

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11) Número de publicación: **2 733 878**

51) Int. Cl.:

**G10L 19/008** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86) Fecha de presentación y número de la solicitud internacional: **11.12.2009 PCT/FR2009/052491**

87) Fecha y número de publicación internacional: **24.06.2010 WO10070225**

96) Fecha de presentación y número de la solicitud europea: **11.12.2009 E 09803838 (3)**

97) Fecha y número de publicación de la concesión europea: **10.04.2019 EP 2374123**

54) Título: **Codificación mejorada de señales de audio digitales multicanales**

30) Prioridad:

**15.12.2008 FR 0858560**

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

**03.12.2019**

73) Titular/es:

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

72) Inventor/es:

**JAILLET, FLORENT y  
VIRETTE, DAVID**

74) Agente/Representante:

**ISERN JARA, Jorge**

**ES 2 733 878 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**

Codificación mejorada de señales de audio digitales multicanales

5 La presente invención se refiere al campo de la codificación/decodificación de señales de audio digital multicanales.  
 Más particularmente, la presente invención se refiere a la codificación/decodificación paramétrica de señales de audio multicanales.

10 Este tipo de codificación/decodificación se basa en la extracción de parámetros de espacialización para que, al decodificar, la percepción espacial del oyente puede ser reconstruida.

Dicha técnica de codificación se conoce con el nombre de "Binaural Cue Coding" en inglés (BCC), cuyo objetivo es, por un lado, extraer y luego codificar los índices de espacialización auditiva y, por otro lado, codificar una señal monofónica o estereofónica procedentes de una masterización de la señal multicanal original.

15 Este enfoque paramétrico es una codificación de baja velocidad. El principal interés de este enfoque de codificación es permitir una mejor tasa de compresión que los métodos de compresión de señales de audio digitales multicanales convencionales al tiempo que asegura la retrocompatibilidad del formato comprimido obtenido con los formatos de codificación y los sistemas de difusiones ya existentes.

20 El estándar MPEG Surround descrito en el documento del estándar MPEG ISO/IEC 23003-1:2007 y en el documento "Breebaart, J. y Hotho, G. y Koppens, J. y Schuijers, E. y Oomen, W. y van de Par, S.", titulado "Background, concept, and architecture for the recent MPEG surround standard on multichannel audio compression" en el Journal of the Audio Engineering Society 55-5 (2007) 331-351, describe una estructura de codificación paramétrica tal como se muestra en la figura 1.

De este modo, la figura 1 describe tal sistema de codificación/decodificación en donde el codificador 100 construye una señal de suma ("downmix" en inglés)  $S_s$  por masterización en 110 de los canales de la señal multicanal original  $S$  y proporciona a través de un módulo de extracción de parámetros 120, un conjunto reducido de parámetros  $P$  que caracterizan el contenido espacial de la señal multicanal original.

En el decodificador 150, la señal multicanal es reconstruida ( $S'$ ) por un módulo de síntesis 160 que tiene en cuenta tanto la señal de suma como los parámetros  $P$  transmitidos.

35 La señal de suma consta de un número reducido de canales. Estos canales pueden ser codificados por un codificador de audio convencional antes de la transmisión o el almacenamiento. Típicamente, la señal de suma consta de dos canales y es compatible con una difusión estéreo convencional. Antes de la transmisión o almacenamiento, esta señal de suma puede, de este modo, ser codificada por cualquier codificador estéreo convencional. La señal de este modo codificada es entonces compatible con los dispositivos que constan del decodificador correspondiente que reconstruye la señal de suma a la vez que ignora los datos espaciales.

40 Este esquema de codificación se basa en una estructura arborescente que permite el procesamiento de solo un número limitado de canales simultáneamente. De este modo, esta técnica es satisfactoria para la codificación y decodificación de señales de complejidad reducida utilizadas en el campo audiovisual, por ejemplo, para señales 5.1. Sin embargo, no permite obtener una calidad satisfactoria para señales multicanales más complejas, por ejemplo, para señales que se originan a partir de tomas de sonido multicanal directas o, incluso, señales de envolventes.

45 En efecto, dicha estructura limita la explotación de la redundancia entre canales que puede existir para señales complejas. Además, las señales multicanales que presentan oposiciones de fase, como, por ejemplo, las señales envolventes, no están bien reconstruidas por estas técnicas del estado de la técnica.

