



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 356 492**

51 Int. Cl.:
G01L 19/14 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06779036 .0**

96 Fecha de presentación : **10.07.2006**

97 Número de publicación de la solicitud: **1907812**

97 Fecha de publicación de la solicitud: **09.04.2008**

54

Título: **Método de conmutación de tasa de transmisión en decodificación de audio escalable en tasa de transmisión y ancho de banda.**

30

Prioridad: **22.07.2005 FR 05 52286**

45

Fecha de publicación de la mención BOPI:
08.04.2011

45

Fecha de la publicación del folleto de la patente:
08.04.2011

73

Titular/es: **FRANCE TELECOM**
6 place d'Alleray
75015 Paris, FR

72

Inventor/es: **Ragot, Stéphane;**
Virette, David y
Kovesi, Balazs

74

Agente: **Lehmann Novo, María Isabel**

ES 2 356 492 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

Método de conmutación de tasa de transmisión en decodificación de audio escalable en tasa de transmisión y ancho de banda

5 La presente invención se refiere a un método de conmutación de tasa de transmisión a la decodificación de una señal de audio codificada por un sistema de codificación de audio multi-tasas de transmisión y más en particular, un sistema de codificación de audio escalable en tasa de transmisión y posiblemente en ancho de banda. Se refiere, además, a una aplicación de dicho método a un sistema de decodificación de audio escalable en tasa de transmisión y en ancho de banda y un decodificador de audio escalable en tasas de transmisión y en ancho de banda.

10 La invención encuentra una aplicación particularmente ventajosa en el campo de la transmisión de señales de voz y/o audio en redes de paquetes, de tipo voz sobre IP, con el fin de proporcionar una calidad modulable en función de la capacidad del canal de transmisión.

15 El método según la invención permite obtener transiciones sin artefactos de comunicación entre las diferentes tasas de transmisión de un codificador/decodificador (*codec*) de audio escalable en tasa de transmisión y en ancho de banda, más en particular en el caso de las transiciones entre la banda telefónica y la banda ancha dentro del contexto de una codificación de audio escalable en tasa de transmisión y en ancho de banda, con un núcleo en banda telefónica y su post-tratamiento, dependiente de la tasa de transmisión, y una o varias capas de mejora en banda ancha.

20 De manera habitual, se entiende por "banda telefónica" o "banda estrecha" la banda de frecuencia situada entre 300 y 3400 Hz, mientras que el término de "banda ancha" se reserva a la banda que se escalona de 50 a 7000 Hz.

Actualmente, existen numerosas técnicas para convertir una señal de audiofrecuencias (voz y/o audio) bajo la forma de una señal digital y para tratar las señales así digitalizadas.

25 Las técnicas más corrientes son los métodos de "codificación de forma de onda", tales como la codificación MIC o MICDA (PCM o ADPCM, en la terminología inglesa), los métodos de "codificación paramétrica por análisis por síntesis" como la codificación CELP ("Code Excited Linear Prediction"), y los métodos de "codificación perceptual en sub-bandas o por transformada". Se hace constar que en la codificación CELP en banda estrecha, se utiliza, en general, un post-tratamiento que sirve para mejorar la calidad. Este post-tratamiento comprende, en una realización típica, un post-filtrado adaptativo y un filtro de paso alto. Estas técnicas clásicas de codificación de las señales de audio-frecuencias se describen, por ejemplo, en la obra de WB. Kleijn and K.K. Paliwal editors, *Speech Coding and Synthesis*, Elsevier, 1995. En la presente invención sólo son de interés las técnicas utilizadas en la transmisión bi-direccional de las señales de audio-frecuencias.

35 En la codificación de voz convencional, el codificador genera un flujo binario de tasa de transmisión fija. Esta limitación de tasa de transmisión fija simplifica la puesta en práctica y la utilización del codificador y del decodificador. Ejemplos de dichos sistemas se proporcionan por la codificación G.711 a 64 kbits/s o la codificación G.729 a 8 kbits/s.

40 En algunas aplicaciones, tal como la telefonía móvil, la voz sobre IP, o las comunicaciones en redes *ad hoc*, es preferible generar un flujo binario de tasa de transmisión variable, tomándose los valores de la tasa de transmisión en un conjunto predefinido. Se distingue varias técnicas de codificación multitasas de transmisión:

- La codificación multi-modos controlada por la fuente y/o el canal, tal como se pone en práctica en los sistemas AMR-NB, AMR-WB, SMV o VMR-WB.

45 - La codificación jerárquica, llamada también codificación "escalable", que genera un flujo binario denominado jerárquico porque comprende una tasa de transmisión núcleo y una o varias capas de mejora. El sistema G.722 a 48, 56 y 64 kbits/s es un ejemplo simple de codificación escalable en tasas de transmisión. El *codec* MPEG-4 CELP es, por sí mismo, escalable en tasa de transmisión y en ancho de banda (T. Nomura et al., *A bitrate and bandwidth scalable CELP coder*, ICASSP 1998).

50 - La codificación con descripciones múltiples (A. Gersho, J.D. Gibson, V. Cuperman, H. Dong, *A multiple description speech coder based on AMR-WB for mobile ad hoc networks*, ICASSP 2004).

En codificación multi-tasas de transmisión, es necesario cerciorarse de que la conmutación de una tasa de transmisión de codificación a otro no implica algunos fallos o artefactos de la comunicación.

La conmutación de tasas de transmisión es fácil de realizar si la codificación se basa, en todas las tasas de transmisión, sobre la representación por un mismo modelo de codificación de una señal de

audio en un mismo ancho de banda. Por ejemplo, en el sistema AMR-NB, la señal se define en banda telefónica (300-3400 Hz) y la codificación se basa en el modelo ACELP ("Algebraic Code Excited Linear Prediction"), salvo para la generación de ruido de confort, la cual, no obstante, se realiza por un modelo de tipo LPC ("Linear Predictive Coding") compatible con el modelo ACELP. Conviene señalar que la codificación AMR-NB utiliza, de forma clásica, un post-tratamiento bajo la forma de un post-filtrado adaptativo y de un filtro de paso alto, dependiendo los coeficientes del post-filtrado adaptativo de la tasa de transmisión de la decodificación. Sin embargo, ha de tomarse alguna precaución para no generar los posibles problemas relacionados con la utilización de parámetros de post-tratamiento, variables según la tasa de transmisión. En cambio, la codificación CELP en banda ancha, de tipo AMR-WB, no utiliza post-tratamiento, esencialmente por razones de complejidad.

La conmutación de tasas de transmisión es todavía más problemática en la codificación de audio escalable en tasa de transmisión y en ancho de banda. En efecto, en este caso, la codificación se basa en modelos y anchos de banda diferentes según la tasa de transmisión.

El concepto básico de la codificación de audio jerárquica se ilustra, por ejemplo, en el artículo de Y. Hiwasaki, T. Mori, H. Ohmuro, J. Ikedo, D. Tokumoto y A. Kataoka, Scalable Speech Coding Technology for High- Quality Ubiquitous Communications, NTT Technical Review, Marzo 2004. En este tipo de codificación, el flujo binario comprende una capa de base y una o varias capas de mejora. La capa de base se genera por un *codec* de baja tasa de transmisión fija, calificado como "codec núcleo", que garantiza la calidad mínima de la codificación. Esta capa se debe recibir por el decodificador para mantener un nivel de calidad aceptable. Las capas de mejora sirven para mejorar la calidad. Si todas ellas se emiten por el codificador, puede darse el caso, sin embargo, de que no sean todas recibidas por el decodificador. El interés principal de la codificación jerárquica es que permite una adaptación de la tasa de transmisión por simple truncado del flujo binario. El número de capas, es decir el número de truncados posibles del flujo binario, define la granularidad de la codificación. Se denomina codificación de granularidad fuerte si el flujo binario comprende pocas capas, del orden de 2 a 4, permitiendo una codificación de granularidad fina un paso del orden de 1 kbit/s.

En este caso, revisten más interés las técnicas de codificación jerárquica que sean escalables en tasa de transmisión y en ancho de banda, con un codificador núcleo de tipo CELP en banda telefónica y una o varias capa(s) de mejora en banda ancha. Ejemplos de tales sistemas se proporcionan en H. Taddéi et al., A Scalable Three Bitrate (8, 14.2 y 24 kbits/s) Audio Coder; 107th Convention AES, 1999 con una granularidad fuerte de 8, 14.2 y 24 kbits/s, y en B. Kovesi, D. Massaloux, A. Sollaud, A scalable speech and audio coding scheme with continuous bitrate flexibility, ICASSP 2004 con granularidad fina de 6,4 a 32 kbits/s o también la codificación MPEG-4 CELP.

