual se divide el factor de ganancia actual por un segundo número que es más grande que el primer número en el caso en que el factor de ganancia actual supera un segundo umbral que es más grande que el primer umbral. Estos rasgos aseguran que las ganancias extremadamente más elevadas disminuyan más rápidamente aun. Todas las ganancias que superen el segundo
20 umbral disminuirán más rápidamente.

[0059] En algunas formas de realización el procedimiento para reducir el factor de ganancia comprende la etapa de ajuste del factor de ganancia actual en el primer umbral en el caso en que el umbral actual, después de su reducción, sea inferior al primer umbral. Mediante estos aspectos se impide que las ganancias disminuidas caigan
25 por debajo del primer umbral.

[0060] Puede verse un ejemplo dentro del pseudocódigo 1:

```

/*limitar ganancia en el caso de múltiples pérdidas de marco*/
30 #DEFINE BWE_GAINDEC 10
si (previousFrameErrorFlag && (gain[k] > BWE_GAINDEC) ) {
/* las ganancias que superan el primer umbral en 50 veces serán disminuidas más rápidamente */
Si (gain[k] > 50* BWE_GAINDEC) {
35   gain[k] /= 6;
}
en caso contrario {
  gain[k] /= 4;
} /* impedir que las ganancias caigan por debajo de BWE_GAINDEC */
si (gain[k] < BWE_GAINDEC) {
40   gain[k] = BWE_GAINDEC;
}}

```

en el que previousFrameErrorFlag es una bandera que indica si hay una pérdida de marco múltiple está presente, BWE_GAINDEC designa el primer umbral, 50* BWE_GAINDEC designa el segundo umbral y gain[k] designa el
45 factor de ganancia actual para la banda de frecuencia k.

[0061] En algunas formas de realización el módulo de extensión de ancho de banda comprende un módulo generador de ruido configurado para añadir ruido al menos a una banda de frecuencia, en la que en el marco de audio actual en la que tiene lugar la pérdida de marco de audio se utiliza una relación entre la energía de señal y la
50 energía de ruido de al menos una banda de frecuencia del marco de audio anterior para calcular la energía de ruido del marco de audio actual.

[0062] En el caso de que haya una característica de ruido de fondo (es decir, componentes de ruido adicionales para conservar el ruido de la señal original) implementado en la extensión del ancho de banda, es
55 necesario adoptar la idea del bloqueo de la ganancia también hacia el ruido de fondo. Para conseguir esto, niveles de energía de ruido de fondo de marcos no ocultos son convertidos en una relación de ruido, para lo cual se tiene en cuenta la energía de las bandas de frecuencia del módulo de extensión de ancho de banda. Se guarda la relación en una memoria intermedia y será la base del nivel de ruido en el caso de una ocultación. La principal ventaja es el mejor acoplamiento del ruido de fondo a la energía del codificador de núcleo debido al cálculo de la relación ratio

prev_noise[k].

[0063] El pseudocódigo 2 muestra esto:

```

5 para (k=bands) {
  si !(frameErrorFlag) {
    prev_noise[k] = nrgHighband[k] / noiseLevel[k];
  } En caso contrario {
    noiseLevel[k] = nrgHighband[k] / prev_noise[k]; } }
10
```

en el que frameErrorFlag es una bandera que indica que hay una pérdida de marco presente y prev_noise[k] es la relación entre la energía nrgHighband[k] de la banda de frecuencia k y el nivel de ruido Level[k] de la banda de frecuencia k.

15 **[0064]** En una forma de realización preferida el decodificador de audio comprende un módulo analizador de espectros configurado para establecer el espectro del marco de audio actual de la señal de audio de banda de núcleo y para derivar la energía estimada para la señal para el marco actual para al menos una banda de frecuencia a partir del espectro del marco de audio actual de la señal de audio de banda de núcleo.

20 **[0065]** En algunas formas de realización el módulo que proporciona el factor de ganancia está configurado de tal manera que, en el caso de que un marco de audio actual, en el que no tiene lugar una pérdida de marco de audio, siga subsiguientemente en un marco de audio anterior, en el que tiene lugar una pérdida de marco de audio, se utiliza el factor de ganancia recibido para el marco de audio actual para el marco vigente, si un retardo entre marcos de audio del módulo de extensión de ancho de banda con respecto a los marcos de audio del módulo de
25 decodificación de banda de núcleo es más pequeño que un umbral de retardo, utilizándose el factor de ganancia tomado del marco de audio anterior para el marco vigente, si el retardo entre marcos de audio del módulo de extensión de ancho de banda con respecto a los marcos de audio del módulo de decodificación de banda de núcleo es mayor que el umbral de retardo.

30 **[0066]** Por encima de la ocultación, en el módulo de extensión de ancho de banda debe prestarse especial atención a la formación de los marcos. Los marcos de audio del módulo de extensión de ancho de banda y los marcos de audio del módulo de decodificación de banda de núcleo con frecuencia no están alineados exactamente, sino que podrían tener un determinado retardo. Por ello puede suceder que un paquete perdido contenga datos de extensión del ancho de banda que se estén retardando, con respecto a la señal de núcleo contenido en el mismo
35 paquete.

[0067] El resultado en este caso es que el primer paquete bueno después de una pérdida puede contener datos de extensión para crear partes de las bandas de frecuencia del módulo de extensión de ancho de banda del módulo de decodificación de banda anterior del marco de audio de núcleo, que ya estaba oculto en el decodificador.

40 **[0068]** Por esta razón, es necesario considerar la formación de marcos durante la recuperación, en función de las respectivas propiedades del núcleo, módulo de decodificación y módulo de extensión de ancho de banda. Esto podría significar tratar el primer marco de audio o partes de él en el módulo de extensión de ancho de banda como erróneos y no aplicar las ganancias más nuevas de una vez sino conservar las ganancias bloqueadas del
45 primer marco de audio para un marco adicional.

[0069] Conservar o no las ganancias bloqueadas para el primer marco bueno depende del retardo. La aplicación experimental a códecs con diferentes retardos mostró diferentes beneficios para códecs con diferentes retardos. Para códecs con retardos bastante pequeños (por ejemplo, de 1ms), es mejor utilizar las ganancias más
50 nuevas para el primer marco de audio bueno.

[0070] En una forma de realización preferida el módulo de extensión de ancho de banda comprende un módulo generador de señales configurado para crear una señal bruta en el dominio de las frecuencias que tiene al menos una banda de frecuencia, que es comunicada al módulo de ajuste de energía, sobre la base de la señal de
55 audio de banda de núcleo y la corriente de bits.

[0071] En una forma de realización preferida el módulo de extensión de ancho de banda comprende un módulo de síntesis de señales configurado de manera que produzca la señal de audio de extensión de ancho de banda a partir de la señal en el dominio de frecuencias.

