

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 724 576**

51 Int. Cl.:

G10L 21/038 (2013.01)

G10L 19/06 (2013.01)

G10L 19/08 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **24.06.2014 PCT/FR2014/051563**

87 Fecha y número de publicación internacional: **31.12.2014 WO14207362**

96 Fecha de presentación y número de la solicitud europea: **24.06.2014 E 14742262 (0)**

97 Fecha y número de publicación de la concesión europea: **13.03.2019 EP 3014611**

54 Título: **Extensión mejorada de la banda de frecuencia en un decodificador de señales de audiofrecuencia**

30 Prioridad:

25.06.2013 FR 1356100

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

12.09.2019

73 Titular/es:

**ORANGE (100.0%)
78, rue Olivier de Serres
75015 Paris, FR**

72 Inventor/es:

**KANIEWSKA, MAGDALENA y
RAGOT, STÉPHANE**

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 724 576 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

Extensión mejorada de la banda de frecuencia en un decodificador de señales de audiofrecuencia

5 La presente invención se refiere al campo de la codificación/decodificación y el tratamiento de señales de audiofrecuencia (tales como señales de voz, de música u otras) para su transmisión o su almacenamiento.

Más particularmente, la invención se refiere a un procedimiento y a un dispositivo de extensión de la banda de frecuencia en un decodificador o un procesador que realiza una mejora de la señal de audiofrecuencia.

10 Existen numerosas técnicas para comprimir (con pérdida) una señal de audiofrecuencia como la voz o la música.

Los métodos clásicos de codificación para las aplicaciones de conversación se clasifican en general en codificación de forma de onda (MIC por "Modulación por Impulso y Codificación", MICDA por "Modulación por Impulsos y Codificación Diferencial Adaptativa", codificación por transformada...), codificación paramétrica (LPC por "Linear Predictive Coding" en inglés, codificación sinusoidal...) y codificación híbrida paramétrica con una cuantificación de los parámetros por "análisis por síntesis" de las que la codificación CELP (por "Code Excited Linear Prediction" en inglés) es el ejemplo más conocido.

20 Para las aplicaciones no de conversación, el estado de la técnica de codificación de señales de audio (mono) está constituido por la codificación perceptual por transformada o en sub-bandas, con una codificación paramétrica de las altas frecuencias por replicación de banda.

Una revisión de los métodos clásicos de codificación de la voz y de audio se encuentra en las obras de W.B. Kleijn y K.K. Paliwal (Eds.), *Speech Coding and Synthesis*, Elsevier, 1995; M. Bosi, R.E. Goldberg, *Introduction to Digital Audio Coding and Standards*, Springer 2002; J. Benesty, M.M. Sondhi, Y. Huang (Eds.), *Handbook of Speech Processing*, Springer 2008.

30 En este caso se pone interés más particularmente en el códec (codificador y decodificador) normalizado 3GPP AMR-WB (por "Adaptive Multi-Rate Wideband" en inglés) que funciona a una frecuencia de entrada/salida de 16 kHz y en el que la señal se divide en dos sub-bandas, la banda base (0-6,4 kHz) que se muestrea a 12,8 kHz y se codifica por un modelo CELP y la banda alta (6,4-7 kHz) que se reconstruye de manera paramétrica por "extensión de banda" (o BWE por "Bandwidth Extension" en inglés) con o sin información suplementaria según el modo de la trama actual. Se puede observar en este caso que la limitación de la banda codificada del códec AMR-WB a 7 kHz está vinculada esencialmente al hecho de que la respuesta en frecuencia en emisión de los terminales en banda ampliada se ha aproximado con ocasión de la normalización (ETSI/3GPP posteriormente UIT-T) según la máscara de frecuencia definida en la norma UIT-T P.341 y más precisamente utilizando un filtro llamado "P341" definido en la norma UIT-T G.191 que corta las frecuencias por encima de 7 kHz (este filtro respeta la máscara definida en P.341). Sin embargo, en teoría, es bien conocido que una señal muestreada a 16 kHz puede tener una banda de audio definida de 0 a 8000 Hz; el códec AMR-WB introduce por tanto una limitación de la banda alta en comparación con la anchura de banda teórica de 8 kHz.

45 El códec de la voz 3GPP AMR-WB se ha normalizado en 2001 principalmente para las aplicaciones de telefonía en modo circuito (CS) sobre GSM (2G) y UMTS (3G). Este mismo códec se ha normalizado también en 2003 en la UIT-T en tanto que recomendación G.722.2 "Wideband coding speech at around 16kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)".

50 Comprende nueve velocidades, llamados modos, de 6,6 a 23,85 kbit/s, y comprende unos mecanismos de transmisión continua (DTX por "Discontinuous Transmission") con detección de actividad vocal (VAD por "Voice Activity Detection") y generación de ruido de confort (CNG por "Comfort Noise Generation") a partir de tramas de descripción del silencio (SID por "Silence Insertion Descriptor"), así como unos mecanismos de corrección de las tramas perdidas (FEC por "Frame Erasure Concealment", a veces llamado PLC por "Packet Loss Concealment").

55 No se repiten aquí los detalles del algoritmo de codificación y de decodificación AMR-WB, se encuentra una descripción detallada de este códec en las especificaciones 3GPP (TS 26.190, 26.191, 26.192, 26.193, 26.194, 26.204) y UIT-TG. 722.2 (y los Anexos y Apéndices correspondientes) así como en el artículo de B. Bessette et al. titulado "The adaptive multirate wideband speech codec (AMR-WB)", *IEEE Transactions on Speech and Audio Processing*, vol. 10, n.º 8, 2002, págs. 620-636 y los códigos de origen de las normas 3GPP y UIT-T asociados.

60 El principio de extensión de banda en el códec AMR-WB es bastante rudimentario. En efecto, la banda alta (6,4-7 kHz) se genera modelando un ruido blanco por medio de una envolvente temporal (aplicada bajo la forma de ganancias por subtrama) y en frecuencia (por la aplicación del filtro de síntesis de predicción lineal o LPC por "Linear Predictive Coding"). Esta técnica de extensión de banda se ilustra en la figura 1.

65 Se genera un ruido blanco, $u_{HB1}(n)$, $n = 0, \dots, 79$, 16 kHz por subtrama de 5 ms mediante generador de congruencia lineal (bloque 100). Este ruido $u_{HB1}(n)$ se modela en el tiempo mediante la aplicación de ganancias por subtrama; esta operación se descompone en dos etapas de tratamiento (bloque 102, 106 o 109):

- Se calcula un primer factor (bloque 101) para poner al ruido blanco $u_{HB1}(n)$ (bloque 102) a un nivel semejante al de la excitación, $u(n)$, $n = 0, \dots, 63$, decodificada a 12,8 kHz en banda base:

$$u_{HB2}(n) = u_{HB1}(n) \sqrt{\frac{\sum_{l=0}^{63} u(l)^2}{\sum_{l=0}^{79} u_{HB1}(l)^2}}$$

5 Se puede observar en este caso que la normalización de las energías se hace comparando unos bloques de tamaño diferente (64 para $u(n)$ y 80 para $u_{HB1}(n)$), sin compensación de las diferencias de frecuencia de muestreo (12,8 o 16 kHz).

- La excitación en la banda alta se obtiene a continuación (bloque 106 o 109) en la forma:

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

15 en la que la ganancia \hat{g}_{HB} se obtiene de modo diferente según la velocidad. Si la velocidad de la trama actual es <23,85 kbit/s, la ganancia \hat{g}_{HB} se estima "a ciegas" (es decir sin información suplementaria); en este caso, el bloque 103 filtra la señal decodificada en banda base mediante un filtro paso alto que tiene una frecuencia de corte a 400 Hz para obtener una señal $\hat{s}_{hp}(n)$, $n = 0, \dots, 63$ —este filtro paso alto elimina la influencia de muchas frecuencias bajas que pueden desviar la estimación realizada en el bloque 104— posteriormente se calcula el "tilt" (indicador de pendiente espectral) indicado como e_{tilt} de la señal $\hat{s}_{hp}(n)$ por auto-correlación normalizada (bloque 104):

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

25 y finalmente se calcula \hat{g}_{HB} en la forma:

$$\hat{g}_{HB} = w_{SP} g_{SP} + (1 - w_{SP}) g_{BG}$$

30 en la que $g_{SP} = 1 - e_{tilt}$ es la ganancia aplicada en las tramas activas de voz (SP por speech), $g_{BG} = 1,25 g_{SP}$ es la ganancia aplicada en las tramas inactivas de voz asociadas a un ruido de fondo (BG por background) y w_{SP} es una función de ponderación que depende de la detección de la actividad de voz (VAD). Se entiende que la estimación de la pendiente (e_{tilt}) permite adaptar el nivel de la banda alta en función de la naturaleza espectral de la señal; esta estimación es particularmente importante cuando la pendiente espectral de la señal decodificada CELP es tal que la energía media disminuye cuando aumenta la frecuencia (caso de una señal de voz en la que e_{tilt} está próxima a 1, por tanto $g_{SP} = 1 - e_{tilt}$ es así reducido). A observar también que el factor \hat{g}_{HB} en la decodificación AMR-WB está delimitado para tomar unos valores en el intervalo $[0, 1, 1, 0]$.

A 23,85 kbit/s, se transmite una información de corrección por el codificador AMR-WB y se decodifica (bloque 107, 108) con el fin de afinar la ganancia estimada por subtrama (4 bits cada 5 ms, es decir 0,8 kbit/s).

40 La excitación artificial $u_{HB}(n)$ se filtra a continuación (bloque 111) mediante un filtro de síntesis LPC (bloque 111) de función de transferencia $1/A_{HB}(z)$ y que funciona a la frecuencia de muestreo de 16 kHz. La realización de este filtro depende de la velocidad de la trama actual:

- A 6,6 kbit/s, el filtro $1/A_{HB}(z)$ se obtiene ponderando por un factor $\gamma = 0,9$ un filtro LPC de orden 20, $1/\hat{A}^{ext}(z)$ que "extrapola" el filtro LPC de orden 16, $1/\hat{A}(z)$, decodificado en la banda base (a 12,8 kHz) —los detalles de la extrapolación en el dominio de los parámetros ISF (por "Imittance Spectral Frequency" en inglés) se describen en la norma G.722.2 en la sección 6.3.2.1—; en este caso,

$$1/A_{HB}(z) = 1/\hat{A}^{ext}(z/\gamma)$$

- A las velocidades > 6,6 kbit/s, el filtro $1/A_{HB}(z)$ es de orden 16 y corresponde simplemente a:

$$1/A_{HB}(z) = 1/\hat{A}(z/\gamma)$$

en la que $\gamma = 0,6$. A observar que en este caso el filtro $1/\hat{A}(z/\gamma)$ se utiliza a 16 kHz, lo que da como resultado un escalonamiento (por homotecia) de la respuesta en frecuencia de este filtro de [0, 6,4 kHz] a [0, 8 kHz].

El resultado, $s_{HB}(n)$, se trata finalmente mediante un filtro paso banda (bloque 112) de tipo FIR (“Finite Impulse Response”), para no mantener más que la banda 6 - 7 kHz; a 23,85 kbit/s, se agrega un filtro paso bajo igualmente de tipo FIR (bloque 113) al tratamiento para atenuar aún más las frecuencias superiores a 7 kHz. La síntesis en altas frecuencias (AF) se adiciona finalmente (bloque 130) a la síntesis en bajas frecuencias (BF) obtenida con los bloques 120 a 123 y se remuestrea a 16 kHz (bloque 123). De ese modo incluso si la banda alta se extiende en teoría de 6,4 a 7 kHz en el códec AMR-WB, la síntesis de AF está más bien comprendida en la banda 6-7 kHz antes de la adición con la síntesis de BF.

Se pueden identificar varios inconvenientes a la técnica de extensión de la banda del códec AMR-WB:

- La señal en la banda alta es un ruido blanco conformado (por ganancias temporales por subtrama, mediante filtrado por $1/A_{HB}(z)$ y filtrado paso banda), lo que no es un buen modelo general de la señal en la banda 6,4-7 kHz. Existen por ejemplo unas señales de música muy armónicas para las que la banda 6,4-7 kHz contiene unos componentes sinusoidales (u otros) y ningún ruido (o poco ruido), para estas señales la extensión de banda del códec AMR-WB degrada fuertemente la calidad.
- El filtro paso bajo a 7 kHz (bloque 113) introduce una desviación de cerca de 1 ms entre las bandas bajas y altas, lo que puede degradar potencialmente la calidad de ciertas señales desincronizando ligeramente las dos bandas a 23,85 kbit/s —esta desincronización puede plantear igualmente problemas durante una conmutación de velocidad de 23,85 kbit/s a otros modos—.
- La estimación de las ganancias por subtrama (bloque 101, 103 a 105) no es óptima. En parte, se basa en una igualación de la energía “absoluta” por subtrama (bloque 101) entre unas señales a unas frecuencias diferentes: la excitación artificial a 16 kHz (ruido blanco) y una señal a 12,8 kHz (excitación ACELP decodificada). Se puede observar en particular que este planteamiento induce implícitamente una atenuación de la excitación en banda alta (con una relación $12,8/16 = 0,8$); de hecho, se observará igualmente que no se efectúa ninguna desacentuación (o de-énfasis) sobre la banda alta en el códec AMR-WB, lo que induce implícitamente una amplificación relativamente próxima a 0,6 (que corresponde al valor de la respuesta en frecuencia de $1/(1-0,68z^{-1})$ a 6400 Hz).

De hecho, los factores de 1/0,8 y 0,6 se compensan aproximadamente.

- En la voz, los ensayos de caracterización del códec 3GPP AMR-WB documentados en el informe 3GPP TR 26.976 han mostrado que el modo a 23,85 kbit/s tiene una calidad menos buena que a 23,05 kbit/s, su calidad es de hecho similar a la del modo a 15,85 kbit/s. Esto muestra en particular que el nivel de señal de AF artificial debe controlarse de manera muy prudente, porque la calidad se degrada a 23,85 kbit/s mientras que los 4 bits por trama se detectan permitiendo aproximar mejor la energía de las altas frecuencias originales.
- La limitación de la banda codificada a 7 kHz da como resultado la aplicación de un modelo estricto de la respuesta en emisión de los terminales acústicos (filtro P.341 en la norma UIT-T G.191). Ahora bien, para una frecuencia de muestreo de 16 kHz, las frecuencias en la banda 7-8 kHz continúan siendo importantes, en particular para las señales de música, para asegurar un buen nivel de calidad.

