



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 333 020**

51 Int. Cl.:
G10L 19/14 (2006.01)
G10L 19/12 (2006.01)
G10L 19/00 (2006.01)
G01L 19/02 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **04805538 .8**
96 Fecha de presentación : **24.11.2004**
97 Número de publicación de la solicitud: **1692689**
97 Fecha de publicación de la solicitud: **23.08.2006**

54 Título: **Procedimiento de codificación múltiple optimizado.**

30 Prioridad: **10.12.2003 FR 03 14490**

45 Fecha de publicación de la mención BOPI:
16.02.2010

45 Fecha de la publicación del folleto de la patente:
16.02.2010

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

72 Inventor/es: **Virette, David;**
Lamblin, Claude y
Benjelloun Touimi, Abdellatif

74 Agente: **Justo Bailey, Mario de**

ES 2 333 020 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

Procedimiento de codificación múltiple optimizado.

5 La presente invención se refiere a la codificación/decodificación de señales digitales, en aplicaciones de transmisión o de almacenamiento de señales multimedia tales como las señales de audio (habla y/o sonidos) o vídeo.

10 Para ofrecer movilidad y continuidad, los servicios de comunicación multimedia modernos e innovadores deben poder funcionar en una gran variedad de condiciones. El dinamismo del sector de la comunicación multimedia, la heterogeneidad de las redes, del acceso y de los terminales han generado una proliferación de formatos de compresión.

15 La presente invención se inscribe en el contexto de una optimización de las técnicas de “codificación múltiple”, puestas en práctica cuando una señal digital, o una parte de esta señal, se codifica según diversas técnicas de codificación. Esta codificación múltiple puede efectuarse de manera simultánea (en una sola pasada) o no. Los tratamientos pueden realizarse sobre la misma señal, o eventualmente sobre versiones derivadas de la misma señal (por ejemplo, según anchos de banda diferentes). Se distingue por tanto la “codificación múltiple” de las “transcodificaciones”, en las que cada codificador realiza la compresión de una versión procedente de la decodificación de la señal comprimida por el codificador anterior.

20 La codificación múltiple se presenta por ejemplo en el caso de un mismo contenido que está codificado según varios formatos y se transmite a continuación a terminales que no soportan los mismos formatos de codificación. Si se trata de una difusión en tiempo real, el tratamiento deberá realizarse de manera simultánea. Si se trata de un acceso a una base de datos, las codificaciones podrán realizarse una tras otra, en diferido. En estos ejemplos, la codificación múltiple permite codificar una misma señal según formatos diferentes utilizando varios codificadores (o eventualmente varias velocidades o varios modos de un mismo codificador), funcionando cada codificador de manera independiente con respecto a los demás codificadores.

30 Otro uso de codificación múltiple se encuentra en estructuras de codificación en las que varios codificadores se encuentran en competición para codificar un segmento de señal, seleccionándose finalmente un único codificador para la codificación de este segmento. La elección del codificador seleccionado puede realizarse como resultado de un tratamiento de este segmento, o incluso con posterioridad (por decisión retardada). En lo sucesivo, se designará como “codificación multimodo” este tipo de estructura (en referencia a la selección de un “modo” de codificación). En estas estructuras multimodo, se hace que varios codificadores que comparten una “pasada común” codifiquen la misma parte de señal. Las técnicas de codificación utilizadas pueden ser diferentes, o proceder de una única estructura de codificación. No obstante, no serán totalmente independientes salvo que se trate de técnicas “sin memoria”. En efecto, en el caso (actual) de técnicas de codificación que ponen en práctica tratamientos recursivos, el tratamiento de un segmento dado de señal depende de la manera en la que se codificó esta señal en el pasado. Por tanto hay una cierta dependencia entre los codificadores, puesto que un codificador deberá tener en cuenta en sus memorias la salida de otro codificador.

40 En estos diferentes contextos se ha introducido la noción de “codificación múltiple” así como las condiciones de uso de tales técnicas. No obstante, la complejidad de puesta en práctica puede resultar insalvable.

45 Por ejemplo, en el caso de los servidores de contenidos que difunden un mismo contenido en varios formatos adaptados a las condiciones de acceso, de redes y terminales de diferentes clientes, esta operación se vuelve extremadamente compleja a medida que aumenta el número de formatos deseado. Si se trata de una difusión en tiempo real, se está limitado rápidamente por los recursos del sistema suponiendo que los diferentes formatos están codificados en paralelo.

50 El segundo caso de uso mencionado se refiere a las aplicaciones de codificación multimodo, que permiten la selección de un codificador de entre un conjunto para cada parte de señal analizada. La selección requiere la definición de un criterio, pretendiendo los más habituales la optimización del equilibrio velocidad-distorsión. Al analizarse la señal en segmentos temporales sucesivos, en cada segmento se evalúan varias codificaciones. Se selecciona a continuación la codificación de velocidad más baja para una calidad dada, o la de la mejor calidad para una velocidad dada. Se observará que pueden utilizarse otras restricciones aparte de la de velocidad/distorsión.

60 En general, en tales estructuras, la selección de la codificación se realiza “a priori” mediante un análisis de la señal en el segmento considerado (la selección según las características de la señal). No obstante, la dificultad de producir una clasificación robusta de la señal para esta selección ha llevado a proponer una selección “a posteriori” del modo óptimo después de la codificación del conjunto de los modos, a costa sin embargo de una gran complejidad.

65 Se han propuesto métodos intermedios que combinan los dos planteamientos para aliviar el coste de cálculo. No obstante, estas estrategias no son óptimas y resultan menos eficaces que la exploración de todos los modos. La exploración de todos los modos o de una gran parte de los modos constituye una aplicación de codificación múltiple que presenta una complejidad potencialmente alta, difícilmente compatible a priori con la codificación en tiempo real, por ejemplo.

ES 2 333 020 T3

Actualmente, la mayor parte de las operaciones de codificación múltiple y de transcodificación no tienen en cuenta las interacciones entre los formatos y entre el formato y su contenido. Se han propuesto algunas técnicas de codificación multimodo, pero la decisión del modo utilizado se hace generalmente *a priori* o bien sobre la señal (mediante clasificación, como por ejemplo el codificador SMV por “*Selectable Modo Vocoder*”, *vocoder* de modo seleccionable),
5 o bien en función de las condiciones de la red (por ejemplo, en los codificadores AMR por “*Adaptive Multi-Rate*”, tasa múltiple adaptativa).

En los documentos:

10 “An overview of variable rate speech coding for cellular networks”, **Gersho, A.; Paksoy, E.**; Wireless Communications, 1992. Conference Proceedings, 1992 IEEE International Conference on Selected Topics, 25-26 de junio de 1992, página(s): 172-175,

15 “A variable rate speech coding algorithm for cellular networks”, **Paksoy, E.; Gersho, A.**; Speech Coding for Telecommunications, 1993. Proceedings, IEEE Workshop 1993, página(s): 109-110,

20 “Variable rate speech coding for multiple access wireless networks”, **Paksoy E.; Gersho A.**; Electrotechnical Conference, 1994, Proceedings, 7th Mediterranean, 12-14 de abril de 1994, página(s): 47-50 vol.1,

se presentan varios modos de selección, en particular una decisión controlada por la fuente y una decisión controlada por la red.

25 En el caso de una decisión controlada por la fuente, la decisión *a priori* se realiza a partir de una clasificación de la señal de entrada. Existen por tanto numerosos métodos de clasificación de la señal.

30 En el caso de una decisión controlada por la red, es más sencillo realizar un codificador multimodo cuya velocidad se selecciona mediante un módulo externo en vez de por la fuente. El método más sencillo consiste en elaborar una familia de codificadores cada uno a velocidad fija, pero cuyas velocidades son diferentes entre codificadores y conmutar entre estas diferentes velocidades para obtener un modo actual deseado.

35 También se han presentado algunos trabajos sobre la posibilidad de combinar varios criterios para seleccionar *a priori* el modo que debe utilizarse, especialmente en los documentos:

“Variable-rate for the basic speech service in UMTS”, **Berruto, E.; Sereno, D.**; Vehicular Technology Conference, 1993 IEEE 43rd, 18-20 de mayo de 1993, página(s): 520-523,

40 “A VR-CELP codec implementation for CDMA mobile communications”, **Cellario, L.; Sereno, D.; Giani, M.; Blocher, P.; Hellwig, K.**; Acoustics, Speech, and Signal Processing, 1994, ICASSP-94, 1994 IEEE International Conference, Volumen: 1, 19-22 de abril de 1994, página(s): I/281-I/284 vol.1.

45 Todos los algoritmos de codificación multimodo con selección del modo de codificación *a priori* padecen un mismo inconveniente, en particular asociado a problemas de robustez de la clasificación *a priori*.

50 Por este motivo se han propuesto técnicas que utilizan una decisión *a posteriori* del modo de codificación. Por ejemplo, en el documento:

“Finite state CELP for variable rate speech coding”, **Vaseghi, S.V.**; Acoustics, Speech, and Signal Processing, 1990, ICASSP-90, 1990 International Conference, 3-6 de abril de 1990, página(s): 37-40 vol.1,

55 el codificador puede conmutar entre diferentes modos mediante optimización de una medida de calidad objetiva, la decisión se toma por tanto *a posteriori* en función de las características de la señal de entrada, de la relación prevista velocidad/SQNR (por “*Signal to Quantization Noise Ratio*”, relación señal a ruido de cuantificación) y del estado actual del codificador. Un esquema de codificación de este tipo permite obtener una mejora de la calidad. No obstante, al realizarse las diferentes codificaciones en paralelo, la complejidad resultante de este tipo de sistema es prohibitiva.

60 Se han propuesto otras técnicas que combinan una decisión *a priori* y una mejora en bucle cerrado. En el documento:

65 “Multimode variable bit rate speech coding: an efficient paradigm for high-quality low-rate representation of speech signal”, **Das, A.; DeJaco, A.; Manjunath, S.; Ananthapadmanabhan, A.; Huang, J.; Choy, E.**; Acoustics, Speech, and Signal Processing, 1999, ICASSP '99 Proceedings, 1999 IEEE International Conference, Volumen: 4, 15-19 de marzo de 1999, página(s): 2307-2310 vol.4,

ES 2 333 020 T3

el sistema propuesto realiza una primera selección (selección en bucle abierto) del modo, en función de las características de la señal. Esta decisión puede realizarse mediante clasificación. A continuación, a partir de una medida de error, si los rendimientos del modo seleccionado no son satisfactorios, se aplica un modo de velocidad más alta y se repite la operación (según una decisión buscada en bucle cerrado).

5

Igualmente, en los documentos:

* “Variable rate speech coding for umts”, **Cellario, L.; Sereno, D.**; Speech Coding for Telecommunications, 1993.
10 Proceedings, IEEE Workshop, 1993, página(s): 1-2,

“Phonetically-based vector excitation coding of speech at 3.6 kbps”, **Wang, S.; Gersho, A.**; Acoustics, Speech, and Signal Processing, 1989. ICASSP-89., 1989 International Conference, 23-26 de mayo de 1989, página(s): 49-52
15 vol.1

** “A modified CS-ACELP algorithm for variable-rate speech coding robust in noisy environments”, **Beritelli, F.**;
IEEE Signal Processing Letters, Volumen: 6 Edición: 2, febrero de 1999, página(s): 31-34,

se han utilizado técnicas similares. Una primera selección en bucle abierto se realiza después de la clasificación de
20 la señal de entrada (clasificación fonética, o sonora/no sonora), a continuación se toma una decisión en bucle cerrado:

- o bien en el codificador completo y, en este caso, se vuelve a codificar todo el segmento del habla;

- o bien sobre una parte de la codificación, como en las referencias anteriores precedidas de un asterisco (*), para
25 las que la elección del diccionario que debe utilizarse se realiza en bucle cerrado.

El conjunto de los estudios mencionados anteriormente tiende a solucionar el problema de la complejidad de la
30 selección óptima del modo mediante la utilización, total o parcial, de una selección o preselección *a priori*, que evita la codificación múltiple o disminuye el número de codificadores que deben ponerse en práctica en paralelo.

El documento US 6 581 032 da a conocer un sistema de compresión del habla que comprende cuatro códecs
activados de manera selectiva según una señal que debe comprimirse y según una clasificación de las señales.

35 El documento US 6 141 638 da a conocer un codificador que utiliza diferentes diccionarios de código según parámetros de la señal que va a codificarse.

Sin embargo, no se ha propuesto ninguna técnica en la técnica anterior que permita reducir la complejidad de las
40 codificaciones realizadas en paralelo.

La presente invención va a mejorar la situación.

Propone para ello un procedimiento de codificación múltiple en compresión tal como se define en la reivindicación
1.
45

En una realización ventajosa, las etapas anteriores se ponen en práctica mediante un producto de programa de
ordenador que comprende instrucciones de programa para ello. A este respecto, la presente invención se refiere también
a un producto de programa de ordenador de este tipo, destinado a almacenarse en una memoria de una unidad de
tratamiento, especialmente de un ordenador o de un terminal móvil, o en un soporte de memoria amovible y destinado
50 a actuar conjuntamente con un lector de la unidad de tratamiento.

La presente invención se refiere también a un dispositivo auxiliar para una codificación en compresión, para la
puesta en práctica del procedimiento según la invención, y que comprende así una memoria adecuada para almacenar
instrucciones de un producto de programa de ordenador del tipo anteriormente mencionado.
55

Otras características y ventajas de la invención se desprenderán del examen de la descripción detallada a continua-
ción, y de los dibujos adjuntos en los que:

60 - la figura 1a ilustra esquemáticamente el contexto de aplicación de la presente invención, con una pluralidad de codificadores puestos en paralelo,

- la figura 1b ilustra esquemáticamente la aplicación de la invención, compartiendo bloques funcionales entre varios
codificadores puestos en paralelo,

65 - la figura 1c ilustra esquemáticamente la aplicación de la invención, compartiendo bloques funcionales en codifi-
cación multimodo,

ES 2 333 020 T3

- la figura 1d ilustra esquemáticamente la aplicación de la invención, en codificación multimodo en cuadrícula,

- la figura 2 representa esquemáticamente los bloques funcionales principales de un codificador de frecuencia perceptivo,

- la figura 3 representa esquemáticamente los bloques funcionales principales de un codificador de análisis por síntesis.