[0072] El objetivo de la invención puede lograrse por medio de un procedimiento para producir una señal de audio desde una corriente de bits que contiene marcos de audio según la reivindicación 14. El objetivo de la invención puede lograrse además mediante un programa informático para llevar a cabo, cuando se ejecuta en un ordenador o procesador, el procedimiento anteriormente descrito, según la reivindicación 15. En lo que sigue se exponen formas de realización preferidas de la invención con respecto a los dibujos adjuntos, en los que:

la Figura 4 ilustra una forma de realización de un decodificador de audio según la invención en una vista esquemática; y
 10 la Figura 5 ilustra la formación de marcos de una forma de realización de un decodificador de audio según la invención.

[0073] La Figura 4 ilustra una forma de realización de un decodificador de audio 1 según la invención en una vista esquemática. El decodificador de audio 1 está configurado para producir una señal de audio, AS (audio signal) de una corriente de bits, BS (bitstream) que contiene marcos de audio AF (audio frames). El decodificador de audio 1 comprende:

un módulo de decodificación de banda de núcleo configurado para derivar una señal de audio de banda de núcleo directamente decodificada CBS a partir de la corriente de bits BS;

20 un módulo de extensión de ancho de banda 2 configurado para derivar una señal de audio de extensión de ancho de banda paramétricamente decodificada BES a partir de la señal de audio de banda de núcleo CBS y a partir de la corriente de bits BS, en el que la señal de audio de extensión de ancho de banda BES está basada en una señal en el dominio de las frecuencias FDS que tiene al menos una banda de frecuencia FB; y

25 un combinador 4 configurado para combinar la señal de audio de banda de núcleo CBS y la señal de audio de extensión de ancho de banda BES de manera que se produzca la señal de audio AS;

en el que el módulo de extensión de ancho de banda 3 comprende un módulo de ajuste de energía 5 que está configurado de tal manera que en un marco de audio actual AF2 en el que tiene lugar una pérdida de marco de audio AFL, se establece una energía de señal ajustada para el marco de audio actual AF2 para al menos una banda de frecuencia FB sobre la base de un factor de ganancia actual CGF para el marco de audio actual AF2, en el que el factor de ganancia actual CGF se deriva a partir de un factor de ganancia tomado de un marco de audio previo AF1 o a partir de la corriente de bits BS, y sobre la base de una energía estimada para la señal EE para al menos una banda de frecuencia FB, donde la energía estimada para la señal EE se deriva a partir de un espectro del marco de audio actual AF2 de la señal de audio de banda de núcleo CBS.

[0074] El decodificador de audio 1 según la invención une el módulo de extensión de ancho de banda 3 al módulo de decodificación de banda de núcleo en términos de energía o, en otras palabras, asegura que el módulo de extensión de ancho de banda 3 siga al módulo de decodificación de banda de núcleo 2 en cuanto a energía durante la ocultación, no importa lo que haga el módulo de decodificación de banda de núcleo.

[0075] La innovación con esta estrategia es que –en el caso de la ocultación– la generación de bandas altas ya no está estrictamente adaptada para envolver energías. Con la técnica del bloqueo de ganancias, las energías de banda alta se adaptan a las energías de baja banda durante la ocultación y por lo tanto ya no se basan solamente sobre los datos transmitidos durante el último marco bueno AF1. Este procedimiento aprovecha la idea de utilizar información de baja banda para la reconstrucción de bandas altas.

[0076] Con esta estrategia, no es necesario transferir datos adicionales (por ejemplo, el factor de desvanecimiento) desde decodificador de núcleo 2 hacia el codificador de extensión de ancho de banda 3. Esto hace que la técnica sea fácilmente aplicable a cualquier codificador 1 con extensión de ancho de banda 3, especialmente a SBR, en el que ya se lleva a cabo el cálculo de la ganancia de manera inherente (ecuación 1).

[0077] La ocultación del decodificador de audio según la invención 1 tiene en cuenta la pendiente de desvanecimiento del módulo de decodificación de banda de núcleo 2. Esto conduce al comportamiento previsto del desvanecimiento como un conjunto. Se evitan las situaciones en las que las energías de las bandas de frecuencia FB del módulo de decodificación de banda de núcleo 2 se desvanecen más lentamente que las energías de las bandas de frecuencia FB del módulo de ampliación de ancho de banda 3, que se harían perceptibles y causarían la impresión desagradable de un banda de la señal limitada.

- [0078]** Por otra parte, también se evitan las situaciones en las que las energías en las bandas de frecuencia FB del módulo de decodificación de banda de núcleo 2 se desvanecen más rápidamente que las energías de las bandas de frecuencia FB del módulo de ampliación de ancho de banda 3, lo que introduciría artefactos porque las 5 bandas de frecuencia FB de los módulos de extensión de ancho de banda 3 se amplifican excesivamente, en comparación con las bandas de frecuencia FB del módulo de decodificación de banda de núcleo 2.
- [0079]** En contraste con un decodificador sin desvanecimiento que tiene una extensión de ancho de banda con niveles de energía predefinidos (como por ejemplo un decodificador CELP/HVXC + SBR decodificador), que 10 conserva solamente la inclinación espectral de un cierto tipo de señal, funciona el decodificador de audio 1 de la invención de forma independiente con respecto a las características espectrales de las señales, de manera que se evita una degradación perceptualmente decodificada de la señal de audio AS.
- [0080]** La técnica propuesta podría ser utilizada con cualquier procedimiento de extensión de ancho de banda 15 (BWE) en la parte superior de un módulo de decodificación de banda de núcleo 2 (en lo que sigue: codificador de núcleo). La mayor parte de la técnica de extensión de ancho de banda se basa en la ganancia por banda entre los niveles originales de energía y los niveles de energía que obtienen después de copiar el espectro de núcleo. La técnica propuesta no actúa sobre las energías del marco de audio previo, como lo hace el estado de la técnica, sino sobre las ganancias del marco de audio previo AF1.
- [0081]** Cuando un marco de audio AF2 se pierde o es ilegible (o en otras palabras, si se produce una pérdida 20 de marco de audio AFL) las ganancias derivadas del último marco bueno se introducen en el proceso de decodificación normal del módulo de decodificación de banda de núcleo 2, que ajusta las energías de las bandas de frecuencia del módulo de extensión de ancho de banda 3 (véase la ecuación 1). Esto forma la ocultación. Cualquier 25 desvanecimiento, que se aplique en el módulo de decodificación de banda de núcleo por una ocultación del módulo de decodificación de banda de núcleo 2, se aplicará de forma automática a las energías de las bandas de frecuencia FB del módulo de ampliación de ancho de banda 3 mediante el bloqueo de las energías entre las bandas alta y baja.
- [0082]** En algunas formas de realización, el módulo de extensión de ancho de banda 3 comprende un módulo 30 que proporciona factor de ganancia 6 configurado para comunicar el factor de ganancia actual CGF al menos en el marco de audio actual AF2 en el que tiene lugar la pérdida de marco de audio AFL hacia el módulo de ajuste de energía 5.
- [0083]** En una forma de realización preferida, el módulo que proporciona el factor de ganancia 6 está 35 configurado de tal manera que en el marco de audio actual AF2 en el que tiene lugar la pérdida de marco de audio AFL el factor de ganancia actual CGF es el factor de ganancia del marco de audio previo AF1.
- [0084]** Esta forma de realización desactiva por completo el desvanecimiento contenido en el módulo de 40 decodificación de la extensión del ancho de banda 3 mediante el mero bloqueo de las ganancias derivadas para la última envoltura en el último tramo bueno. En otra forma de realización preferida, el módulo que proporciona el factor de ganancia 6 está configurado de tal manera que en el marco de audio actual AF2 en el que tiene lugar la pérdida de marco AFL el factor de ganancia actual CGS se calcula a partir del factor de ganancia del marco de audio previo y a partir de una clase de señal del marco de audio previo.
- [0085]** Esta forma de realización utiliza un clasificador de señales para calcular ganancias GCS sobre la base 45 de las ganancias pasadas y también adaptativamente sobre la clase de señal del marco anteriormente recibido AF1. Las clases de señal pueden referirse a clases de sonidos del habla tales como: obstruyente (con subclases: stop, africtivo, fricativo), sonoras (estas subclases: nasal, flap approximant, vocal), lateral, trill.
- [0086]** En una forma de realización preferida, el módulo que proporciona el factor de ganancia 6 está 50 configurado para calcular un número de marcos de audio subsiguientes en los que tiene lugar pérdidas de marco de audio AFL y configurado para llevar a cabo un procedimiento reductor del factor de ganancia en el que el número de marcos de audio subsiguientes en los que tiene lugar una pérdida de marco de audio AFL supera un número predefinido.
- [0087]** Si tiene lugar un fricativo inmediatamente antes de una pérdida de marco en forma de ráfaga 55 (múltiples pérdidas de marco AFL en marcos de audio subsiguientes AF), el desvanecimiento inherente por defecto del módulo de decodificación de banda de núcleo 2 puede ser demasiado lento para asegurar un sonido agradable y natural en combinación con el bloqueo de ganancia. El resultado percibido de esto puede ser un fricativo prolongado