Entre las referencias más pertinentes relacionadas con el problema de la conmutación de tasa de transmisión en el contexto de la codificación de audio escalable en tasa de transmisión y en ancho de banda, se puede citar las solicitudes internacionales WO 01/48931 y WO 02/060075.

Sin embargo, las técnicas descritas en estos dos documentos sólo tratan problemas de interoperabilidad entre redes de comunicación que utilizan codificaciones en banda telefónica y en banda ancha.

En particular, la solicitud internacional WO 02/060075 describe un sistema optimizado de decimación que permite la conversión de la banda ancha hacia la banda telefónica.

El método propuesto en la solicitud internacional WO 01/48931 es, de hecho, una técnica de extensión de banda que consiste en generar una señal en banda pseudo-ancha a partir de una señal en banda telefónica, en particular por extracción de un "perfil espectral". Las técnicas similares conocidas de la técnica anterior dan respuesta principalmente a los problemas asociados a la conmutación de la banda ancha hacia la banda telefónica tratando de evitar la reducción de banda mediante la utilización de una técnica de extensión de banda, sin transmisión de información, que permite generar una señal en banda ancha a partir de la señal recibida en banda telefónica. Conviene resaltar que estos métodos no buscan controlar verdaderamente la transición entre anchos de banda y que presentan, por otro lado, el inconveniente de basarse en técnicas o de extensión de banda, cuya calidad es muy variable y que, por lo tanto, no puede garantizar una calidad estable a la salida. En el documento US2001/044712, se realiza un post-tratamiento para la decodificación, en el momento de las transiciones, con el fin de simular una variación continua del ancho de banda.

Además, el problema técnico a resolver por el objeto de la presente invención es dar a conocer un método de conmutación de tasas de transmisión a la decodificación de una señal de audio codificada por un sistema de codificación de audio multi-tasas de transmisión, comprendiendo dicha decodificación al menos una etapa de post-tratamiento dependiente de la tasa de transmisión, que permitiría tratar las transiciones entre tasas de transmisión diferentes para los cuales se utilizan post-tratamientos según la tasa de transmisión de decodificación, de manera que se eliminen los artefactos de comunicación

particularmente sensibles en el momento de variaciones rápidas de tasa de transmisión en la decodificación. En efecto, un post-tratamiento introduce un defasaje en la señal, y la utilización de dos post-tratamientos diferentes plantea problemas de continuidad de fase en el momento de las transiciones.

5 La solución al problema técnico planteado, según la presente invención, se describe en la reivindicación 1.

La invención se refiere, además, a un programa de ordenador que comprende instrucciones de código para la puesta en práctica del método, según la invención, cuando dicho programa se ejecuta por un ordenador.

10 La invención se refiere, además, a una aplicación del método, según la invención a un sistema de decodificación de audio escalable en tasa de transmisión.

15 La invención se refiere, además, a una aplicación del método según la invención a un sistema de decodificación de audio escalable en tasa de transmisión y ancho de banda, en donde la tasa de transmisión inicial se obtiene por al menos una primera capa de decodificación en una primera banda de frecuencia y la tasa de transmisión final se obtiene por al menos una segunda capa de decodificación, denominada capa de extensión de la denominada primera banda de frecuencia en una segunda banda de frecuencia, aplicándose la etapa de post-tratamiento a la decodificación realizada al tasa de transmisión inicial.

20 La invención se refiere, además, a una aplicación del método según la invención a un sistema de decodificación de audio escalable en tasa de transmisión y ancho de banda en donde la tasa de transmisión final se obtiene por al menos una primera capa de decodificación en una primera banda de frecuencia, y la tasa de transmisión inicial se obtiene por al menos una segunda capa de decodificación, denominada capa de extensión de la denominada primera banda de frecuencia en una segunda banda de frecuencia, siendo la etapa de post-tratamiento aplicada a la decodificación realizada al tasa de transmisión final.

25 Un ejemplo particular de "banda extendida" es el de la "banda ancha" definida con anterioridad, siendo la denominada primera banda, en este caso, la banda telefónica.

La invención se refiere, además, a un decodificador de audio multi-tasas de transmisión, según se reivindica en la reivindicación 10.

30 La descripción siguiente, que hace referencia a los dibujos adjuntos, dados a título de ejemplos no limitativos, hará comprender mejor en qué consiste la invención y cómo se puede realizar.

La Figura 1 es un diagrama de un codificador escalable en tasa de transmisión y en ancho de banda con cuatro capas.

La Figura 2 es un diagrama de un decodificador, según la invención, asociado al codificador de la Figura 1.

35 La Figura 3 ilustra una estructura del tren binario asociado al codificador de la Figura 1.

La Figura 4 es un organigrama de un método de conmutación entre una señal post-tratada y una señal no post-tratada, en la banda telefónica del decodificador, según la invención.

La Figura 5 es un organigrama del método de conmutación conforme a la invención entre una banda telefónica y una banda ancha con extensión de banda.

40 La Figura 6 es un organigrama del método de conmutación conforme a la invención entre una banda telefónica y una banda ancha con una capa de decodificación predictiva por transformada.

La Figura 7 es un organigrama de la gestión del conteo de tramas recibidas en banda ancha para la conmutación entre tasas de transmisión y entre bandas, de conformidad con el método según la invención.

45 La Figura 8 es una tabla que resume el funcionamiento del organigrama de la Figura 7.

La Figura 9 es una tabla que proporciona los coeficientes de atenuación adaptativa en el momento de una conmutación desde la banda telefónica a la banda ancha.

50 La invención se describe ahora dentro del marco de una codificación de audio escalable en tasa de transmisión y en ancho de banda. La estructura de la codificación escalable en tasa de transmisión y en ancho de banda considerada aquí tiene como codificación núcleo un codificador de tipo CELP en banda telefónica, cuyo caso particular utiliza el codificador G.729A tal como se describe en ITU-T G729

Recommandation, Coding of Speech at 8 kbit/s using Conjugate Structure Algebraic Code Excited Linear Prediction (CS-ACELP), Marzo 1996 y en R. Salami et al., Description of ITU-T Recommendation G.729 Annex A: Reduced complexity 8 kbit/s CS-ACELP codec, ICASSP 1997.

5 A la codificación núcleo CELP se añaden tres etapas de mejora, a saber, una mejora de la codificación CELP en banda telefónica, una extensión de banda y una codificación predictiva por transformada.

Las conmutaciones de tasa de transmisión aquí consideradas se referirán a conmutaciones entre la banda telefónica y la banda ancha y viceversa.

La Figura 1 representa un diagrama del codificador utilizado.

10 Una señal de audio de banda útil de 50-7000 Hz y muestreada a 16 kHz se secciona en tramas de 320 muestras, o sea 20 ms. Un filtro de paso alto 101, de frecuencia de corte a 50 Hz, se aplica a la señal de entrada. La señal obtenida, denominada S^{WB} , se reutiliza en varias ramas del codificador.

15 Ante todo, en una primera rama, un filtro de paso bajo y un sub-muestreo por dos, 102, de 16 a 8 kHz se aplican a la señal S^{WB} . Esta operación permite obtener una señal en banda telefónica muestreada a 8 kHz. Esta señal es tratada por el codificador núcleo 103, según una codificación de tipo CELP. Esta codificación corresponde aquí al codificador G.729A, el cual genera el núcleo del tren binario con una tasa de transmisión de 8 kbits/s.

20 A continuación, una primera capa de mejora introduce una segunda etapa 103 de codificación CELP. Esta segunda etapa consiste en un diccionario innovador que realiza un enriquecimiento de la excitación CELP y ofrece una mejora de la calidad, particularmente en los sonidos no vocalizados. La tasa de transmisión de esta segunda etapa de codificación es de 4 kbits/s y los parámetros asociados son las posiciones y los signos de los impulsos así como la ganancia del diccionario innovador asociado a cada sub-trama de 40 muestras (5 ms a 8 kHz).

25 Les decodificaciones del codificador núcleo y de la primera capa de mejora se realizan para obtener la señal de síntesis 104 en banda telefónica a 12 kbits/s. Un sobre-muestreo por dos de 8 a 16 kHz y un filtro de paso bajo 105 permiten obtener la versión muestreada a 16 kHz de las dos primeras etapas del codificador.

La tercera capa de mejora permite pasar a la banda ancha 106. La señal de entrada S^{WB} puede ser pre-tratada por un filtro de pre-énfasis.