El algoritmo de decodificación AMR-WB se ha mejorado en parte con el desarrollo del códec escalable UIT-T G.718 que se ha normalizado 2008.

La norma UIT-T G.718 comprende un modo llamado de interoperabilidad, para el que la codificación núcleo es compatible con la codificación G.722.2 (AMR-WB) a 12,65 kbit/s; además, el decodificador G.718 tiene la particularidad de poder decodificar un tren binario AMR-WB/G.722.2 a todas las velocidades posibles del códec AMR-WB (de 6,6 a 23,85 kbit/s).

El decodificador interoperable G.718 en modo bajo retardo (“low delay” en inglés) (G.718-LD) se ilustra en la figura 2. Se listan a continuación las mejoras aportadas a la funcionalidad de decodificación del tren binario AMR-WB en el decodificador G.718, con referencias a la figura 1 cuando es necesario:

- La extensión de banda (descrita por ejemplo en la cláusula 7.13.1 de la recomendación G.718, bloque 206) es idéntica a la del decodificador AMR-WB, salvo que el filtro paso banda 6-7 kHz y el filtro de síntesis $1/A_{HB}(z)$ (bloques 111 y 112) están en orden inverso. Además, a 23,85 kbit/s los 4 bits transmitidos por subtrama por el codificador AMR-WB no se utilizan en el decodificador G.718 interoperable; la síntesis de las altas frecuencias (AF) a 23,85 kbit/s es por tanto idéntica a 23,05 kbit/s lo que evita el problema conocido de calidad de la decodificación AMR-WB a 23,85 kbit/s. Con más razón, no se utiliza el filtro paso bajo a 7 kHz (bloque 113), y se omite la decodificación específica de modo 23,85 kbit/s (bloques 107 a 109).
- Se implementa un pos-tratamiento de la síntesis a 16 kHz (véase la cláusula 7.14 de G.718) en G.718 por “noise gate” en el bloque 208 (para “mejorar” la calidad de los silencios por reducción del nivel), filtrado paso alto (bloque 209), pos-filtro de bajas frecuencias (llamado “bass postfilter”) en el bloque 210 atenuando el ruido interno

armónico en bajas frecuencias y una conversión en enteros de 16 bits con control de saturación (con control de la ganancia o AGC) en el bloque 211.

5 Sin embargo la extensión de banda en los códecs AMR-WB y/o G.718 (modo interoperable) permanece también limitada en varios aspectos.

- En particular, la síntesis de altas frecuencias por ruido blanco conformado (mediante un planteamiento temporal de tipo fuente-filtro de LPC) es un modelo muy limitado de la señal en la banda de las frecuencias superiores a 6,4 kHz.
- 10 • Solo la banda 6,4-7 kHz se vuelve a sintetizar de manera artificial, mientras que en la práctica es posible en teoría una banda mayor (hasta 8 kHz) a la frecuencia de muestreo de 16 kHz, lo que puede mejorar potencialmente la calidad de las señales, si no se tratan previamente mediante un filtro de tipo P.341 (50-7000 Hz) tal como se define en la *Software Tool Library* (norma G.191) de la UIT-T.

15 Otro ejemplo de extensión de banda a ciegas se divulga por la técnica anterior US 2003/050786 A1 que comprende un análisis LPC de la señal en banda estrecha, la utilización de estos coeficientes con el fin de obtener una envolvente de la señal en banda ampliada, la extensión de la señal de excitación en banda estrecha filtrada, un filtrado que permite obtener una señal en banda alta que se combina finalmente con la señal en banda estrecha para proporcionar una señal de banda ampliada.

20 Existe por tanto una necesidad de mejorar la extensión de banda en un códec de tipo AMR-WB o una versión interoperable de este códec o más generalmente para mejorar la extensión de banda de una señal de audio.

La presente invención se dirige a la mejora de la situación.

25 La invención proporciona con este fin, un procedimiento de extensión de banda de frecuencia de una señal de audiofrecuencia durante un proceso de decodificación o de mejora que incluye una etapa de decodificación o de extracción, en una primera banda de frecuencia llamada banda baja, de una señal de excitación y de unos coeficientes de un filtro de predicción lineal. El procedimiento es tal que incluye las siguientes etapas:

- obtención de una señal extendida en al menos una segunda banda de frecuencia superior a la primera banda de frecuencia a partir de la señal de excitación sobre-muestreada y extendida en la al menos una segunda banda de frecuencia;
- escalado de la señal extendida mediante una ganancia definida por subtrama en función de una relación entre la relación entre la energía por subtrama y la energía por trama de la señal de excitación de la banda baja y la relación entre la energía por subtrama y la energía por trama de la señal extendida;
- 35 - filtrado de dicha señal extendida escalada mediante un filtro de predicción lineal cuyos coeficientes se derivan de los coeficientes del filtro de la banda baja.

40 De ese modo, la consideración de la señal de excitación (procedente de la decodificación de la banda baja o de una extracción de la señal en banda baja) permite efectuar la extensión de banda con un modelo de señal más adaptado para ciertos tipos de señales tales como las señales de música.

45 En efecto, la señal de excitación decodificada o estimada en la banda baja incluye en ciertos casos unos armónicos que, cuando existen, pueden traspasarse a alta frecuencia de tal manera que esto permite asegurar un cierto nivel de armonía en la banda de alta reconstruida.

La extensión de banda según el procedimiento permite por tanto mejorar la calidad para este tipo de señales.

50 Además, la extensión de banda según el procedimiento se realiza extendiendo inicialmente una señal de excitación y aplicando a continuación una etapa de filtrado de síntesis; este planteamiento aprovecha el hecho de que la excitación decodificada en la banda baja es una señal cuyo espectro es relativamente plano, lo que evita los tratamientos de blanqueo de la señal decodificada que puedan existir en los métodos conocidos de extensión de banda en el campo de la frecuencia en el estado de la técnica.

55 Se observará que incluso aunque la invención está motivada por la mejora de la calidad de la extensión de banda en el contexto de la codificación AMR-WB interoperable, los diferentes modos de realización se aplican al caso más general de la extensión de banda de una señal de audio, en particular en un dispositivo de mejora que efectúa un análisis de la señal de audio para extraer los parámetros necesarios para la extensión de banda.

60 El hecho de tener en cuenta la energía a nivel de la trama actual y la de la subtrama en la señal en banda baja (primera banda de frecuencia) permite ajustar la relación entre la energía por subtrama y energía por trama en la banda alta (segunda banda de frecuencia) y ajustar de ese modo las relaciones de energía más que las energías absolutas. Esto permite mantener en la banda alta la misma relación de energía entre subtrama y trama que en la banda baja, lo que es particularmente beneficioso cuando la energía de las subtramas varía mucho, por ejemplo en el caso de sonidos transitorios, de ataques.

Los diferentes modos particulares de realización mencionados a continuación pueden añadirse independientemente o en combinación entre ellos, a las etapas del procedimiento de extensión definido anteriormente.

5 En un modo de realización, el procedimiento incluye además una etapa de filtrado paso banda adaptativo en función de la velocidad de decodificación de la trama actual.

Este filtrado adaptativo permite optimizar la anchura de banda extendida en función de la velocidad y por tanto la calidad de la señal reconstruida después la extensión de la banda. En efecto, para las bajas velocidades (típicamente a 6,6 y 8,85 kbit/s para AMR-WB), al no ser muy buena la calidad general de la señal decodificada en banda baja (mediante el códec AMR-WB o una versión interoperable), es preferible no extender demasiado la banda decodificada y por tanto limitar la extensión de banda adaptando la respuesta en frecuencia del filtro paso banda asociado para cubrir por ejemplo una banda aproximada de 6 a 7 kHz; esta limitación es tanto más ventajosa cuanto que la señal de excitación en sí misma esté relativamente mal codificada y es preferible no utilizar una sub-banda demasiado grande para la extensión de las altas frecuencias. En oposición, para las velocidades superiores (12,65 kbit/s y superiores para AMR-WB), la calidad puede mejorarse con una síntesis de AF que cubra una banda mayor, por ejemplo aproximadamente de 6 a 7,7 kHz. El límite alto de 7,7 kHz (en lugar de 8 kHz) es un ejemplo de realización, que se podrá ajustar a unos valores próximos a 7,7 kHz. Este límite se justifica aquí por el hecho de que la extensión se realiza en la invención sin información auxiliar y que una extensión hasta 8 kHz (incluso aunque es posible en teoría) podría dar como resultado artificiosidades para unas señales particulares. Además, esta limitación a 7,7 kHz tiene en cuenta el hecho de que típicamente los filtros anti-repliegue en conversión analógica/digital y los filtros de re-muestreo entre 16 kHz y otras frecuencias no son perfectos e introducen típicamente un rechazo a las frecuencias inferiores a 8 kHz.

25 En un modo posible de realización el procedimiento incluye una etapa de transformada tiempo-frecuencia de la señal de excitación, efectuándose entonces la etapa de obtención de una señal extendida en el campo de la frecuencia y una etapa de transformada tiempo-frecuencia inversa de la señal extendida antes de las etapas de escalado y de filtrado.

30 La implementación de la extensión de banda (de la señal de excitación) en el campo de la frecuencia permite obtener una finura de análisis en frecuencia de la que no se dispone con un planteamiento temporal, y permite también tener una resolución en frecuencia suficiente para detectar unos armónicos y transponer en altas frecuencias unos armónicos de la señal (en la banda baja) para mejorar la calidad en tanto que se respeta la estructura de la señal.

35 En un modo de realización detallado, la etapa de generación de una señal de excitación sobremuestreada y extendida se efectúa según la siguiente ecuación:

$$U_{HB1}(k) = \begin{cases} 0 & k = 0, \dots, 199 \\ U(k) & k = 200, \dots, 239 \\ U(k + start_band - 240) & k = 240, \dots, 319 \end{cases}$$

40 siendo k el índice de la muestra, $U_{HB1}(k)$ el espectro de la señal de excitación extendida, $U(k)$ el espectro de la señal de excitación obtenida después de la etapa de transformada y *start_band* una variable predefinida.

45 De ese modo, esta función comprende claramente un re-muestreo de la señal de excitación añadiendo unas muestras al espectro de esta señal.

En la banda de frecuencia correspondiente a las muestras que van de 200 a 239, se conserva el espectro original, para poder aplicar en él una respuesta de atenuación progresiva del filtro paso alto en esta banda de frecuencia y también para no introducir los defectos 9 audibles durante la etapa de adición de la síntesis de baja frecuencia a la síntesis de alta frecuencia.

50 En un modo de realización particular, el procedimiento incluye una etapa de filtrado de desacentuación de la señal extendida al menos en la segunda banda de frecuencia.

55 De ese modo, la señal en la segunda banda de frecuencia se trae a un campo coherente con la señal en la primera banda de frecuencia.

En un modo particular de realización, el procedimiento incluye además una etapa de generación de una señal de ruido al menos en la segunda banda de frecuencia obteniéndose la señal extendida por combinación de la señal de excitación extendida y de la señal de ruido.

60 En efecto, es suficiente tener unas características procedentes de la señal de excitación sobre-muestreada y extendida en al menos una segunda banda de frecuencia para tener un modelo de señal adaptado a ciertos tipos de

señales. Esto puede combinarse con otra señal, por ejemplo un ruido generado para obtener la señal extendida que tiene un modelo de señal adaptado.

5 En un modo de realización la etapa de combinación se efectúa por mezcla aditiva adaptativa con una ganancia de igualación del nivel entre la señal de excitación extendida y la señal de ruido.

La aplicación de esta ganancia de igualación permite a la etapa de combinación adaptarse a las características de la señal para optimizar la proporción relativa de ruido en la mezcla.

10 La presente invención se dirige igualmente a un dispositivo de extensión de banda de frecuencia de una señal de audiofrecuencia que incluye una etapa de decodificación o de extracción, en una primera banda de frecuencia llamada banda baja, de una señal de excitación y de los coeficientes de un filtro de predicción lineal. El dispositivo es tal que incluye:

- 15 - un módulo de obtención de una señal extendida ($U_{HB2}(k)$, 503) en al menos una segunda banda de frecuencia superior a la primera banda de frecuencia a partir de la señal de excitación sobre-muestreada y extendida en la al menos una segunda banda de frecuencia ($U_{HB1}(k)$);
- 20 - escalado (507) de la señal extendida mediante una ganancia definida por subtrama en función de una relación entre la relación entre la energía por subtrama y la energía por trama de la señal de excitación de la banda baja y la relación entre la energía por subtrama y la energía por trama de la señal extendida;
- un módulo de filtrado (510) de dicha señal extendida escalada mediante un filtro de predicción lineal cuyos coeficientes se derivan de los coeficientes del filtro de la banda baja.

Este dispositivo presenta las mismas ventajas que el procedimiento descrito anteriormente, que él implementa.

25 La invención se dirige a un decodificador que incluye un dispositivo tal como se ha descrito.

Se dirige a un programa informático que incluye unas instrucciones de código para la implementación de las etapas del procedimiento de extensión de banda tal como se ha descrito, cuando estas instrucciones se ejecutan por un procesador.

30 Finalmente la invención se refiere a un soporte de almacenamiento, legible por un procesador, integrado o no en el dispositivo de extensión de banda, eventualmente extraíble, que memoriza un programa informático que implementa el procedimiento de extensión de banda al como se ha descrito anteriormente.

35 Surgirán más claramente otras características y ventajas de la invención con la lectura de la descripción siguiente, dada únicamente a título de ejemplo no limitativo, y realizada con referencia a los dibujos adjuntos, en los que:

- 40 - la figura 1 ilustra una parte de un decodificador de tipo AMR-WB que implementa unas etapas de extensión de banda de frecuencia del estado de la técnica tal como se ha descrito anteriormente;
- la figura 2 ilustra un decodificador de tipo interoperable G.718-LD a 16 kHz según el estado de la técnica y tal como se ha descrito anteriormente;
- 45 - la figura 3 ilustra un decodificador interoperable con la codificación AMR-WB y que integra un dispositivo de extensión de banda según un modo de realización de la invención;
- la figura 4 ilustra en forma de organigrama, las etapas principales de un procedimiento de extensión de banda según un modo de realización de la invención;
- la figura 5 ilustra un primer modo de realización en el campo de la frecuencia de un dispositivo de extensión de banda según la invención;
- 50 - la figura 6 ilustra un ejemplo de respuesta en frecuencia de un filtro paso banda utilizado en un modo particular de la invención;
- la figura 7 ilustra un segundo modo de realización en el campo temporal de un dispositivo de extensión de banda según la invención y
- la figura 8 ilustra una realización material de un dispositivo de extensión de banda según la invención.