con mucha energía en las bandas de frecuencia FB del módulo de extensión de ancho de banda 3. Por esta razón es posible que se lleve a cabo una verificación de múltiples pérdidas de marco AFL. Si esta verificación es positiva, puede llevarse a cabo un procedimiento para disminuir el factor de ganancia.

5 **[0088]** En una forma de realización preferida, el procedimiento para reducir el factor de ganancia comprende la etapa de reducción del factor de ganancia actual dividiendo el factor de ganancia actual por un primer número en el caso en que el factor de ganancia actual supera un primer umbral. Mediante estos rasgos se reducen las ganancias que superen un primer umbral (que puede determinarse empíricamente).

10 **[0089]** En una forma de realización preferida, el procedimiento para reducir el factor de ganancia comprende la etapa de reducir el factor de ganancia actual para lo cual se divide el factor de ganancia actual por un segundo número que es más grande que el primer número en el caso en que el factor de ganancia actual supera un segundo umbral que es mayor que el primer umbral. Estos rasgos aseguran que las ganancias extremadamente más elevadas disminuyan más rápidamente incluso. Todas las ganancias que superen el segundo umbral disminuirán
15 más rápidamente.

[0090] En algunas formas de realización el procedimiento para reducir el factor de ganancia comprende la etapa de ajuste del factor de ganancia actual en el primer umbral en el caso en que el umbral actual, después de su reducción, es inferior al primer umbral. Mediante estos aspectos se impide que las ganancias disminuidas caigan por
20 debajo del primer umbral.

[0091] En algunas formas de realización el módulo de extensión de ancho de banda 3 comprende un módulo generador de ruido 7 configurado para añadir ruido NOI al menos a una banda de frecuencia FB, en el que, en el marco de audio actual AF2 en el que tiene lugar la pérdida de marco de audio AFL, se utiliza una relación entre la
25 energía de la señal y la energía del ruido de al menos una banda de frecuencia FB del marco de audio previo AF1 para calcular la energía del ruido del marco de audio actual AF2.

[0092] En el caso en que haya un rasgo de ruido de fondo (es decir, componentes de ruido adicionales para conservar el ruido de la señal original) implementado en la extensión del ancho de banda 3, es necesario adoptar la
30 idea del bloqueo de la ganancia también hacia el ruido de fondo. Para lograr esto, niveles de energía de ruido de fondo de marcos no ocultos son convertidos en una relación de ruido, para lo cual se tiene en cuenta la energía de las bandas de frecuencia del módulo de extensión de ancho de banda. Se guarda la relación en una memoria intermedia y será la base del nivel de ruido en el caso de una ocultación. La principal ventaja es el mejor acoplamiento del ruido de fondo a la energía del codificador de núcleo debido a un cálculo de la relación.

[0093] En una forma de realización preferida el decodificador de audio 1 comprende un módulo para analizar espectros 8 configurado para establecer el espectro del marco de audio actual AF2 de la señal de audio de banda de núcleo CBS y para derivar la energía estimada para la señal EE para el marco vigente AF2 para al menos una
40 banda de frecuencia FB a partir del espectro del marco de audio actual AF2 de la señal de audio de banda de núcleo CBS.

En una forma de realización preferida el módulo de extensión de ancho de banda 3 comprende un módulo generador de señales 9 configurado para crear una señal en el dominio de las frecuencias en bruto RFS que tiene al menos una banda de frecuencia FB, que se comunica al módulo de ajuste de energía 5, sobre la base de la señal de audio de banda de núcleo CBS y de la corriente de bits BS.

45 En una forma de realización preferida el módulo de extensión de ancho de banda 3 comprende un módulo de síntesis de señales 10 configurado de manera que se produzca la señal de audio de extensión de ancho de banda BES a partir de la señal en el dominio de las frecuencias FDS.

La Figura 5 ilustra la formación de marcos de una forma de realización de un decodificador de audio 1 según la invención.

50

[0094] En algunas formas de realización el módulo que proporciona el factor de ganancia 6 está configurado de tal manera que, en el caso en que un marco de audio actual AF2, en el que no tiene lugar una pérdida de marco AFL, sigue sustancialmente un marco de audio previo AF1, en el que tiene lugar una pérdida de marco de audio AFL, el factor de ganancia recibido para el marco de audio actual AF2 se utiliza para el marco vigente AF2, si un
55 retardo DEL entre marcos de audio AF del módulo de extensión de ancho de banda 3 con respecto a los marcos de audio AF' del módulo decodificador de banda de núcleo 2 es más pequeño que un retardo umbral, mientras que el factor de ganancia tomado del marco de audio previo AF1 se utiliza para el marco vigente AF 2, si el retardo DEL entre marcos de audio AF del módulo de extensión de ancho de banda 3 con respecto a los marcos de audio AF' del módulo de decodificación de banda de núcleo 3 es mayor que el umbral de retardo.