30 Este filtro permite representar mejor las altas frecuencias a partir del filtro de predicción lineal en banda ancha. Para compensar el efecto del filtro de pre-énfasis, un filtro inverso de de-énfasis se utiliza entonces en la síntesis. Una alternativa a esta estructura de codificación y de decodificación no utilizará ningún filtro de pre-énfasis ni de de-énfasis.

35 La etapa siguiente consiste en calcular y en cuantificar los filtros de predicción lineal en banda ancha. El orden del filtro de predicción lineal es de 18, pero en una variante, se elegirá un orden de predicción más pequeño, por ejemplo 16. El filtro de predicción lineal se puede calcular por el método de la auto-correlación y el algoritmo de Levinson-Durbin.

40 Este filtro de predicción lineal $A_{WB}(Z)$, en banda ancha, se cuantifica utilizando una predicción de estos coeficientes a partir del filtro $\hat{A}_{NB}(Z)$ procedente del codificador núcleo en banda telefónica. Los coeficientes se pueden cuantificar, a continuación, utilizando, por ejemplo, una cuantificación vectorial multi-etapas y empleando los parámetros LSF ("Line Spectrum Frequency") descuantificados del codificador núcleo, en banda telefónica, según se describe en H. Ebara, T. Morii, M. Oshikiri y K. Yoshida, Predictive VQ for bandwidth scalable LSP quantization, ICASSP 2005.

45 La excitación, en banda ancha, se obtiene a partir de los parámetros de la excitación en banda telefónica del codificador núcleo: el retardo del periodo fundamental o "pitch", la ganancia asociada así como las excitaciones algébricas del codificador núcleo y de la primera capa de enriquecimiento de la excitación CELP y las ganancias asociadas. Esta excitación se genera utilizando una versión sobre-muestreada de los parámetros de la excitación de las etapas en banda telefónica.

50 Esta excitación, en banda ancha, es a continuación puesta en forma por el filtro de síntesis $\hat{A}_{WB}(Z)$ calculado anteriormente. En el caso de que una pre-énfasis fuera aplicada a la señal de entrada, se aplica el filtro de de-énfasis a la señal de salida del filtro de síntesis. La señal obtenida es una señal en banda ancha que no se ajusta en energía. Para el cálculo de la ganancia, que permite la puesta al nivel de la energía de la banda alta (3400-7000 Hz), un filtrado de paso alto se aplica a la señal de síntesis en banda ancha. Paralelamente, el mismo filtro de paso alto se aplica a la señal de error correspondiente a la diferencia entre la señal original retardada y la señal de síntesis de las dos etapas precedentes. Estas dos

55

señales se utilizan, a continuación, para el cálculo de la ganancia a aplicar a la señal de síntesis de la banda alta. Esta ganancia se calcula por una relación de energía entre las dos señales. La ganancia g_{WB} cuantificada se aplica, a continuación, a la señal S_{14}^{WB} por sub-trama de 80 muestras (5 ms a 16 kHz), siendo añadida la señal así obtenida a la señal de síntesis de la etapa precedente para crear la señal en banda ancha correspondiente al tasa de transmisión de 14 kbits/s.

La sucesión de la codificación se realiza, en el dominio de la frecuencia, utilizando un sistema de codificación predictiva por transformada. Las señales de entrada retardadas 108 y de síntesis a 14 kbits/s, 107, se filtran por un filtro 109, 111 de ponderación perceptual de tipo $A_{WB}(z/y)^*(1-\mu z)$, típicamente $y=0,92$ y $\mu=0,68$. Estas señales son a continuación codificadas por el esquema de codificación por transformada, con recubrimiento de tipo TDAC ("Time Domain Aliasing Cancellation") (Y. Mahieux et J.P. Petit, Transform coding of audio signals at 64 kbit/s, IEEE GLOBECOM 1990).

Una transformada en coseno discreto modificada (o MDCT en terminología inglesa) se aplica, de una parte, 110, en bloques de 640 muestras de la señal de entrada ponderada con un recubrimiento del 50% (regeneración del análisis MDCT cada 20 ms) y, de otra parte, 112, en la señal de síntesis ponderada procedente de la etapa precedente de extensión de banda a 14 kbits/s (misma longitud de bloque y misma tasa de recubrimiento). El espectro MDCT a codificar, 113, corresponde a la diferencia entre la señal de entrada ponderada y la señal de síntesis a 14 kbits/s para la banda de 0 a 3400 Hz y a la señal de entrada ponderada de 3400 Hz a 7000 Hz. Se limita el espectro a 7000 Hz poniéndolo a cero los 40 últimos coeficientes (sólo los 280 primeros coeficientes son codificados). El espectro está dividido en 18 bandas: una banda de 8 coeficientes y 17 bandas de 16 coeficientes. Para cada banda del espectro, se calcula la energía de los coeficientes MDCT (factores de escala). Los 18 factores de escala constituyen la envolvente espectral de la señal ponderada que es, a continuación, cuantificada, codificada y transmitida en la trama. La Figura 3 representa el formato del tren binario.

La asignación dinámica de los bits se basa en la energía de las bandas del espectro a partir de la versión descuantificada de la envolvente espectral. Esto permite tener una compatibilidad entre la asignación binaria del codificador y del decodificador. Los coeficientes MDCT normalizados (estructura fina), en cada banda son, a continuación, cuantificados mediante cuantificadores vectoriales utilizando diccionarios imbricados en tamaño y en dimensión, estando los diccionarios compuestos por una unión de códigos de permutación, tales como los descritos en C. Lamblin et al., Cuantificación vectorial en dimensión y resolución variables, patente PCT FR 04 00219, 2004. Por último, las informaciones sobre el codificador núcleo, la etapa de enriquecimiento CELP en banda telefónica, la etapa CELP en banda ancha y por último, la envolvente espectral y los coeficientes normalizados codificados son multiplexados y transmitidos en trama.

La Figura 2 representa un diagrama de bloques del decodificador asociado al codificador de la Figura 1.

El módulo 201 realiza la demultiplexación de los parámetros contenidos en el tren binario. Existen varios casos de decodificación en función del número de bits recibidos para una trama, los cuatro casos se describen a partir de la Figura 2:

1. El primer caso se refiere a la recepción del número de bits mínimo por el decodificador, para una tasa de transmisión recibida de 8 kbits/s. En este caso, sólo la primera etapa es decodificada. Por lo tanto, solamente el tren binario relativo al decodificador núcleo 202 de tipo CELP (G.729A+) es recibido y decodificado. Esta síntesis se puede tratar por el post-filtrado adaptativo 203 y el post-tratamiento de tipo filtro de paso alto 204 del decodificador G.729. Se denominará, en este ejemplo de realización, "post-tratamiento" a la combinación de estas dos operaciones. Sin embargo, es evidente que el término de "post-tratamiento" puede hacer también referencia únicamente al post-filtrado adaptativo o al post-tratamiento de tipo filtro de paso alto. Esta señal es sobre-muestreada, 206, y filtrada, 207, para proporcionar una señal muestreada a 16 kHz.

2. El segundo caso se refiere a la recepción del número de bits relativo a las primera y segunda etapas de decodificación únicamente, para una tasa de transmisión recibida de 12 kbits/s. En este caso, se decodifican el decodificador núcleo así como la primera etapa de enriquecimiento de la excitación CELP. Esta síntesis se puede tratar mediante el post-tratamiento 203, 204 del decodificador G.729. Como anteriormente, esta señal es, a continuación, sobre-muestreada, 206, y filtrada, 207 para proporcionar una señal muestreada a 16 kHz.

3. El tercer caso corresponde a la recepción del número de bits relativos a las tres primeras etapas de la decodificación, para una tasa de transmisión recibida de 14 kbits/s. En este caso, las dos primeras etapas de decodificación son, ante todo, realizadas como en el caso 2, salvo el hecho de que no se realiza el post-tratamiento aplicado a la salida de decodificación CELP, luego el módulo de extensión de banda genera una señal muestreada a 16 kHz después de la decodificación de los parámetros de los pares de rayas espectrales (WB-LSF) en banda ancha, 209, así como ganancias asociadas a la

excitación, 213. La excitación en banda ancha se genera a partir de los parámetros del codificador núcleo y de la primera etapa de enriquecimiento de la excitación CELP 208. Esta excitación es a continuación filtrada por el filtro 210 de síntesis y ocasionalmente, por el filtro 211 de de-énfasis en el caso en donde se utilizó pre-énfasis al codificador. Se aplica un filtro de paso alto 212 a la señal obtenida y se adapta la energía de la señal de extensión de banda con la ayuda de las ganancias asociadas 214 cada 5 milisegundos. Esta señal se añade, a continuación, a la señal en banda telefónica muestreada a 16 kHz, obtenida a partir de las dos primeras etapas 215 de decodificación. Con el objeto de obtener una señal limitada a 7000 Hz, esta señal se filtra en el dominio transformado por la puesta a 0 de los 40 últimos coeficientes MDCT, antes del paso por la MDCT inversa 220 y el filtro de síntesis ponderado 221.