55 La figura 3 ilustra un ejemplo de decodificador, compatible con la norma AMR-WB/G.722.2 en el que se encuentra un pos-tratamiento similar al introducido en G.718 y descrito con referencia a la figura 2 y una extensión de banda mejorada según el procedimiento de extensión de la invención, implementado mediante el dispositivo de extensión de banda ilustrado por el bloque 309.

60 Contrariamente a la decodificación AMR-WB que funciona con una frecuencia de muestreo de salida de 16 kHz y a la decodificación G.718 que funciona a 8 o 16 kHz, se considera en este caso un decodificador que puede funcionar con una señal de salida (síntesis) a la frecuencia de $f_s = 8, 16, 32$ o 48 kHz. A observar que se supone en este caso que la codificación se ha efectuado según el algoritmo AMR-WB con una frecuencia interna de $12,8$ kHz para la codificación CELP en banda baja y $23,85$ kbit/s una codificación de ganancia por subtrama a la frecuencia de 16 kHz; aunque la invención se describe en este caso en la decodificación, se supone en este caso que la codificación puede funcionar también con una señal de entrada a la frecuencia $f_s = 8, 16, 32$ o 48 kHz y se implementan unas

operaciones de re-muestreo adecuadas, que sobrepasan el marco de la invención, en la codificación en función del valor de f_s . Se puede observar que cuando $f_s = 8$ kHz, en el caso de una decodificación compatible con AMR-WB, no es necesario extender la banda baja 0-6,4 kHz, porque la banda de audio reconstruida a la frecuencia f_s está limitada a 0-4000 Hz.

En la figura 3, la decodificación CELP (BF para bajas frecuencias) funciona siempre a la frecuencia interna de 12,8 kHz, como en AMR-WB y G.718, y la extensión de banda (AF para altas frecuencias) que forma el objeto de la invención funciona a la frecuencia de 16 kHz, la síntesis de BF y AF se combinan (bloque 312) a la frecuencia f_s después del re-muestreo adecuado (bloque 306 y tratamiento interno del bloque 311). En unas variantes de la invención, la combinación de las bandas baja y alta podrá hacerse a 16 kHz, después de haber re-muestreado la banda baja de 12,8 a 16 kHz, antes de re-muestrear la señal extendida a la frecuencia f_s .

La decodificación según la figura 3 depende del modo (o velocidad) AMR-WB asociado a la trama actual recibida. A título indicativo y sin que esto impacte en el bloque 309, la decodificación de la parte CELP en banda baja incluye las siguientes etapas:

- Demultiplexado de los parámetros codificados (bloque 300) en caso de trama correctamente recibida ($bfi=0$ en la que bfi es el "bad frame indicator" que vale 0 para una trama recibida y 1 para una trama perdida)
- Decodificación de los parámetros ISF con interpolación y conversión en coeficientes LPC (bloque 301) como se describe en la cláusula 6.1 de la norma G.722.2.
- Decodificación de la excitación CELP (bloque 302) con una parte adaptativa y fija para reconstruir la excitación (exc o $u'(n)$) en cada subtrama de longitud 64 a 12,8 kHz:

$$u'(n) = \hat{g}_p v(n) + \hat{g}_c c(n) , n = 0, \dots, 63$$

siguiendo las notaciones de la cláusula 7.1.2.1 de G.718 que se refieren a la decodificación CELP, en la que $v(n)$ y $c(n)$ son respectivamente las palabras de códigos de los diccionarios adaptativo y fijo, y \hat{g}_p y \hat{g}_c son las ganancias decodificadas asociadas, esta excitación $u'(n)$ se utiliza en el diccionario adaptativo de la subtrama siguiente; se post-trata a continuación y se distingue como en G.718 la excitación $u'(n)$ (también indicada por exc) de su versión post-tratada modificada $u(n)$ (también indicada por exc2) que sirve de entrada al filtro de síntesis, $1/\hat{A}(z)$, en el bloque 303. En unas variantes que pueden implementarse para la invención, los pos-tratamientos aplicados a la excitación pueden modificarse (por ejemplo, la dispersión de fase puede mejorarse) o pueden extenderse estos pos-tratamientos (por ejemplo, puede implementarse una reducción de ruido inter-armónico), sin afectar a la naturaleza del procedimiento de extensión de banda según la invención.

- Filtrado de síntesis mediante $1/\hat{A}(z)$ (bloque 303) en el que el filtro LPC decodificado $\hat{A}(z)$ es de orden 16.
- Pos-tratamiento en banda estrecha (bloque 304) según la cláusula 7.3 de G.718 si $f_s = 8$ kHz.
- Desacentuación (bloque 305) por el filtro $1/(1-0,68z^{-1})$.
- Pos-tratamiento de las bajas frecuencias (bloque 306) tal como se describe en la cláusula 7.14.1.1 de G.718. Este tratamiento introduce un retardo que se tiene en cuenta en la decodificación de la banda alta (>6,4 kHz).
- Re-muestreo de la frecuencia interna de 12,8 kHz a la frecuencia de salida f_s (bloque 307). Son posibles varias realizaciones. Sin pérdida de generalidad, se considera en este caso a título de ejemplo que si $f_s = 8$ o 16 kHz, el re-muestreo descrito en la cláusula 7.6 de G.718 se retoma en este caso, y si $f_s = 32$ o 48 kHz, se utilizan unos filtros de respuesta a impulsos finita (FIR) suplementarios.
- Cálculo de los parámetros de la "noise gate" (bloque 308) que se realiza de manera preferente como se describe en la cláusula 7.14.3 de G.718.

Se puede observar que la utilización de los bloques 306, 308, 314 es opcional.

Se observa igualmente que la decodificación de la banda baja descrita anteriormente supone una trama actual llamada "activa" con una velocidad entre 6,6 y 23,85 kbit/s. De hecho, cuando está activado el modo DTX (transmisión continua en español), pueden codificarse ciertas tramas como "inactivas" y en este caso se puede transmitir un descriptor de silencio (sobre 35 bits) o bien no transmitir nada. En particular, se recuerda que la trama SID describe varios parámetros: parámetros ISF medios sobre 8 tramas, energía media sobre 8 tramas, marcador de dispersión para la reconstrucción de ruido no estacionario. En todos los casos, en el decodificador, se encuentra el mismo modelo de decodificación que para una trama activa, con una reconstrucción de la excitación y un filtro LPC para la trama actual, lo que permite aplicar la invención incluso sobre tramas inactivas. Se aplica la misma constante para la decodificación de "tramas perdidas" (o FEC, PLC) en las que se aplica el modelo LPC.

Contrariamente a la decodificación AMR-WB o G.718, el decodificador según la invención permite extender la banda baja decodificada (50-6400 Hz teniendo en cuenta el filtrado paso alto a 50 Hz en el decodificador, 0-6400 Hz en el caso general) a una banda extendida cuya anchura varía, siendo aproximadamente de 50-6900 Hz a 50-7700 Hz en función del modo implementado en la trama actual. Se puede hablar así de una primera banda de frecuencia de 0 a 6400 Hz y una segunda banda de frecuencia de 6400 a 8000 Hz. En realidad, en el modo de realización preferente,

la extensión de la excitación se realiza en el campo de la frecuencia en una banda de 5000 a 8000 Hz, para permitir un filtrado paso banda de anchura 6000 a 6900 o 7700 Hz.

5 En el modo de realización preferente, a 23,85 kbit/s, como en el decodificador G.718 descrito con referencia a la figura 2, la información de corrección de ganancia de AF (0,8 kbit/s) transmitida a 23,85 kbit/s se ignora en este caso. De ese modo en la figura 3, no se utiliza ningún bloque específico a 23,85 kbit/s.

10 La parte de decodificación de banda alta se realiza en el bloque 309 que representa el dispositivo de extensión de banda según la invención y que se detalla en la figura 5 en un primer modo de realización y en la figura 7 en un segundo modo de realización.

15 Este dispositivo comprende el menos un módulo de obtención de una señal extendida en al menos una segunda banda de frecuencia superior a la primera banda de frecuencia a partir de una señal de excitación sobre-muestreada y extendida en al menos una segunda banda de frecuencia ($U_{NB1}(k)$), un módulo de escalado de la señal extendida mediante una ganancia definida por subtrama en función de una relación de energía por trama y subtrama de la señal de audiofrecuencia en la primera banda de frecuencia y un módulo de filtrado de dicha señal extendida escalada mediante un filtro de predicción lineal cuyos coeficientes se derivan de los coeficientes de filtro de la banda baja.

20 Con el fin de alinear las bandas bajas y altas decodificadas, se introduce un retardo (bloque 310) en el primer modo de realización para sincronizar las salidas de los bloques 306 y 307 y la banda alta sintetizada a 16 kHz se re-muestra a 16 kHz a la frecuencia f_s (salida del bloque 311). Por ejemplo, cuando $f_s = 16$ kHz el retardo $T = 30$ muestras, que corresponde al retardo de re-muestreo de 12,8 a 16 kHz de 15 muestras + retardo del pos-tratamiento de las bajas frecuencias de 15 muestras. El valor del retardo T deberá adaptarse para los otros casos ($f_s = 32, 48$ kHz) en función de los tratamientos implementados. Se recuerda que cuando $f_s = 8$ kHz, no es necesario aplicar los bloques 309 a 311 porque la banda de la señal en la salida del decodificador está limitada a 0-4000 Hz.

30 A observar que el procedimiento de extensión de la invención implementado en el bloque 309 según el primer modo de realización no introduce de manera preferente ningún retardo suplementario con relación a la banda baja reconstruida a 12,8 kHz; sin embargo, en unas variantes de la invención (por ejemplo utilizando una transformación tiempo/frecuencia con recubrimiento), se podría introducir un retardo. De ese modo, de manera general el valor de T en el bloque 310 deberá ajustarse en función de la implementación específica. Por ejemplo en el caso en que no se utiliza el pos-tratamiento de las bajas frecuencias (bloque 306), el retardo a introducir para $f_s = 16$ kHz podrá fijarse en $T = 15$ muestras; igualmente si la invención se realiza según la variante del modo de realización descrito en la figura 7, el valor de T se reduce para compensar el retardo introducido por el pos-tratamiento de las bajas frecuencias (bloque 306) si se utiliza.

40 Se combinan (añaden) a continuación las bandas baja y alta en el bloque 312 y la síntesis obtenida se pos-trata mediante filtrado paso alto a 50 Hz (de tipo IIR) de orden 2 cuyos coeficientes dependen de la frecuencia f_s (bloque 313) y pos-tratamiento de salida con aplicación opcional de la "noise gate" de manera similar a G.718 (bloque 314).

45 El dispositivo de extensión de banda según la invención, ilustrado por el bloque 309 según el modo de realización del decodificador de la figura 3, implementa un procedimiento de extensión de banda descrito ahora con referencia a la figura 4.

50 Este dispositivo de extensión puede ser igualmente independiente del decodificador y puede implementar el procedimiento descrito en la figura 4 para efectuar una extensión de banda de una señal de audio existente almacenada o transmitida al dispositivo, con un análisis de la señal de audio para extraer de ella una excitación y un filtro LPC.

55 Este dispositivo recibe en la entrada una señal de excitación en una primera banda de frecuencia llamada banda baja $u(n)$ en el caso de una implementación en el campo temporal o $U(k)$ en el caso de una implementación en el dominio de la frecuencia para el que se aplica entonces una etapa de transformada tiempo frecuencia.

En el caso de una aplicación en un decodificador, esta señal de excitación recibida es una señal decodificada.

En el caso de un dispositivo de mejora independiente del decodificador, la señal de excitación de banda baja se extrae por análisis de la señal de audio.

60 En un modo posible de realización, la señal de audio de banda baja se re-muestra antes de la etapa de extracción de la excitación, aunque la excitación extraída de la señal de audio por predicción lineal estimada a partir de la señal de banda baja (o de parámetros LPC asociados a la banda baja) ya está re-muestreada. Un ejemplo de realización en este caso consiste en tomar una señal de banda baja muestreada a 12,8 kHz de la que se dispone un filtro LPC de banda baja que describe la envolvente espectral a corto plazo para la trama actual, sobre-muestrearla a 16 kHz, y filtrarla mediante un filtro de predicción LPC obtenido extrapolando el filtro LPC. Otro ejemplo de realización consiste en tomar una señal de banda baja muestreada a 12,8 kHz de la que no se dispone de modelo LPC, sobre-

muestrearla a 16 kHz, efectuar un análisis LPC sobre esta señal a 16 kHz, y filtrar esta señal mediante un filtro de predicción LPC obtenido mediante este análisis.

5 Se efectúa una etapa E401 de generación de una señal de excitación sobre-muestreada extendida ($u_{ext}(n)$ o $U_{HB1}(k)$) en una segunda banda de frecuencia superior a la primera banda de frecuencia. Esta etapa de generación debe incluir a la vez una etapa de re-muestreo y una etapa de extensión o simplemente una etapa de extensión en función de la señal de excitación obtenida en la entrada.

10 Esta etapa se detalla posteriormente en los modos de realización descritos con referencia a las figuras 5 y 7.

Esta señal de excitación sobre-muestreada extendida se utiliza para obtener una señal extendida ($U_{HB2}(k)$) en una segunda banda de frecuencia. Esta señal extendida posee entonces un modelo de señal adaptado a ciertos tipos de señales gracias a las características de la señal de excitación extendida.

15 Esta señal extendida pueda obtenerse tras la combinación de la señal de excitación sobre-muestreada y extendida con otra señal, por ejemplo una señal de ruido.