- [0095]** Además de la ocultación, en el módulo de extensión de ancho de banda 3 debe prestarse especial atención a la formación de marcos. Los marcos de audio AF del módulo de extensión de ancho de banda y los marcos de audio AF' del módulo de decodificación de banda de núcleo 3 frecuentemente no están exactamente alineados pero podrían tener un determinado retardo DEL. Por ello puede suceder que un paquete perdido contenga datos de extensión de ancho de banda que estén retardados, con respecto a la señal de núcleo contenida en el mismo paquete.
- [0096]** El resultado en este caso es que el primer paquete bueno después de una pérdida puede contener datos de extensión para crear partes de las bandas de frecuencia FB del módulo de extensión de ancho de banda 3 del módulo de decodificación de banda previo del marco de audio de núcleo AF', que ya estaba oculto en el decodificador 2.
- [0097]** Por esta razón, es necesario considerar la formación de marcos durante la recuperación, en función de las respectivas propiedades del núcleo, módulo de decodificación y módulo de extensión de ancho de banda. Esto podría significar tratar el primer marco de audio o partes de él en el módulo de extensión de ancho de banda 3 como erróneos y no aplicar las ganancias más nuevas de una vez sino conservar las ganancias bloqueadas del primer marco de audio para un marco adicional.
- [0098]** Conservar o no las ganancias bloqueadas para el primer marco bueno depende del retardo. La aplicación experimental a códecs con diferentes retardos mostró diferentes beneficios para códecs con diferentes retardos. Para códecs con retardos más bien pequeños (por ejemplo, de 1 ms), es mejor utilizar las ganancias más nuevas para el primer marco de audio bueno.
- [0099]** Aunque algunos aspectos han sido descritos en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del procedimiento correspondiente, en el que un recuadro o dispositivo corresponde a una etapa de procedimiento o a un rasgo de una etapa del procedimiento. De manera análoga, los aspectos descritos en el contexto de una etapa de procedimiento también representan una descripción de un recuadro, elemento o característica, correspondientes de un aparato correspondiente. Algunas de las etapas de procedimiento, o todas ellas, pueden ser llevadas a cabo mediante un aparato de hardware (o mediante la utilización de un aparato de hardware), tal como por ejemplo un microprocesador, un ordenador programable o un circuito electrónico. En algunas formas de realización, algunas, una o más de las etapas más importantes del procedimiento pueden ser llevadas a cabo mediante un aparato de este tipo.
- [0100]** En función de determinados requisitos para la implementación, las formas de realización de la invención pueden ser implementadas en hardware o en software. La implementación puede llevarse a cabo mediante un medio de almacenamiento digital, por ejemplo un floppy disk, un DVD, un Blu-Ray, un CD, un ROM; un PROM, un EPROM, un EEPROM, o una memoria FLASH, que tienen señales de control electrónicamente legibles almacenadas en ellos, que cooperan (o que son capaces de cooperar) con un sistema de ordenador programable de tal manera que se lleve a cabo el respectivo procedimiento. Por ello, el medio de almacenamiento digital puede ser legible por ordenador.
- [0101]** Algunas formas de realización según la invención comprenden un soporte de datos no transitorio que tiene señales de control electrónicamente legibles, que son capaces de cooperar con un sistema de ordenador programable, de tal manera que se lleve a cabo uno de los procedimientos descritos en esta invención.
- [0102]** En términos generales, las formas de realización de la presente invención pueden implementarse como un producto de programa informático con un código de programa, siendo el código del programa operativo para llevar a cabo uno de los procedimientos cuando el producto de programa informático es ejecutado en un ordenador. El código de programa puede almacenarse por ejemplo en un soporte legible por máquina.
- [0103]** Otras formas de realización comprenden el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención, almacenado en un soporte legible por máquina.
- [0104]** En otras palabras, una forma de realización del procedimiento inventivo consiste, por tanto, en un programa informático que tiene un código de programa para llevar a cabo uno de los procedimientos descritos presenten esta invención, cuando el programa informático se ejecuta en un ordenador.
- [0105]** Por lo tanto, otra forma de realización de los procedimientos inventivos consiste en un soporte de

datos (o en un medio de almacenamiento digital, o en un medio legible por ordenador) que comprende, grabado en él, el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención. El soporte de datos, el medio de almacenamiento digital o el medio grabado son típicamente tangibles y/o no transitorios.

5 **[0106]** Por lo tanto, otra forma de realización del procedimiento inventivo consiste en una corriente de datos o en una secuencia de señales que representa el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención. La corriente de datos o la secuencia de señales pueden estar configuradas por ejemplo para ser transferida por medio de una conexión para la comunicación de datos, por ejemplo por intermedio de Internet.

10

[0107] Una forma de realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador, o un dispositivo lógico programable, configurado o adaptado para llevar a cabo uno de los procedimientos descritos en esta invención.

15 **[0108]** Otra forma de realización comprende un ordenador que tiene instalada en él el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención.

[0109] Otra forma de realización según la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, electrónicamente u ópticamente) un programa informático para llevar a cabo uno de los procedimientos descritos en esta invención en un receptor. El receptor puede ser por ejemplo un ordenador, un dispositivo móvil, un dispositivo de memoria o similar. El aparato o sistema puede comprender por ejemplo un servidor de archivos para transferir el programa informático hacia el receptor.

20 **[0110]** En algunas formas de realización, puede utilizarse un dispositivo lógico programable (por ejemplo un field programmable gate array) para llevar a cabo algunas de las funcionalidades, o todas ellas, de los procedimientos descritos en esta invención. En algunas formas de realización, un field programmable gate array puede cooperar con un microprocesador con el fin de llevar a cabo uno de los procedimientos descritos en esta invención. En términos generales es preferible que los procedimientos sean llevados a cabo mediante cualquier aparato de hardware.

30

[0111] Las formas de realización anteriormente descritas son meramente ilustrativas de los principios de la presente invención. Se entiende que para la persona experta en la técnica le serán evidentes modificaciones y variaciones en cuanto a las disposiciones y detalles descritos en esta invención. Por ello la finalidad es que la invención quede limitada solamente por los alcances de las reivindicaciones de patente pendientes y no por los detalles específicos presentados a modo de descripción y explicación de las formas de realización en esta invención.