50 El documento WO 2007/104882 A1 se refiere a un sistema y procedimiento de codificación por análisis de componente principal de una señal de audio multicanal que consta de las siguientes etapas: - descomponer al menos dos canales de dicha señal de audio en una pluralidad de subbandas de frecuencia, - calcular al menos un parámetro de transformación en función de al menos una parte de dicha pluralidad de subbandas de frecuencia, transformar al menos una parte de dicha pluralidad de subbandas de frecuencia en una pluralidad de subcomponentes de frecuencia en función de dicho al menos un parámetro de transformación, dicha pluralidad de subcomponentes de frecuencia, incluidos los subcomponentes principales de frecuencia, - combinar al menos una parte de dichos subcomponentes principales de frecuencia para formar un componente principal y - definir una señal de audio codificada que represente dicha señal de audio multicanal, constando dicha señal de audio codificada de dicho componente principal y dicho al menos un parámetro de transformación. De este modo, cada subbanda de frecuencia puede corresponder a un parámetro de transformación que define un ángulo de rotación correspondiente a la posición de la fuente dominante de la subbanda de frecuencia.

65

El documento US 2007/0269063 A1 se refiere a la codificación de una señal de audio multicanal en el campo de frecuencia en función de la escena sonora percibida en lugar del contenido de los canales. En un modo de realización, los vectores de dirección espacial de tiempo-frecuencia se usan como índices para describir la escena sonora.

5 Por lo tanto, existe la necesidad de una técnica de codificación/decodificación paramétrica de señales de audio multicanales de gran complejidad que permita gestionar tanto las señales que presentan oposiciones de fase como tener en cuenta las redundancias entre canales entre las señales y, al mismo tiempo, ser compatibles con codificación de baja velocidad.

10 La presente invención viene a mejorar la situación.

Para tal efecto, propone un procedimiento de codificación de una señal de audio multicanal que representa una escena sonora que consta de una pluralidad de fuentes sonoras. El procedimiento es tal que consta de una etapa de descomposición (T) de la señal multicanal en bandas de frecuencia y las siguientes etapas por banda de frecuencia:

- 15
- obtención de informaciones de patrón por fuente sonora de la escena sonora, siendo las informaciones de patrón representativas de al menos una dirección y de un ancho angular de la fuente sonora en la escena sonora;
  - selección de un conjunto de fuentes sonoras de la escena sonora que constituyen unas fuentes principales;
  - 20 - masterización de las fuentes principales seleccionadas para obtener una señal de suma con un número reducido de canales;
  - codificación de las informaciones de patrón y formación de un flujo binario que consta de las informaciones de patrones codificadas, siendo el flujo binario adecuado para transmitirse en paralelo a la señal de suma.

25 De este modo, las informaciones de patrón asociadas con una fuente dan no solo la dirección de la fuente sino también la forma o la distribución espacial, de la fuente, es decir, la interacción que esta fuente puede tener con las otras fuentes de la escena sonora.

30 El conocimiento de estas informaciones de patrones asociadas con la señal de suma permitirá al decodificador obtener una señal de mejor calidad que tenga en cuenta las redundancias entre canales de manera global y las probables oposiciones de fase entre canales.

35 Al codificar por separado las informaciones de patrones y las fuentes sonoras por banda de frecuencia, se explota el hecho de que el número de fuentes activas en una banda de frecuencia es generalmente bajo, lo que aumenta los rendimientos de codificación.

Además, la señal de suma resultante de la codificación según la invención se puede decodificar mediante un decodificador estándar tal como se conoce en el estado de la técnica, aportando así la interoperabilidad con los decodificadores existentes.

40 Los diferentes modos particulares de realización mencionados a continuación pueden añadirse de independientemente o en combinación los unos con los otros, a las etapas del procedimiento de codificación definido anteriormente.

45 En un modo de realización particular de la invención, el procedimiento que consta, además, de una etapa de codificación de fuentes secundarias de entre las fuentes no seleccionadas de la escena sonora y de inserción de informaciones de codificación de las fuentes secundarias en el flujo binario.

50 La codificación de las fuentes secundarias permitirá, de este modo, aportar una precisión adicional en la señal decodificada, en particular, para señales complejas del tipo, por ejemplo, envolventes.

Las informaciones de codificación de las fuentes secundarias pueden ser, por ejemplo, envolturas espectrales codificadas o envolturas temporales codificadas que pueden constituir representaciones paramétricas de las fuentes secundarias.

55 En una variante de realización, la codificación de las fuentes secundarias consta de las siguientes etapas:

- construcción de pseudofuentes que representen al menos una parte de las fuentes secundarias, por descorrelación con al menos una fuente principal y/o al menos una fuente secundaria codificada;
- codificación de las pseudofuentes construidas; e
- 60 - inserción en el flujo binario de un índice de fuente utilizada y de un índice de descorrelador utilizado para la etapa de construcción.

65 Esto se aplica más particularmente en el caso donde la señal multicanal es de gran complejidad, pudiendo una parte de las fuentes secundarias o fuentes difusas luego ser representadas por pseudofuentes. En el caso de esta figura, entonces es posible codificar esta representación sin aumentar la velocidad de codificación.

En un modo de realización, la codificación de las informaciones de patrones se efectúa mediante un método de representación paramétrica.

5 Este método es de baja complejidad y se adapta particularmente al caso de la escena sonora de síntesis que representa una situación de codificación ideal.