- la figura 4a representa esquemáticamente los bloques funcionales principales de un codificador TDAC,

- la figura 4b representa esquemáticamente el formato del flujo binario codificado por el codificador de la figura 4a,

- la figura 5 representa esquemáticamente la aplicación de la invención a una pluralidad de codificadores TDAC en paralelo, según una realización ventajosa,

- la figura 6a representa esquemáticamente los bloques funcionales principales de un codificador MPEG-1 (capa I y II),

- la figura 6b representa esquemáticamente el formato del flujo binario codificado por el codificador de la figura 6a,

- la figura 7 representa esquemáticamente la aplicación de la invención a una pluralidad de codificadores MPEG-1 (capa I y II) puestos en paralelo, según una realización ventajosa, y

- la figura 8 representa más en detalle los bloques funcionales de un codificador de análisis par síntesis, en este caso de tipo NBAMR según la norma 3GPP.

Se hace referencia en primer lugar a la figura 1a en la que se ha representado una pluralidad de codificadores C0, C1,..., CN, en paralelo y que reciben cada uno una señal s_0 de entrada. Cada codificador comprende bloques BF1 a BF n funcionales para poner en práctica etapas sucesivas de codificación y suministrar finalmente un flujo BS0, BS1,..., BSN binario codificado. Se indica además que en una aplicación en codificación multimodo, las salidas de los codificadores C0 a CN están conectadas a un módulo MM de selección del modo óptimo y se transmite el flujo Binario BS del codificador óptimo (flechas en trazos discontinuos de la figura 1a).

Por motivos de simplicidad, todos los codificadores del ejemplo de la figura 1a tienen el mismo número de bloques funcionales, aunque por supuesto todos estos bloques funcionales no están previstos necesariamente en todos los codificadores, en la práctica.

Ciertos bloques BF i funcionales son a veces idénticos entre un modo (o un codificador) y otro, otros sólo difieren en cuanto a los cuantificadores. También existen relaciones aprovechables cuando se utilizan codificadores procedentes de una misma familia de codificación, que utilizan modelos similares o que calculan parámetros asociados físicamente a la señal.

Son estas relaciones las que la presente invención propone aprovechar, con el fin de reducir la complejidad de las operaciones de codificación múltiple.

En un primer momento, la invención propone identificar los bloques funcionales que componen cada uno de los codificadores. Se aprovechan a continuación las similitudes técnicas entre los codificadores considerando los bloques funcionales cuyas funciones son equivalentes o próximas. Para cada uno de estos bloques, la invención propone:

- por un lado, definir operaciones denominadas “*comunes*”, y realizarlas una sola vez para el conjunto de los codificadores;

- por otro lado, poner en práctica métodos de cálculo específicos de cada codificador y utilizando especialmente los resultados de estos cálculos comunes. Estos métodos de cálculo dan un resultado eventualmente diferente al producido por una codificación completa. El objetivo es, por tanto, de hecho, acelerar los tratamientos aprovechando las informaciones disponibles y proporcionadas especialmente por los cálculos comunes. Tales procedimientos que permiten acelerar los cálculos se ponen en práctica, por ejemplo, en técnicas destinadas a reducir la complejidad de las operaciones de transcodificación (denominadas técnicas de “*transcodificación inteligente*”).

La figura 1b ilustra la solución propuesta. En el ejemplo representado, las operaciones “*comunes*” anteriormente mencionadas se realizan una sola vez para una parte al menos de los codificadores y, preferiblemente, para el conjunto de los codificadores, en un módulo MI independiente que redistribuirá los resultados obtenidos a una parte al menos de los codificadores, o preferiblemente a todos estos codificadores. Se trata por tanto de compartir entre una parte al

ES 2 333 020 T3

menos de todos los codificadores C0 a CN(o “mutualización” en lo sucesivo) los resultados obtenidos. Un módulo MI independiente de este tipo puede formar parte de un dispositivo auxiliar para una codificación múltiple en compresión tal como se definió anteriormente.

5 En una variante ventajosa, en lugar de recurrir a un módulo MI de cálculo externo, se utiliza el o los bloques BF1 a BF_n funcionales existentes de un mismo o de varios codificadores distintos, eligiéndose este o estos codificadores según criterios que se describirán más adelante.

10 La presente invención puede poner en práctica varias estrategias que, por supuesto, pueden diferir según el papel del bloque funcional considerado.

Una primera estrategia consiste en utilizar los parámetros del codificador cuya velocidad es más baja para focalizar la búsqueda de los parámetros para todos los demás modos.

15 A la inversa, una segunda estrategia consiste en utilizar los parámetros del codificador cuya velocidad es más alta, y después “degradar” progresivamente hasta el codificador cuya velocidad es la más baja.

20 Por supuesto, si se desea dar prioridad a un codificador particular, es posible codificar un segmento de señal utilizando este codificador, y después, aplicando las dos estrategias anteriores, alcanzar los codificadores de velocidad superior e inferior.

25 Por supuesto, pueden utilizarse otros criterios aparte de la velocidad para controlar la búsqueda. Por ejemplo, para ciertos bloques funcionales, puede favorecerse el codificador cuyos parámetros se prestan mejor a una extracción (o a un análisis) y/o a una codificación eficaces de los parámetros similares de los demás codificadores, pudiendo considerarse la eficacia según la complejidad, la calidad o un equilibrio de ambas.

30 También puede preverse crear un módulo de codificación independiente, no presente en los codificadores, pero que permite una codificación más eficaz de los parámetros del bloque funcional considerado, para el conjunto de los codificadores.

35 Estas diversas estrategias de puesta en práctica son particularmente interesantes en el caso de la codificación multimodo. En este contexto ilustrado en la figura 1c, la presente invención permite reducir la complejidad de los cálculos preliminares para la selección *a posteriori* de un codificador realizado en la última etapa, por ejemplo, por el último módulo MM antes de la transmisión del flujo Binario BS.

40 En este caso particular de la codificación multimodo, una variante de la presente invención, representada en el ejemplo de la figura 1c, propone introducir un módulo MSP_i de selección parcial (con $i = 1, 2, \dots, N$) en cada etapa de codificación (es decir, después de los bloques funcionales BF_{i1} a BF_{iN1} en competición y cuyo resultado del o de los bloques BF_{icc} seleccionados va a utilizarse a continuación). Así, las similitudes entre los diferentes modos se aprovechan para acelerar el cálculo de cada bloque funcional. No se evaluarán por tanto obligatoriamente todos los esquemas de codificación.

45 Se propone ahora una variante más sofisticada de la estructura multimodo que se basa en el corte en bloques funcionales descrita anteriormente, en referencia a la figura 1d. La estructura multimodo de la figura 1d se denomina “en cuadrícula”, con varios caminos posibles en la cuadrícula. En efecto, en la figura 1d, se han representado todos los caminos posibles de la cuadrícula de modo que se presenta en una forma arborescente. Se indica en particular que cada camino de la cuadrícula está definido por una combinación de modos de funcionamiento de los bloques funcionales, alimentando cada bloque funcional varias variantes posibles del bloque funcional siguiente.

50 Así, cada modo de codificación procede de la combinación de modos de funcionamiento de los bloques funcionales: el bloque funcional 1 tiene N_1 modos de funcionamiento, el bloque funcional 2 tiene N_2 , y así sucesivamente hasta el bloque P. El conjunto de las $NN = N_1 \times N_2 \times \dots \times N_p$ combinaciones posibles está representado por tanto por una cuadrícula de NN ramas, que describen, de extremo a extremo, un codificador multimodo completo de NN modos. Eventualmente pueden suprimirse *a priori* ciertas ramas de la cuadrícula y definir así una arborescencia que comprende un número reducido de ramas. Una primera particularidad de esta estructura es que prevé, para un bloque funcional dado, un módulo de cálculos comunes mediante la salida del bloque funcional anterior. Estos módulos de cálculos comunes realizan las mismas operaciones, pero basándose en señales diferentes ya que proceden de bloques anteriores diferentes. De manera ventajosa, los módulos de cálculos comunes de un mismo nivel están mutualizados: los resultados de un módulo dado que pueden aprovecharse por los módulos siguientes se proporcionan a estos módulos siguientes. Por otro lado, una selección parcial, realizada a la salida del tratamiento de cada bloque funcional, permite ventajosamente suprimir las ramas menos eficaces según el criterio elegido. Puede por tanto reducirse el número de ramas de la cuadrícula que ha de evaluarse.

65 Una aplicación ventajosa de esta estructura multimodo en cuadrícula es la siguiente. Cuando los bloques funcionales son susceptibles de operar a velocidades respectivas diferentes y utilizando parámetros respectivos propios de dichas velocidades, para un bloque funcional dado, el camino de la cuadrícula elegido es el que atraviesa el bloque funcional de velocidad más baja, o también el bloque funcional de velocidad más alta según el contexto de codificación, y los resultados obtenidos del bloque funcional de velocidad más baja (o más alta) se adaptan a las velocidades de

ES 2 333 020 T3

una parte al menos de los demás bloques funcionales mediante una búsqueda focalizada de parámetros para una parte al menos de todos los demás bloques funcionales, hasta el bloque funcional de velocidad más alta (o respectivamente más baja).

5

En una variante, se elige un bloque funcional de velocidad dada y se adapta progresivamente al menos una parte de los parámetros propios de este bloque funcional:

10

- hasta el bloque funcional que puede operar a la velocidad más baja, mediante búsqueda focalizada, y
- hasta el bloque funcional que puede operar a la velocidad más alta, mediante búsqueda focalizada.

15

De forma general, se reduce así la complejidad asociada a la codificación múltiple.

La invención se aplica a cualquier esquema de compresión que ponga en práctica la codificación múltiple de un contenido multimedia. Tres ejemplos de realización se presentan a continuación, en el campo de la compresión de audio (habla y sonido). Los dos primeros ejemplos de realización se sitúan en el contexto de la familia de los codificadores por transformada, para el que puede darse el siguiente documento a modo de referencia:

20

“Perceptual Coding of Digital Audio”, **Painter, T.; Spanias, A.**; Proceedings of the IEEE, Vol. 88, N.º 4, abril de 2000.

25

El tercer ejemplo de realización se sitúa en el contexto de los codificadores CELP, para el que puede darse el documento siguiente a modo de referencia:

30

“Code Excited Linear Prediction (CELP): High quality speech at very low bit rates” **Schroeder M.R.; Atal B.S.**; Acoustics, Speech, and Signal Processing, 1985. Proceedings. 1985 IEEE International Conference, página(s): 937-940.

Un recuerdo de las principales características de estas dos familias de codificación se presenta en primer lugar a continuación.

35

* Los codificadores por transformada o en subbandas

40

Se trata de codificadores en compresión por transformada o en subbandas basados en criterios psicoacústicos. Este tipo de codificador actúa mediante transformación sobre bloques de la señal temporal para obtener un conjunto de coeficientes. Las transformaciones son del tipo tiempo-frecuencia, siendo una de las transformaciones más utilizadas la transformada en coseno discreta modificada (denominada “MDCT”, del inglés “Modified Discrete Cosine Transform”). Antes de la cuantificación de estos coeficientes, un algoritmo procede a la asignación de los bits de modo que el ruido de cuantificación sea lo menos audible posible. La asignación binaria y la cuantificación de los coeficientes ponen en práctica una curva de enmascaramiento, obtenida con ayuda de un modelo psicoacústico que permite evaluar, para cada línea espectral considerada, un umbral de enmascaramiento representativo de la amplitud necesaria para que un sonido a esta frecuencia sea audible. La figura 2 da el esquema básico de un codificador de frecuencia. Se observa que la estructura en forma de bloques funcionales está bien representada. En referencia a la figura 2, los bloques funcionales principales son:

50

- un bloque 21 de transformación tiempo/frecuencia de la señal s_0 de audio digital de entrada,
- un bloque 22 de determinación de un modelo perceptivo a partir de la señal transformada,
- un bloque 23 de cuantificación y codificación, a partir del modelo perceptivo, y
- un bloque 24 de formateo del flujo binario para obtener una trama s_{ic} de audio codificada.

55

* Los codificadores de análisis por síntesis (codificación CELP)

60

En los codificadores de análisis por síntesis, el modelo de síntesis de la señal reconstruida se utiliza en el codificador para extraer los parámetros que modelan las señales que van a codificarse. Estas señales pueden muestrearse a la frecuencia de 8 kHz (banda de telefonía 300-3400 Hz) o a una frecuencia más alta, por ejemplo, a 16 kHz para la codificación en banda ancha (ancho de banda de 50 Hz a 7 kHz). Según la aplicación y la calidad deseada, la tasa de compresión varía de 1 a 16. Estos codificadores funcionan a velocidades de 2 a 16 kbit/s en banda de telefonía, y a velocidades de 6 a 32 kbit/s en banda ancha. El dispositivo de codificación digital de tipo CELP, codificador de análisis por síntesis más utilizado en la actualidad, se presenta en la figura 3 en forma de bloques funcionales principales. La señal s_0 de habla se muestrea y se convierte en una sucesión de tramas de un número L de muestras. Cada trama se sintetiza filtrando una forma de onda extraída de un repertorio (llamado “diccionario”), multiplicada por una ganancia, a través de dos filtros que varían en el tiempo. El diccionario de excitación fija es un conjunto finito de formas de onda de las L muestras. El primer filtro es un filtro de predicción a largo plazo. Un análisis “LTP” (por

ES 2 333 020 T3

“*Long Term Prediction*”, predicción a largo plazo) permite evaluar los parámetros de este predictor a largo plazo que aprovecha la periodicidad de los sonidos sonoros, modelándose esta componente armónica en forma de un diccionario adaptativo (bloque 32). El segundo filtro es un filtro de predicción a corto plazo. Los métodos de análisis “*LPC*” (por “*Linear Prediction Coding*”, codificación de predicción lineal) permiten obtener estos parámetros de predicción a corto plazo, representativos de la función de transferencia del conducto vocal y características de la envolvente del espectro de la señal. El procedimiento utilizado para determinar la secuencia de innovación es el método de análisis por síntesis que se resume como sigue. En el codificador, se filtra un gran número de secuencias de innovación del diccionario de excitación fijas por el filtro LPC (filtro de síntesis del bloque 34 funcional de la figura 3). Previamente, la excitación adaptativa se ha obtenido de manera similar. La forma de onda seleccionada es la que produce la señal sintética más próxima a la señal original (minimización del error al nivel del bloque 35 funcional), según un criterio de ponderación perceptiva (bloque 36 funcional) que se conoce en general por el nombre de criterio “*CELP*”.