20 De ese modo, en un modo de realización, se efectúa una etapa E402 de generación de una señal de ruido ($u_{HB}(n)$ o $U_{HB}(k)$) al menos en la segunda banda de frecuencia. La segunda banda de frecuencia es por ejemplo una banda de alta frecuencia que va de 6000 a 8000 Hz. Por ejemplo, este ruido puede generarse de manera pseudoaleatoria mediante un generador de congruencia lineal. En unas variantes de la invención, se podrá sustituir esta generación de ruido por otros métodos, por ejemplo se podrá definir una señal de amplitud constante (de valor arbitrario, tal como 1) y aplicar unas señales aleatorias a cada raya de frecuencia generada.

25 La señal de excitación extendida se combina a continuación con la señal de ruido en la etapa E403 para obtener la señal extendida que se podrá denominar igualmente señal combinada ($u_{HB1}(n)$ o $U_{HB2}(k)$) en la banda de frecuencia extendida correspondiente a toda la banda de frecuencia que incluye la primera y la segunda banda de frecuencia. De ese modo, la combinación de estos dos tipos de señales permite obtener una señal combinada con unas características más adaptadas a ciertos tipos de señales como unas señales musicales.

30 En efecto, la señal de excitación decodificada o estimada en la banda baja incluye en ciertos casos unos armónicos más próximos a unas señales musicales que la señal de ruido sola. Los armónicos de baja frecuencia, si existen, pueden transponerse a la alta frecuencia de tal manera que su mezcla con el ruido permita asegurar un cierto nivel de armonicidad o de nivel relativo de ruido o de plenitud espectral ("spectral flatness" en inglés) en la banda alta reconstruida.

35 La extensión de banda según el procedimiento mejora la calidad para este tipo de señales con relación a AMR-WB.

40 La señal combinada (o extendida) se filtra a continuación en E404 mediante un filtro de predicción lineal cuyos coeficientes se derivan de los coeficientes del filtro de la banda baja ($\hat{A}(z)$) decodificados u obtenidos mediante análisis y extracción a partir de la señal de banda baja o una versión sobre-muestreada de esta. La extensión de banda según el procedimiento se realiza por tanto extendiendo inicialmente una señal de excitación y aplicando a continuación una etapa de filtrado de síntesis mediante predicción lineal (LPC); este planteamiento aprovecha el hecho de que la excitación LPC decodificada en la banda baja es una señal cuyo espectro es relativamente plano, lo que evita los tratamientos suplementarios de blanqueado de la señal decodificada en la extensión de banda.

45 De manera ventajosa, los coeficientes de este filtro pueden obtenerse por ejemplo a partir de los parámetros decodificados del filtro de predicción lineal (LPC) en banda baja. Si el filtro LPC utilizado en banda alta muestreado a 16 kHz es de la forma $1/\hat{A}(z/\gamma)$, en la que $1/\hat{A}(z)$ es el filtro decodificado en banda baja, y γ un factor de ponderación, la respuesta en frecuencia del filtro $1/\hat{A}(z/\gamma)$ corresponde a un escalonamiento de la respuesta en frecuencia del filtro decodificado en banda baja. En una variante se podrá extender el filtro $1/\hat{A}(z)$ a un orden superior (como a 6,6 kbit/s en el bloque 111) para evitar un escalonamiento así.

50 De manera preferida pero opcional, pueden efectuarse unas etapas suplementarias de filtrado paso banda adaptativo en E405 y/o de escalado en E407 para por un lado mejorar la calidad de la señal de extensión según la velocidad de decodificación y por otro lado para asegurar mantener la misma relación de energía entre una subtrama y una trama de señal combinada que en la banda de baja de frecuencia.

55 Estas etapas se explicarán más en detalle en los modos de realización de las figuras 5 y 7.

60 En un primer modo de realización, el dispositivo de extensión de banda se describe ahora con referencia a la figura 5. Este dispositivo implementa el procedimiento de extensión de banda descrito anteriormente con referencia a la figura 4.

De ese modo, a la entrada de este dispositivo, se recibe una señal de excitación en banda baja decodificada o estimada por análisis ($u(n)$). La extensión de banda utiliza en este caso la excitación decodificada a 12,8 kHz (exc2 o $u(n)$) en la salida del bloque 302.

5 Se observará que en este modo de realización, la generación de la excitación sobre-muestreada y extendida se efectúa en una banda de frecuencia que va de 5 a 8 kHz incluyendo por tanto una segunda banda de frecuencia (6,4-8 kHz) superior a la primera banda de frecuencia (0-6,4 kHz).

De ese modo, la generación de la señal de excitación extendida se efectúa al menos sobre la segunda banda de frecuencia pero también sobre una parte de la primera banda de frecuencia.

10 Por supuesto, los valores que definen estas bandas de frecuencia pueden ser diferentes según el decodificador o el dispositivo de tratamiento en el que se aplica la invención.

Para este ejemplo de realización, esta señal se transforma para obtener un espectro de señal de excitación $U(k)$ por el módulo de transformación tiempo-frecuencia 500.

15 En un modo de realización particular, la transformada utiliza una DCT-IV (por “*Discrete Cosine Transform*” – *Type IV* en inglés) (bloque 500) sobre la trama actual de 20 ms (256 muestras), sin creación de ventanas, lo que se convierte en transformar directamente $u(n)$ con $n = 0, \dots, 255$ según la siguiente fórmula:

$$20 \quad U(k) = \sum_{n=0}^{N-1} u(n) \cos\left(\frac{\pi}{N}\left(n + \frac{1}{2}\right)\left(k + \frac{1}{2}\right)\right)$$

en la que $N = 256$ y $k = 0, \dots, 255$.

25 Se observa en este caso que la transformación sin creación de ventanas (o de manera equivalente con una creación de ventana rectangular implícita de la longitud de la trama) es posible porque el tratamiento se efectúa en el campo de la excitación, y no en el dominio de la señal, aunque no sea audible ninguna artificiosidad (efectos de bloque), lo que constituye una ventaja importante de este modo de realización de la invención.

30 En este modo de realización, la transformación DCT-IV se implementa mediante FFT según el algoritmo llamado “Evolved DCT (EDCT)” descrito en el artículo de D.M. Zhang, H.T. Li, A Low Complexity Transform - Evolved DCT, IEEE 14th International Conference on Computational Science and Engineering (CSE), agosto de 2011, págs. 144-149, e implementado en las normas UIT-T G.718 Anexo B y G.729.1 Anexo E.

35 En unas variantes de la invención y sin pérdida de generalidad, la transformación DCT-IV se podrá sustituir por otras transformaciones tiempo-frecuencia a corto plazo de la misma longitud y en el campo de la excitación, como una FFT (por “*Fast Fourier Transform*” en inglés) o una DCT-II (*Discrete Cosine Transform - Type II*). De manera alternativa, se podrá sustituir la DCT-IV sobre la trama por una transformación con recubrimiento-adición y creación de ventanas de longitud superior a la longitud de la trama actual, por ejemplo utilizando una MDCT (por “*Modified Discrete Cosine Transform*” en inglés). En este caso el retardo T en el bloque 310 de la figura 3, deberá ajustarse (reducirse) de manera adecuada en función del retardo adicional debido al análisis/síntesis mediante esta transformada.

45 El espectro DCT, $U(k)$, de 256 muestras que cubren la banda 0-6400 Hz (a 12,8 kHz), se extiende a continuación (bloque 501) en un espectro de 320 muestras que cubren la banda 0-8000 Hz (a 16 kHz) en la forma siguiente:

$$U_{HB1}(k) = \begin{cases} 0 & k = 0, \dots, 199 \\ U(k) & k = 200, \dots, 239 \\ U(k + start_band - 240) & k = 240, \dots, 319 \end{cases}$$

en la que se toma de manera preferente $start_band = 160$.

50 El bloque 501 funciona como un módulo de generación de una señal de excitación sobre-muestreada y extendida y realiza la etapa E401 que incluye un re-muestreado de 12,8 a 16 kHz en el campo de la frecuencia, añadiendo $\frac{1}{4}$ de muestras ($k = 240, \dots, 319$) al espectro, siendo la relación entre 16 y 12,8 de 5/4.

55 Además, el bloque 501 realiza un filtrado paso alto implícito en la banda 0-5000 Hz puesto que las 200 primeras muestras de $U_{HB1}(k)$ se ponen a cero; como se explica más adelante, este filtrado paso alto se completa igualmente por una parte de atenuación progresiva de los valores espectrales de índices $k = 200, \dots, 255$ en la banda 5000-6400 Hz, esta atenuación progresiva se implementa en el bloque 504 pero se podría realizar por separado fuera del bloque 504. De manera equivalente y en unas variantes de la invención, la implementación del filtrado paso alto separada en bloques de coeficientes de índice $k = 0, \dots, 199$ puestos a cero, de coeficientes $k = 200, \dots, 255$ atenuados, en el dominio de la transformada, se podrá efectuar por tanto en una única etapa.

En este ejemplo de realización y según la definición de $U_{HB1}(k)$, se remarca que la banda 5000-6000 Hz de $U_{HB1}(k)$ (que corresponde a los índices $k = 200, \dots, 239$) se copia a partir de la banda 5000-6000 Hz de $U(k)$. Este planteamiento permite conservar el espectro original en esta banda y evita introducir unas distorsiones en la banda 5000-6000 Hz durante la adición de la síntesis de AF con la síntesis de BF —en particular se preserva la fase de la señal (implícitamente representada en el dominio DCT-IV) en esta banda—.

La banda 6000-8000 Hz de $U_{HB1}(k)$ se define en este caso copiando la banda 4000-6000 Hz de $U(k)$ puesto que el valor de *start_band* se fija preferentemente a 160.

En una variante del modo de realización, el valor de *start_band* podrá convertirse en adaptativo alrededor del valor de 160, sin modificar la naturaleza de la invención. Los detalles de la adaptación del valor de *start_band* no se describen aquí porque sobrepasan el marco de la invención sin cambiar en ella el alcance.

Para ciertas señales en banda ampliada (muestreadas a 16 kHz), la banda alta (>6 kHz) puede ser con ruido, armónica o incluir una mezcla de ruido y de armónicos. Además, el nivel de armonicidad en la banda 6000-8000 Hz se correlaciona en general con el de las bandas de frecuencia inferiores. De ese modo, en un modo particular de realización, el bloque 502 de generación de ruido, implementa la etapa E402 de la figura 4 y realiza una generación de ruido en el dominio de la frecuencia, $U_{HBN}(k)$ para $k = 240, \dots, 319$ (80 muestras) correspondientes a la segunda banda de frecuencia llamada de alta frecuencia con el fin de combinar a continuación este ruido con el espectro $U_{HB1}(k)$ en el bloque 503.

En un modo de realización particular, el ruido (en la banda 6000-8000 Hz) se genera de manera pseudoaleatoria con un generador de congruencia lineal sobre 16 bit:

$$U_{HBN}(k) = \begin{cases} 0 & k = 0, \dots, 239 \\ 31821U_{HBN}(k-1) + 13849 & k = 240, \dots, 319 \end{cases}$$

con la convención de que $U_{HBN}(239)$ en la trama actual corresponde al valor $U_{HBN}(319)$ de la trama precedente. En unas variantes de la invención, se podrá sustituir esta generación de ruido por otros métodos.

El bloque 503 de combinación puede realizarse de diferentes maneras. De manera preferente, se considera una mezcla aditiva adaptativa de la forma:

$$U_{HB2}(k) = \beta U_{HB1}(k) + \alpha G_{HBN} U_{HBN}(k), \quad k = 240, \dots, 319$$

en la que G_{HBN} es un factor de normalización que sirve para igualar el nivel de energía entre las dos señales,

$$G_{HBN} = \sqrt{\frac{\sum_{k=240}^{319} U_{HB1}(k)^2 + \varepsilon}{\sum_{k=240}^{319} U_{HBN}(k)^2 + \varepsilon}}$$

siendo $\varepsilon=0,01$, y se ajusta el coeficiente α (comprendido entre 0 y 1) en función de parámetros estimados a partir de la banda baja decodificada y el coeficiente β (comprendido entre 0 y 1) depende de α .

En un modo de realización preferente se calcula la energía del ruido en tres bandas: 2000-4000 Hz, 4000-6000 Hz y 6000-8000 Hz, con

$$E_{N2-4} = \sum_{k \in N(80,159)} U^2(k)$$

$$E_{N4-6} = \sum_{k \in N(160,239)} U^2(k)$$

$$E_{N4-6} = \sum_{k \in N(240,319)} U^2(k)$$

en las que

$$U'(k) = \begin{cases} \sqrt{\frac{\sum_{k=160}^{239} U^2(k)}{\sum_{k=80}^{159} U^2(k)}} U(k) & k = 80, \dots, 159 \\ U(k) & k = 160, \dots, 239 \\ \sqrt{\frac{\sum_{k=160}^{239} U^2(k)}{\sum_{k=240}^{319} U_{HB1}^2(k)}} U_{HB1}(k) & k = 240, \dots, 319 \end{cases}$$

y $N(k_1, k_2)$ es el conjunto de los índices k para los que el coeficiente de índice k se clasifica como asociado al ruido. Este conjunto puede obtenerse por ejemplo detectando los picos locales en $U'(k)$ que verifican $|U'(k)| \geq |U'(k-1)|$ y $|U'(k)| \geq |U'(k+1)|$ y considerando que estas rayas no están asociadas al ruido, es decir (aplicando la negación de la condición anterior):

$$N(a, b) = \{a \leq k \leq b \mid |U'(k)| < |U'(k-1)| \text{ o } |U'(k)| < |U'(k+1)|\}$$

Se puede observar que son posibles otros métodos de cálculo de la energía del ruido, por ejemplo tomando el valor medio del espectro sobre la banda considerada o aplicando un alisado a cada raya de frecuencia antes de calcular la energía por banda.

Se fija α de tal manera que la relación entre la energía del ruido en las bandas 4-6 kHz y 6-8 kHz sea la misma que entre las bandas 2-4 kHz y 4-6 kHz:

$$\alpha = \sqrt{\frac{\rho - E_{N6-8}}{\sum_{k=160}^{239} U^2(k) - E_{N6-8}}}$$

en la que

$$E_{N4-6} = \max(E_{N4-6}, E_{N2-4}), \quad \rho = \frac{E_{N4-6}^2}{E_{N2-4}}, \quad \rho = \max(\rho, E_{N6-8})$$

en la que $\max(\dots)$ es la función que da el máximo de los dos argumentos.