Estas representaciones paramétricas pueden constar de, por ejemplo, informaciones de dirección de llegada, para la reconstrucción de un patrón simulando una onda plana o índices de selección de forma de patrón en un diccionario de formas de patrones.

10 En otro modo de realización, la codificación de las informaciones de patrón se efectúa mediante un método de análisis de componente principal que entrega vectores de patrón de base asociados con ganancias que permiten la reconstrucción de los patrones iniciales.

15 Esto permite, de este modo, codificar los patrones de escenas sonoras complejas cuya codificación no puede ser representada fácilmente por un modelo.

En todavía otro modo de realización, la codificación de las informaciones de patrón se efectúa mediante una combinación de un método de análisis de componente principal y de un método de representación paramétrica.

20 De este modo, por ejemplo, es posible efectuar en paralelo la codificación mediante los dos métodos y elegir el que cumpla un criterio de optimización de la velocidad de codificación, por ejemplo.

25 También es posible efectuar estos dos métodos en cascada para simplemente codificar una parte de los patrones mediante el método de codificación paramétrica y para lo que no están modelados, efectuar una codificación utilizando el método de análisis de componente principal, para representar en el mejor de los casos, todos los patrones. La distribución de la velocidad entre los dos modelos de codificación patrones se puede elegir según un criterio de minimización del error de reconstrucción de los patrones.

30 La presente invención se refiere también a un procedimiento de decodificación de una señal de audio multicanal que representa una escena sonora que consta de una pluralidad de fuentes sonoras, a partir de un flujo binario y de una señal de suma. El procedimiento es tal que consta de las siguientes etapas:

- 35 - extracción en el flujo binario y decodificación de informaciones de patrones representativos de al menos una dirección y de un ancho angular de las fuentes en la escena sonora;
- demasterización de la señal de suma para obtener un conjunto de fuentes principales;
- reconstrucción de la señal de audio multicanal mediante espacialización de al menos las fuentes principales con las informaciones de patrones decodificadas.

40 El método de decodificación permite de este modo reconstruir la señal multicanal de alta calidad para una restitución fiel del sonido espacializado teniendo en cuenta las redundancias entre canales a nivel global y las probables oposiciones de fase entre canales.

45 En un modo particular de realización del procedimiento de decodificación, éste que consta, además, de las etapas siguientes:

- extracción del flujo binario, de informaciones de codificación de fuentes secundarias codificadas;
- decodificación de las fuentes secundarias a partir de las informaciones de codificación extraídas;
- 50 - reagrupación de las fuentes secundarias con las fuentes principales para la espacialización.

La decodificación de fuentes secundarias aporta entonces más precisión a la escena sonora.

En una variante de realización, el procedimiento consta, además, de la etapa siguiente:

- 55 - decodificación de las fuentes secundarias mediante uso de una fuente efectivamente transmitida y de un correlador predefinido para reconstruir unas pseudofuentes representativas de al menos una parte de las fuentes secundarias.

En otra variante de realización, el procedimiento consta, además, de las etapas siguientes:

- 60 - extracción del flujo binario, de un índice de fuente principal y/o de al menos una fuente secundaria codificada y de un índice de un descorrelador a aplicar a esta fuente;
- decodificación de las fuentes secundarias mediante uso de la fuente y del índice de descorrelador para reconstruir pseudofuentes representativas de al menos una parte de las fuentes secundarias.

65 Esto hace posible encontrar pseudofuentes que representan una parte de las fuentes secundarias originales sin degradar la reproducción sonora de la escena sonora decodificada.

La presente invención se refiere también a un codificador de una señal de audio multicanal que representa una escena sonora que consta de una pluralidad de fuentes sonoras. El codificador es tal que consta, además, de:

- 5 - un módulo de descomposición de la señal multicanal en banda de frecuencia;
- un módulo de obtención de informaciones de patrón adecuado para obtener estas informaciones por fuente sonora de la escena sonora y por banda de frecuencia, siendo las informaciones de patrón representativas de al menos una dirección y de un ancho angular de la fuente sonora en la escena sonora;
- 10 - un módulo de selección de un conjunto de fuentes sonoras de la escena sonora que constituyen unas fuentes principales;
- un módulo de masterización de las fuentes principales procedentes del módulo de selección para obtener una señal de suma con un número reducido de canales;
- un módulo de codificación de las informaciones de patrón y un módulo de formación de un flujo binario que consta de las informaciones de patrones codificadas, siendo el flujo binario adecuado para transmitirse en paralelo a la
- 15 señal de suma.

Se refiere también a un decodificador de una señal de audio multicanal que representa una escena sonora que consta de una pluralidad de fuentes sonoras, que recibe a la entrada un flujo binario y una señal de suma. Este decodificador es tal que consta, además, de:

- 20 - un módulo de extracción y de decodificación de informaciones de patrones representativas de al menos una dirección y de un ancho angular de las fuentes en la escena sonora;
- un módulo de demasterización de la señal de suma para obtener un conjunto de fuentes principales;
- 25 - un módulo de reconstrucción de la señal de audio multicanal mediante espacialización de al menos las fuentes principales con las informaciones de patrones decodificadas.