En el esquema básico del codificador CELP dado en la figura 3, la extracción de la frecuencia fundamental de los sonidos sonoros (o “*pitch*”, altura tonal), aplicada sobre la señal que resulta del análisis LPC del bloque 31, permite a continuación extraer la correlación a largo plazo al nivel del bloque 32, denominada componente armónica o excitación adaptativa (E.A.). La señal residual se modela finalmente de manera clásica mediante algunos impulsos, cuyo conjunto de las posiciones está predefinido en un repertorio, llamado repertorio de excitación fija (E.F) en el bloque 33.

La decodificación es, por su parte, mucho menos compleja que la codificación. El flujo binario generado por el codificador permite al decodificador, después de la demultiplexación, obtener el índice de cuantificación de cada parámetro. La decodificación de los parámetros y la aplicación del modelo de síntesis permiten entonces reconstruir la señal.

Se describen a continuación los tres ejemplos de realización anteriormente mencionados, empezando en primer lugar con un codificador por transformada del tipo representado en la figura 2.

* *Primer ejemplo de realización: aplicación en un codificador “TDAC”*

El primer ejemplo de realización se refiere al codificador de frecuencia perceptivo denominado “*TDAC*” y descrito especialmente en el documento publicado US-2001/027393. Este codificador TDAC se utiliza para codificar señales de audio digitales muestreadas a 16 kHz (banda ancha). La figura 4a ilustra los bloques funcionales principales de este codificador. Una señal $x(n)$ de audio limitada en banda a 7 kHz y muestreada a 16 kHz se corta en tramas de 320 muestras (20 ms). Una transformada en coseno discreta modificada (o “*MDCT*”) se aplica (bloque funcional 41) sobre las tramas de la señal de entrada de 640 muestras con una recuperación del 50%, por tanto con una actualización del análisis MDCT cada 20 ms. Se limita el espectro a 7225 Hz poniendo a cero los 31 últimos coeficientes (sólo los 289 primeros coeficientes son diferentes de 0). Se determina una curva de enmascaramiento (bloque 42) a partir de este espectro y todos los coeficientes enmascarados se ponen a cero. El espectro se divide en 32 bandas de anchuras desiguales. Las eventuales bandas enmascaradas se determinan en función de los coeficientes transformados de las señales. Para cada banda del espectro, se calcula la energía de los coeficientes MDCT (para obtener factores de escala). Los 32 factores de escala constituyen la envolvente espectral de la señal que se cuantifica a continuación y después se codifica mediante una codificación por entropía (bloque funcional 43), y finalmente se transmite en la trama codificada s_c .

La asignación dinámica de los bits (bloque 44 funcional) se basa en una curva de enmascaramiento por banda (bloque funcional 42) calculada a partir de la versión decodificada y decuantificada de la envolvente espectral. Esta medida permite tener una compatibilidad entre la asignación binaria del codificador y del decodificador. Los coeficientes MDCT normalizados en cada banda se cuantifican a continuación (bloque 45 funcional) mediante cuantificadores vectoriales que utilizan diccionarios intercalados por tamaño, estando compuestos los diccionarios por una unión de códigos de permutación de tipo II. Finalmente, en referencia a la figura 4b, las informaciones sobre la tonalidad (codificadas en este caso sobre un bit B_1) y la sonoridad (codificadas en este caso sobre un bit B_0), así como la envolvente espectral $e_q(i)$ y los coeficientes codificados $y_q(j)$ se multiplexan (bloque 46 de la figura 4a) y se transmiten en tramas.

Al poder funcionar este codificador a varias velocidades, se pretende realizar un codificador de múltiples velocidades, por ejemplo, a 16, 24 y 32 kbit/s. En este esquema de codificación, los bloques funcionales siguientes pueden ponerse en común entre los diferentes modos:

- Transformada MDCT (bloque 41),
- Detección de sonoridad (bloque 47 funcional de la figura 4a) y detección de tonalidad (bloque 48 funcional de la figura 4a),
- Cálculo, cuantificación y codificación por entropía de la envolvente espectral (bloque 43),
- Cálculo de una curva de enmascaramiento, coeficiente por coeficiente, y de una curva de enmascaramiento por banda (bloque 42).

ES 2 333 020 T3

Estos diferentes bloques constituyen un 61,5% de la complejidad del tratamiento en el proceso de codificación. Su factorización es, por tanto, de gran interés para reducir esta complejidad durante la generación de varios flujos binarios correspondientes a velocidades diferentes.

5 Los resultados de estos bloques funcionales permiten ya obtener una primera parte común a todos los flujos binarios de salida que contienen los bits de información sobre la sonoridad, la tonalidad y la envolvente espectral codificada.

En una primera variante de este ejemplo de realización, es posible realizar las operaciones de asignación de los bits y de cuantificación para cada uno de los flujos binarios de salida correspondientes a cada una de las velocidades binarias consideradas. Estas dos operaciones se realizan exactamente de la misma forma en que se realizan habitualmente en un codificador TDAC.

En una segunda variante más avanzada tal como se ilustra en la figura 5, pueden ponerse en práctica técnicas de transcodificación “*inteligente*” (como se describe en el documento publicado US-2001/027393 mencionado anteriormente) para reducir aún más la complejidad y mutualizar ciertas operaciones, especialmente:

- la asignación de bits (bloque 44 funcional),
- y también la cuantificación de los coeficientes (bloques 45_i funcionales), como se verá más adelante.

En la figura 5, los bloques funcionales compartidos entre los codificadores (o “*mutualizados*”) llevan la misma referencia que los de un solo codificador TDAC tal como se representa en la figura 4a. Se trata de los bloques 41, 42, 47, 48, 43 y 44. En particular, el bloque 44 de asignación de los bits se utiliza en varias pasadas, y el número de bits asignados se ajusta para la transcuantificación que realiza cada codificador (bloques 45₁,..., 45_(K-2), 45_(K-1)), como se verá más adelante. Se observa además que estas transcuantificaciones utilizan los resultados obtenidos por el bloque 45₀ funcional de cuantificación para un codificador seleccionado, de índice 0 (el codificador de velocidad más baja en el ejemplo descrito). Finalmente, los únicos bloques funcionales de los codificadores que actúan sin interacción real son los bloques 46₀, 46₁,..., 46_(K-2), 46_(K-1) de multiplexación, aunque utilizan todas las mismas informaciones de sonoridad y de tonalidad, así como la misma envolvente espectral codificada. A este respecto, se indica simplemente que también en este caso puede efectuarse una mutualización parcial de la multiplexación.

Para los dos bloques funcionales de asignación de bits y de cuantificación, la estrategia empleada consiste en aprovechar los resultados de los dos bloques funcionales de asignación de los bits y de cuantificación realizadas para el flujo (0) binario, a la velocidad más baja D_0 , para acelerar las operaciones de los dos bloques funcionales correspondientes para los $K-1$ otros flujos (k) binarios ($1 \leq k < K$). También puede considerarse el esquema de codificación de múltiples velocidades que utiliza un bloque funcional de asignación de bits por flujo binario (sin factorización prevista para este bloque) pero que mutualiza una parte de las operaciones de cuantificación a continuación.

Las técnicas de codificación múltiple presentadas a continuación se basan, ventajosamente, en una transcodificación inteligente utilizada para la reducción de la velocidad binaria del flujo de audio codificado, generalmente situado en un nodo de la red.

A continuación, los flujos k binarios, $0 \leq k < K$, están clasificados según un orden creciente de velocidades ($D_0 < D_1 < \dots < D_{K-1}$). Así, el flujo 0 binario corresponde a la velocidad binaria más baja.

* *Asignación de bits*

La asignación de bits en el codificador TDAC se realiza en dos fases. En primer lugar se realiza un primer cálculo del número de bits que han de asignarse a cada banda preferiblemente según la fórmula siguiente:

$$b_{opt}(i) = \frac{1}{2} \log_2 \left[\frac{e_q^2(i)}{S_b(i)} \right] + C, \quad 0 \leq i \leq M-1,$$

en la que $C = \frac{B}{M} - \frac{1}{2M} \sum_{l=0}^{M-1} \log_2 [e_q^2(l)/S_b(l)]$ es una constante,

B es el número total de bits disponibles,

M es el número de bandas,

$e_q(i)$ es el valor decodificado y decuantificado de la envolvente espectral sobre la banda i , y

$S_b(i)$ es el umbral de enmascaramiento para esta banda.

ES 2 333 020 T3

Cada uno de los valores obtenidos se redondea al entero natural más próximo. Si la velocidad total asignada no es exactamente igual a la disponible, se utiliza una segunda fase para realizar el reajuste. Esta etapa se hace preferiblemente mediante una sucesión de operaciones iterativas basadas en un criterio perceptivo que añade o retira bits en las bandas.

5 Así, si el número total de bits distribuidos es inferior al disponible, la adición de bits se hace en las bandas en las que la mejora perceptiva es más importante. Esta mejora perceptiva se mide por la variación de la relación ruido a máscara entre la asignación inicial y final de las bandas. La velocidad se aumenta para la banda en la que esta variación es más grande. En el caso contrario en el que el número total de bits distribuidos es superior al disponible, la extracción de bits sobre las bandas se hace de manera doble en este último procedimiento.

10 En el esquema de codificación de múltiples velocidades correspondiente al codificador TDAC, es posible factorizar ciertas operaciones para la asignación de bits. Así, la primera fase de determinación mediante la fórmula anterior puede hacerse una sola vez basándose en la velocidad binaria D_0 más baja. La fase de reajuste que añade bits puede hacerse a continuación de manera continua. Una vez que el número total de bits distribuido alcanza el número correspondiente a una velocidad binaria de un flujo k binario, $k=1,2,\dots,K-1$, la distribución actual se considera como aquella que se utiliza para la cuantificación de los vectores de coeficientes normalizados por banda de este flujo binario.

* Cuantificación de los coeficientes

20 Por lo que respecta a la cuantificación de los coeficientes, el codificador TDAC utiliza una cuantificación vectorial que utiliza diccionarios intercalados por tamaño, estando compuestos los diccionarios por una unión de códigos de permutación de tipo II. Este tipo de cuantificación se aplica sobre cada uno de los vectores de los coeficientes MDCT sobre una banda. Un vector de este tipo está normalizado previamente utilizando el valor decuantificado de la envolvente espectral sobre esta banda. Se observa:

- $C(b_i, d_i)$ el diccionario correspondiente al número de bits b_i y a la dimensión d_i ,
- $N(b_i, d_i)$ el número de elementos en este diccionario,
- 30 - $CL(b_i, d_i)$ el conjunto de sus *leaders* (*cabeceras de guía*), y
- $NL(b_i, d_i)$ el número de *leaders*.

35 El resultado de cuantificación para cada banda i de la trama es una palabra de código m_i transmitida en el flujo binario. Representa el índice del vector cuantificado en el diccionario y calculado a partir de las informaciones siguientes:

- 40 • el número L_i , en el conjunto $CL(b_i, d_i)$ de *leaders* del diccionario $C(b_i, d_i)$, del vector de *leader* cuantificado $\tilde{Y}_q(i)$ vecino más próximo de un *leader* actual $\tilde{Y}(i)$,
- el rango r_i de $Y_q(i)$ en la clase del *leader* $\tilde{Y}_q(i)$,
- 45 • y la combinación de signos $sign_q(i)$ que ha de aplicarse a $Y_q(i)$ (o a $\tilde{Y}_q(i)$), donde se precisan las notaciones siguientes:
- $Y(i)$ es el vector de los valores absolutos de los coeficientes normalizados de la banda i ,
- 50 • $sign(i)$ es el vector de los signos de los coeficientes normalizados de la banda i ,
- $\tilde{Y}(i)$ es el vector de *leader* del vector $Y(i)$ anteriormente mencionado, obtenido mediante ordenación decreciente de sus componentes (la permutación correspondiente se indica como $perm(i)$), y
- 55 • $Y_q(i)$ es el vector cuantificado de $Y(i)$ (o “*el vecino más próximo*” de $Y(i)$ en el diccionario $C(b_i, d_i)$).

60 En lo sucesivo, la notación $\alpha^{(k)}$, con un exponente k , indica el parámetro utilizado en el tratamiento realizado para obtener el flujo binario del codificador k . Los parámetros sin este exponente se calculan de una vez para el flujo 0 binario. Son independientes de la velocidad (o del modo) considerada.

La propiedad “*de intercalado*” de los diccionarios anteriormente mencionada se expresa según la relación:

$$65 \quad C(b_i^{(0)}, d_i) \subseteq \dots \subseteq C(b_i^{(k-1)}, d_i) \subseteq C(b_i^{(k)}, d_i) \dots \subseteq C(b_i^{(K-1)}, d_i)$$

con también:

$$CL(b_i^{(0)}, d_i) \subseteq \dots \subseteq CL(b_i^{(k-1)}, d_i) \subseteq CL(b_i^{(k)}, d_i) \dots \subseteq CL(b_i^{(K-1)}, d_i)$$

Se indica como $CL(b_i^{(k)}, d_i) \setminus CL(b_i^{(k-1)}, d_i)$ el complementario de $CL(b_i^{(k-1)}, d_i)$ en $CL(b_i^{(k)}, d_i)$.

Su cardinal es igual a $NL(b_i^{(k)}, d_i) - NL(b_i^{(k-1)}, d_i)$.

La obtención de las palabras de código $m_i^{(k)}$ (con $0 \leq k < K$), resultado de la cuantificación del vector de los coeficientes de la banda i para cada uno de los flujos k binarios, se hace como sigue.

• Para el flujo binario $k = 0$, la operación de cuantificación se hace clásicamente de la manera habitual en el codificador TDAC. Permite obtener los parámetros, $sign_q^{(0)(i)}$, $L_i^{(0)}$ y $r_i^{(0)}$ que permiten construir la palabra de código. Se determinan por otro lado en esta misma etapa los vectores $\tilde{Y}(i)$ y $sign(i)$. Se almacenan en memoria, así como la permutación correspondiente $perm(i)$, para utilizarse, dado el caso, en las etapas siguientes relativas a los otros flujos binarios.