Signos de referencia:

40 **[0112]**

- 1 Decodificador de audio
- 2 Módulo de decodificación de banda de núcleo
- 3 Módulo de extensión de ancho de banda
- 4 Combinador
- 45 5 Módulo de ajuste de la energía
- 6 Módulo que proporciona factor de ganancia
- 7 Módulo generador de ruido
- 8 Módulo de análisis de espectros
- 9 Módulo generador de señales
- 50 10 Módulo de síntesis de señales
- AS Señal de audio
- BS Corriente de bits
- AF Marco de audio
- CBS Señal de audio de banda de núcleo
- 55 BES Señal de audio de extensión de ancho de banda
- FDS Señal de dominio de frecuencias
- FB Banda de frecuencia
- AFL Pérdida de marco de audio
- CGF Factor de ganancia actual

EE Energía de señal estimada
 NOI Ruido
 DEL Retardo
 RFS Señal de dominio de frecuencias en bruto

5

Referencias:

[0113]

- 10 [3GP09] 3GPP; Technical Specification Group Services and System Aspects, Extended adaptive multi-rate – wideband (AMR-WB+) codec, 3GPP TS 26.290, 3rd Generation Partnership Project, 2009.
 [3GP12a] General audio codec audio processing functions; Enhanced aacPlus general audio codec; additional decoder tools (release 11), 3GPP TS 26.402, 3rd Generation Partnership Project, Sep 2012.
 [3GP12b] Speech codec speech processing functions; adaptive multi-rate – wideband (AMRWB) speech codec;
 15 error concealment of erroneous or lost frames, 3GPP TS 26.191, 3rd Generation Partnership Project, Sep 2012.
 [EBU10] EBU/ETSI JTC Broadcast, Digital audio broadcasting (DAB); transport of advanced audio coding (AAC) audio, ETSI TS 102 563, European Broadcasting Union, May 2010.
 [EBU12] Digital radio mondiale (DRM); system specification, ETSI ES 201 980, ETSI, Jun 2012.
 [ISO09] ISO/IEC JTC1/SC29/WG11, Information technology – coding of audio–visual objects – part 3: Audio,
 20 ISO/IEC IS 14496–3, International Organization for Standardization, 2009.
 [ITU08] ITU–T, G.718: Frame error robust narrow–band and wideband embedded variable bit–rate coding of speech and audio from 8–32 kbit/s, Recommendation ITU–T G.718, Telecommunication Standardization Sector of ITU, Jun 2008.
 [RR06] Sang–Uk Ryu and Kenneth Rose, Frame loss concealment for audio decoders employing spectral
 25 band replication, Convention Paper 6962, Electrical and Computer Engineering, University of California, Oct 2006, AES.
 [SKE06] Andreas Schneider, Kurt Krauss, and Andreas Ehret, Evaluation of real–time transport protocol configurations using aacplus, Convention paper 6789, AES, May 2006, Presented at the 120th Convention 2006 May 20–23.

30

REIVINDICACIONES

1. Decodificador de audio configurado para producir una señal de audio (AS) desde una corriente de bits (BS) que contiene marcos de audio (AF), comprendiendo el decodificador de audio (1):
- 5 un módulo de decodificación de banda de núcleo (2) configurado para derivar una señal de audio de banda de núcleo directamente decodificada (CBS) a partir de la corriente de bits (BS);
- un módulo de extensión de ancho de banda (3) configurado para derivar una señal de audio de extensión de ancho de banda paraméricamente decodificada (BES) a partir de la señal de audio de banda de núcleo (CBS) y a partir de la corriente de bits (BS), donde la señal de audio de extensión de ancho de banda (BES) está basada en una señal en el dominio de las frecuencias (FDS) que tiene al menos una banda de frecuencia (FB); y
- 10 un combinador (4) configurado para combinar la señal de audio de banda de núcleo (CBS) y la señal de audio de extensión de ancho de banda (BES) de manera que se produzca la señal de audio (AS);
- 15 en el que el módulo de extensión de ancho de banda (3) comprende un módulo de ajuste de energía (5) que está configurado de tal manera que en un marco de audio actual (AF2) en el que tiene lugar una pérdida de marco de audio (AFL), se establece una energía de señal ajustada para el marco de audio actual (AF2) para al menos una banda de frecuencia (FB);
- 20 sobre la base de un factor de ganancia actual (CGF) para el marco de audio actual (AF2), en el que el factor de ganancia actual (CGF) se deriva a partir de un factor de ganancia de un marco de audio previo (AF1) o a partir de la corriente de bits (BS), y
- 25 sobre la base de una energía estimada para la señal (EE) para al menos una banda de frecuencia, en la que la energía estimada para la señal (EE) se deriva a partir de un espectro del marco de audio actual (AF2) de la señal de audio de banda de núcleo (CBS).
- 30 2. Decodificador de audio según la reivindicación anterior, en el que el módulo de extensión de ancho de banda (3) comprende un módulo que proporciona factor de ganancia (6) configurado para comunicar el factor de ganancia actual (CGF) al menos en el marco de audio actual (AF2) en el que tiene la pérdida de marco de audio (AFL), al módulo de ajuste de energía (5).
- 35 3. Decodificador de audio según la reivindicación anterior, en el que el módulo que proporciona el factor de ganancia (6) está configurado de tal manera que en el marco de audio actual (AF2) en el que tiene lugar la pérdida de marco de audio (AFL) el factor de ganancia actual (CGF) es el factor de ganancia del marco de audio previo (AF1).
- 40 4. Decodificador de audio según la reivindicación 2 o 3, en el que el módulo que proporciona el factor de ganancia (6) está configurado de tal manera que en el marco de audio actual (AF2) en el que tiene lugar la pérdida de marco (AFL) el factor de ganancia actual (CGF) se calcula a partir del factor de ganancia del marco de audio previo (AF1) y a partir de una clase de señal del marco de audio previo (AF1).
- 45 5. Decodificador de audio según una de las reivindicaciones 2 a 4, en el que el módulo que proporciona el factor de ganancia (6) está configurado para calcular un número de marcos de audio subsiguientes en los que tienen lugar pérdidas de marco de audio (AFL) y configurado para llevar a cabo un procedimiento reductor de la ganancia en el caso en que el número de marcos de audio subsiguientes en los que tienen lugar pérdidas de marcos de audio (AFL) supere un número predefinido.
- 50 6. Decodificador de audio según la reivindicación anterior, en el que el procedimiento para reducir el factor de ganancia comprende la etapa de reducción del factor de ganancia actual para lo que se divide el factor de ganancia actual por un primer número en el caso en que el factor de ganancia actual supere un primer umbral.
- 55 7. Decodificador de audio según la reivindicación 5 o 6, en el que el procedimiento para reducir el factor de ganancia comprende la etapa de reducción del factor de ganancia actual para lo que se divide el factor de ganancia actual por un segundo número que es mayor que el primer número en el caso en que el factor de ganancia actual supere un segundo umbral que es mayor que el primer umbral.

8. Decodificador de audio según una de las reivindicaciones 5 a 7, en el que el procedimiento para reducir el factor de ganancia comprende la etapa de ajuste del factor de ganancia actual al primer umbral en el caso en que el umbral actual, después de su reducción, sea inferior al primer umbral.

5 9. Decodificador de audio según una de las reivindicaciones anteriores, en el que el módulo de extensión de ancho de banda (3) comprende un módulo generador de ruido (7) configurado para añadir ruido (NOI) al menos a una banda de frecuencia (FB), en el que, en el marco de audio actual (AF2) en el que tiene lugar la pérdida de marco de audio (AFL), se utiliza una relación entre la energía de la señal y la energía del ruido de al menos una banda de frecuencia (FB) del marco de audio previo (AF1) para calcular la energía del ruido del marco de audio
10 actual (AF2).