Se refiere, finalmente, a un programa informático que consta de instrucciones de código para la implementación de las etapas de un procedimiento de codificación tal como se describe y/o un procedimiento de decodificación tal como se describe, cuando estas instrucciones se ejecutan por un procesador.

30 De manera más general, un medio de almacenamiento, legible por un ordenador o por un procesador, integrado o no en el codificador, posiblemente amovible, memoriza un programa informático que implementa un procedimiento de codificación y/o un procedimiento de decodificación según la invención.

35 Otras características y ventajas de la invención se pondrán de manifiesto de manera más clara con la lectura de la siguiente descripción, dada únicamente a título de ejemplo no limitativo y hecha con referencia a los dibujos adjuntos, en los que:

- 40 - la figura 1 ilustra un sistema de codificación/decodificación del estado de la técnica del tipo de sistema estandarizado MPEG Surround;
- la figura 2 ilustra un codificador y un procedimiento de codificación según un modo de realización de la invención;
- la figura 3a ilustra un primer modo de realización de la codificación de los patrones según la invención;
- la figura 3b ilustra un segundo modo de realización de la codificación de los patrones según la invención;
- 45 - la figura 4 representa ejemplos de patrones utilizados por la invención;
- la figura 5 ilustra un decodificador y un procedimiento de decodificación según un modo de realización de la invención;
- la figura 6 representa una variante de realización de un codificador y de un procedimiento de codificación según la invención;
- la figura 7 representa una variante de realización de un decodificador y de un procedimiento de decodificación
- 50 según la invención; e
- las figuras 8a y 8b representan respectivamente un ejemplo de dispositivo que comprende un codificador y un ejemplo de dispositivo que comprende un decodificador según la invención.

55 La figura 2 ilustra en forma de diagrama de bloques, un codificador según un modo de realización de la invención, así como las etapas de un procedimiento de codificación según un modo de realización de la invención.

El conjunto de los procesamientos en este codificador se efectúa por trama temporal. Por razones de simplificación, la representación y la descripción del codificador tal como se representa en la figura 2 se realiza considerando el procesamiento efectuado en una trama temporal fija, sin mostrar la dependencia temporal en las anotaciones.

60 Sin embargo, un mismo procesamiento se aplica sucesivamente al conjunto de las tramas temporales de la señal.

El codificador de este modo ilustrado consta de un módulo de transformada de tiempo-frecuencia 210 que recibe a la entrada una señal multicanal original que representa una escena sonora que consta de una pluralidad de fuentes sonoras.

Por lo tanto, este módulo efectúa una etapa T de cálculo de la transformada de tiempo-frecuencia de la señal multicanal original. Esta transformada se realiza, por ejemplo, mediante una transformada de Fourier a corto plazo.

5 Para ello, cada uno de los canales  $n_x$  de la señal original se muestra en una ventana en la trama temporal actual, luego, la transformada de Fourier F de la señal de ventana se calcula utilizando un algoritmo de cálculo rápido en puntos  $n_{FFT}$ . De este modo, se obtiene una matriz X compleja de tamaño  $n_{FFT} \times n_x$  que contiene los coeficientes de la señal multicanal original en el espacio de frecuencia.

10 Los procesamientos efectuados después por el codificador se efectúan por banda de frecuencia. Para ello, la matriz de los coeficientes X se corta en un conjunto de submatrices  $X_j$ , cada una de las cuales contiene los coeficientes de frecuencia en la banda en la posición j.

15 Son posibles diferentes opciones para el corte de frecuencia de las bandas. Para garantizar que el procesamiento se aplica a señales reales, a corto plazo se eligen bandas simétricas con respecto a la frecuencia cero en la transformada de Fourier. Además, con el fin de optimizar la eficacia de codificación, se prefiere la elección de bandas de frecuencias que se acerquen a escalas de frecuencia perceptivas, por ejemplo, eligiendo anchos de banda constantes en las escalas ERB (para "Equivalent Rectangular Bandwidth" en inglés) o Bark.

20 Por razones de simplificación, la descripción de las etapas de codificación efectuadas por el codificador se realizará para una banda de frecuencia determinada. Por supuesto, las etapas se llevan a cabo para cada una de las bandas de frecuencia a procesar.

En la salida del módulo 210, de este modo, se obtiene la señal para una banda de frecuencia  $S_{ij}$  determinada.

25 Un módulo de obtención de informaciones de patrones 220, permite determinar por una etapa OBT, por una parte, los patrones asociados con cada una de las fuentes de la escena sonora y, en segundo lugar, determinar las fuentes de la escena sonora para la banda de frecuencia determinada.