• Para los flujos binarios $1 \leq k < K$, se procede de manera incremental, de $k = 1$ hasta $k = K-1$, utilizando preferiblemente las etapas siguientes:

Si $(b_i^{(k)} = b_i^{(k-1)})$ entonces:

1. la palabra de código, sobre la banda i , de la trama del flujo k binario es la misma que la de la trama del flujo $(k-1)$ binario:

$$m_i^{(k)} = m_i^{(k-1)}$$

Si no, es decir $(b_i^{(k)} > b_i^{(k-1)})$:

2. Se busca entre las $(NL(b_i^{(k)}, d_i) - NL(b_i^{(k-1)}, d_i))$ *leaders* de $CL(b_i^{(k)}, d_i) \setminus CL(b_i^{(k-1)}, d_i)$ el vecino más próximo de $\tilde{Y}(i)$,

3. Con el resultado de la etapa 2 y conociendo el vecino más próximo de $\tilde{Y}(i)$ en $CL(b_i^{(k-1)}, d_i)$, se comprue-

ba si el vecino más próximo de $\tilde{Y}(i)$ en $CL(b_i^{(k)}, d_i)$ es en $CL(b_i^{(k-1)}, d_i)$ (caso "Flag=0" más adelante)

o $CL(b_i^{(k)}, d_i) \setminus CL(b_i^{(k-1)}, d_i)$ (caso "Flag=1" más adelante),

4. Si Flag=0 ((el *leader* más próximo de $\tilde{Y}(i)$ en $CL(b_i^{(k-1)}, d_i)$ es también su vecino más próximo

en $CL(b_i^{(k)}, d_i)$) entonces:

$$m_i^{(k)} = m_i^{(k-1)}$$

Si Flag=1 (el *leader* más próximo de $\tilde{Y}(i)$ en $CL(b_i^{(k)}, d_i) \setminus CL(b_i^{(k-1)}, d_i)$ encontrado en la etapa 2 es

5 también su vecino más próximo en $CL(b_i^{(k)}, d_i)$, sea $L_i^{(k)}$ su número (con entonces $L_i^{(k)} \geq NL(b_i^{(k-1)}, d_i)$,
entonces se realizan las etapas siguientes:

10 a. Búsqueda del rango $r_i^{(k)}$ de $Y_q^{(k)(i)}$ (nuevo vector cuantificado de $Y(i)$ en la clase del *leader* $\tilde{Y}_q^{(k)(i)}$) por ejemplo mediante el algoritmo de Schalkwijk utilizando $perm(i)$,

15 b. Determinación de $sign_q^{(k)}(i)$ utilizando $sign(i)$ y $perm(i)$,

c. Determinación de la palabra de código $m_i^{(k)}$ a partir de $L_i^{(k)}$, $r_i^{(k)}$ y $sign_q^{(k)}(i)$.

20 * Segundo ejemplo de realización: aplicación en un codificador por transformada de tipo MPEG-1 capa I&II

El codificador MPEG-1 capa I&II, presentado en la figura 6a, utiliza un banco de filtros de 32 subbandas uniformes (bloque 61 de la figura 6a) para realizar la transformación tiempo/frecuencia de la señal s_0 de audio de entrada. Las muestras de salida de cada subbanda se reagrupan, y después se normalizan mediante un factor de escala común (determinado por el bloque 67 funcional) antes de cuantificarse (bloque 62). El número de niveles del cuantificador escalar uniforme utilizado para cada subbanda es el resultado de un procedimiento de asignación dinámica de los bits (realizada por el bloque 63). Este procedimiento utiliza un modelo psicoacústico (bloque 64) para determinar la distribución de los bits que hace el ruido de cuantificación lo menos perceptible posible. Los modelos de audición propuestos en la norma se basan en la estimación del espectro obtenido por una transformada de Fourier rápida (FFT) de la señal de tiempo de entrada (realizada por el bloque 65). En referencia a la figura 6b, la trama s_c , multiplexada por el bloque 66 de la figura 6a y que se transmite finalmente, contiene, en un campo de cabecera H_D , el conjunto de las muestras de las subbandas E_{SB} cuantificadas, que representan la información principal, y una información complementaria utilizada para la operación de decodificación constituida por los factores F_E de escala y la asignación A_i de bits.

40 A partir de este esquema de codificación, la construcción de un codificador de múltiples velocidades, en una aplicación de la invención, puede realizarse poniendo en común los bloques funcionales siguientes, en referencia a la figura 7:

- Banco 61 de filtros de análisis
- Determinación de los factores de escala 67
- 45 • Cálculo 65 de la transformada de Fourier FFT
- Determinación de los umbrales de enmascaramiento según un modelo psicoacústico 64.

50 Los dos bloques 64 y 65 proporcionan ya las relaciones señal a máscara (flechas SMR de las figuras 6a y 7), utilizadas para el procedimiento de asignación de bits (bloque 70 de la figura 7).

55 En este ejemplo de realización tal como se representa en la figura 7, es posible beneficiarse del procedimiento utilizado para la asignación de bits para ponerlo también en común, pero añadiendo no obstante algunas modificaciones en la asignación (bloque 70 de asignación de los bits de la figura 7). Sólo el bloque 62_0 a 62_(K-1) funcional de cuantificación es por tanto específico de cada flujo binario correspondiente a una velocidad D_k , $0 \leq k \leq K-1$. Lo mismo sucede para el bloque 66_0 a 66_(K-1) de multiplexación.

60 *Asignación de los bits*

En el codificador MPEG-1 capa I&II, la asignación se hace preferiblemente mediante una sucesión de etapas iterativas como sigue.

65 *Etapas 0:* Inicialización a cero del número de bits b_i de cada una de las subbandas i , $0 \leq i < M$.

ES 2 333 020 T3

Etapas 1: Puesta al día de la función de distorsión $NMR(i)$ (denominada “relación ruido a máscara” del inglés “Noise to Mask Ratio”) sobre cada una de las subbandas:

$$NMR(i) = SMR(i) - SNR(b_i),$$

donde $SNR(b_i)$ es la relación señal a ruido correspondiente al cuantificador que tiene un número de bits b_i , y $SMR(i)$ la relación señal a máscara proporcionada por el modelo psicoacústico.

Etapas 2: aumento del número de bits b_{i_0} de la subbanda i_0 donde esta distorsión es máxima:

$$b_{i_0} = b_{i_0} + \varepsilon, \quad i_0 = \arg \max_i [NMR(i)]$$

donde ε es un valor entero positivo que depende de la banda, en general igual a 1.

Las etapas 1 y 2 se repiten de manera iterativa hasta que se distribuye el número total de bits disponibles, correspondientes a la velocidad de funcionamiento. El resultado entonces es un vector de distribución de bits $(b_0, b_1, \dots, b_{M-1})$.

En el esquema de codificación de múltiples velocidades, estas etapas se ponen en común con algunas otras modificaciones, especialmente:

- el bloque funcional que tiene como salida K vectores de distribuciones de bits $(b_0^{(k)}, b_1^{(k)}, \dots, b_{M-1}^{(k)})$ (con

$0 \leq k \leq K-1$), se obtiene un $(b_0^{(k)}, b_1^{(k)}, \dots, b_{M-1}^{(k)})$ cuando se distribuye el número total de bits disponibles que se corresponden con la velocidad D_k binaria del flujo k binario, en la iteración de las etapas 1 y 2.

- La detención de la iteración de las etapas 1 y 2 se hace cuando se ha distribuido totalmente el número total de bits disponibles que se corresponden con la velocidad binaria más alta D_{K-1} (recuérdese que los flujos binarios se ordenan siguiendo un orden ascendente de velocidades).

Se observará que los vectores de distribución de bits se obtienen sucesivamente a partir de $k = 0$ hasta $k = K-1$. Las K salidas de este bloque de asignación de bits alimentan entonces los bloques de cuantificación para cada uno de los flujos binarios a una velocidad dada.

** Tercer ejemplo de realización: aplicación en un codificador de tipo CELP*

El último ejemplo de realización se refiere a la codificación del habla multimodo de decisión *a posteriori* a partir del codificador 3GPP NB-AMR (por “Narrow-Band Adaptive Multi-Rate”, tasa múltiple adaptativa de banda estrecha) que es un codificador adaptativo del habla en banda de telefonía de múltiples velocidades, según una norma 3GPP. Este codificador, que pertenece a la familia ampliamente conocida de los codificadores CELP cuyo principio se describió brevemente con anterioridad, comprende ocho modos (o velocidades) que van de 12,2 kbit/s a 4,75 kbit/s, basados todos en la técnica ACELP (por “Algebraic Code Excited Linear Prediction”, predicción lineal con excitación por código algebraico). La figura 8 ofrece el esquema de codificación en bloques funcionales de este codificador. Esta estructura se ha aprovechado con el fin de realizar un codificador multimodo de decisión *a posteriori*, basado en 4 modos del codificador NB-AMR (7,4; 6,7; 5,9; 5,15).

En una variante, sólo se aprovecha la mutualización de los bloques funcionales idénticos (los resultados de las cuatro codificaciones son entonces idénticos a los de las cuatro codificaciones en paralelo).

En una segunda variante, la complejidad se reduce aún más. Los cálculos de bloques funcionales no idénticos para determinados modos se aceleran aprovechando los de otro modo o de un módulo de tratamiento común, como se verá a continuación. Los resultados de las cuatro codificaciones así mutualizadas son entonces diferentes de los de las cuatro codificaciones en paralelo.

En aún otra variante, los bloques funcionales de estos cuatro modos se utilizan para una codificación multimodo en cuadrícula, como se vio anteriormente con referencia a la figura 1d.

Se recuerdan brevemente a continuación los cuatro modos (7,4; 6,7; 5,9; 5,15) del codificador 3GPP NB-AMR.

ES 2 333 020 T3

El codificador 3GPP NB-AMR trabaja sobre una señal de habla limitada en banda a 3,4 kHz y muestreada a 8 kHz cortada en tramas de 20 ms (160 muestras). Cada trama comprende 4 subtramas de 5 ms (40 muestras) agrupadas 2 a 2 en "super subtramas" de 10 ms (80 muestras). Para todos los modos, los mismos tipos de parámetros se extraen de la señal aunque con variantes de modelado y/o de cuantificación de estos parámetros. En el codificador NB-AMR se van a analizar y codificar cinco tipos de parámetros. Los parámetros LSP (por "Line Spectral Pairs", pares de líneas espectrales) se tratan una vez por trama para todos los modos, excepto para el modo 12,2 (es decir una vez por super subtrama). Los demás parámetros (especialmente el retardo LTP, la ganancia de la excitación adaptativa, le excitación fija, la ganancia de la excitación fija) se tratan una vez por subtrama.

Los cuatro modos considerados en este caso (7,4; 6,7; 5,9; 5,15) se distinguen esencialmente por las cuantificaciones de sus parámetros. La asignación binaria de estos 4 modos se resume en la tabla 1 a continuación.

TABLA 1

Asignación binaria de los 4 modos (7,4; 6,7; 5,9; 5,15) del codificador 3GPP NB-AMR

Modo (kbit/s)	7,4	6,7	5,9	5,15
LSP	26=(8+9+9)	26=(8+9+9)	26=(8+9+9)	23=(8+8+7)
Retardos LTP	8/5/8/5	8/4/8/4	8/4/8/4	8/4/8/4
Excitación fija	17/17/ 17/17	14/14/ 14/14	11/11/ 11/11	9/9/ 9/9
Ganancias de las excitaciones fija y adaptativa	7/7//7/7	7-/7/7/7	6/6/6/6	6/6/6/6
Total por trama	148	134	118	103

Estos 4 modos del codificador NB-AMR (7,4; 6,7; 5,9; 5,15) tienen módulos idénticos como, por ejemplo, el pretratamiento, el análisis de los coeficientes de predicción lineal, el cálculo de señal ponderada. El pretratamiento de la señal es un filtrado paso alto de frecuencia de corte de 80 Hz para eliminar las componentes continuas combinado con una división por dos de las señales de entrada para evitar desbordamientos. El análisis LPC comprende submódulos de aplicación de función ventana, de cálculo de las autocorrelaciones, de puesta en práctica del algoritmo de Levinson-Durbin, de transformación $A(z) \rightarrow LSP$, de cálculo de los parámetros de LSP_i no cuantificados para cada subtrama ($i=0, \dots, 3$) por interpolación entre los LSP de la trama anterior y los de la trama actual, y de transformación inversa ($LSP_i \rightarrow A_i(z)$).

El cálculo de la señal de habla ponderada reside en un filtrado por el filtro de ponderación perceptiva ($W_i(z) = A_i(z/\gamma_1)/A_i(z/\gamma_2)$) donde $A_i(z)$ es el filtro no cuantificado de la subtrama de índice i con $\gamma_1 = 0,94$ y $\gamma_2 = 0,6$.

Otros bloques funcionales sólo son idénticos para tres de estos modos (7,4; 6,7; 5,9). Por ejemplo, la búsqueda del retardo LTP en bucle abierto realizada sobre la señal ponderada una vez por super subtrama para estos tres modos. Para el modo a 5,15, ésta se realiza en cambio sólo una vez por trama.