10. Decodificador de audio según una de las reivindicaciones anteriores, en el que el decodificador de audio (1) comprende un módulo de análisis de espectros (8) configurado para establecer el espectro del marco de audio actual (AF2') de la señal de audio de banda de núcleo (CBS) y para derivar la energía estimada para la señal
15 para el marco vigente (AF2) para al menos una banda de frecuencia (FB) a partir del espectro del marco de audio actual (AF2') de la señal de audio de banda de núcleo (CBS).

11. Decodificador de audio según una de las reivindicaciones 2 a 10, en el que el módulo que proporciona el factor de ganancia (6) está configurado de tal manera que, en el caso en que un marco de audio actual, en el que
20 no tiene lugar una pérdida de marco, siga sustancialmente un marco de audio previo, en el que tiene lugar una pérdida de marco, el factor de ganancia recibido para el marco de audio actual es utilizado para el marco vigente, si un retardo (DEL) entre marcos de audio (AF1, AF2) del módulo de extensión de ancho de banda (3) con respecto a los marcos de audio (AF1', AF2') del módulo de decodificación de banda de núcleo (2) es menor que un retardo umbral, mientras que el factor de ganancia a partir del marco de audio previo se utiliza para el marco vigente, si el
25 retardo (DEL) entre marcos de audio del módulo de extensión de ancho de banda con respecto a los marcos de audio del módulo de decodificación de banda de núcleo es mayor que el umbral de retardo.

12. Decodificador de audio según una de las reivindicaciones anteriores, en el que el módulo de extensión de ancho de banda (3) comprende un módulo generador de señales (9) configurado para crear una señal en el
30 dominio de las frecuencias en bruto (RFS) que tiene al menos una banda de frecuencia (FB), que es comunicada al módulo de ajuste de energía (5), sobre la base de la señal de audio de banda de núcleo (CBS) y de la corriente de bits (BS).

13. Decodificador de audio según una de las reivindicaciones anteriores, en el que el módulo de extensión de ancho de banda (3) comprende un módulo de síntesis de señales (10) configurado de manera que se produzca la
35 señal de audio de extensión de ancho de banda (BES) a partir de la señal en el dominio de las frecuencias (FDS).

14. Procedimiento para producir una señal de audio (AS) a partir de una corriente de bits (BS) que contiene marcos de audio (AF), comprendiendo el procedimiento las siguientes etapas:

40 derivación de una señal de audio de banda de núcleo directamente decodificada (CBS) a partir de la corriente de bits (BS);

45 derivación de una señal de audio de extensión de ancho de banda paraméricamente decodificada (BES) a partir de la señal de audio de banda de núcleo (CBS) y a partir de la corriente de bits (BS), en la que la señal de audio de extensión de ancho de banda (BES) se basa sobre una señal en el dominio de las frecuencias (FDS) que tiene al menos una banda de frecuencia (FB); y

50 combinación de la señal de audio de banda de núcleo (CBS) y la señal de audio de extensión de ancho de banda (BES) de manera que se produzca la señal de audio (AS);

en el que, en un marco de audio actual (AF2) en el que tiene una pérdida de marco de audio (AFL), se establece una energía de señal ajustada para el marco de audio actual (AF2) para al menos una banda de frecuencia (FB) sobre la base de un factor de ganancia actual (CGF) para el marco de audio actual (AF2), en el que el factor de ganancia
55 actual (CGF) se deriva a partir de un factor de ganancia a partir de un marco de audio previo (AF1), y

sobre la base de una energía estimada para la señal para al menos una banda de frecuencia (FB), en la que la energía estimada para la señal se deriva a partir de un espectro del marco de audio actual (AF2') de la señal de audio de banda de núcleo (CBS).

15. Programa informático adaptado para llevar a cabo, cuando se ejecuta en un ordenador o procesador, el procedimiento según la reivindicación 14.