4. Este último caso corresponde a la decodificación de todas las etapas del decodificador, para una tasa de transmisión recibida superior o igual a 16 kbits/s. La última etapa está constituida por un decodificador predictivo por transformada. La etapa 3, anteriormente descrita, se realiza en primer lugar. Luego, en función del número de bits suplementarios recibidos, el esquema de decodificación y predictivo por transformada está adaptado para:

* En el caso de que el número de bits solamente corresponda a una parte o a la totalidad de la envolvente espectral, pero que no se recibe la estructura fina, la envolvente espectral parcial o completa se utiliza para ajustar la energía de las bandas de coeficientes MDCT, 216 y 217, entre 3400 Hz y 7000 Hz 218, correspondiente a la señal generada por la etapa 215 de extensión de banda. Este sistema permite obtener una mejora progresiva de la calidad de audio en función del número de bits recibido.

* En el caso de que el número de bits corresponda a la totalidad de la envolvente espectral y a una parte o a la totalidad de la estructura fina, la asignación binaria se realiza de la misma manera que con el codificador. En las bandas en donde se reciba la estructura fina, los coeficientes MDCT decodificados se calculan a partir de la envolvente espectral y de la estructura fina descuantificadas. En las bandas espectrales entre 3400 Hz y 7000 Hz, en donde la estructura fina no ha sido recibida, se utiliza el procedimiento del apartado anterior, es decir, que los coeficientes MDCT calculados sobre la señal obtenida por la extensión de banda, 216 y 217, se ajustan en energía a partir de la envolvente espectral recibida 218. El espectro MDCT utilizado para la síntesis está, por lo tanto, constituida, de una parte, por la señal de síntesis de las dos primeras etapas de la decodificación añadida a la señal de error decodificado en las bandas entre 0 y 3400 Hz; de otra parte, para las bandas comprendidas entre 3400 Hz y 7000 Hz, de los coeficientes MDCT decodificados en las bandas, en donde fue recibida la estructura fina y coeficientes MDCT de la etapa de extensión de banda ajustados en energía para las demás bandas espectrales.

Una MDCT inversa se aplica, a continuación, a los coeficientes MDCT decodificados, 220, y un filtrado por el filtro 221 de síntesis ponderada permite obtener la señal de salida.

El método de conmutación conforme a la invención se expondrá ahora dentro del marco del decodificador de la Figura 2.

El bloque 205 representa un módulo de "fundido encadenado". Cuando el número de bits recibidos por el decodificador sólo permite decodificar la primera o la primera y la segunda etapas, es decir para una tasa de transmisión recibida de 8 o 12 kbits/s, la banda pasante efectiva de la salida final del decodificador es la banda telefónica. En estos casos, para mejorar la calidad de la señal sintetizada, el post-tratamiento 203, 204 en sentido amplio que forma parte del decodificador G.729A se aplica, en banda telefónica, antes del sobre-muestreado.

Por el contrario, si se realiza también la decodificación de las etapas en banda ancha, para una tasa de transmisión recibida superior o igual a 14 kbits/s, este post-tratamiento no se activa porque, en el codificador, la codificación de las etapas superiores fue calculada a partir de la versión sin post-tratamiento de la banda telefónica.

El post-tratamiento, 203 y 204, introduce un desfase de la señal. En el momento de la conmutación entre modos, sin y con post-tratamiento, hay que garantizar, por lo tanto, una transición suave. La Figura 4 describe la realización del bloque 205 que garantiza esta transición lenta entre la señal en banda telefónica post-tratada y no post-tratada, aplicando los denominados fundidos encadenados.

La etapa 401 examina si la trama corriente es una trama en banda telefónica o no, es decir que se comprueba si la tasa de transmisión de la trama corriente está a 8 o 12 kbits/s. En caso de respuesta negativa, se llama una etapa 402 para comprobar si la trama precedente fue post-tratada, o no, en la banda telefónica (lo que consiste en verificar si la tasa de transmisión de la trama precedente era de 8-12 kbits/s o no). En caso de respuesta negativa, en la etapa 403, la señal no post-tratada S_1 se copia en la señal S_3 . Al contrario, en caso de respuesta positiva a la prueba 402, en la etapa 404, la señal S_3 contendrá el resultado de una fundido encadenada, en donde aumenta el peso de la componente no post-tratada S_1 mientras que disminuye el peso de la componente post-filtro S_2 . La etapa 404 va seguida por la etapa 405 que actualiza la bandera prevPF con el valor 0.

En el caso de una respuesta positiva a la etapa 401, en la etapa 406, se comprueba si, en la trama precedente, el post-tratamiento estaba activo, o no, en la banda telefónica. En caso de respuesta positiva, en la etapa 408, la señal post-tratada S_2 es copiada en la señal S_3 . Cuando, al contrario, la respuesta es negativa a la etapa 406, se calcula la señal S_3 , en la etapa 407, como el resultado de un fundido encadenado, en donde esta vez disminuye el peso de la componente no post-tratada S_1 mientras que aumenta el peso de la componente post-tratada S_2 . Después de la etapa 407, se llama la etapa 409 para reactualizar la bandera prevPF con el valor 1.

En una variante de esta forma de realización, cuando el número de bits recibidos por el decodificador sólo permite decodificar la primera o las primera y segunda etapas, es decir, para una tasa de transmisión recibida de 8 o 12 kbits/s, la banda pasante efectiva de la salida final del decodificador es la banda telefónica (señal S_1). En estos casos, para mejorar la calidad de la señal sintetizada, se aplica un post-tratamiento en banda telefónica, antes del sobre-muestreo.

Por el contrario, si se realiza también la decodificación de las etapas en banda ancha, para una tasa de transmisión recibida superior o igual a 14 kbits/s, se activa un post-tratamiento diferente (señal S_2) al codificador, la codificación de las etapas superiores fue calculada a partir de la versión con este post-tratamiento de la banda telefónica.

El post-tratamiento utilizado para las tasas de transmisión de 8 o 12 kbits/s y el post-tratamiento empleado para las tasas de transmisión superiores o iguales a 14 kbits/s introducen desfases de la señal diferentes entre sí. En el momento de la conmutación entre modos con los diferentes post-tratamientos, es preciso, por lo tanto, garantizar una transición suave. Esta transición lenta entre las señales en banda telefónica, con los diferentes post-tratamientos, se realiza aplicando fundidos encadenados (que proporcionan la señal S_3).

Se examina si la trama corriente es una trama en banda telefónica o no lo es. En caso de respuesta negativa, se comprueba si la trama precedente era una trama en banda telefónica. En caso de respuesta negativa, la señal post-tratada S_1 se copia en la señal S_3 . Al contrario, en caso de respuesta positiva, la señal S_3 contendrá el resultado de un fundido encadenado, en donde aumenta el peso de la componente post-tratada S_1 , mientras que disminuye el peso de la componente post-tratada S_2 .

En el caso de una respuesta positiva, se comprueba si la trama precedente era una trama en banda telefónica. En caso de respuesta positiva, la señal post-tratada S_2 se copia en la señal S_3 . Cuando, por el contrario, la respuesta es negativa, se calcula la señal S_3 como el resultado de un fundido encadenado, en donde, esta vez, disminuye el peso de la componente post-tratada S_1 , mientras que aumenta el peso de la componente post-tratada S_2 .

El bloque 209 calcula los filtros de predicción lineal, en banda ancha, que se necesitan para las etapas de extensión de banda y decodificación predictiva por transformada. Este cálculo se necesita en el caso de que solamente se reciba la parte, en banda telefónica, del tren binario de una trama, después de haber recibido una trama en banda ancha y que se desee realizar una extensión de banda, con el fin de mantener el efecto de banda. Un juego de LSF se extrapola a partir de los LSF del decodificador núcleo en banda telefónica. Se puede, por ejemplo, distribuir uniformemente 8 LSF en la banda comprendida entre el último LSF emitido desde la banda telefónica y la frecuencia de Nyquist. Ello permite hacer tender el filtro de predicción lineal hacia un filtro de respuesta en amplitud plana para las altas frecuencias.

El bloque 213 realiza la adaptación de la ganancia utilizada para la extensión de banda según la presente invención. Los organigramas que corresponden a este bloque se describen en las Figuras 5 y 7.