Asimismo, si los cuatro modos utilizan una cuantificación vectorial ponderada predictiva MA (por "Moving Average", promedio variable), de orden 1 con media eliminada y por producto cartesiano de los parámetros LSP en el dominio de frecuencia normalizado, la cuantificación de los parámetros LSP del modo a 5,15 kbit/s se hace sobre 23 bits, la de los otros tres modos sobre 26 bits. Tras la transformación en el dominio de frecuencia normalizado, la cuantificación vectorial por producto cartesiano (denominada "split VQ") de los parámetros LSP divide los 10 parámetros

ES 2 333 020 T3

LSP en 3 subvectores, de dimensión 3, 3 y 4. El primer subvector compuesto por los 3 primeros LSP se cuantifica sobre 8 bits por el mismo diccionario para los cuatro modos. El segundo subvector compuesto por los 3 LSP siguientes se cuantifica para los 3 modos de alta velocidad por un diccionario de tamaño 512 (9 bits) y para el modo a 5,15 por la mitad de este diccionario (un vector de cada 2). El tercer y último subvector compuesto por los 4 últimos LSP se cuantifica para los 3 modos de alta velocidad por un diccionario de tamaño 512 (9 bits) y para el modo de velocidad más baja por un diccionario de tamaño 128 (7 bits). La transformación en el dominio de frecuencia normalizado, el cálculo de los pesos del criterio de error cuadrático y la predicción MA (por "Moving Average", promedio variable) del residuo LSP que va a cuantificarse son idénticos para los 4 modos. Los tres modos de alta velocidad que utilizan los mismos diccionarios para cuantificar los LSP, pueden compartir, además del mismo módulo de cuantificación vectorial, la transformación inversa (para volver del dominio de frecuencia normalizado al dominio en coseno), así como el cálculo de los LSP_i^Q cuantificados para cada subtrama ($i=0,\dots,3$) mediante interpolación entre los LSP cuantificados de la trama anterior y los de la trama actual, y finalmente la transformación inversa $LSP_i^Q \rightarrow A_i^Q(z)$.

Las búsquedas en bucle cerrado de las excitaciones adaptativa y fija se hacen secuencialmente y necesitan previamente el cálculo de la respuesta impulsional del filtro de síntesis ponderada, y después de las señales objetivo. La respuesta impulsional del filtro de síntesis ponderada ($A_i(z/\gamma_1)/[A_i^Q(z)A_i(z/\gamma_2)]$) es idéntica para los 3 modos de alta velocidad (7,4; 6,7; 5,9). Para cada subtrama, el cálculo de la señal objetivo para la excitación adaptativa depende de la señal ponderada (independientemente del modo); del filtro cuantificado $A_i^Q(z)$ (idéntico para 3 de los modos) y del pasado de la subtrama (diferente para cada subtrama aparte de la primera subtrama). Para cada subtrama, la señal objetivo para la excitación fija se obtiene retirando en la señal objetivo anterior la contribución de la excitación adaptativa filtrada de esta subtrama (que es diferente de un modo a otro excepto para la primera subtrama de los 3 primeros modos).

Se utilizan tres diccionarios adaptativos. El primer diccionario, para las subtramas pares ($i=0$ y 2) de los modos (7,4; 6,7; 5,9) y para la primera subtrama del modo a 5,15, comprende 256 retardos absolutos fraccionarios, de resolución $1/3$ en el intervalo $[19 + 1/3, 84 + 2/3]$ y de resolución entera en el intervalo $[85, 143]$. La búsqueda en este diccionario de retardos absolutos está focalizada entorno al retardo encontrado en bucle abierto (intervalo de ± 5 para el modo a 5,15, de ± 3 para los demás modos). Para la primera subtrama de los modos (7,4; 6,7; 5,9), al ser la señal objetivo y el retardo en bucle abierto idénticos, el resultado de esta búsqueda en bucle cerrado también lo es. Los otros dos diccionarios son de tipo diferencial y permiten codificar la diferencia entre el retardo actual y el retardo entero T_{i-1} más próximo al retardo fraccionario de la subtrama anterior. El primer diccionario diferencial sobre 5 bits, utilizado para las subtramas impares del modo a 7,4, es de resolución $1/3$ entorno al retardo entero T_{i-1} en el intervalo $[T_{i-1}-5 + 2/3, T_{i-1} + 4 + 2/3]$. El segundo diccionario diferencial sobre 4 bits, incluido en el primero, se utiliza para las subtramas impares de los modos a 6,7 y 5,9 así como para las tres últimas subtramas del modo a 5,15. Este segundo diccionario es de resolución entera entorno al retardo entero T_{i-1} en el intervalo $[T_{i-1}-5, T_{i-1}+4]$ más una resolución de $1/3$ en el intervalo $[T_{i-1}-1 + 2/3, T_{i-1} + 2/3]$.

Los diccionarios fijos pertenecen a la familia ampliamente conocida de los diccionarios ACELP. La estructura de un repertorio ACELP se basa en el concepto de ISPP ("Interleaved Single-Pulse Permutation", permutación entrelazada de pulsos simples) que consiste en dividir el conjunto de las L posiciones en K pistas entrelazadas, estando ubicados cada uno de los N impulsos en determinadas pistas predefinidas. Los 4 modos (7,4; 6,7; 5,9; 5,15) utilizan el mismo corte de las 40 muestras de una subtrama en 5 pistas de longitud 8 entrelazadas, como muestra la tabla 2a. La tabla 2b muestra, por su parte, para los 3 modos (7,4; 6,7; 5,9) la velocidad del diccionario, el número de impulsos y su distribución en las pistas. La distribución de los 2 impulsos del diccionario ACELP de 9 bits del modo a 5,15 es todavía más restrictiva.

TABLA 2a

Corte en pistas entrelazadas de las 40 posiciones de una subtrama del codificador 3GPP NB-AMR

Pista	Posiciones
P_0	0, 5, 10, 15, 20, 25, 30, 35
P_1	1, 6, 11, 16, 21, 26, 31, 36,
P_2	2, 7, 12, 17, 22, 27, 32, 37
P_3	3, 8, 13, 18, 23, 28, 33, 38
P_4	4, 9, 14, 19, 24, 29, 34, 39

ES 2 333 020 T3

TABLA 2b

Distribución de los impulsos en las pistas para los modos 7,4; 6,7; 5,9 del codificador 3GPP NB-AMR

5	Modo (kbit/s)	7,4	6,7	5,9
10	Velocidad del diccionario ACELP (posiciones + amplitudes)	17 (13+4)	14 (11+3)	11 (9+2)
15	Número de impulsos	4	3	2
20	Pistas potenciales para i_0	p_0	p_0	p_1, p_3
25	Pistas potenciales para i_1	p_1	p_1, p_3	p_0, p_1, p_2, p_4
30	Pistas potenciales para i_2	p_2	p_2, p_4	-
35	Pistas potenciales para i_3	p_3, p_4	-	-

Las ganancias de las excitaciones adaptativa y fija se cuantifican sobre 7 ó 6 bits (con una predicción MA aplicada a la ganancia de la excitación fija) mediante una cuantificación vectorial conjunta que minimiza el criterio CELP.

* Codificación multimodo de decisión *a posteriori* que sólo aprovecha la mutualización de los bloques funcionales idénticos

A partir de este esquema de codificación, la construcción de un codificador multimodo de decisión *a posteriori* puede realizarse poniendo en común los bloques funcionales siguientes.

Haciendo referencia a la figura 8, para los 4 modos, se realiza en común:

- 50 • el pretratamiento (bloque 81),
- el análisis de los coeficientes de predicción lineal (aplicación de función ventana y cálculo de las autocorrelaciones 82, puesta en práctica del algoritmo de Levinson-Durbin 83, transformación $A(z) \rightarrow$ LSP 84, interpolación de los LSP y transformación inversa 862),
- 55 • el cálculo de la señal de entrada ponderada 87,
- la transformación de los parámetros LSP en el dominio de frecuencia normalizado, el cálculo de los pesos del criterio de error cuadrático para la cuantificación vectorial de los LSP, la predicción MA del residuo LSP, la cuantificación vectorial de los 3 primeros LSP (en el bloque 85).

Para todos estos bloques, su complejidad acumulada se divide así por 4. Para los 3 modos de velocidad más alta (7,4; 6,7; 5,9), se realiza:

- 65 • la cuantificación vectorial de los 7 últimos LSP (una vez por trama) (en el bloque 85 de la figura 8),
- la búsqueda del retardo LTP en bucle abierto (2 veces por trama) (bloque 88),

ES 2 333 020 T3

- la interpolación de los LSP cuantificados (861) y la transformación inversa hacia los filtros A_i^Q (para cada subtrama),

- el cálculo de la respuesta impulsional 89 del filtro de síntesis ponderada (para cada subtrama).

5

Para estos bloques, los cálculos ya no se realizan 4 veces sino 2 veces, una vez para los 3 modos de velocidad más alta y una vez para el modo de velocidad baja. Su complejidad se divide por tanto por 2.

Se puede mutualizar también, para estos 3 modos de velocidad más alta, para la primera subtrama el cálculo de las señales objetivo para la excitación fija (bloque 91 en la figura 8) y adaptativa (bloque 90), así como la búsqueda LTP en bucle cerrado (bloque 881). Debe observarse que la mutualización de estas operaciones para la primera subtrama sólo produce resultados idénticos en el contexto de la codificación múltiple de tipo multimodo de decisión *a posteriori*. En el contexto general de codificación múltiple, el pasado de la primera subtrama es, como para las 3 otras subtramas, diferente según las velocidades, estas operaciones conducen generalmente entonces a resultados diferentes.

15

* *Codificación multimodo de decisión a posteriori avanzada*

Bloques funcionales no idénticos pueden acelerarse aprovechando los de otro modo o de un módulo de tratamiento común. Según las restricciones de la aplicación (en cuanto a calidad y/o complejidad), pueden utilizarse variantes diferentes. Algunos ejemplos se describen a continuación. También es posible apoyarse en técnicas de transcodificación inteligente entre codificadores CELP.

20

* *La cuantificación vectorial del segundo subvector de LSP*

Se puede aprovechar, como en el caso del modo de realización para el codificador TDAC, el intercalado de determinados diccionarios para acelerar los cálculos. Así, al estar incluido el diccionario del segundo subvector de LSP del modo a 5,15 en el de los 3 otros modos, la cuantificación de este subvector Y por los 4 modos puede combinarse así de manera ventajosa:

25

- Etapa 1: buscar su vecino más próximo Y_l en el diccionario más pequeño (que corresponde a la mitad del diccionario grande),

30

- Y_l cuantifica Y para el modo a 5,15;

- Etapa 2: Buscar el vecino más próximo Y_h en el complementario en el diccionario grande (es decir, la otra mitad del diccionario);

35

- Etapa 3: Probar si el vecino más próximo de Y en el diccionario de 9 bits es Y_l (caso “Flag=0”) o Y_h (caso “Flag=1”),

40

- caso “Flag=0”: Y_l cuantifica también Y para los modos a 7,4; 6,7 y 5,9

- si no, (caso “Flag=1”), Y_h cuantifica Y para los modos a 7,4; 6,7 y 5,9

Esta puesta en práctica da un resultado idéntico al de la codificación multimodo no optimizada. Si se desea reducir más la complejidad de la cuantificación, se puede detener en la etapa 1 y tomar Y_l como vector cuantificado para los modos de alta velocidad si se considera que este vector está suficiente mente próximo a Y. Esta simplificación puede dar por tanto un resultado diferente de una búsqueda exhaustiva.

45

* *Aceleración de la búsqueda LTP en bucle abierto*

50

La búsqueda del retardo LTP en bucle abierto del modo a 5,15 puede aprovechar los resultados de la de otros modos. Si los dos retardos en bucle abierto encontrados sobre las 2 super subtramas están suficientemente próximos para permitir una codificación diferencial, la búsqueda en bucle abierto del modo a 5,15 no se realiza. Se utilizan en su lugar los resultados de los modos superiores. Si no, se puede:

55

- realizar la búsqueda clásica,

- o focalizar la búsqueda en bucle abierto sobre toda la trama entorno a dos retardos en bucle abierto encontrados por los modos superiores.

60

A la inversa, también se puede realizar en primer lugar la búsqueda del retardo en bucle abierto en el modo a 5,15 y focalizar las dos búsquedas del retardo en bucle abierto de los modos superiores entorno al valor determinado por el modo a 5,15.

65

En una tercera variante más avanzada, ilustrada en la figura 1d, se propone realizar un codificador multimodo en cuadrícula que permite varias combinaciones de bloques funcionales, presentando cada bloque funcional al menos dos modos de funcionamiento (o velocidades). Se construye este nuevo codificador a partir de las cuatro velocidades del

ES 2 333 020 T3

codificador NB-AMR mencionadas anteriormente (5,15; 5,90; 6,70; 7,40). En este codificador, se distinguen cuatro bloques funcionales: el bloque LPC, el bloque LTP, el bloque de excitación fija y el bloque de ganancias. En referencia a la tabla 1 presentada anteriormente, la tabla 3a a continuación recapitula, para cada uno de estos bloques funcionales, su número de velocidades y sus velocidades.

TABLA 3a

Número de velocidades y velocidades de los bloques funcionales para los cuatro modos (5,15; 5,90; 6,70; 7,40) del codificador NB-AMR

Bloque funcional	Número de velocidades	Velocidades de los bloques funcionales
LPC (LSP)	2	26 y 23
Retardo LTP	3	26, 24 y 20
Excitación fija	4	68, 56, 44 y 36
Ganancias	2	28 y 24

Hay entonces P=4 bloques funcionales y $2 \times 3 \times 4 \times 2 = 48$ combinaciones posibles. En el ejemplo particular de realización, se elige no considerar la alta velocidad del bloque funcional 2 (LTP velocidad 26 bits/trama). Por supuesto, es posible otra elección.

El codificador de múltiples velocidades obtenido de este modo presenta una gran granularidad de velocidades, con 32 modos posibles indicados en la tabla 3b. Sin embargo, se indica que el codificador así obtenido no es interoperable con el codificador NB-AMR anteriormente mencionado. En la tabla 3b, los modos correspondientes a las tres velocidades del NB-AMR (5,15; 5,90; 6,70) se presentan en **negrita**, eliminando la exclusión de la velocidad más alta del bloque funcional LTP la velocidad de 7,40.