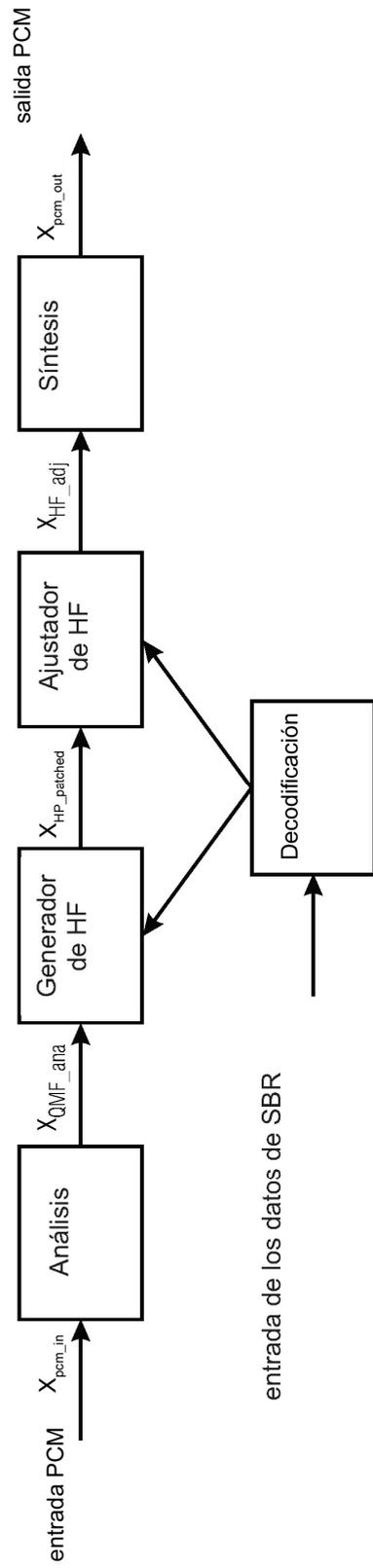


FIG 1

Decodificador de SBR del Estado de la Técnica

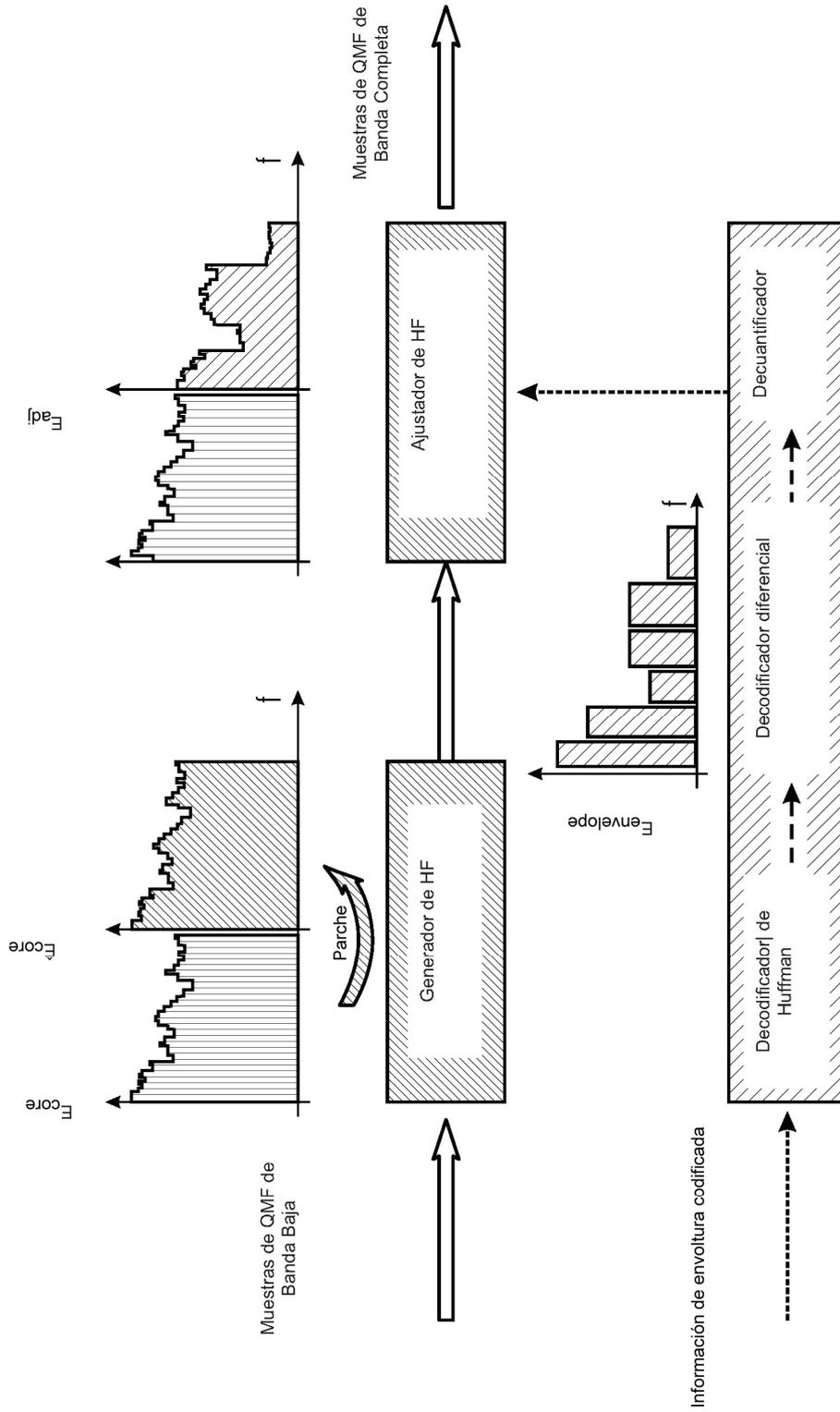


FIG 2

Decodificación de SBR: Generación y Ajuste de Señales de SBR fuera de la Información de envoltura codificada y Señal del Codificador de Núcleo

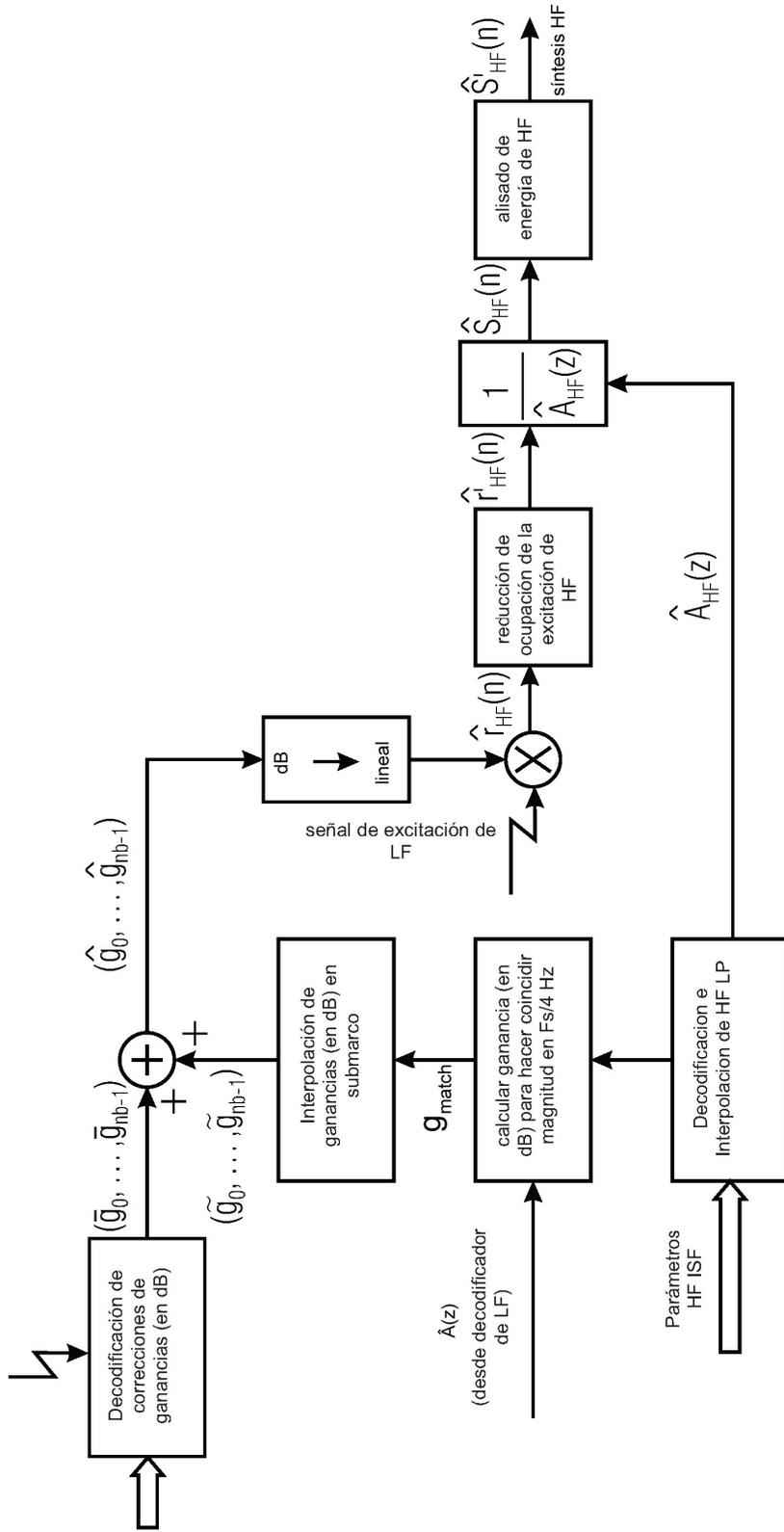


FIG 3

Extensión de Ancho de banda en AMR – WB + decodificador (Estado de la Técnica)