El principio de la atenuación adaptativa de la ganancia aplicado a la banda alta se describe en la Figura 5. Ante todo, se efectúa el cálculo de la ganancia de la primera capa de decodificación en banda ancha, 501, según dos posibilidades. En el caso de que se haya recibido el tren binario correspondiente a esta capa de extensión de banda, la ganancia se obtiene por decodificación, 503. Por el contrario, en el caso de que esta ganancia no se haya recibido en el tren binario, se realiza una extrapolación de la ganancia asociada a esta capa de decodificación, 502. Se puede, por ejemplo, realizar un cálculo de la ganancia mediante alineación de la energía de la banda baja de la etapa de decodificación, en banda ancha, con la decodificación real de la banda telefónica precedentemente realizada.

A continuación, se actualiza un contador del número de tramas en banda ancha precedentemente recibidas, 504, según el principio descrito en la Figura 7.

Por último, este contador se utiliza para definir los parámetros de la atenuación aplicada a la ganancia de la primera etapa de decodificación en banda ancha, 505.

La Figura 7 representa el organigrama de la gestión del conteo del número de tramas en banda ancha recibidas. La actualización del contador se hace de la forma siguiente. Si la trama corriente es una trama en banda ancha, en tal caso fue recibida si la ganancia asociada a la primera etapa de

5 decodificación en banda ancha (bloque 501 de la Figura 5) y la trama precedente era también una trama en banda ancha, entonces el contador se incrementa en 1 y satura el valor MAX_COUNT_RCV. Este valor corresponde al número de tramas durante las cuales la señal decodificada en banda ancha se atenúa en el momento de una conmutación entre una tasa de transmisión en banda telefónica hacia una tasa de transmisión en banda ancha.

10 Por el contrario, si la trama corriente recibida es una trama en banda telefónica, existen varios comportamientos posibles. Si la trama precedente era también una trama en banda telefónica, el contador se posiciona en 0. En el caso contrario, si la trama precedente era una trama en banda ancha y el contador presenta un valor inferior a MAX_COUNT_RCV, se posiciona también el contador a 0. En todos los demás casos, el contador permanece en el valor precedente.

15 El funcionamiento de este organigrama se resume en la tabla de la Figura 8. Los valores tomados por el coeficiente de atenuación se proporcionan en la tabla de la Figura 9 en el caso de que MAX_COUNT_RCV tome el valor de 100, esta tabla se proporciona a título de ejemplo. Se puede constatar que, hasta la trama 65, el coeficiente de atenuación se mantiene a 0, que corresponde a una fase de prolongación de la decodificación en la banda telefónica. La fase de transición propiamente dicha se realiza a partir de la trama 66 mediante un aumento progresivo del coeficiente de atenuación.

El bloque 219 efectúa la atenuación adaptativa de las capas de mejora mediante codificación predictiva por transformada, según la presente invención, tal como se describe en la Figura 6.

20 En esta figura se ilustra el organigrama del procedimiento de atenuación adaptativa de la capa de decodificación predictiva por transformada. Ante todo, se comprueba si la envolvente espectral de esta capa ha sido totalmente recibida, 601. Si tal es el caso, entonces se realiza una atenuación de los coeficientes MDCT de corrección de la banda base 0-3500 Hz, 602, el contador de tramas en banda ancha recibidas y la tabla de atenuación definida en la Figura 9.

25 A continuación, en ambos casos, se controla el número de tramas en banda ancha recibidas. Si este número es inferior a MAX_COUNT_RCV, los coeficientes MDCT que corresponden a la primera etapa de decodificación en banda ancha con extensión de banda con transmisión de información se utilizan para la etapa de decodificación predictiva por transformada. Por el contrario, si el contador tiene el valor máximo, se realiza el procedimiento de puesta al nivel de la energía de las bandas de la decodificación predictiva por transformada con la envolvente espectral decodificada.

REIVINDICACIONES

- 5 1.- Método de conmutación de tasas de transmisión a la decodificación de una señal de audio codificada por un sistema de codificación de audio multi-tasas de transmisión, a partir de una señal de código, dos señales, denominadas primera señal (S1) y segunda señal (S2), se proporcionan a la entrada de un módulo de fundido encadenado, siendo al menos una de las señales post-tratada en el momento de una etapa de post-tratamiento, formando el post-tratamiento parte de un conjunto de post-tratamientos adaptados a conjuntos de tasas de transmisión diferentes y porque:
- 10 - para la detección (401,406) de una conmutación de tasa de transmisión entre una trama corriente a una tasa de transmisión comprendido en un primer conjunto de tasas de transmisión y una trama precedente a una tasa de transmisión comprendido en un segundo conjunto de tasas de transmisión, siendo las tasas de transmisión del primer conjunto superiores a los del segundo conjunto, se realiza un fundido encadenado (407) por ponderación, disminuyendo el peso de la segunda señal, post-tratada o no, según el post-tratamiento adaptado al segundo conjunto de tasas de transmisión y aumentando el peso de la primera señal, post-tratada o no, según el post-tratamiento adaptado al primer conjunto de tasas de transmisión, para obtener una señal de salida (S3) y
- 15 - para la detección (401,402) de una conmutación de tasa de transmisión entre una trama corriente a una tasa de transmisión comprendido en un segundo conjunto de tasas de transmisión y una trama precedente a una tasa de transmisión comprendido en un primer conjunto de tasas de transmisión, se realiza un fundido encadenado (404) por ponderación, disminuyendo el peso de la primera señal, post-tratada o no, según el post-tratamiento adaptado al primer conjunto de tasas de transmisión y aumentando el peso de la segunda señal, post-tratada o no, según el post-tratamiento adaptado al segundo conjunto de tasas de transmisión, para obtener una señal de salida (S3).
- 20 2.- Método según la reivindicación 1, caracterizado porque uno de los post-tratamientos es un filtro de paso alto(204).
- 25 3.- Método según la reivindicación 1, caracterizado porque un de los post-tratamientos es un post-filtrado adaptativo (203).
- 4.- Método según la reivindicación 1, caracterizado porque uno de los post-tratamientos es una combinación de un filtro de paso alto y de un post-filtro adaptativo.
- 5.- Método según la reivindicación 1, caracterizado porque una sola señal a la entrada del módulo de fundido encadenado es post-tratada.
- 30 6.- Método según la reivindicación 1, caracterizado porque las dos señales a la entrada del módulo de fundido encadenado son post-tratadas con post-tratamientos diferentes adaptados a conjuntos de tasas de transmisión diferentes.
- 35 7.- Programa de ordenador que comprende instrucciones de código para la puesta en práctica del método según una cualquiera de las reivindicaciones 1 a 6 cuando dicho programa se ejecute por un ordenador.
- 8.- Aplicación del método según una cualquiera de las reivindicaciones 1 a 6 a un sistema de decodificación de audio escalable en tasas de transmisión.
- 40 9.- Aplicación del método según una cualquiera de las reivindicaciones 1 a 6 a un sistema de decodificación de audio escalable en tasas de transmisión y ancho de banda en donde una primera tasa de transmisión se obtiene por al menos una primera capa de decodificación en una primera banda de frecuencias, y una segunda tasa de transmisión se obtiene por una segunda capa de decodificación, denominada capa de extensión de la denominada primera banda de frecuencias en una segunda banda de frecuencias.
- 45 10.- Decodificador de audio multi-tasas de transmisión, caracterizado porque comprende un módulo de fundido encadenado (205) que recibe, a la entrada, una primera señal (S1) y una segunda señal(S2) obtenidas a partir de una señal decodificada habiendo sufrido, al menos una de las dos señales, un post-tratamiento (203,204) entre un conjunto de post-tratamientos adaptados a conjuntos de tasas de transmisión diferentes, siendo el módulo de fundido encadenado adecuado:
- 50 - para la detección (401,406) de una conmutación de tasa de transmisión entre una trama corriente a una tasa de transmisión, comprendida en un primer conjunto de tasas de transmisión y una trama precedente a una tasa de transmisión, comprendida en un segundo conjunto de tasas de transmisión, se realiza un fundido encadenado (407) por ponderación, disminuyendo el peso de la segunda señal, post-tratada o no, según el post-tratamiento adaptado al segundo conjunto de tasas de transmisión y aumentando el peso de la primera señal, post-tratada o no, según el post-tratamiento

adaptado al primer conjunto de tasas de transmisión, para obtener la señal de salida (S3) del módulo de fundido encadenado y

5 - para la detección (401,402) de una conmutación de tasa de transmisión entre una trama corriente a una tasa de transmisión, comprendida en un segundo conjunto de tasas de transmisión, y una trama precedente a una tasa de transmisión comprendido en un primer conjunto de tasas de transmisión, se realiza un fundido encadenado (404) por ponderación, disminuyendo el peso de la primera señal, post-tratada o no, según el post-tratamiento adaptado al primer conjunto de tasas de transmisión y aumentando el peso de la segunda señal, post-tratada o no, según el post-tratamiento adaptado al
10 segundo conjunto de tasas de transmisión, para obtener la señal de salida (S3) del módulo de fundido encadenado.