En unas variantes de la invención, el cálculo de α podrá sustituirse por otros métodos. Por ejemplo, en una variante, se podrán extraer (calcular) diferentes parámetros (o "features" en inglés) que caracterizan la señal en banda baja, entre ellos un parámetro "tilt" similar al calculado en el códec AMR-WB, y se estimará el factor α en función de una regresión lineal a partir de estos diferentes parámetros limitando su valor entre 0 y 1. La regresión lineal podrá estimarse por ejemplo de manera supervisada estimando el factor α proporcionándose la banda alta original en una base de aprendizaje. Se observará que el modo de cálculo de α no limita la naturaleza de la invención.

En un modo de realización preferente, se toma

$$\beta = \sqrt{1 - \alpha^2}$$

con el fin de preservar la energía de la señal extendida después de la mezcla.

En una variante los factores β y α podrán adaptarse para tener en cuenta el hecho de que un ruido inyectado en una banda dada de la señal es percibido en general como más fuerte que la señal armónica de la misma energía en la misma banda. Así se podrán modificar los factores β y α como sigue:

$$\begin{aligned} \beta &\leftarrow \beta \cdot f(\alpha) \\ \alpha &\leftarrow \alpha \cdot f(\alpha) \end{aligned}$$

en las que $f(\alpha)$ es una función decreciente de α , por ejemplo $f(\alpha) = b - a\sqrt{\alpha}$, $f(\alpha) b = 1,1$, $a = 1,2$, $f(\alpha)$ limitado de 0,3 a 1. Es necesario remarcar que después de la multiplicación por $f(\alpha)$, $\alpha^2 + \beta^2 < 1$ aunque la energía de la señal $U_{HB2}(k) = \beta U_{HB1}(k) + \alpha G_{HBN}U_{HBN}(k)$ es más baja que la energía de $U_{HB1}(k)$ (la diferencia de energía depende de α , cuanto más ruido se añade, más energía se atenúa). En otras variantes de la invención se podrá tomar:

5

$$\beta = 1 - \alpha$$

lo que permite preservar el nivel de amplitud (cuando las señales combinadas son del mismo signo); sin embargo esta variante tiene la desventaja de dar como resultado una energía global (a nivel de $U_{HB2}(k)$) que no es monótona en función de α .

10

Se remarca por tanto aquí que el bloque 503 realiza el equivalente al bloque 101 de la figura 1 para normalizar el ruido blanco en función de una excitación que está por el contrario en este caso en el dominio de la frecuencia, ya extendida a la cadencia de 16 kHz; además, la mezcla está limitada a la banda 6000-8000 Hz.

15

En una variante simple, se puede considerar una realización del bloque 503, en el que los espectros, $U_{HB1}(k)$ o $G_{HBN}U_{HBN}(k)$, se seleccionan (conmutan) de manera adaptativa, lo que se convierte en no permitir más que los valores 0 o 1 para α ; este planteamiento se convierte en clasificar el tipo de excitación a generar en la banda 6000-8000 Hz.

20

El bloque 504 realiza de manera opcional, una doble operación de aplicación de respuesta en frecuencia del filtro paso banda y de filtrado de desacentuación (o de-énfasis) en el dominio de la frecuencia.

En una variante de la invención, el filtrado de desacentuación se podrá realizar en el dominio temporal, después del bloque 505 incluso antes del bloque 500; sin embargo, en este caso, el filtrado paso banda realizado en el bloque 504 puede dejar ciertos componentes de baja frecuencia de niveles muy reducidos que se ven amplificar por desacentuación, lo que puede modificar de manera ligeramente perceptible la banda baja decodificada. Por esta razón, se prefiere en este caso realizar la desacentuación en el dominio de la frecuencia. En el modo de realización preferente, los coeficientes de índice $k = 0, \dots, 199$ se ponen a cero, así la desacentuación se limita a los coeficientes superiores.

25

30

La excitación se desacentúa inicialmente según la ecuación siguiente:

$$U_{HB2}'(k) = \begin{cases} 0 & k = 0, \dots, 199 \\ G_{deénf}(k-200)U_{HB2}(k) & k = 200, \dots, 255 \\ G_{deénf}(55)U_{HB2}(k) & k = 256, \dots, 319 \end{cases}$$

35

en la que $G_{deénf}(k)$ es la respuesta en frecuencia del filtro $1/(1-0,68z^{-1})$ sobre una banda de frecuencia discreta restringida. Teniendo en cuenta las frecuencias discretas (impares) de la DCT-IV, se define en este caso $G_{deénf}(k)$ como:

$$G_{deénf}(k) = \frac{1}{|e^{j\theta_k} - 0,68|}, \quad k = 0, \dots, 255$$

40

en la que

$$\theta_k = \frac{256 - 80 + k + \frac{1}{2}}{256}$$

45

En el caso en el que se utilice otra transformación distinta a la DCT-IV, la definición de θ_k podrá ajustarse (por ejemplo para unas frecuencias pares).

Se observa que la desacentuación se aplica en dos fases para $k = 200, \dots, 255$ correspondiente a la banda de frecuencia 5000-6400 Hz, en la que la respuesta $1/(1-0,68z^{-1})$ se aplica, 12,8 kHz, y para $k = 256, \dots, 319$ correspondiente a la banda de frecuencia 6400-8000 Hz, en la que la respuesta se extiende a 16 kHz en este caso en un valor constante en la banda 6,4-8 kHz.

50

Se puede observar que en el códec AMR-WB la síntesis de AF no está desacentuada. En el modo de realización aquí presentado, la señal de altas frecuencias se desacentúa por el contrario de manera que se la lleve a un dominio coherente con la señal de bajas frecuencias (0-6,4 kHz) que sale del bloque 305. Esto es importante para la estimación y el ajuste posterior de la energía de la síntesis en AF.

55

ES 2 724 576 T3

En una variante del modo de realización, con el fin de reducir la complejidad, se podrá fijar $G_{deénf}(k)$ en un valor constante independiente de k , tomando por ejemplo $G_{deénf}(k) = 0,6$ lo que corresponde aproximadamente al valor medio de $G_{deénf}(k)$ para $k = 200, \dots, 319$ en las condiciones del modo de realización anteriormente descritas.

5 En otra variante del modo de realización del dispositivo de extensión, la desacentuación se podrá realizar de manera equivalente en el dominio temporal después de la DCT inversa. Una realización de ese tipo se implementa en la figura 7 descrita más adelante.

10 Además de la desacentuación, se aplica un filtrado paso banda con dos partes separadas: la una paso alto fija, la otra paso bajo adaptativa (función de la velocidad).

Este filtrado se efectúa en el dominio de la frecuencia, y su respuesta en frecuencia se ilustra en la figura 6. Las frecuencias de corte a 3 dB son 6000 Hz para la parte baja y para la parte alta aproximadamente 6900, 7300, 7600 Hz a 6,6, 8,86 y a las velocidades superiores a 8,85 kbit/s (respectivamente).

15 En el modo de realización preferente, se calcula la respuesta parcial del filtro paso bajo en el dominio de la frecuencia como sigue:

$$G_{lp}(k) = 1 - 0,999 \frac{k}{N_{lp} - 1}$$

20 en la que $N_{lp} = 60$ a 6,6 kbit/s, 40 a 8,85 kbit/s, 20 a velocidades $>8,85$ bit/s. A continuación se aplica un filtro paso banda en la forma:

$$U_{HB3}(k) = \begin{cases} 0 & k = 0, \dots, 199 \\ G_{hp}(k - 200)U_{HB2}'(k) & k = 200, \dots, 255 \\ U_{HB2}'(k) & k = 256, \dots, 319 - N_{lp} \\ G_{lp}(k - 320 - N_{lp})U_{HB2}'(k) & k = 320 - N_{lp}, \dots, 319 \end{cases}$$

25 La definición de $G_{hp}(k)$, $k = 0, \dots, 55$ se da por ejemplo en la tabla 1 siguiente.

Tabla 1

K	$g_{hp}(k)$	K	$g_{hp}(k)$	K	$g_{hp}(k)$	k	$g_{hp}(k)$
0	0,001622428	14	0,114057967	28	0,403990611	42	0,776551214
1	0,004717458	15	0,128865425	29	0,430149896	43	0,800503267
2	0,008410494	16	0,144662643	30	0,456722014	44	0,823611104
3	0,012747280	17	0,161445005	31	0,483628433	45	0,845788355
4	0,017772424	18	0,179202219	32	0,510787115	46	0,866951597
5	0,023528982	19	0,197918220	33	0,538112915	47	0,887020781
6	0,030058032	20	0,217571104	34	0,565518011	48	0,905919644
7	0,037398264	21	0,238133114	35	0,592912340	49	0,923576092
8	0,045585564	22	0,259570657	36	0,620204057	50	0,939922577
9	0,054652620	23	0,281844373	37	0,647300005	51	0,954896429
10	0,064628539	24	0,304909235	38	0,674106188	52	0,968440179
11	0,075538482	25	0,328714699	39	0,700528260	53	0,980501849
12	0,087403328	26	0,353204886	40	0,726472003	54	0,991035206
13	0,100239356	27	0,378318805	41	0,751843820	55	1,000000000

30 Se observará que en las variantes de la invención, los valores de $G_{hp}(k)$ podrán modificarse mientras se mantiene una atenuación progresiva. Igualmente el filtrado paso bajo de anchura de banda variable, $G_p(k)$, se podrá ajustar con unos valores o un soporte de frecuencia diferentes, sin cambiar el principio de esta etapa de filtrado.

35 Se observará también que el ejemplo de filtro paso banda ilustrado en la figura 6 podrá adaptarse definiendo una única etapa de filtrado que combina los filtrados paso alto y paso bajo.

En otro modo de realización, el filtrado paso banda podrá realizarse de manera equivalente en el dominio temporal (como en el bloque 112 de la figura 1) con diferentes coeficientes de filtro según la velocidad, después de la etapa

de DCT inversa. Una realización así se implementa en la figura 7 descrita más adelante. Sin embargo, se observará que es ventajoso realizar esta etapa directamente en el dominio de la frecuencia porque el filtrado se efectúa en el dominio de la excitación LPC y por tanto los problemas de convolución circular y de efectos de borde son muy limitados en este dominio.

5 El bloque 505 de transformada inversa realiza una DCT inversa sobre 320 muestras para encontrar la excitación de alta frecuencia muestreada a 16 kHz. Su implementación es idéntica al bloque 500, porque la DCT-IV es ortonormal, salvo que la longitud de la transformada es de 320 en lugar de 256, y se obtiene:

$$10 \quad u_{HB}(n) = \sum_{k=0}^{N_{16k}-1} U_{HB3}(k) \cos\left(\frac{\pi}{N_{16k}}\left(k + \frac{1}{2}\right)\left(n + \frac{1}{2}\right)\right)$$

en la que $N_{16k} = 320$ y $k = 0, \dots, 319$.

15 Esta excitación muestreada a 16 kHz se escala a continuación mediante unas ganancias definidas por subtrama de 80 muestras (bloque 507).

En un modo de realización preferido, se calcula inicialmente (bloque 506) una ganancia $g_{HB1}(m)$ por subtrama mediante unas relaciones de energía de las subtramas tal que en cada subtrama de índice $m = 0, 1, 2$ o 3 de la trama actual:

$$20 \quad g_{HB1}(m) = \sqrt{\frac{e_3(m)}{e_2(m)}}$$

en la que

$$25 \quad e_1(m) = \sum_{n=0}^{63} u(n + 64m)^2 + \varepsilon$$

$$e_2(m) = \sum_{n=0}^{79} u_{HB}(n + 80m)^2 + \varepsilon$$

$$e_3(m) = e_1(m) \frac{\sum_{n=0}^{319} u_{HB}(n)^2 + \varepsilon}{\sum_{n=0}^{255} u(n)^2 + \varepsilon}$$

30 siendo $\varepsilon = 0,01$. Se puede escribir la ganancia por subtrama $g_{HB1}(m)$ en la forma:

$$g_{HB1}(m) = \sqrt{\frac{\frac{\sum_{n=0}^{63} u(n + 64m)^2 + \varepsilon}{\sum_{n=0}^{255} u(n)^2 + \varepsilon}}{\frac{\sum_{n=0}^{79} u_{HB}(n + 80m)^2 + \varepsilon}{\sum_{n=0}^{319} u_{HB}(n)^2 + \varepsilon}}}$$

35 Lo que muestra que se asegura en la señal u_{HB} la misma relación entre la energía por subtrama y energía por trama que en la señal $u(n)$.

El bloque 507 efectúa el escalado de la señal combinada (o extendida) (etapa E406 de la figura 4) según la ecuación siguiente:

$$40 \quad u_{HB}'(n) = g_{HB1}(m)u_{HB}(n), \quad n = 80m, \dots, 80(m+1) - 1$$

Se observará que la realización del bloque 506 difiere de la del bloque 101 de la figura 1, porque se tiene en cuenta la energía a nivel de la trama actual además de aquella de la subtrama. Esto permite tener la relación de la energía de cada subtrama con relación a la energía de la trama. Se comparan por tanto unas relaciones de energía (o energías relativas) más que las energías absolutas entre banda baja y banda alta.

De ese modo, esta etapa de escalado permite conservar en la banda alta la relación de energía entre la subtrama y la trama de la misma manera que en la banda baja.

De manera opcional, el bloque 509 efectúa a continuación el escalado de la señal (etapa E407 de la figura 4) según la ecuación siguiente:

$$u_{HB}''(n) = g_{HB2}(m)u_{HB}'(n), \quad n = 80m, \dots, 80(m+1) - 1$$

en la que la ganancia $g_{HB2}(m)$ se obtiene a partir del bloque 508 ejecutando los bloques 103, 104 y 105 del códec AMR-WB (siendo la entrada del bloque 103 la excitación decodificada en banda baja, $u(n)$). Los bloques 508 y 509 son útiles para ajustar el nivel de filtro de síntesis LPC (bloque 510), en este caso en función del tilt de la señal. Son posibles otros métodos de cálculo de la ganancia $g_{HB2}(m)$ sin cambiar la naturaleza de la invención.