30 Los patrones son vectores de la misma dimensión que el número  $n_s$  de canales de la señal multicanal  $S_m$ .

Cada fuente está asociada a un vector de patrón.

35 Para una señal multicanal, el vector de patrón asociado con una fuente corresponde a la función de ponderación a aplicar a esta fuente antes de reproducirla en un altavoz, para reproducir, en el mejor de los casos, una dirección de llegada y un ancho de fuente. Es fácil comprender que, para un gran número de altavoces espaciados regularmente, el vector de patrón permitirá representar fielmente la radiación de una fuente sonora.

40 En presencia de una señal envolvente, el vector de patrón se obtendrá mediante la aplicación de una transformada de Fourier esférica inversa en los componentes de las órdenes envolventes. En efecto, las señales envolventes corresponden a una descomposición en armónicos esféricos, de ahí la correspondencia directa con el patrón de las fuentes.

45 El conjunto de vectores de patrón, por lo tanto, constituye una gran cantidad de datos que serían demasiado costosos de transmitir directamente para aplicaciones de baja velocidad de codificación. Para reducir la cantidad de información a transmitir, se pueden utilizar dos métodos de representación de los patrones, por ejemplo.

El módulo 230 de codificación Cod.Di de las informaciones de patrones puede implementar de este modo uno de los dos métodos descritos a continuación o una combinación de los dos métodos.

50 Un primer método es un método de modelado paramétrico que permite explotar los conocimientos *a priori* en el formato de señal utilizado. Consiste en transmitir solo un número muy reducido de parámetros y reconstruir los patrones en función de los modelos de codificación conocidos.

55 Por ejemplo, se trata de explotar los conocimientos sobre la codificación de las ondas planas para que las señales de tipo envolvente transmitan solo el valor de la dirección (acimut y elevación) de la fuente. Con estas informaciones, entonces es posible reconstruir el patrón correspondiente a una onda plana que proviene de esta dirección.

60 Por ejemplo, para un orden envolvente definido, el patrón asociado se conoce como una función de la dirección de llegada de la fuente sonora. Existen varios métodos para de estimación de los parámetros del modelo. De este modo, una búsqueda de picos en el diagrama de patrón (por analogía con el análisis sinusoidal, como se explica, por ejemplo, en el documento "Modelado informático de sonido musical (análisis, transformación, síntesis)" de Sylvain Marchand, Tesis doctoral, Universidad de Burdeos 1, permite detectar la dirección de llegada relativamente fiel.

65 Otros métodos como el "matching pursuit", como se presenta en S. Mallat, Z. Zhang, Matching pursuit with time-frequency dictionaries, IEEE Transactions on Signal Processing 41 (1993) 3397-3415 o el análisis espectral paramétrico también se pueden usar en este contexto.

Una representación paramétrica también puede utilizar un diccionario de forma simple para representar los patrones. A modo de ejemplo, la figura 4 da algunas formas simples de patrones (en acimut) que se pueden usar. Al codificar los patrones, se asocia con un elemento del diccionario, el acimut correspondiente y una ganancia que permite reproducir en la amplitud de este vector de patrón del diccionario. De este modo, se puede, a partir de un diccionario de forma de patrón, deducir la mejor forma o la combinación de formas que permitirá reconstruir mejor el patrón inicial.

Para la implementación de este primer método, el módulo 230 de codificación de patrones consta de un módulo de modelado paramétrico que genera los parámetros de patrones P a la salida. Estos parámetros se cuantifican a continuación mediante el módulo de cuantificación 240.

Este primer método permite obtener un muy buen nivel de compresión cuando la escena corresponde a una codificación ideal. Este será particularmente el caso de las escenas sonoras de síntesis.

Sin embargo, para escenas complejas o de grabaciones de sonido microfónico, es necesario utilizar modelos de codificación más genéricos, implicando la transmisión de una mayor cantidad de informaciones.

El segundo método descrito a continuación hace posible superar este inconveniente. En este segundo método, la representación de las informaciones de patrón se efectúa en forma de una combinación lineal de un número limitado de patrones de base. Este método se basa en el hecho de que el conjunto de patrones en un momento dado generalmente tiene una dimensión reducida. En efecto, solo un pequeño número de fuentes están activas en un momento dado y el patrón para cada fuente varía poco con la frecuencia.

Por lo tanto, es posible representar el conjunto de los patrones en un grupo de bandas de frecuencias a partir de un número muy reducido de patrones de base bien elegidos. Los parámetros transmitidos son entonces los vectores de patrón de base para el grupo de bandas consideradas y para que se codifique cada patrón, los coeficientes a aplicar a los patrones de base para reconstruir el patrón considerado.