11.- Decodificador según la reivindicación 10, caracterizado porque al menos uno de los post-tratamientos es un filtro de paso alto.

12.- Decodificador según la reivindicación 10, caracterizado porque al menos uno de los post-tratamientos es un post-filtrado adaptativo.

15 13.- Decodificador según la reivindicación 10, caracterizado porque al menos uno de los post-tratamientos es una combinación de un filtro de paso alto y de un post-filtrado adaptativo.

14.- Decodificador según la reivindicación 10, caracterizado porque una sola señal a la entrada del módulo de fundido encadenado es post-tratada.

20 15.- Decodificador según la reivindicación 10, caracterizado porque las dos señales a la entrada del módulo de fundido encadenado son post-tratadas con post-tratamientos diferentes adaptados a conjuntos de tasas de transmisión diferentes.

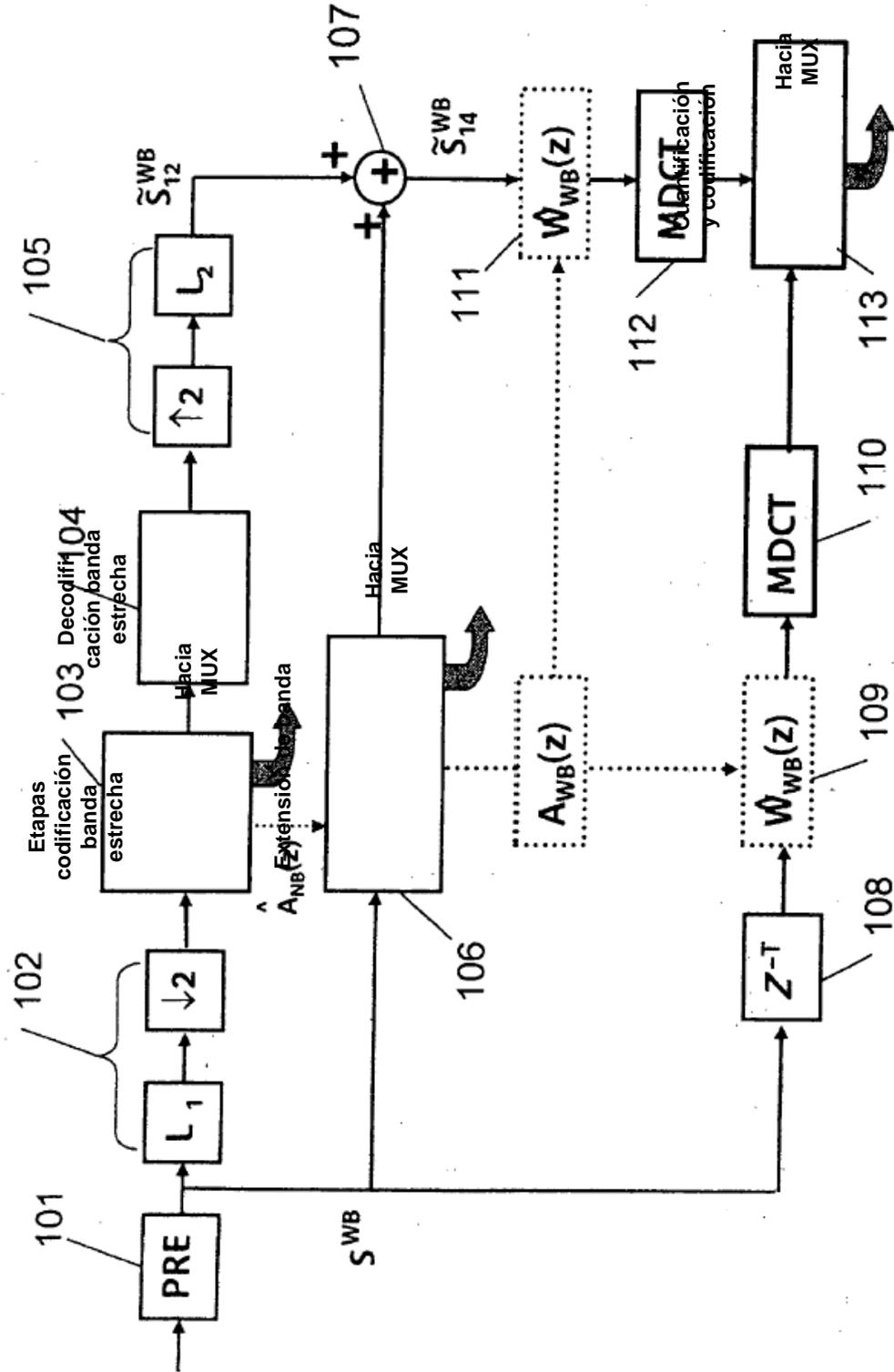


Figura 1

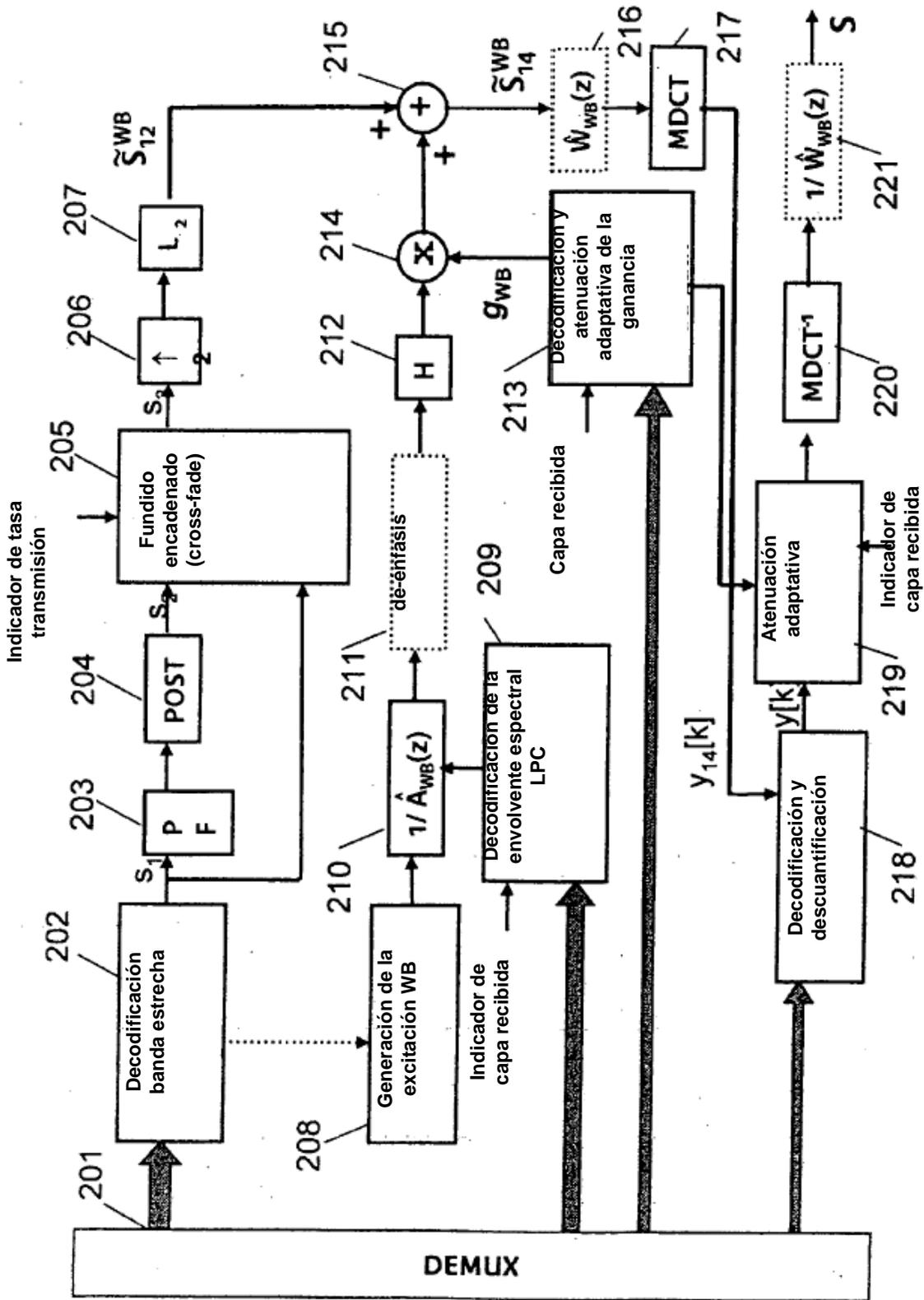


Figura 2

<p>CELP banda estrecha</p>	<p>Extensión de banda</p>	<p>Envolvente espectral (3500 - 7000 Hz)</p>	<p>Envolvente espectral (0-3500 Hz)</p>	<p>Estructura fina (por orden de importancia perceptual)</p>
-----------------------------------	----------------------------------	---	--	---