Finalmente, se filtra la excitación, $u_{HB}'(n)$ o $u_{HB}''(n)$, (etapa E404 de la figura 4) por el módulo de filtrado 510 que puede realizarse en este caso tomando como función de transferencia $1/\hat{A}(z/\gamma)$, en la que $\gamma = 0,9$ a 6,6 kbit/s y $\gamma = 0,6$ a las otras velocidades, lo que limita el orden del filtro al orden 16.

En una variante, este filtrado se podrá realizar de la misma manera que la que se describe para el bloque 111 de la figura 1 del decodificador AMR-WB, sin embargo el orden del filtro pasa a 20 a la velocidad de 6,6, lo que no cambia de manera significativa la calidad de la señal sintetizada. En otra variante, se podrá efectuar el filtrado de síntesis LPC en el dominio de la frecuencia, después de haber calculado la respuesta en frecuencia del filtro implementado en el bloque 510.

En unas variantes de realización de la invención, la codificación de la banda baja (0-6,4 kHz) podrá sustituirse por un codificador CELP distinto al utilizado en AMR-WB, como por ejemplo el codificador CELP en G.718 a 8 kbit/s. Sin pérdida de generalidad podrían utilizarse otros codificadores en banda ampliada o funcionando a frecuencias superiores a 16 kHz, en los que la codificación de la banda baja funciona a una frecuencia interna de 12,8 kHz. Por otro lado, la invención puede adaptarse de manera diferente a otras frecuencias de muestreo distintas de 12,8 kHz, cuando un codificador de bajas frecuencias funciona a una frecuencia de muestreo inferior a la de la señal original o reconstruida. Cuando la decodificación en banda baja no utiliza predicción lineal, no se dispone de una señal de excitación a extender, en este caso se podrá realizar un análisis LPC de la señal reconstruida en la trama actual y se calculará una excitación LPC de manera que se pueda aplicar la invención.

Finalmente, en otra variante de la invención, la excitación ($u(n)$) se re-muestrea, por ejemplo mediante interpolación lineal o "spline" cúbica, de 12,8 a 16 kHz antes de la transformación (por ejemplo DCT-IV) de longitud 320. Esta variante tiene el defecto de ser más compleja, porque la transformada (DCT-IV) de la excitación se calcula entonces sobre una longitud mayor y el re-muestreo no se efectúa en el dominio de la transformada.

Además, en unas variantes de la invención, todos los cálculos necesarios para la estimación de las ganancias (G_{HBN} , $g_{HB1}(m)$, $g_{HB2}(m)$, g_{HBN} , ...) se podrán efectuar en un dominio logarítmico.

Con referencia a la figura 7, se describe ahora un segundo modo de realización del dispositivo de extensión de banda. Este modo de realización funciona en el dominio temporal. Este segundo modo de realización no forma parte de la invención.

Como en el modo de realización de la figura 5, se conserva el principio del modo de realización con mezcla de una señal extendida 16 kHz y de una señal de ruido, sin embargo esta mezcla se realiza esta vez en el dominio temporal y esta vez la generación principal de la excitación se realiza por subtrama y no por trama.

La señal de excitación $u(n)$, $n = 0, \dots, 255$, procedente de la decodificación de baja frecuencia en la trama actual se re-muestrea inicialmente sin retardo (etapa E401 de la figura 4) a 16 kHz (bloque 700) y en un modo de realización particular, se utiliza una interpolación lineal para obtener la señal de excitación extendida en una segunda banda de frecuencia, $u_{ext}(n)$, $n = 0, \dots, 319$. En una variante de realización, se podrán utilizar otros métodos de re-muestreo, por ejemplo mediante "splines" o mediante filtrado multi-cadencias.

Se asegura que la energía de la señal $u_{ext}(n)$ tiene un nivel similar a la excitación $u(n)$ con los bloques 701 y 702 de la manera siguiente:

$$u'_{ext}(n) = u_{ext}(n) \sqrt{\frac{\sum_{l=0}^{63} u(l)^2}{\sum_{l=0}^{79} u_{ext}(l)^2}}$$

En una variante de realización se podrá multiplicar $u'_{ext}(n)$ por 5/4 para compensar la atenuación por la relación 12,8/16, provocada por diferentes frecuencias de muestreo de la señal $u_{ext}(n)$ y $u(n)$.

5 El generador de ruido en el bloque 703 implementa la etapa E402 de la figura 4 y puede realizarse como en el bloque 502 descrito en la figura 5, salvo que la señal a la salida corresponde a una subtrama temporal, $u_{HBN}(n)$, $n = 0, \dots, 319$. El bloque 704 de combinación puede realizarse de diferentes maneras. De manera preferente, se considera una mezcla aditiva adaptativa *por subtrama* de la forma:

10 $u_{HB1}(n + 80m) = \beta u_{ext}(n + 80m) + \alpha g_{HBN} u_{HBN}(n + 80m)$, $n = 0, \dots, 79$ en la que g_{HBN} es un factor de normalización que sirve para igualar el nivel de armonicidad de las dos señales combinadas

$$g_{HBN} = \sqrt{\frac{\sum_{k=0}^{79} u_{ext}(n)^2 + \mathcal{E}}{\sum_{k=0}^{79} u_{HBN}(n)^2 + \mathcal{E}}}$$

15 y m es el índice de la subtrama y los factores α y β se calculan como en el primer modo de realización. Se remarca por tanto en este caso que el bloque 704 realiza el equivalente al bloque 101 de la figura 1. Además, el cálculo del factor α obliga a calcular la transformada de la señal de excitación decodificada (o la señal decodificada en sí misma según el dominio de cálculo del nivel relativo de ruido o de planitud espectral llamado "spectral flatness" en inglés) en banda baja si este cálculo se basa sobre la planitud espectral; en unas variantes, cuya utilización de una regresión lineal se ha descrito anteriormente, no es necesaria una transformada de ese tipo.

20 A continuación se desacentúa la señal temporal (bloque 705) mediante un filtro de la forma $g_{deénf}(1-0,68z^{-1})$, en la que $g_{deénf}$ se calcula de manera que se prolongue el filtro $1/(1-0,68z^{-1})$ (definido a 12,8 kHz) a la frecuencia de muestreo de 16 kHz $g_{deénf} = |(1-0,68e^{j2\pi 6000/16000}) / (1-0,68e^{j2\pi 6000/12800})|$, posteriormente tratado mediante un filtrado paso banda de anchura de banda variable (bloque 706) cuyo orden es fijo (de valor 30) pero los coeficientes cambian en función de la velocidad de decodificación de la trama actual.

30 Un ejemplo de realización de un filtro paso banda adaptativo así de tipo FIR se da en las tablas siguientes que definen la respuesta a impulsos del filtro FIR según la velocidad.

Tabla 2a (6,6 kbit/s)

n	$h(n)$	n	$h(n)$	n	$h(n)$	n	$h(n)$
0	-0,0002581	8	0,0306285	16	-0,1451668	24	-0,0114595
1	0,0003791	9	-0,0716116	17	0,0626279	25	0,0090482
2	0,0002581	10	0,0995869	18	0,0286124	26	-0,0029758
3	-0,0002177	11	-0,0885791	19	-0,0885791	27	-0,0002177
4	-0,0029758	12	0,0286124	20	0,0995869	28	0,0002581
5	0,0090482	13	0,0626279	21	-0,0716116	29	0,0003791
6	-0,0114595	14	-0,1451668	22	0,0306285	30	-0,0002581
7	0	15	0,1783678	23	0	-	-

Tabla 2b (8,85 kbit/s)

n	$h(n)$	n	$h(n)$	n	$h(n)$	n	$h(n)$
0	0,0019706	8	0,0312161	16	-0,1720177	24	-0,0030672
1	-0,0064291	9	-0,0709664	17	0,0817478	25	-0,0041966
2	0,0124179	10	0,0980678	18	0,0181018	26	0,0132058
3	-0,0160589	11	-0,0842625	19	-0,0842625	27	-0,0160589
4	0,0132058	12	0,0181018	20	0,0980678	28	0,0124179
5	-0,0041966	13	0,0817478	21	-0,0709664	29	-0,0064291

<i>n</i>	<i>h(n)</i>	<i>n</i>	<i>h(n)</i>	<i>n</i>	<i>h(n)</i>	<i>n</i>	<i>h(n)</i>
6	-0,0030672	14	-0,1720177	22	0,0312161	30	0,0019706
7	-0,0036671	15	0,2083360	23	-0,0036671	-	

Tabla 2c (velocidades >8,85 kbit/s)

<i>n</i>	<i>h(n)</i>	<i>n</i>	<i>h(n)</i>	<i>n</i>	<i>h(n)</i>	<i>n</i>	<i>h(n)</i>
0	0,0013312	8	0,0606146	16	-0,1916778	24	0,0221682
1	-0,0047346	9	-0,0860005	17	0,1093354	25	-0,0180046
2	0,0098657	10	0,0924138	18	-0,0129187	26	0,0171709
3	-0,0147045	11	-0,0607694	19	-0,0607694	27	-0,0147045
4	0,0171709	12	-0,0129187	20	0,0924138	28	0,0098657
5	-0,0180046	13	0,1093354	21	-0,0860005	29	-0,0047346
6	0,0221682	14	-0,1916778	22	0,0606146	30	0,0013312
7	-0,0360130	15	0,2240719	23	-0,0360130	-	-

La etapa de escalado (E407 en la figura 4) se efectúa mediante los bloques 508 y 509 idénticos a la figura 5.

5 La etapa de filtrado (E404 de la figura 4) se efectúa mediante el módulo de filtrado (bloque 510) idéntico al descrito con referencia a la figura 5.

10 No es útil en este caso implementar una etapa de escalado como se efectúa en el modo de realización de la figura 5 por los bloques 506 y 507 puesto que la excitación se genera por subtramas. La coherencia de la relación de energía a nivel de la trama ya está asegurada.

15 En unas variantes de la invención, la excitación en banda baja $u(n)$ y el filtro LPC $1/\hat{A}(z)$ se estimarán por trama, mediante análisis LPC de una señal en banda baja cuya banda debe extenderse. La señal de excitación en banda baja se extrae entonces por análisis de la señal de audio.

En un modo posible de realización de esta variante, la señal de audio en banda baja se re-muestrea antes de la etapa de extracción de la excitación, aunque la excitación extraída de la señal de audio (por predicción lineal) ya está re-muestreada.

20 La invención ilustrada en la figura 5, o el segundo modo de realización que no forma parte de la invención y se describe en la figura 7, se aplica en este caso a una banda baja que no está decodificada sino analizada.

La figura 8 representa un ejemplo de realización material de un dispositivo de extensión de banda 800 según la invención. Este puede formar parte integrante de un decodificador de señal de audiofrecuencia o de un equipo que recibe unas señales de audiofrecuencia decodificadas o no.

Este tipo de dispositivo incluye un procesador PROC que coopera con un bloque de memoria BM que incluye una memoria de almacenamiento y/o de trabajo MEM.

30 Un dispositivo de ese tipo incluye un módulo de entrada E adecuado para recibir una señal de audio de excitación decodificada o extraída en una primera banda de frecuencia llamada banda baja ($u(n)$ o $U(k)$) y los parámetros de un filtro de síntesis de predicción lineal ($\hat{A}(z)$). Incluye un módulo de salida S adecuado para transmitir la señal de alta frecuencia de síntesis (AF_síntesis) por ejemplo en un módulo de aplicación a un retardo como el bloque 310 de la figura 3 o a un módulo de re-muestreo como el módulo 311.

35 El bloque de memoria puede incluir ventajosamente un programa informático que incluye unas instrucciones de código para la implementación de las etapas del procedimiento de extensión de banda en el sentido de la invención, cuando estas instrucciones se ejecutan por el procesador PROC y principalmente las etapas de obtención de una señal extendida en al menos una segunda banda de frecuencia superior a la primera banda de frecuencia a partir de una señal de excitación sobre-muestreada y extendida en al menos una segunda banda de frecuencia, de escalado de la señal extendida mediante una ganancia definida por subtrama en función de una relación de energía de una trama y de una subtrama y de filtrado de dicha señal extendida escalada mediante un filtro de predicción lineal cuyos coeficientes se derivan de los coeficientes del filtro de la banda baja.

40 Típicamente, la descripción de la figura 4 retoma las etapas de un algoritmo de un programa informático de ese tipo. El programa informático puede almacenarse igualmente en un soporte de memoria legible por un lector del dispositivo o es descargable en el espacio de memoria de este.

45 La memoria MEM registra de manera general, todos los datos necesarios para la implementación del procedimiento.

En un modo posible de realización, el dispositivo así descrito puede incluir igualmente las funciones de decodificación de banda baja y otras funciones de tratamiento descritas por ejemplo en la figura 3 además de las funciones de extensión de banda según la invención.

REIVINDICACIONES

1. Procedimiento de extensión de banda de frecuencia de una señal de audiofrecuencia durante un proceso de decodificación o de mejora que incluye una etapa de decodificación o de extracción, en una primera banda de frecuencia llamada banda baja, de una señal de excitación y de unos coeficientes de un filtro de predicción lineal, estando el procedimiento caracterizado por que incluye las etapas siguientes:

- obtención de una señal extendida ($U_{HB2}(k)$, E403) en al menos una segunda banda de frecuencia superior a la primera banda de frecuencia a partir de la señal de excitación sobre-muestreada y extendida en la al menos una segunda banda de frecuencia ($U_{HB1}(k)$, E401);
- escalado (E406) de la señal extendida mediante una ganancia definida por subtrama en función de una relación entre la relación entre la energía por subtrama y la energía por trama de la señal de excitación de la banda baja y la relación entre la energía por subtrama y la energía por trama de la señal extendida;
- filtrado (E404) de dicha señal extendida escalada mediante un filtro de predicción lineal cuyos coeficientes se derivan de los coeficientes del filtro de la banda baja.

2. Procedimiento según la reivindicación 1, caracterizado por que incluye además una etapa de filtrado (E405) paso banda adaptativo en función de la velocidad de decodificación de la trama actual de la señal extendida antes de la etapa de escalado.

3. Procedimiento según la reivindicación 1, caracterizado por que incluye una etapa de transformada tiempo-frecuencia de la señal de excitación, efectuándose entonces la etapa de obtención de una señal extendida en el campo de la frecuencia y una etapa de transformada tiempo-frecuencia inversa de la señal extendida antes de las etapas de escalado y de filtrado.