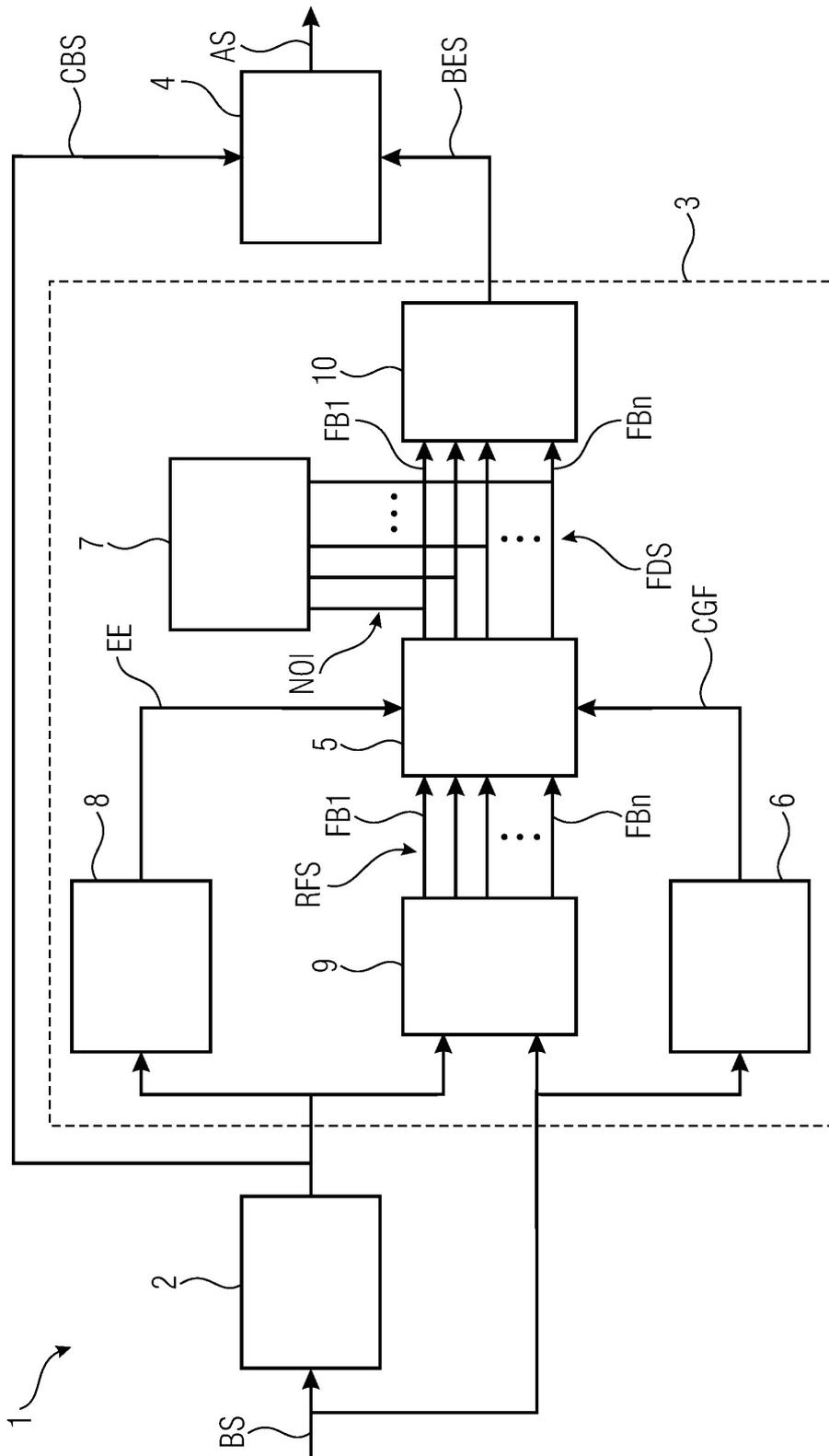


FIG 4

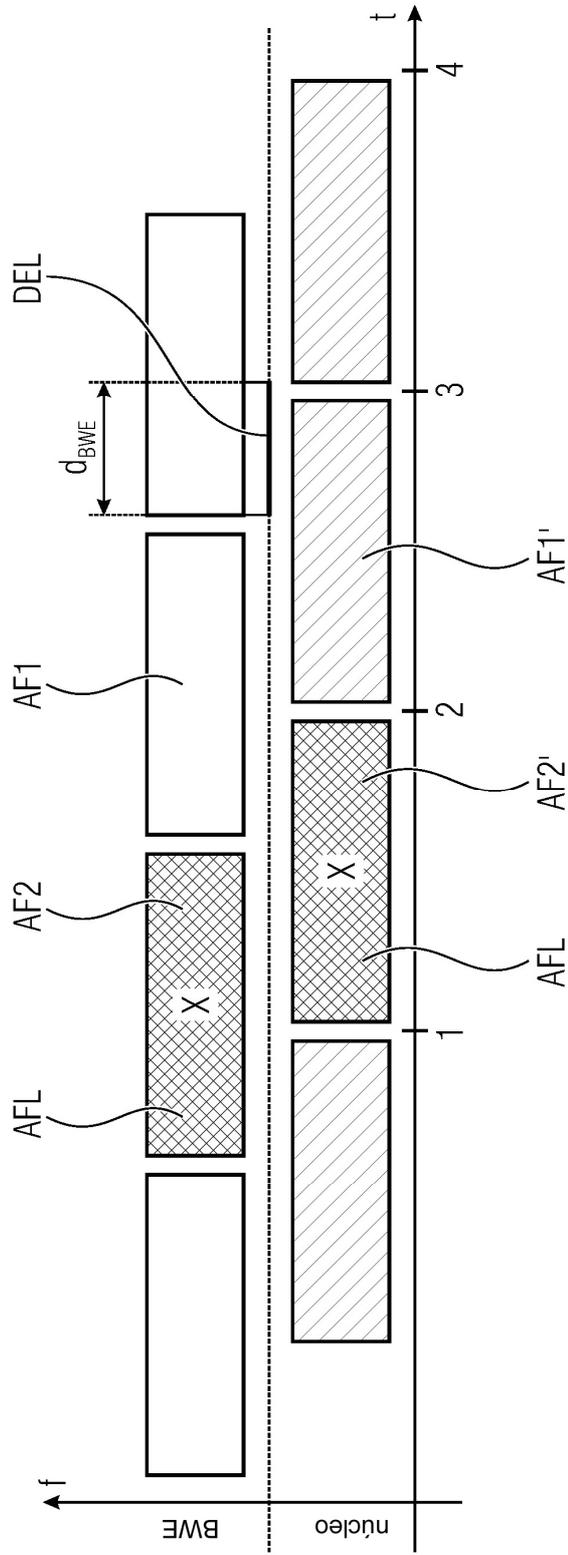


FIG 5