Este método se basa en un método de análisis de componente principal (ACP o PCA en inglés para "Principal Component Analysis"). Esta herramienta es desarrollada en gran parte por I.T. Jolliffe en "Principal Component Analysis", Springer, 2002. La aplicación del análisis de componente principal a la codificación de los patrones se efectúa de la siguiente manera: primero, se forma una matriz de los patrones iniciales  $D_i$ , cuyo número de líneas corresponde al número total de fuentes de la escena sonora y el número de columna corresponde al número de canales de la señal multicanal original. A continuación, el análisis del componente principal, que corresponde a la diagonalización de la matriz de covarianza y que proporciona la matriz de los vectores propios, se efectúa adecuadamente. Finalmente, se seleccionan los vectores propios que llevan la información más importante y que corresponden a los valores propios del valor más alto. El número de vectores propios a conservar puede ser fijo o variable en el tiempo, en función de la velocidad disponible. Esta nueva base, por lo tanto, da la matriz  $D_B^T$ . Los coeficientes de ganancia asociados a esta base se calculan fácilmente a partir de  $G_D = D_i D_B^T$ .

En este modo de realización, la representación de los patrones se efectúa a partir del patrón de base. La matriz de patrones  $D_i$  se escribe como la combinación lineal de estos patrones de base. Así que se puede escribir  $D_i = G_D D_B$ , donde  $D_B$  es la matriz de los patrones de base para el conjunto de las bandas y  $G_D$  la matriz de las ganancias asociadas. El número de líneas de esta matriz representa el número total de fuentes de la escena sonora y el número de columnas representa el número de vectores de patrón de base.

En una variante de este modo de realización, se envían patrones de base por grupo de bandas consideradas, para representar más fielmente los patrones. Es posible, por ejemplo, proporcionar dos grupos de base de patrón: uno para bajas frecuencias y uno para las altas frecuencias. El límite entre estos dos grupos puede elegirse, por ejemplo, entre 5 y 7 kHz.

Para cada banda de frecuencia, se transmite, de este modo, el vector de ganancia asociado a los patrones de base.

Para este modo de realización, el módulo de codificación 230 comprende un módulo de análisis de componente principal que entrega vectores de patrones de base coeficientes o vectores de ganancia asociados  $G_D$ .

De este modo, después ACP, se codificará y transmitirá un número limitado de vectores de patrón. Para ello, se utiliza una cuantificación escalar efectuada por el módulo de cuantificación 240, coeficientes y vectores de patrones de base. El número de vectores de base a transmitir podrá ser fijo o bien seleccionado por el codificador utilizando, por ejemplo, un umbral en el error cuadrático medio entre el patrón original y el patrón reconstruido. De este modo, si el error es inferior al umbral, el o los vectores de base seleccionados hasta ahora son suficientes, no es necesario codificar un vector de base adicional.

En unas variantes de realización, la codificación de los patrones se realiza mediante una combinación de las dos representaciones enumeradas anteriormente. La figura 3a ilustra de forma detallada, el bloque de codificación de patrones 230, en una primera variante de realización.

Este modo de codificación utiliza los dos esquemas de representación de los patrones. De este modo, un módulo 310 efectúa un modelado paramétrico como se explicó anteriormente para proporcionar parámetros de patrones (P).

5 Un módulo 320 efectúa el análisis de componente principal para proporcionar tanto vectores de patrones de base ( $D_B$ ) y de los coeficientes asociados ( $G_D$ ).

En esta variante, un módulo de selección 330 elige una banda de frecuencia por banda de frecuencia, el mejor modo de codificación para el patrón al elegir el mejor compromiso de reconstrucción de los patrones/velocidad.

10 Para cada patrón, la elección de la representación retenida (representación paramétrica o por combinación lineal de patrones de base) se realiza para optimizar la eficacia de la compresión.

15 Un criterio de selección es, por ejemplo, la minimización del error cuadrático medio. Es posible que se utilice una ponderación perceptiva para la elección del modo de codificación de patrón. Esta ponderación tiene por objeto, por ejemplo, favorecer la reconstrucción de los patrones en la zona frontal, por la cual el oído es más sensible. En este caso, la función de error a minimizar en el caso del modelo de codificación por ACP puede ser de la siguiente forma:

$$E = (W(D_i - G_D D_B))^2$$

20 Con  $D_i$ , los patrones originales y  $W$ , la función de ponderación perceptiva.

Los parámetros de patrones procedentes del módulo de selección se cuantifican entonces por una etapa Q por el módulo de cuantificación 240 de la figura 2.

25 En una segunda variante del bloque de codificación 230, ambos modos de codificación están en cascada. La Figura 3b ilustra en detalle este bloque de codificación. De este modo, en esta variante de realización, un módulo 340 de modelado paramétrico efectúa un modelado para un cierto número de patrones y proporciona a la salida al mismo tiempo parámetros de patrones (P) para los patrones modelados y patrones no modelados o patrones residuales  $D_iR$ .