Figura 3

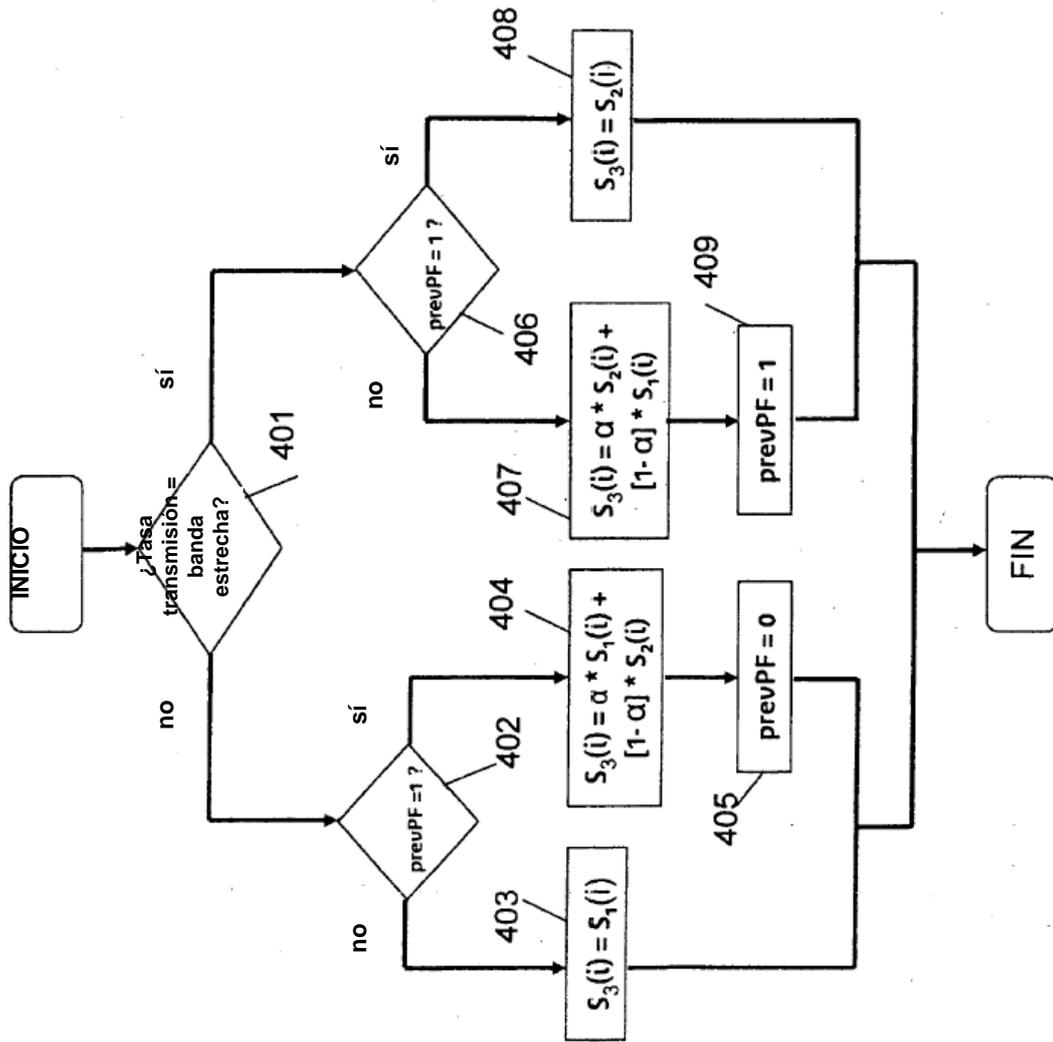


Figura 4

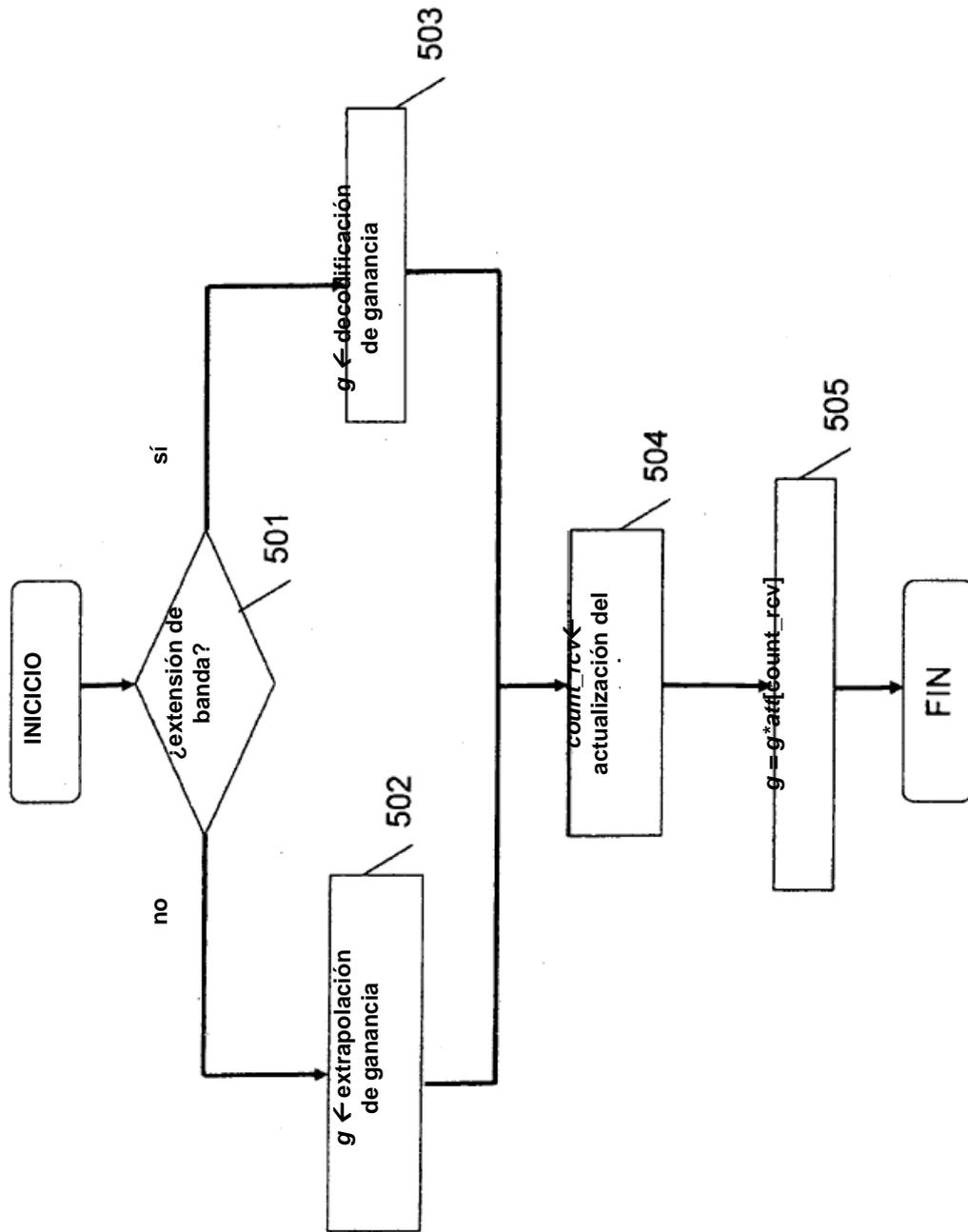


Figura 5

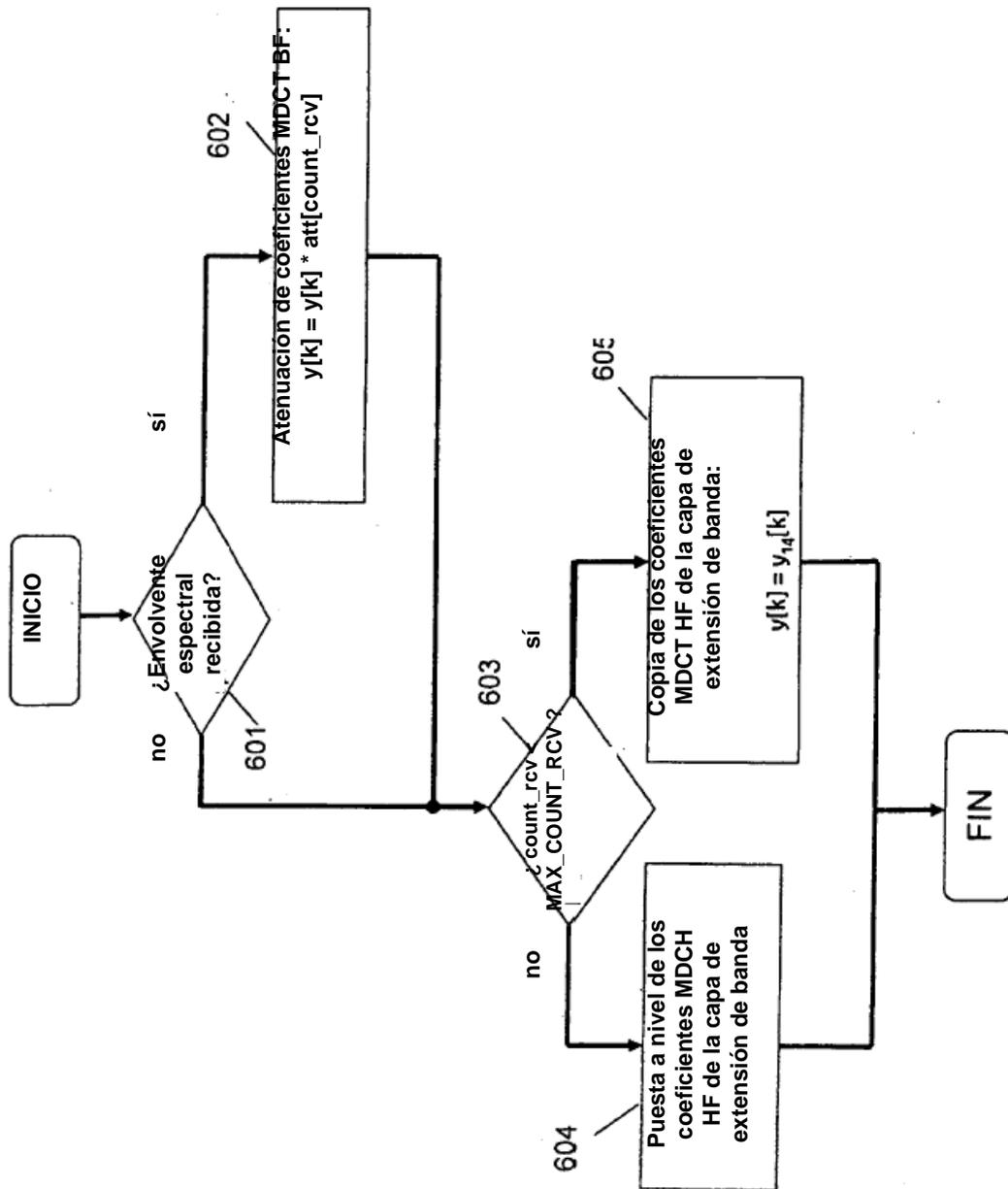


Figura 6

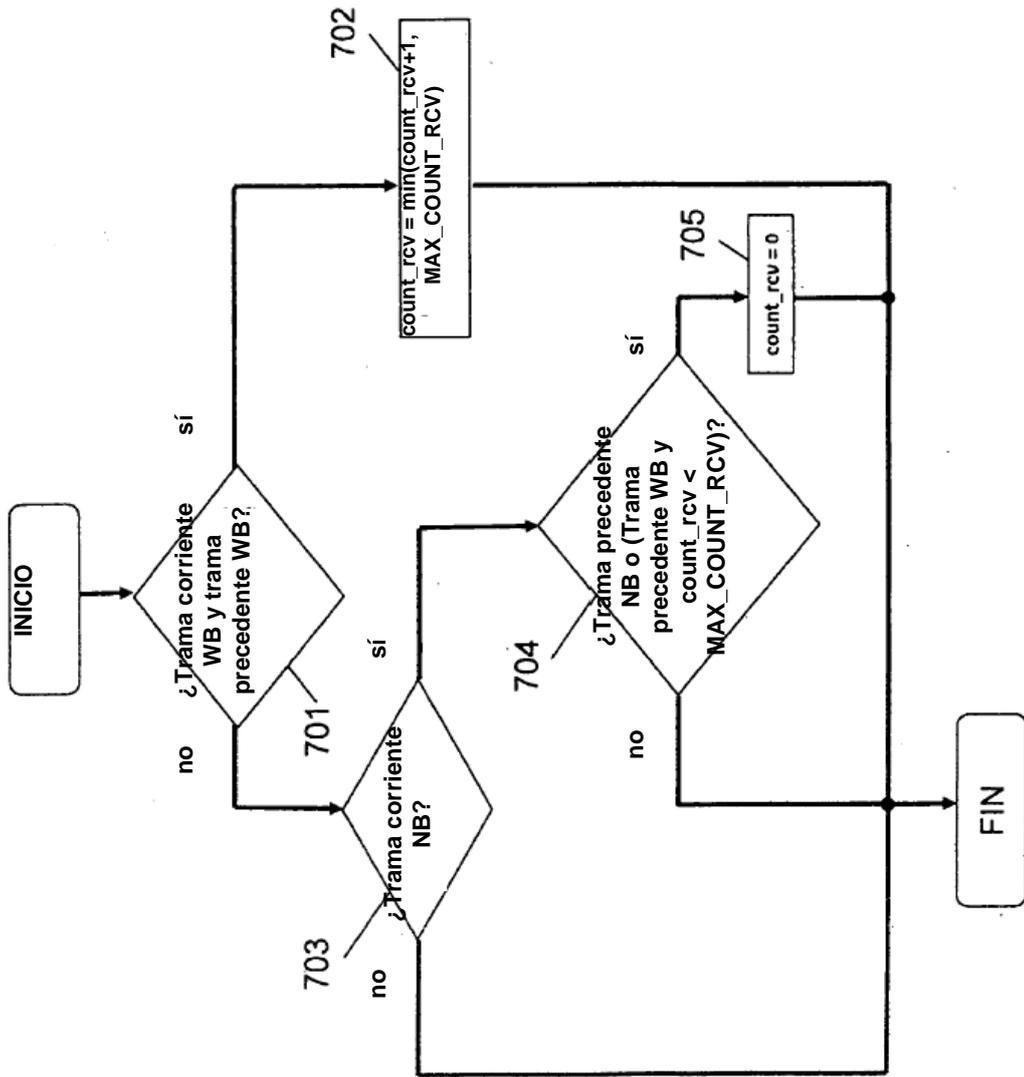


Figura 7

Trama corriente en banda ancha	Trama precedente en banda ancha	Comportamiento del decodificador
si	sí	Contador = $\max(\text{compteur}+1, \text{MAX_COUNT_RCV})$
Si	No	
no	sí	Si $\text{contador} < \text{MAX_COUNT_RCV}$ entonces contador = 0, si no, extensión de banda sin transmisión de información
no	no	Contador=0

Figura 8

n	Att(n)	n	Att(n)	n	Att(n)	n	Att(n)
0	0	25	0	50	0	75	0,01
1	0	26	0	51	0	76	0,01
2	0	27	0	52	0	77	0,01
3	0	28	0	53	0	78	0,01
4	0	29	0	54	0	79	0,01
5	0	30	0	55	0	80	0,1
6	0	31	0	56	0	81	0,1
7	0	32	0	57	0	82	0,1
8	0	33	0	58	0	83	0,1
9	0	34	0	59	0	84	0,1
10	0	35	0	60	0	85	0,31
11	0	36	0	61	0	86	0,31
12	0	37	0	62	0	87	0,31
13	0	38	0	63	0	88	0,31
14	0	39	0	64	0	89	0,31
15	0	40	0	65	0,0001	90	0,5
16	0	41	0	66	0,0001	91	0,5
17	0	42	0	67	0,0001	92	0,5
18	0	43	0	68	0,0001	93	0,5
19	0	44	0	69	0,0001	94	0,5
20	0	45	0	70	0,001	95	0,707
21	0	46	0	71	0,001	96	0,707
22	0	47	0	72	0,001	97	0,707
23	0	48	0	73	0,001	98	0,707
24	0	49	0	74	0,001	99	0,707
						100	1,0

Figura 9