4. Procedimiento según la reivindicación 3, caracterizado por que la etapa de generación de una señal de excitación sobre-muestreada y extendida se efectúa según la ecuación siguiente:

$$U_{HB1}(k) = \begin{cases} 0 & k = 0, \dots, 199 \\ U(k) & k = 200, \dots, 239 \\ U(k + start_band - 240) & k = 240, \dots, 319 \end{cases}$$

siendo k el índice de la muestra, $U_{HB1}(k)$ el espectro de la señal de excitación extendida, U(k) el espectro de la señal de excitación obtenida después de la etapa de transformada y *start_band* una variable predefinida.

5. Procedimiento según una de las reivindicaciones 1 a 4, caracterizado por que incluye una etapa de filtrado de desacentuación de la señal extendida al menos en la segunda banda de frecuencia.

6. Procedimiento según la reivindicación 1, caracterizado por que incluye además una etapa de generación (E402) de una señal de ruido al menos en la segunda banda de frecuencia obteniéndose la señal extendida ($U_{HB2}(k)$) por combinación (E403) de la señal de excitación extendida y de la señal de ruido.

7. Procedimiento según la reivindicación 6, caracterizado por que la etapa de combinación se efectúa por mezcla aditiva adaptativa con una ganancia de igualación del nivel entre la señal de excitación extendida y la señal de ruido.

8. Dispositivo de extensión de banda de frecuencia de una señal de audiofrecuencia que incluye una etapa de decodificación o de extracción, en una primera banda de frecuencia llamada banda baja, de una señal de excitación y de los coeficientes de un filtro de predicción lineal, estando el dispositivo caracterizado por que incluye:

- un módulo de obtención de una señal extendida ($U_{HB2}(k)$, 503) en al menos una segunda banda de frecuencia superior a la primera banda de frecuencia a partir de la señal de excitación sobre-muestreada y extendida en la al menos una segunda banda de frecuencia ($U_{HB1}(k)$);
- escalado (507) de la señal extendida mediante una ganancia definida por subtrama en función de una relación entre la relación entre la energía por subtrama y la energía por trama de la señal de excitación de la banda baja y la relación entre la energía por subtrama y la energía por trama de la señal extendida;
- un módulo de filtrado (510) de dicha señal extendida escalada mediante un filtro de predicción lineal cuyos coeficientes se derivan de los coeficientes del filtro de la banda baja.

9. Un decodificador de señal de audiofrecuencia caracterizado por que incluye un dispositivo de extensión de banda de frecuencia de acuerdo con la reivindicación 8.

10. Programa informático que incluye unas instrucciones de código para la implementación de las etapas del procedimiento de extensión de banda de frecuencia según una de las reivindicaciones 1 a 7, cuando estas instrucciones se ejecutan por un procesador.

11. Soporte de almacenamiento legible por un dispositivo de extensión de banda de frecuencia en el que se registra un programa informático que comprende unas instrucciones de código para la ejecución de las etapas del procedimiento de extensión de banda de frecuencia según una de las reivindicaciones 1 a 7.