30 Estos patrones residuales ( $D_iR$ ) están codificados por un módulo de análisis de componente principal 350 que proporciona a la salida vectores de patrones de base ( $D_B$ ) y coeficientes asociados ( $G_D$ ).

35 Los parámetros de patrones, los vectores de patrones de base, así como los coeficientes, se proporcionan a la entrada del módulo de cuantificación 240 de la figura 2.

40 La cuantificación Q se efectúa reduciendo la precisión en función de los datos en la percepción, luego, aplicando una codificación entrópica. También, las posibilidades de explotación de la redundancia entre bandas de frecuencia o entre tramas sucesivas pueden permitir reducir la velocidad. Por lo tanto, se pueden utilizar predicciones intratrama o intertrama en los parámetros. De manera general, se podrán utilizar métodos de cuantificación convencionales. Por otra parte, siendo los vectores a cuantificar ortonormales, esta propiedad se podrá explotar durante la cuantificación escalar de los componentes del vector. En efecto, para un vector de dimensión N, solo se deberán cuantificar los componentes de  $N-1$ , pudiéndose el último componente recalcular.

45 Volviendo a la descripción de la figura 2, a la salida del cuantificador 240, un módulo de construcción de un flujo binario 250 inserta estas informaciones de patrones codificadas en un flujo binario  $F_b$  según la etapa  $Con.F_b$ .

50 El codificador tal como se describe aquí comprende, además, un módulo de selección 260 adecuado para seleccionar en la etapa Select las fuentes principales ( $S_{princ}$ ) de entre las fuentes de la escena sonora a codificar ( $S_{tot}$ ).

Para ello, un modo de realización particular utiliza un método de análisis de componente principal, ACP, en cada banda de frecuencia en el bloque 220 para extraer todas las fuentes de la escena sonora ( $S_{tot}$ ). Este análisis permite clasificar las fuentes en subbandas en orden de importancia según el nivel de energía, por ejemplo.

55 Las fuentes de mayor importancia (por lo tanto, de mayor energía) son seleccionadas por el módulo 260 para constituir las fuentes principales ( $S_{princ}$ ), que luego son estampadas en la etapa M por el módulo 270 para construir una señal de suma ( $S_{sfi}$ ) (o "downmix" en inglés).

60 El número de fuentes principales ( $S_{princ}$ ) se elige en función del número de canales de la señal de suma. Este número se elige inferior o igual al número de canales. Preferentemente, se elige un número de fuentes principales igual al número de canales de la señal de suma. La matriz M es entonces una matriz cuadrada predefinida.

Esta señal de suma por banda de frecuencia sufre una transformada de tiempo-frecuencia inversa  $T^{-1}$  por el módulo de transformada inversa 290 con el fin de proporcionar una señal de suma temporal ( $S_s$ ). Esta señal de suma es luego

codificada por un codificador de voz o un codificador de audio del estado de la técnica (por ejemplo: G.729.1 o MPEG-4 AAC).

Las fuentes secundarias ( $S_{sec}$ ) pueden codificarse por un módulo de codificación 280 y agregarse al flujo binario en el módulo de construcción del flujo binario 250.

5 Para estas fuentes secundarias, es decir, las fuentes que no se transmiten directamente en la señal de suma, existen diferentes alternativas de procesamientos.

Como estas fuentes se consideran no esenciales para la escena sonora, no pueden ser transmitidas.

10 Sin embargo, es posible codificar algunas o todas estas fuentes secundarias mediante el módulo de codificación 280, que en un modo de realización puede ser un módulo de codificación de transformada de Fourier a corto plazo. Estas fuentes se pueden codificar por separado utilizando los codificadores de audio o voz antes mencionados.

15 En una variante de esta codificación, los coeficientes de la transformada de estas fuentes secundarias pueden codificarse directamente solo en las bandas que se consideran importantes.

Las fuentes secundarias pueden ser codificadas por representaciones paramétricas; estas representaciones pueden ser en forma de envoltura espectral o envoltura temporal.

20 Estas representaciones se codifican en la etapa Cod. $S_{sec}$  del módulo 280 y se insertan en la etapa Con.Fb en el flujo binario. Estas representaciones paramétricas constituyen entonces informaciones de codificación de las fuentes secundarias.

25 Este procedimiento de codificación de una señal multicanal tal como se describe es particularmente interesante por el hecho de que el análisis se realiza en ventanas que pueden ser de corta longitud. De este modo, este modelo de codificación provoca un bajo retardo algorítmico que permite su uso en aplicaciones donde el control del retardo es importante.

30 En el caso de ciertas señales multicanales, en particular del tipo envolvente, el codificador, tal como se describe, implementa una etapa de procesamiento previo adicional P por un módulo 215 de procesamiento previo.

Este módulo efectúa una etapa de cambio de base para expresar la escena sonora utilizando la descomposición de onda plana del campo acústico.