TABLA 3b

Velocidad por bloque funcional y global del codificador multimodo en cuadrícula

Parámetros	LSP	Retardo LTP	Excitación fija	Ganancias de las excitaciones fija y adaptativa	Total
Velocidad por trama	23	20	36	24	103
	23	20	36	28	107
	23	20	44	24	111
	23	20	44	28	115
	23	20	56	24	123
	23	20	56	28	127
	23	20	68	24	135
	23	20	68	28	139
	23	24	36	24	107
	23	24	36	28	111
	23	24	44	24	115
	23	24	44	28	119

ES 2 333 020 T3

5	23	24	56	24	127
	23	24	56	28	131
	23	24	68	24	139
	23	24	68	28	143
10	26	20	36	24	106
	26	20	36	28	110
	26	20	44	24	114
15	26	20	44	28	118
	26	20	56	24	126
	26	20	56	28	130
20	26	20	68	24	138
	26	20	68	28	142
	26	24	36	24	110
25	26	24	36	28	114
	26	24	44	24	118
30	26	24	44	28	122
	26	24	56	24	130
	26	24	56	28	134
35	26	24	68	24	142
	26	24	68	28	146
40					

Este codificador que presenta 32 velocidades posibles, son necesarios 5 bits para identificar el modo utilizado. Como en la variante anterior, se aprovecha la mutualización de bloques funcionales. Se aplican estrategias de codificación diferentes para los diferentes bloques funcionales. Por ejemplo, para el bloque funcional 1 que comprende la cuantificación de los LSP, se da prioridad a la velocidad baja como se mencionó anteriormente de la manera siguiente:

- El primer subvector compuesto por los 3 primeros LSP se cuantifica sobre 8 bits por el mismo diccionario para las dos velocidades asociadas a este bloque funcional.

- El segundo subvector compuesto por los 3 LSP siguientes se cuantifica sobre 8 bits por el diccionario de velocidad más pequeña. Al corresponder este diccionario a la mitad del diccionario de velocidad más alta, sólo se realiza la búsqueda en la otra mitad del diccionario si la distancia entre los 3 LSP y el elemento elegido en el diccionario supera un umbral determinado.

- El tercer y último subvector compuesto por los 4 últimos LSP se cuantifica por un diccionario de tamaño 512 (9 bits) y por un diccionario de tamaño 128 (7 bits).

En cambio, como se mencionó anteriormente en la segunda variante (correspondiente a la codificación multimodo de decisión *a posteriori* avanzada), se elige dar prioridad a la alta velocidad para el bloque funcional 2 (retardo LTP). En el codificador NB-AMR, la búsqueda del retardo LTP en bucle abierto se realiza dos veces por trama para el retardo LTP de 24 bits y se realiza una sola vez por trama para el de 20 bits. Para este bloque funcional, se desea favorecer la alta velocidad. Por tanto, el cálculo del retardo LTP en bucle abierto se realiza de la manera siguiente:

- Se calculan dos retardos en bucle abierto sobre las 2 super subtramas. Si son suficientemente próximos para permitir una codificación diferencia; la búsqueda en bucle abierto sobre la trama entera no se realiza. Se utilizan en su lugar los resultados de las dos super subtramas.

ES 2 333 020 T3

- Si no, se realiza una búsqueda en bucle abierto sobre toda la trama focalizándola entorno a los dos retardos en bucle abierto encontrados anteriormente. Una variante que reduce la complejidad retiene únicamente el retardo en bucle abierto de la primera.

5

Tras determinados bloques funcionales, es posible realizar una selección parcial que permite reducir el número de combinación que han de examinarse. Por ejemplo, tras el bloque funcional 1 (LPC), pueden eliminarse las combinaciones con 26 bits para este bloque si el comportamiento de la velocidad de 23 bits está suficientemente próximo o, inversamente, eliminar el modo a 23 bits si su comportamiento está demasiado degradado con respecto al modo a 26 bits.

10

Así, la presente invención permite proporcionar una solución eficaz al problema de la complejidad de las codificaciones múltiples, mediante la mutualización y la aceleración de los cálculos puestos en práctica por los diferentes codificadores. Las estructuras de codificación pueden por tanto representarse con la ayuda de bloques funcionales que describen las diferentes operaciones realizadas en el transcurso de un tratamiento. Los bloques funcionales de las diferentes codificaciones puestas en práctica en una codificación múltiple presentan relaciones sólidas que se aprovechan en el sentido de la presente invención. Estas relaciones son particularmente sólidas cuando las diferentes codificaciones corresponden a diferentes modos de una misma estructura.

15

20

Se indica finalmente que la presente invención es flexible desde el punto de vista de la complejidad. Es posible, en efecto, decidir *a priori* la complejidad máxima de la codificación múltiple y adaptar el número de codificadores examinados en función de esta complejidad.

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

5 1. Procedimiento de codificación múltiple en compresión, en el que una señal de entrada está destinada a alimentar en paralelo al menos un primer codificador y un segundo codificador, comprendiendo cada uno de los codificadores primero y segundo una sucesión de bloques funcionales, con vistas a una codificación en compresión de dicha señal de entrada por cada uno de los codificadores primero y segundo,

10 operando al menos una parte de dichos bloques funcionales cálculos para suministrar parámetros respectivos que sirven para la codificación de dicha señal de entrada por cada codificador,

comprendiendo los codificadores primero y segundo respectivamente al menos un primer y un segundo bloque funcional dispuestos para realizar operaciones comunes, **caracterizado** porque:

- 15 - se opera, en una misma etapa y en un único bloque, cálculos para suministrar un mismo juego de parámetros al primer bloque y al segundo bloque, y
- 20 - en el caso en el que el primer y/o el segundo codificador funcione a una velocidad diferente de dicho único bloque, se adapta dicho juego de parámetros a la velocidad del primer y/o segundo codificador para utilizarse por el primer y/o segundo bloque, respectivamente.

25 2. Procedimiento según la reivindicación 1, **caracterizado** porque dicho un único bloque está constituido por uno o varios bloques de uno de los codificadores primero y segundo.

3. Procedimiento según la reivindicación 1, **caracterizado** porque comprende las etapas de preparación siguientes:

- 30 a) identificar los bloques funcionales que forman cada codificador, así como una o varias funciones realizadas por cada bloque,
- b) localizar, entre dichas funciones, funciones que son comunes de un codificador al otro, y
- 35 c) ejecutar dichas funciones comunes, de una vez, para una parte al menos de todos los codificadores, en al menos un mismo módulo de cálculo.

40 4. Procedimiento según la reivindicación 3, **caracterizado** porque, para cada función ejecutada en la etapa c), se utiliza al menos un bloque funcional de un codificador seleccionado de entre dicha pluralidad de codificadores, y porque el bloque de dicho codificador seleccionado está dispuesto para suministrar resultados parciales a los demás codificadores, para una codificación eficaz, en dichos otros codificadores, verificando un criterio óptimo entre la complejidad y la calidad de la codificación.

45 5. Procedimiento según la reivindicación 4, en el que los codificadores son susceptibles de operar a velocidades respectivas diferentes, **caracterizado** porque el codificador seleccionado es el codificador de velocidad más baja, y porque los resultados obtenidos, como resultado de la ejecución de la función en la etapa c) con parámetros propios del codificador seleccionado, se adaptan a las velocidades de una parte al menos de los demás codificadores mediante una búsqueda focalizada de parámetros para una parte al menos de todos los demás modos, hasta el codificador de velocidad más alta.

50 6. Procedimiento según la reivindicación 4, en el que los codificadores son susceptibles de operar a velocidades respectivas diferentes, **caracterizado** porque el codificador seleccionado es el codificador de velocidad más alta, y porque los resultados obtenidos, como resultado de la ejecución de la función en la etapa c) con parámetros propios del codificador seleccionado, se adaptan a las velocidades de una parte al menos de los demás codificadores mediante una búsqueda focalizada de parámetros para una parte al menos de todos los demás modos, hasta el codificador de velocidad más baja.

60 7. Procedimiento según la reivindicación 5, en combinación con la reivindicación 6, **caracterizado** porque, para una velocidad dada, se utiliza el bloque funcional de un codificador que opera con dicha velocidad dada, como módulo de cálculo, y se adapta progresivamente al menos una parte de los parámetros propios de este codificador:

- 65 - hasta el codificador de velocidad más alta, mediante búsqueda focalizada, y
- hasta el codificador de velocidad más baja, mediante búsqueda focalizada.

ES 2 333 020 T3

8. Procedimiento según la reivindicación 2, en el que los bloques funcionales de los diferentes codificadores están dispuestos en cuadrícula, con varios caminos posibles en la cuadrícula, **caracterizado** porque cada camino de la cuadrícula está definido por una combinación de modos de funcionamiento de los bloques funcionales, alimentando cada bloque funcional varias variantes posibles del bloque funcional siguiente.
9. Procedimiento según la reivindicación 8, **caracterizado** porque está previsto un módulo de selección parcial, después de cada etapa de codificación realizada por uno o varios bloques funcionales, que puede seleccionar los resultados proporcionados por uno o varios de estos bloques funcionales, para etapas siguientes de codificación.
10. Procedimiento según la reivindicación 8, en el que los bloques funcionales son susceptibles de operar a velocidades respectivas diferentes y utilizando parámetros respectivos propios de dichas velocidades, **caracterizado** porque, para un bloque funcional dado, el camino de la cuadrícula seleccionada es el que atraviesa el bloque funcional de velocidad más baja, y porque los resultados obtenidos de dicho bloque funcional de velocidad más baja se adaptan a las velocidades de una parte al menos de los demás bloques funcionales mediante una búsqueda focalizada de parámetros para una parte al menos de todos los demás bloques funcionales, hasta el bloque funcional de velocidad más alta.
11. Procedimiento según la reivindicación 8, en el que los bloques funcionales son susceptibles de operar a velocidades respectivas diferentes y utilizando parámetros respectivos propios de dichas velocidades, **caracterizado** porque, para un bloque funcional dado, el camino de la cuadrícula seleccionada es el que atraviesa el bloque funcional de velocidad más alta, y porque los resultados obtenidos de dicho bloque funcional de velocidad más alta se adaptan a las velocidades de una parte al menos de los demás bloques funcionales mediante una búsqueda focalizada de parámetros para una parte al menos de todos los demás bloques funcionales, hasta el bloque funcional de velocidad más baja.
12. Procedimiento según la reivindicación 10, en combinación con la reivindicación 11, **caracterizado** porque, para una velocidad dada asociada a los parámetros de un bloque funcional de un codificador, se utiliza el bloque funcional que opera a dicha velocidad dada, como módulo de cálculo, y se adapta progresivamente al menos una parte de los parámetros propios de este bloque funcional:
- hasta el bloque funcional que puede operar a la velocidad más baja, mediante búsqueda focalizada, y
 - hasta el bloque funcional que puede operar a la velocidad más alta, mediante búsqueda focalizada.
13. Procedimiento según la reivindicación 3, **caracterizado** porque dicho módulo de cálculo es un módulo independiente de dichos codificadores, y dispuesto para redistribuir los resultados obtenidos en la etapa c) a todos los codificadores.
14. Procedimiento según la reivindicación 13, en combinación con la reivindicación 3, **caracterizado** porque el módulo independiente y el o los bloques de al menos uno de los codificadores están dispuestos para intercambiar mutuamente los resultados obtenidos en la etapa c), y porque el módulo de cálculo está dispuesto para realizar una transcodificación de adaptación entre bloques funcionales de codificadores diferentes.
15. Procedimiento según una de las reivindicaciones 13 y 14, **caracterizado** porque el módulo independiente comprende un bloque de codificación al menos parcial y un bloque de transcodificación de adaptación.
16. Procedimiento según una de las reivindicaciones anteriores, en el que los codificadores en paralelo están dispuestos para operar en codificación multimodo, **caracterizado** porque está previsto un módulo de selección *a posteriori*, que puede seleccionar un codificador de entre los codificadores.
17. Procedimiento según la reivindicación 16, **caracterizado** porque está previsto un módulo de selección parcial, después de cada etapa de codificación realizada por uno o varios bloques funcionales, independiente de los codificadores y que puede seleccionar uno o varios codificadores.
18. Procedimiento según una de las reivindicaciones anteriores, en el que los codificadores son de tipo por transformada, **caracterizado** porque el módulo de cálculo comprende un bloque de asignación de bits, compartido entre todos los codificadores, estando cada asignación de bits realizada para un codificador seguida de una adaptación a este codificador especialmente en función de su velocidad.
19. Procedimiento según la reivindicación 18, **caracterizado** porque el procedimiento comprende además una etapa de cuantificación, cuyos resultados se proporcionan a todos los codificadores.
20. Procedimiento según la reivindicación 19, **caracterizado** porque comprende además etapas comunes a todos los codificadores de entre:
- una transformada tiempo-frecuencia (MDCT),
 - una detección de sonoridad en la señal de entrada,

ES 2 333 020 T3

- una detección de tonalidad,
- la determinación de una curva de enmascaramiento, y
- 5 - una codificación de envolvente espectral.

21. Procedimiento según la reivindicación 18, en el que los codificadores realizan una codificación en subbanda (MPEG-1), **caracterizado** porque el procedimiento comprende además etapas comunes a todos los codificadores de entre:

- la aplicación de un banco de filtros de análisis,
- una determinación de factores de escala,
- 15 - un cálculo de transformada espectral (FFT), y
- la determinación de umbrales de enmascaramiento según un modelo psicoacústico.

22. Procedimiento según una de las reivindicaciones 1 a 17, en el que los codificadores son del tipo de análisis por síntesis (CELP), **caracterizado** porque el procedimiento comprende etapas comunes a todos los codificadores de entre al menos:

- un pretratamiento,
- el análisis de coeficientes de predicción lineal,
- un cálculo de señal de entrada ponderada, y
- 30 - una cuantificación para al menos una parte de los parámetros.

23. Procedimiento según la reivindicación 22, en combinación con la reivindicación 17, **caracterizado** porque el módulo de selección parcial se pone en práctica después de una etapa compartida de cuantificación vectorial para parámetros a corto plazo (LPC).

24. Procedimiento según la reivindicación 22, en combinación con la reivindicación 17, **caracterizado** porque el módulo de selección parcial se pone en práctica después de una etapa compartida de búsqueda de parámetro a largo plazo (LTP) en bucle abierto.

25. Producto de programa de ordenador destinado a almacenarse en una memoria de una unidad de tratamiento, especialmente de un ordenador o de un terminal móvil, o en un soporte de memoria amovible y destinado a actuar conjuntamente con un lector de la unidad de tratamiento, **caracterizado** porque comprende las instrucciones para la puesta en práctica del procedimiento de transcodificación según una de las reivindicaciones anteriores.