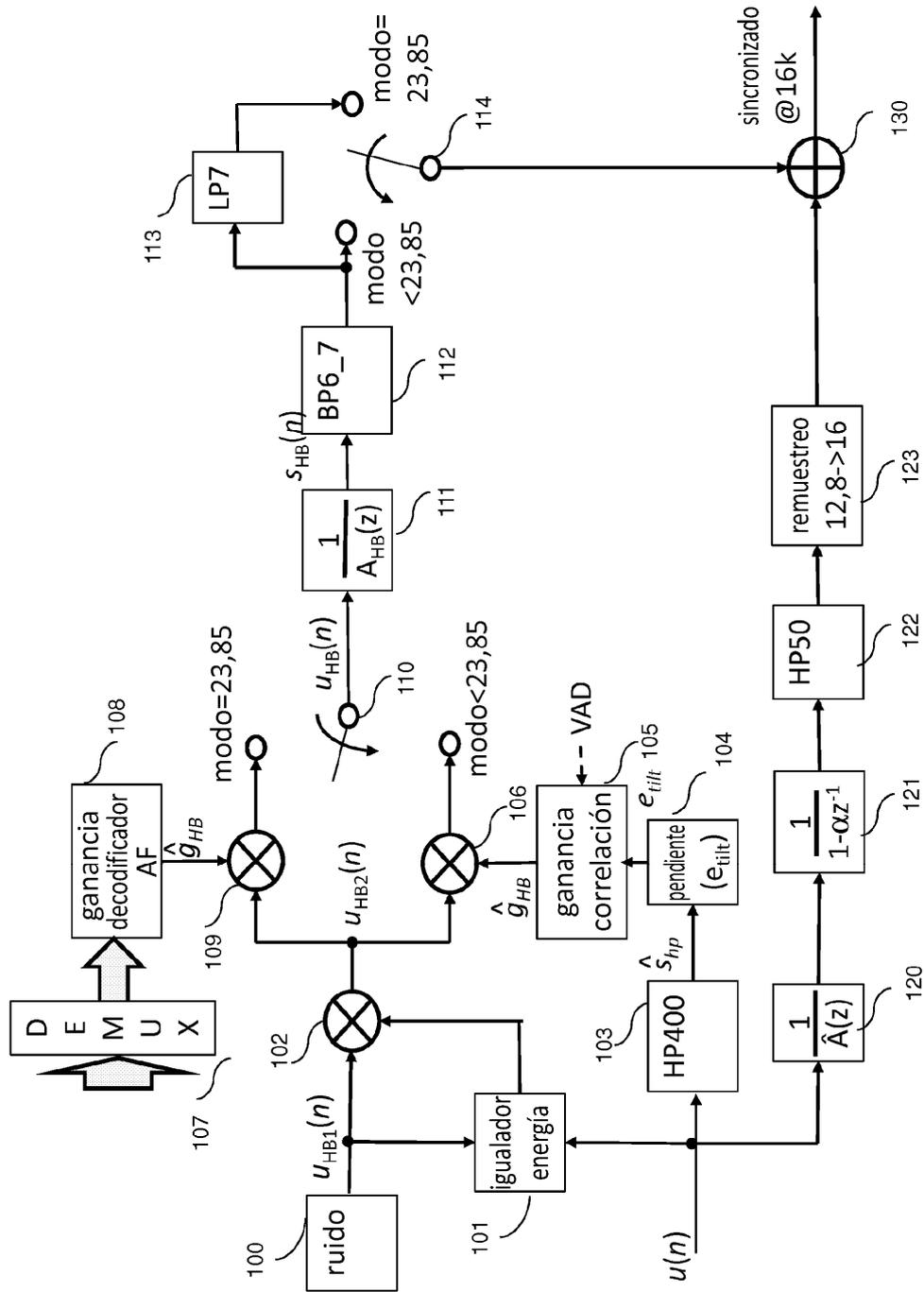


Fig.1

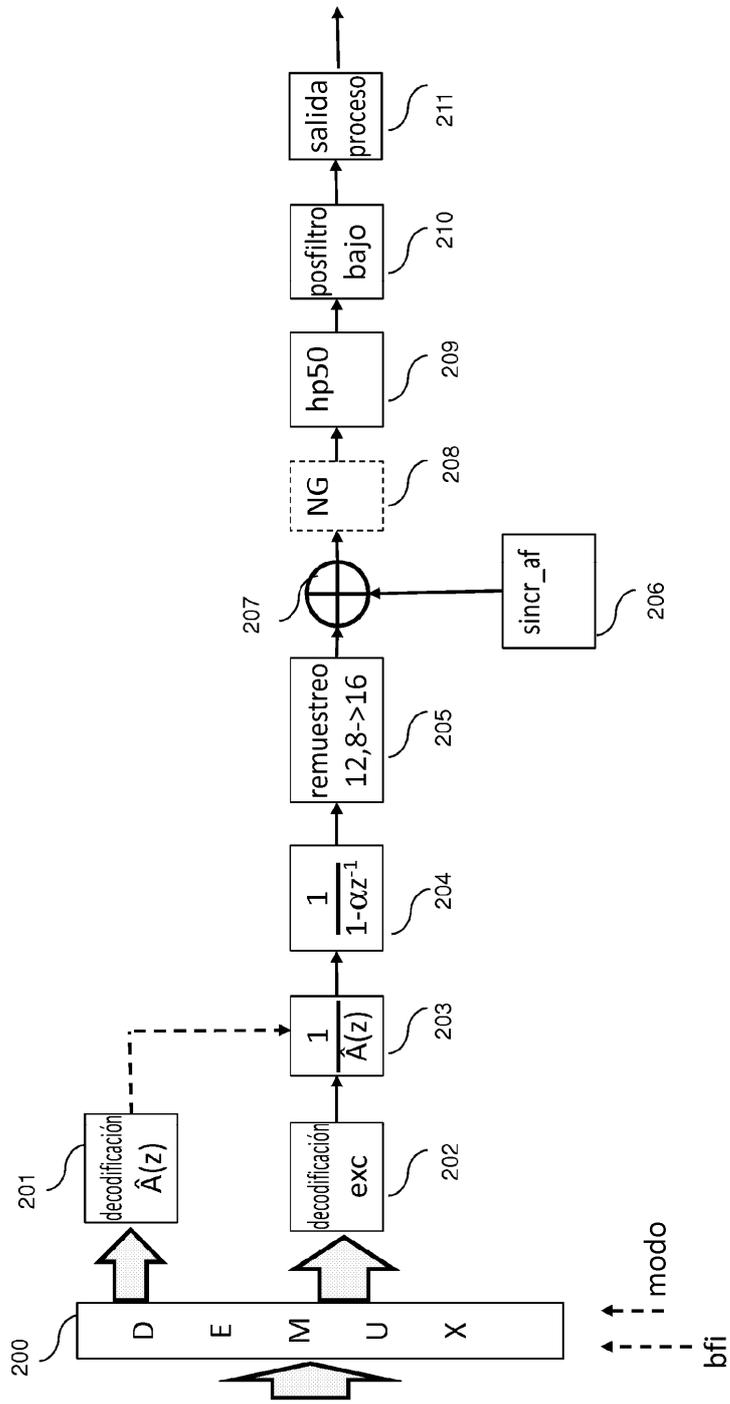


Fig.2

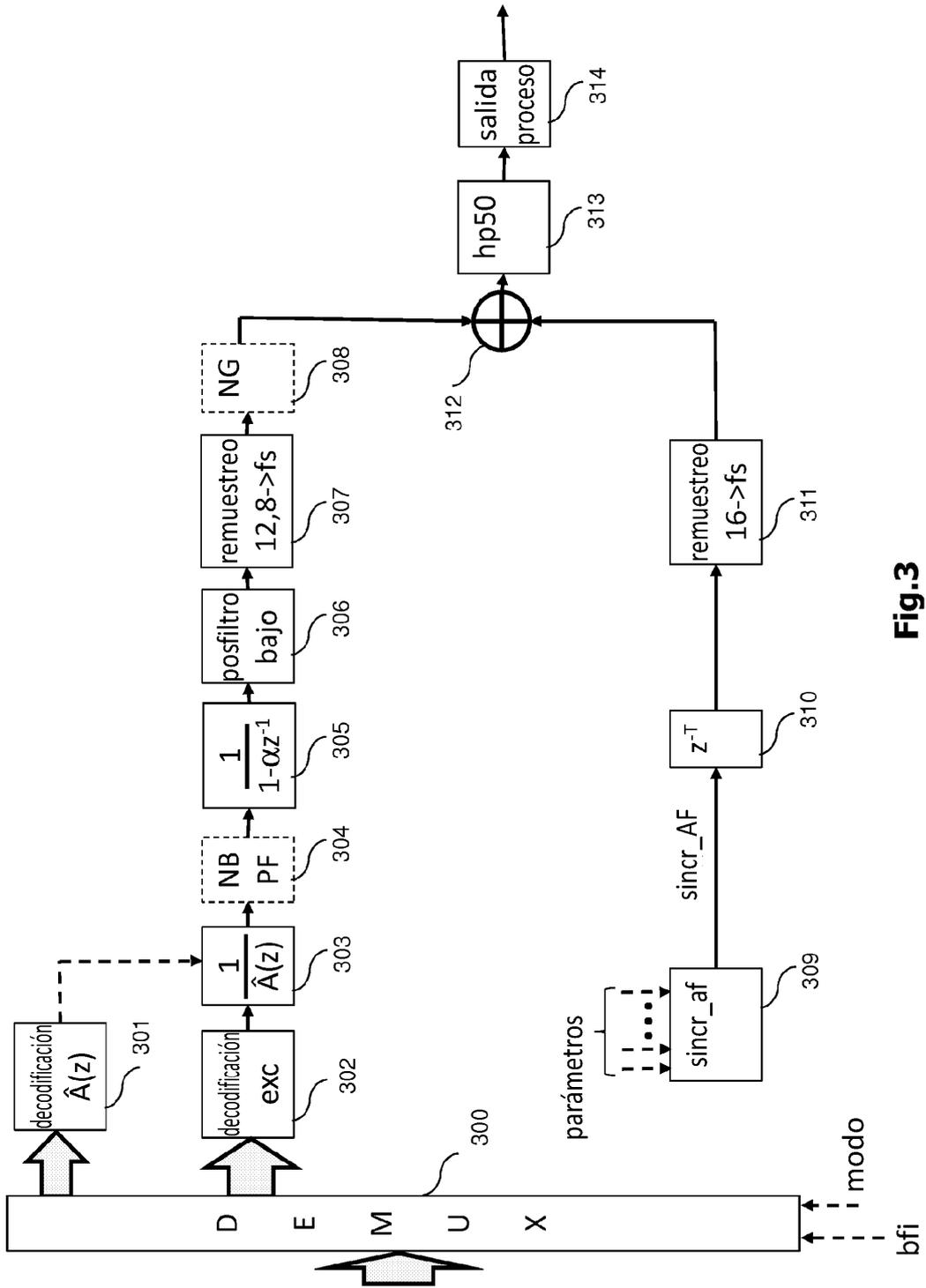


Fig.3

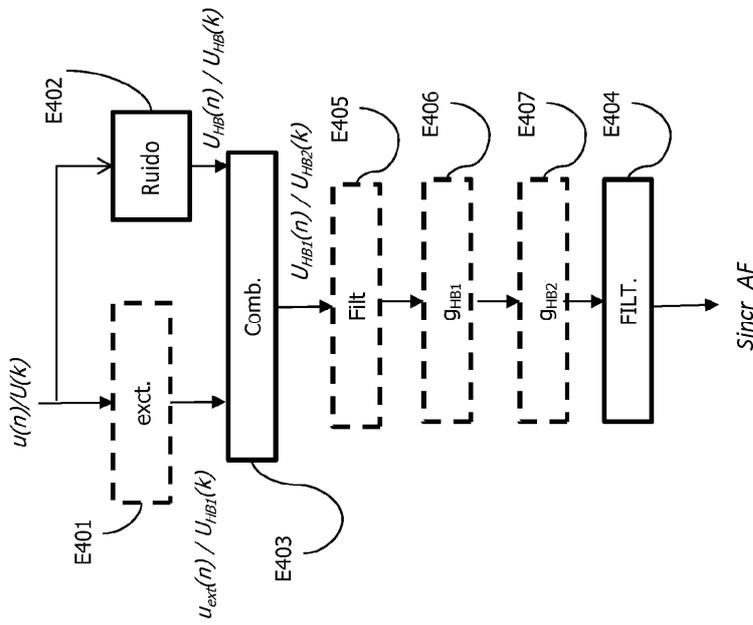


Fig.4

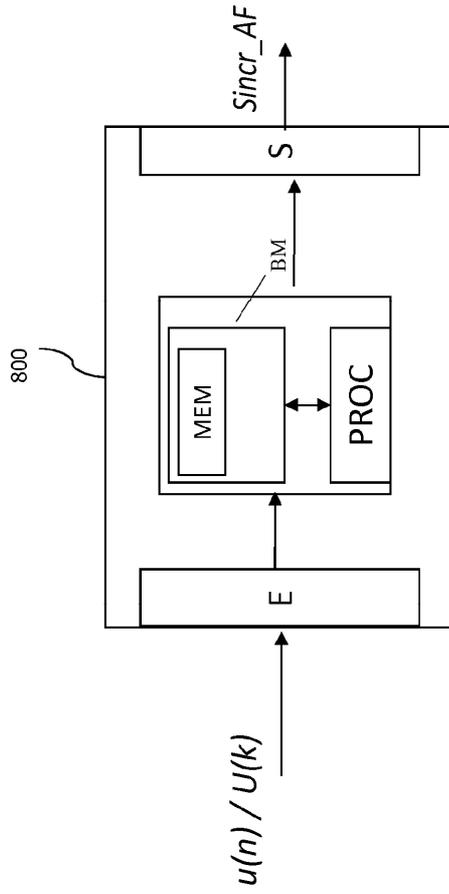


Fig.8

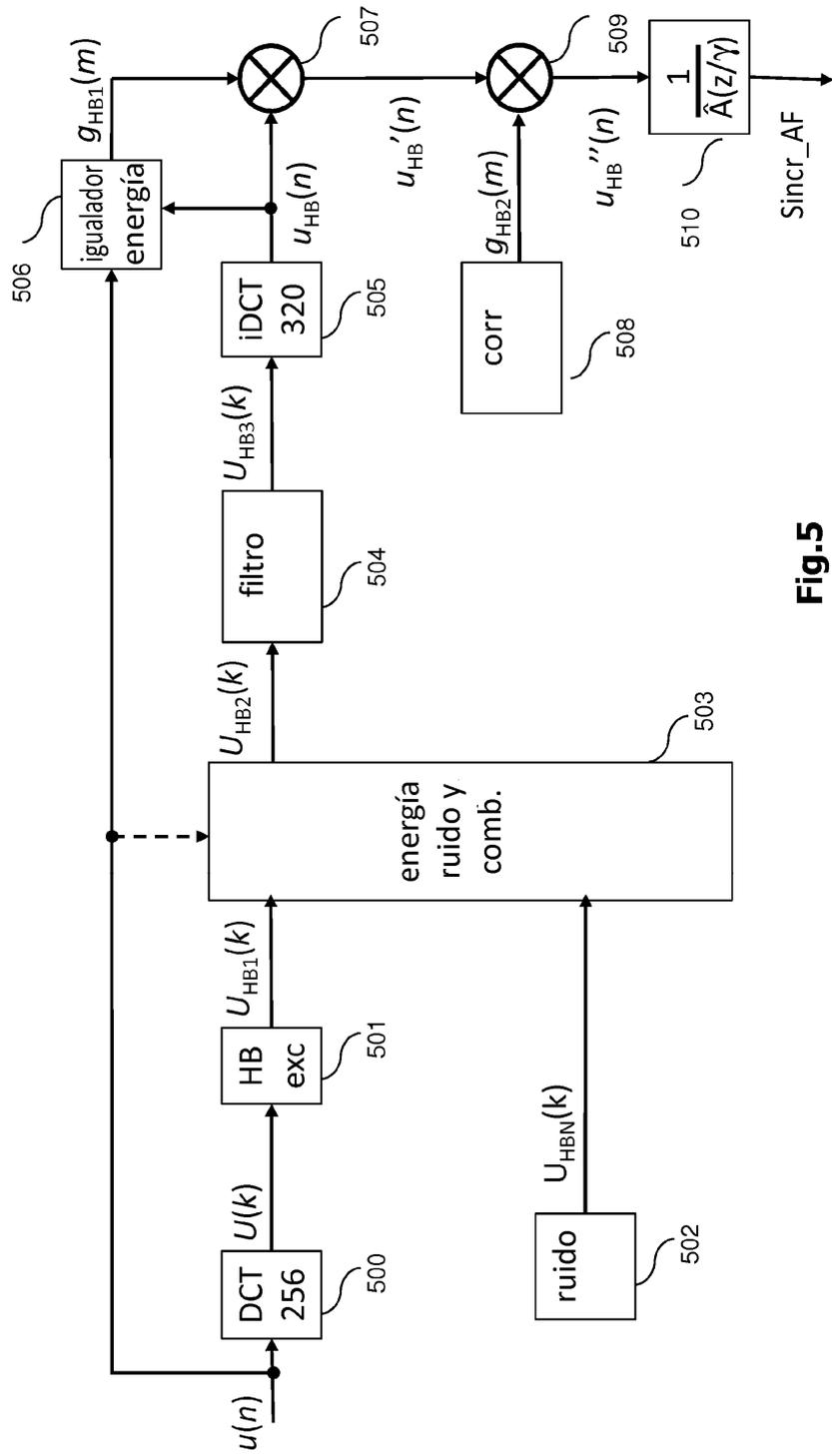


Fig.5

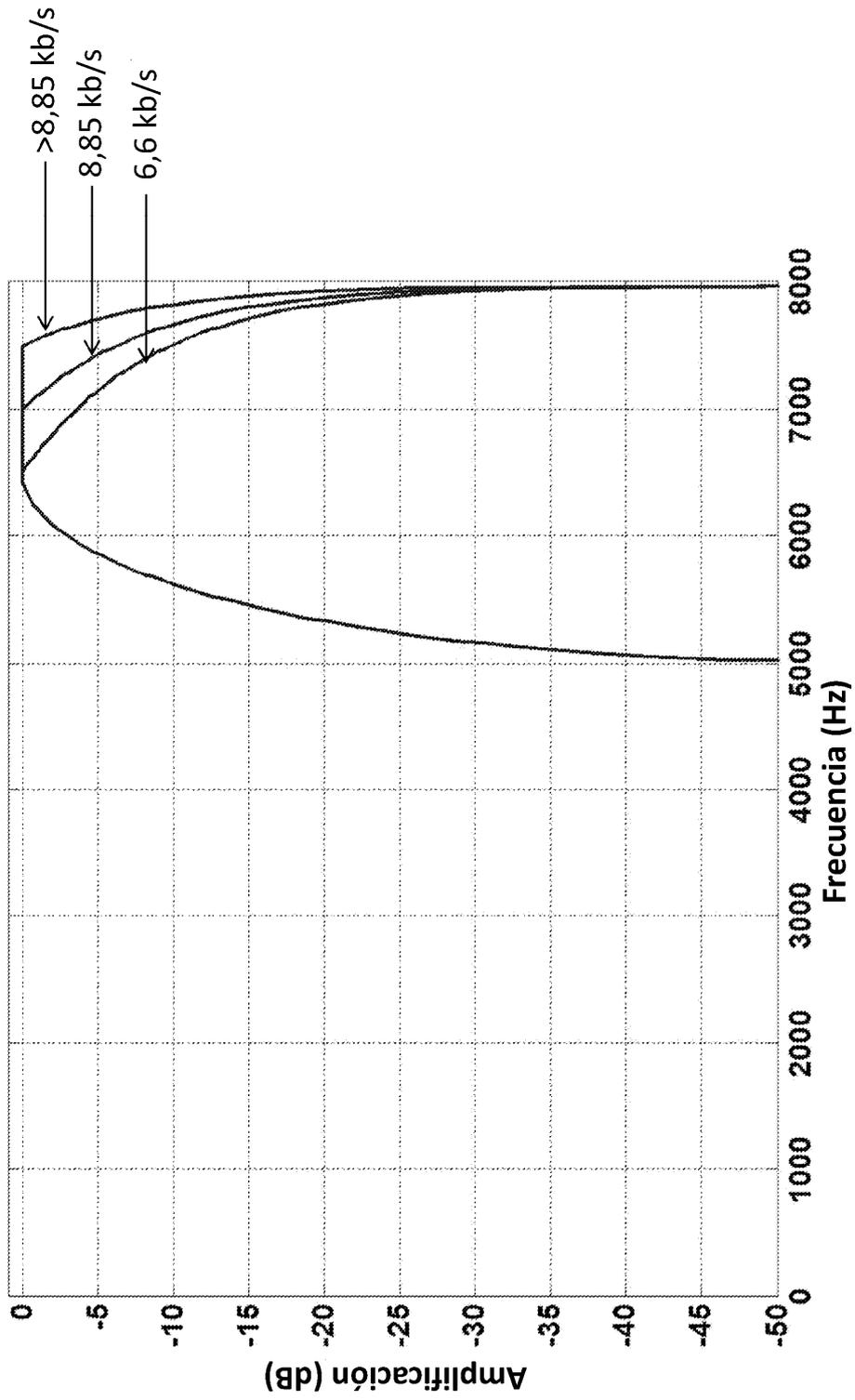


Fig.6

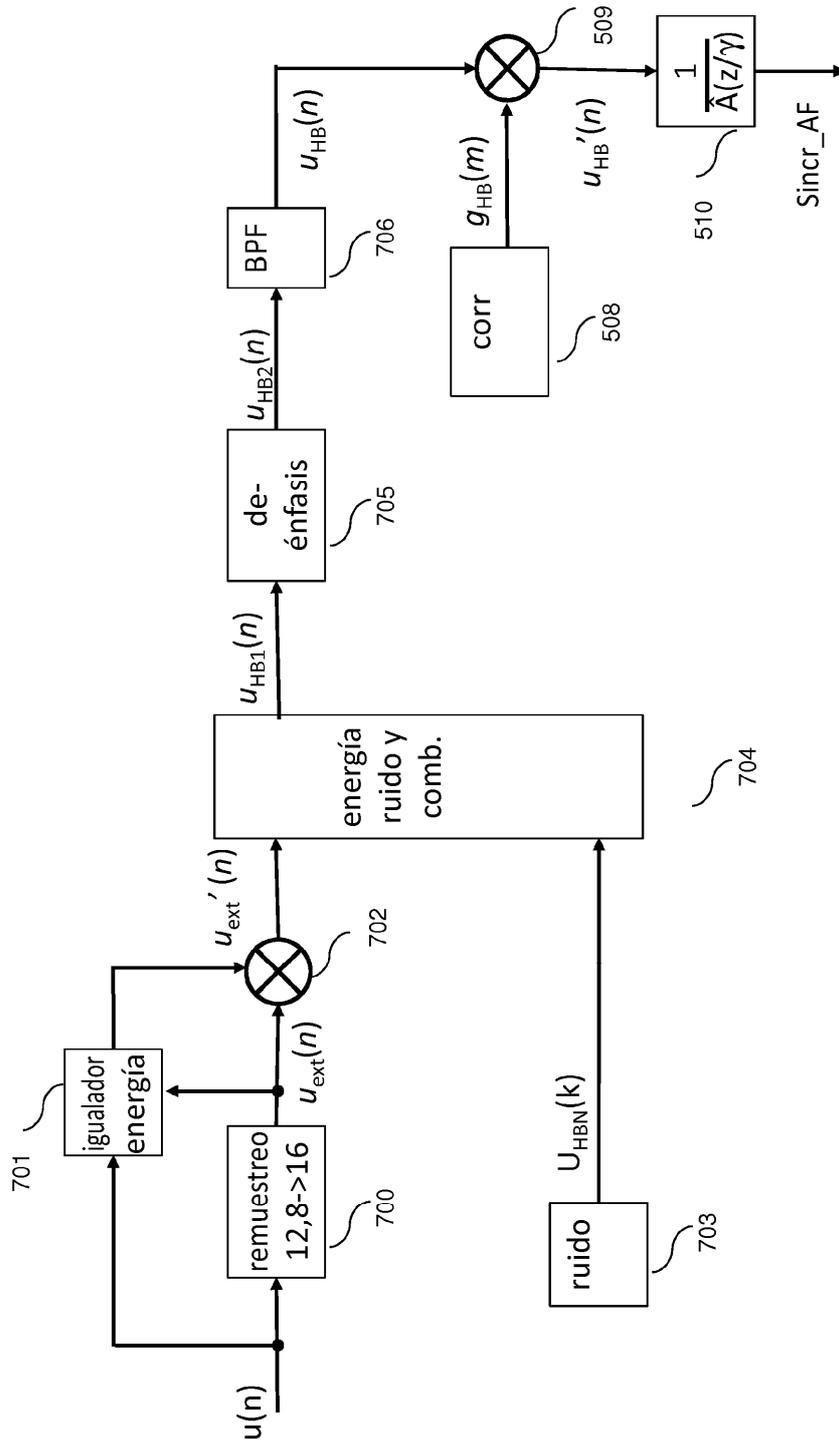


Fig.7