26. Dispositivo auxiliar para una codificación múltiple en compresión, codificación en la que una señal de entrada está destinada a alimentar en paralelo una pluralidad de codificadores que comprenden cada uno una sucesión de bloques funcionales, con vistas a una codificación en compresión de dicha señal por cada codificador, **caracterizado** porque comprende una memoria que almacena las instrucciones de un producto de programa de ordenador según la reivindicación 25.

27. Dispositivo según la reivindicación 26, **caracterizado** porque comprende además un módulo (MI) de cálculo independiente para la puesta en práctica del procedimiento según una de las reivindicaciones 13 a 17 y 23, 24.

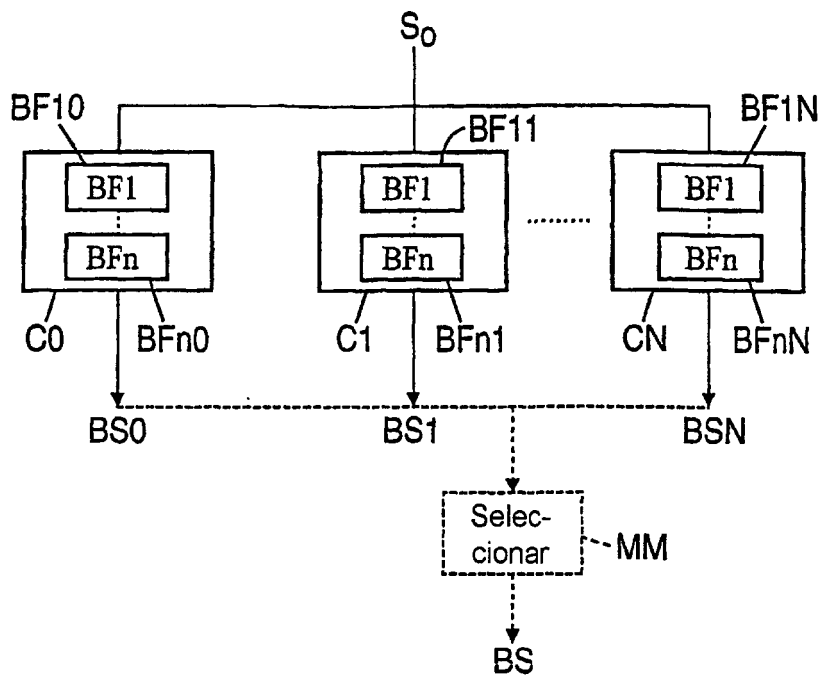


FIG. 1a

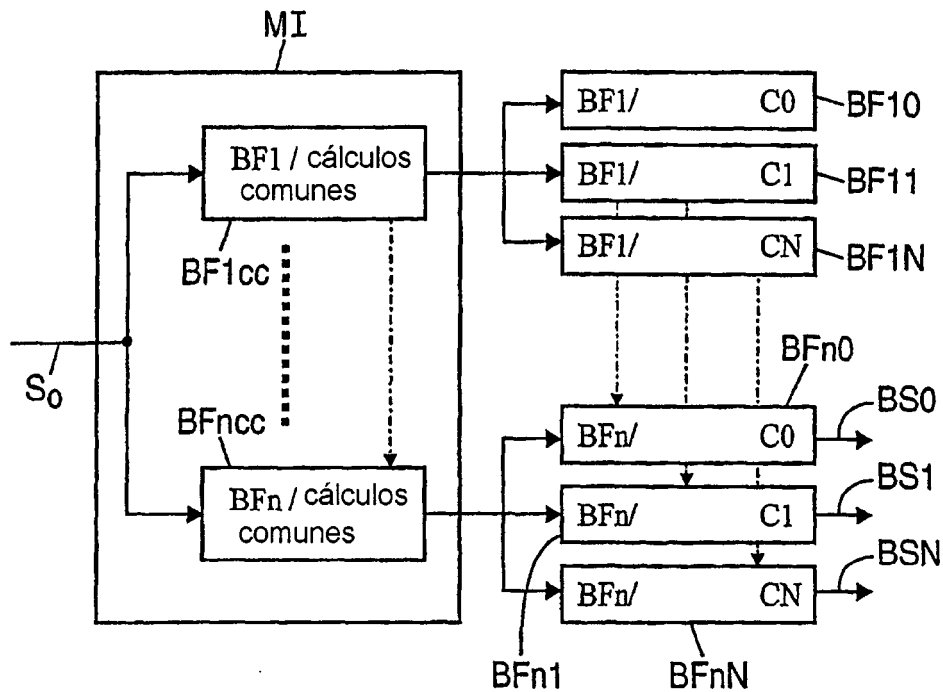


FIG. 1b

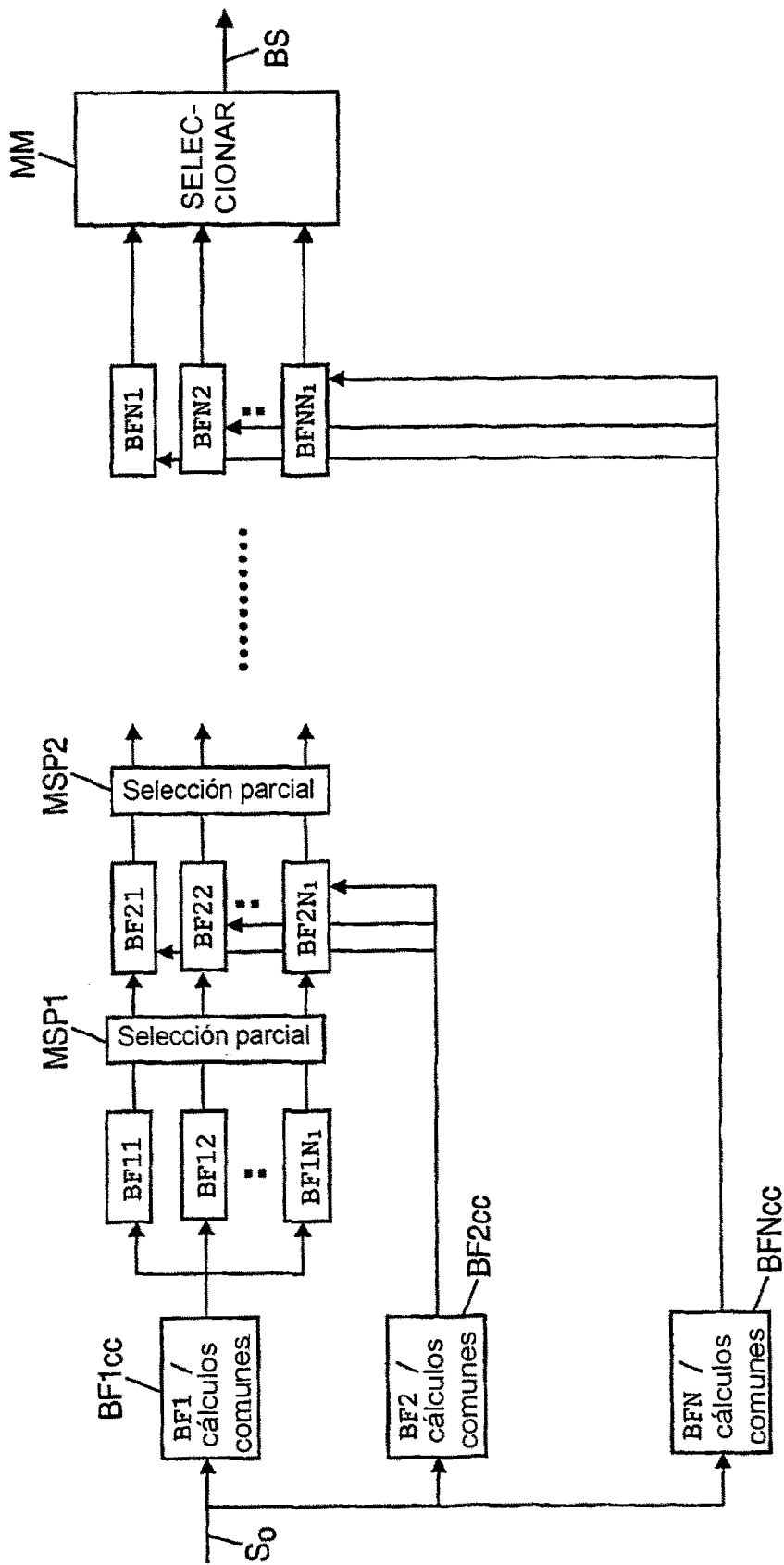


FIG. 1c

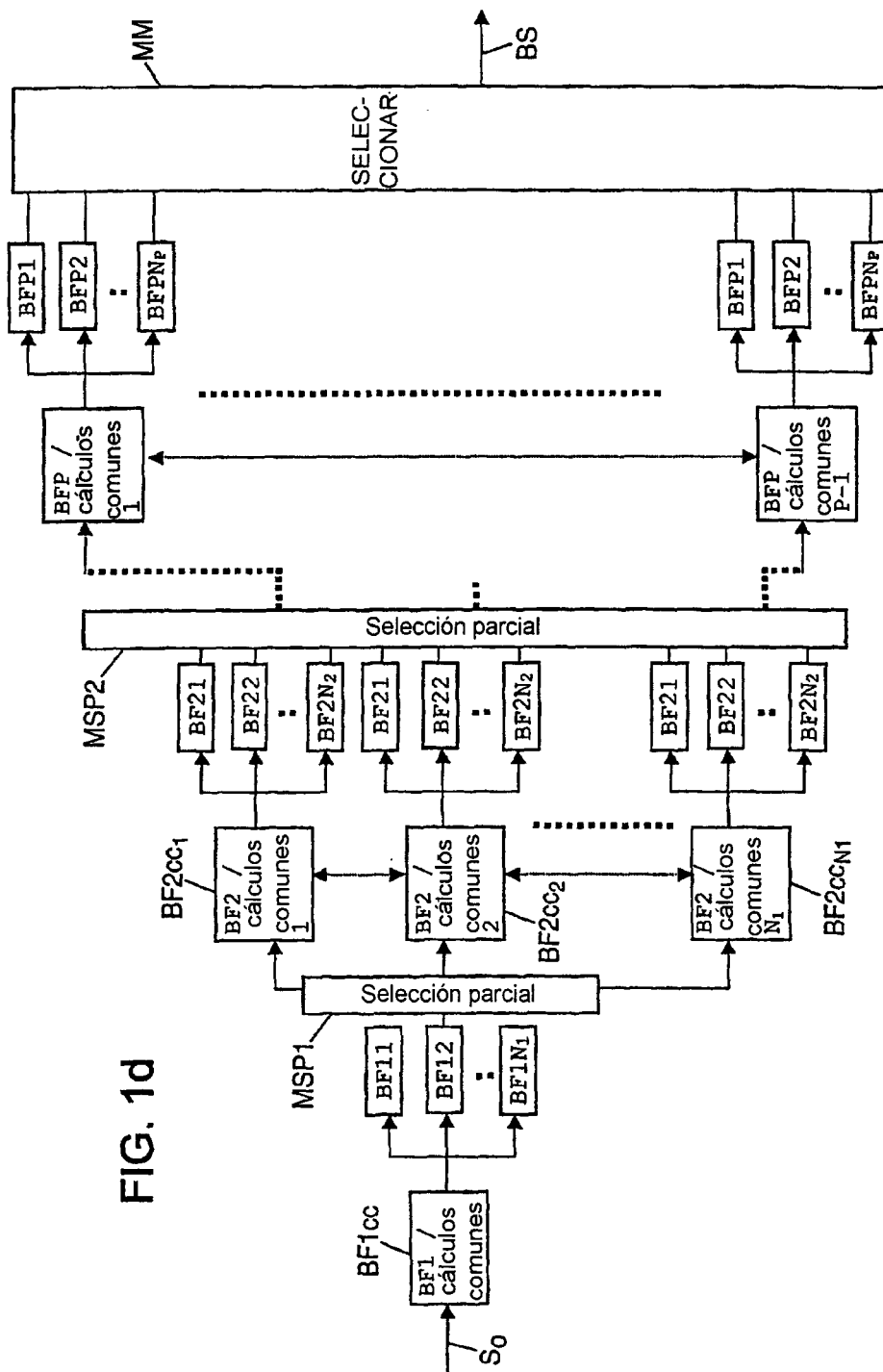


FIG. 1d

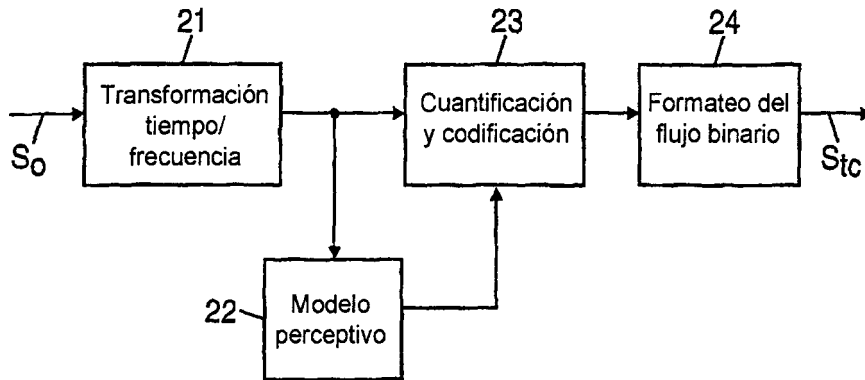


FIG. 2

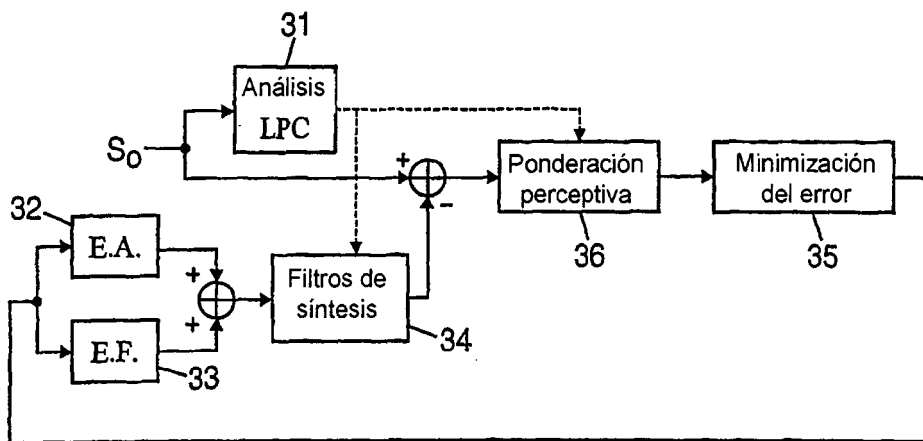


FIG. 3

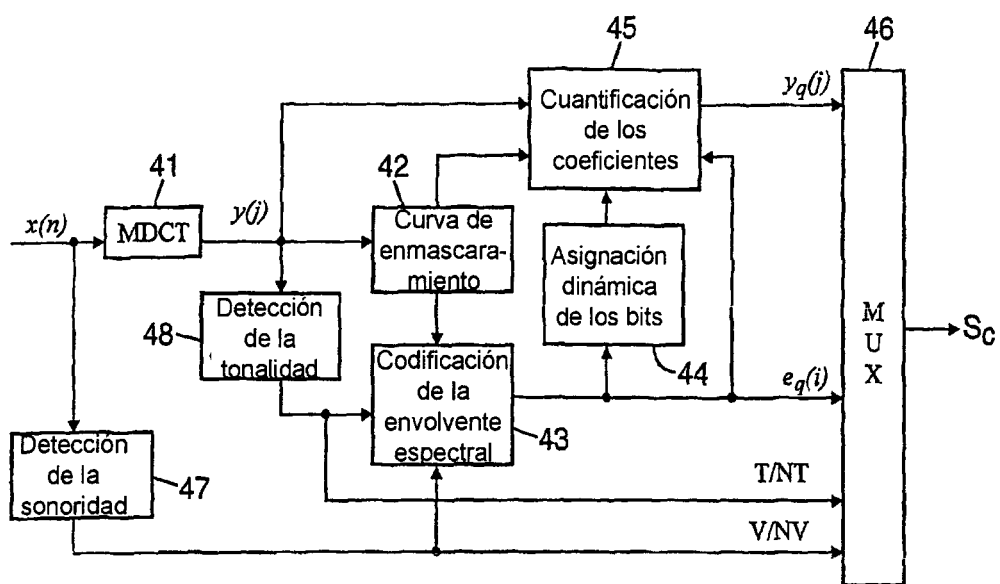


FIG. 4a

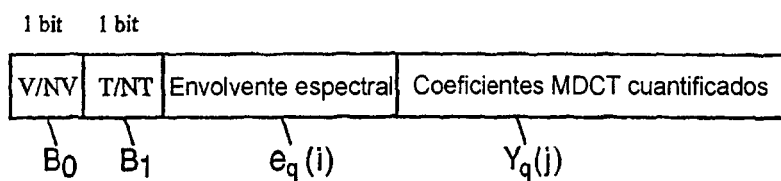


FIG. 4b

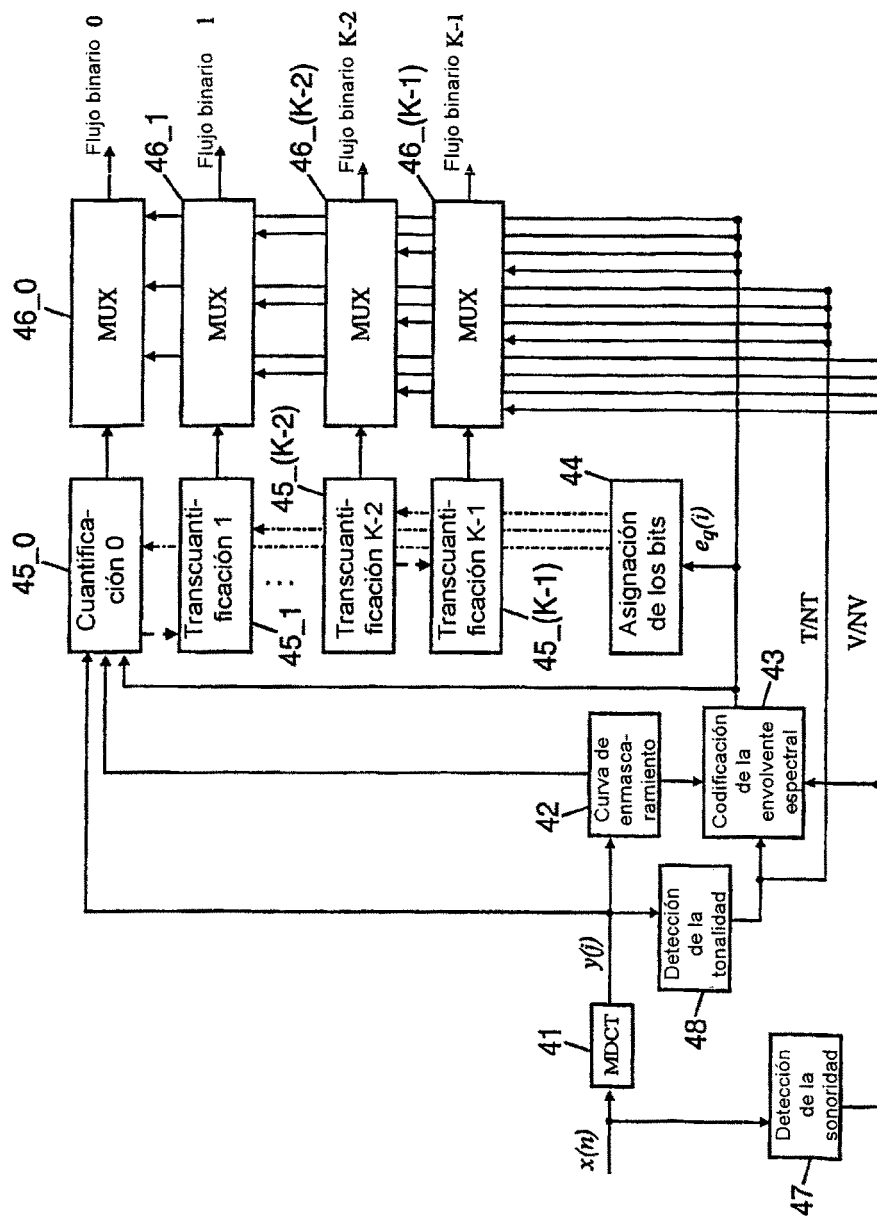


FIG. 5

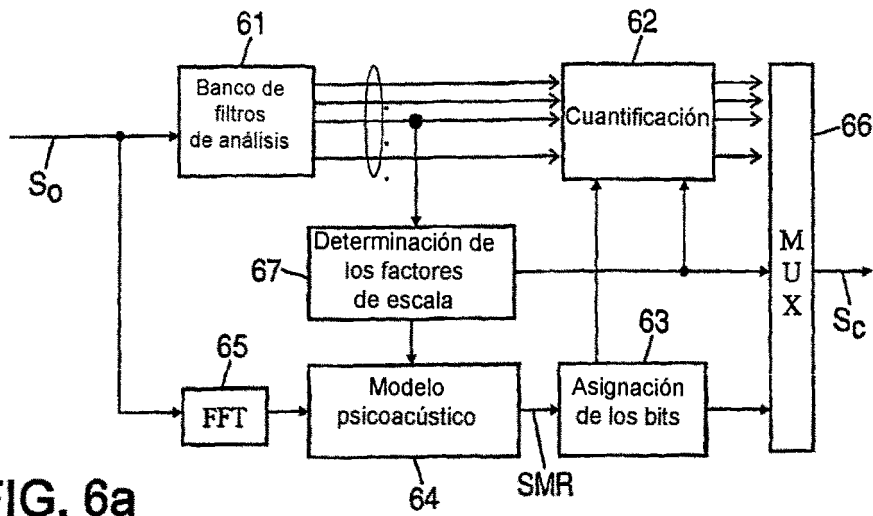


FIG. 6a

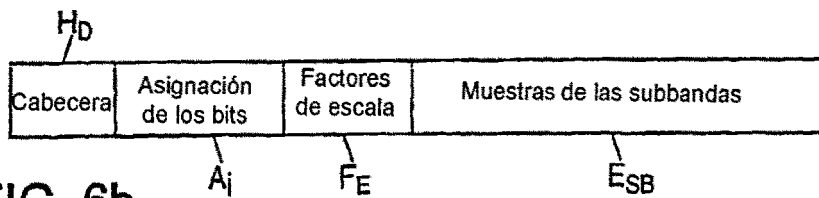


FIG. 6b

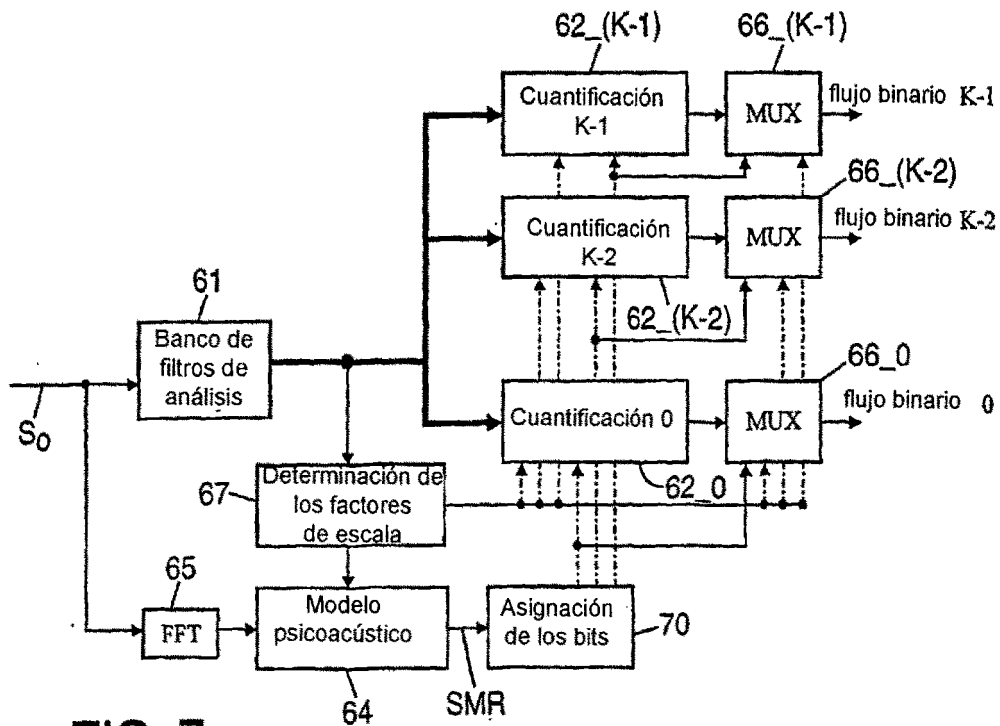


FIG. 7

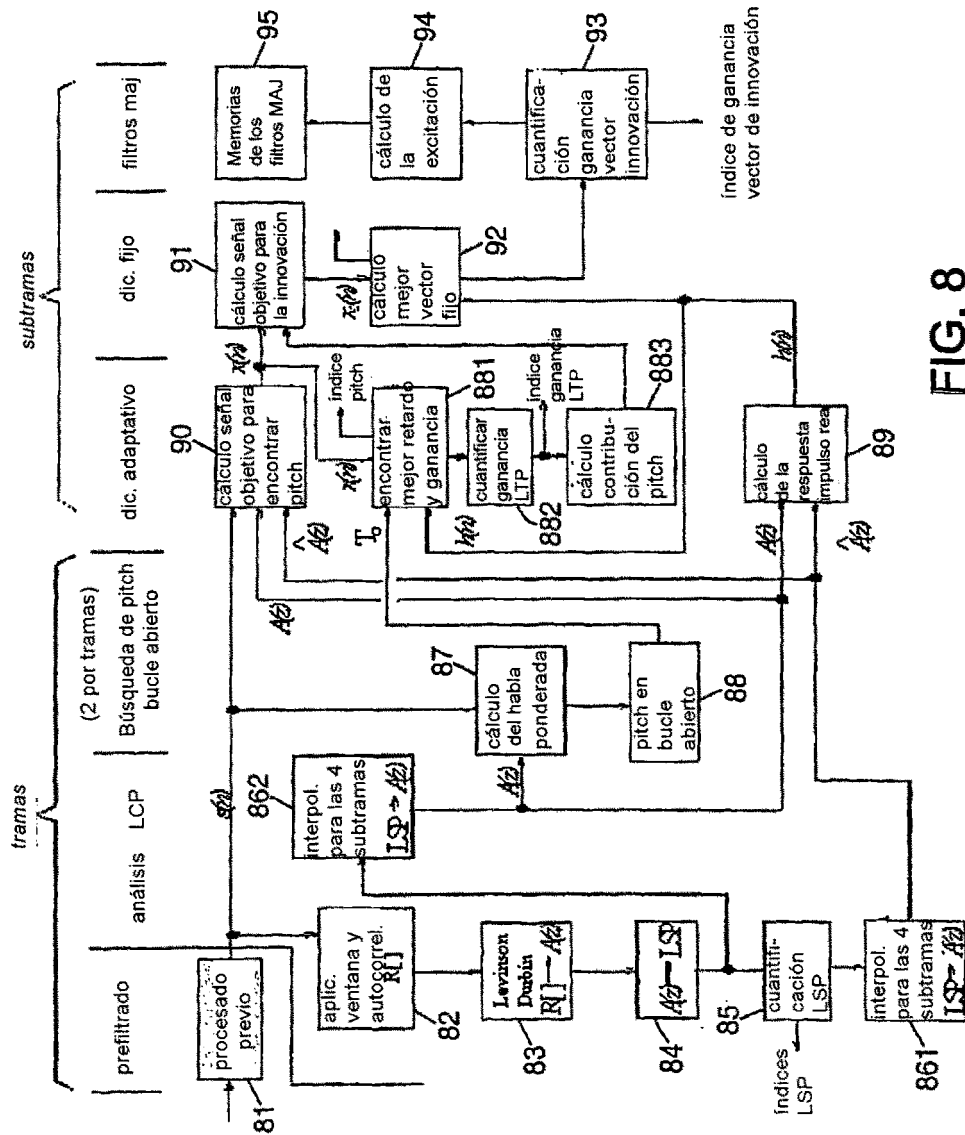


FIG. 8