35 La señal envolvente original se ve como la transformada de Fourier angular de un campo sonoro. Así, los diferentes componentes representan los valores para las diferentes frecuencias angulares. La primera operación de descomposición en ondas planas corresponde, por lo tanto, a tomar el componente omnidireccional de la señal envolvente como representando la frecuencia angular cero (este componente es, por lo tanto, un componente real).  
40 A continuación, los siguientes componentes envolventes (orden 1,2, 3, etc.) se combinan para obtener los coeficientes complejos de la transformada de Fourier angular.

45 Para una descripción más precisa del formato envolvente, se podrá referir a la tesis de Jérôme Daniel, titulada "Representación de campos acústicos, aplicación a la transmisión y a la reproducción de escenas sonoras complejas en un contexto multimedia" 2001, París 6.

De este modo, para cada orden envolvente superior a 1 (en 2 dimensiones), el primer componente representa la parte real y el segundo componente representa la parte imaginaria. Para una representación bidimensional, para un orden  $O$ , se obtiene  $O + 1$  componentes complejos. Luego se aplica una transformada de Fourier a Corto Plazo (en la dimensión temporal) para obtener las transformadas de Fourier (en el campo de frecuencia) de cada armónico angular esta etapa luego integra la etapa de transformación T del módulo 210. A continuación, la transformada angular completa se construye recreando los armónicos de frecuencias negativas mediante la simetría hermitiana. Finalmente, se efectúa una transformada de Fourier inversa en la dimensión de las frecuencias angulares para pasar al campo de los patrones.

55 Esta etapa de procesamiento previo le permite al codificador trabajar en un espacio de señales cuya interpretación física y perceptiva se simplifica, lo que permite explotar de manera más eficaz el conocimiento de la percepción auditiva espacial y, por lo tanto, mejorar los rendimientos de la codificación. Sin embargo, la codificación de las señales envolventes sigue siendo posible sin esta etapa de tratamiento previo.

60 Para señales de sonido no procedentes de técnicas envolventes, esta etapa es innecesaria. Para estas señales, el conocimiento del sistema de captura o la restitución asociado con la señal permite interpretar directamente las señales como una descomposición de ondas planas del campo acústico.

65 La figura 5 describe ahora un decodificador y un procedimiento de decodificación en un modo de realización de la invención.

Este decodificador recibe a la entrada, el flujo binario  $F_b$  tal como lo construido por el codificador descrito anteriormente, así como la señal de suma  $S_s$ .

5 De la misma manera que para el codificador, el conjunto de los procesamientos se efectúa por trama temporal. Para simplificar las anotaciones, la descripción del decodificador que sigue únicamente describe el procesamiento efectuado en una trama temporal fija y no muestra la dependencia temporal en las anotaciones. En el decodificador, sin embargo, este mismo procesamiento se aplica sucesivamente al conjunto de las tramas temporales de la señal.

10 Para encontrar las fuentes sonoras, la primera etapa de decodificación consiste en realizar la transformada de tiempo-frecuencia  $T$  de la señal de suma  $S_s$  por el módulo de transformada 510 para obtener una señal de suma por banda de frecuencia,  $S_{sfi}$ .

15 Esta transformada se realiza usando, por ejemplo, la transformada de Fourier a corto plazo. Cabe señalar que también se pueden usar otras transformadas o bancos de filtros y en particular bancos de filtros no uniformes según una escala de percepción (por ejemplo, Bark). Cabe señalar que, para evitar discontinuidades durante la reconstrucción de la señal de esta transformada, se utiliza un método de adición de recubrimiento.

20 Para la trama temporal considerada, la etapa de cálculo de la transformada de Fourier a corto plazo consiste en abrir una ventana en cada uno de los canales  $n_f$  de la señal de suma  $S_s$  utilizando una ventana  $w$  de longitud superior a la trama temporal, luego, en calcular la transformada de Fourier de la señal de ventana utilizando un algoritmo de cálculo rápido en puntos  $n_{FFT}$ . De este modo, se obtiene una matriz  $F$  compleja de tamaño  $n_{FFT} \times n_f$  que contiene los coeficientes de la señal de suma en el espacio de frecuencia.

25 En lo siguiente, el conjunto del procesamiento se efectúa por bandas de frecuencia. Para ello, la matriz de los coeficientes  $F$  se corta en un conjunto de submatrices  $F_j$ , cada una de las cuales contiene los coeficientes de frecuencia en la banda en la posición  $j$ . Son posibles diferentes opciones para el corte de frecuencia de las bandas. Para garantizar que el procesamiento se aplica a señales reales, a corto plazo se eligen bandas simétricas con respecto a la frecuencia cero en la transformada de Fourier. Además, con el fin de optimizar la eficacia de decodificación, se prefiere la elección de bandas de frecuencias que se acerquen a escalas de frecuencia perceptivas, por ejemplo, eligiendo anchos de banda constantes en las escalas ERB o Bark.

30 Por razones de simplificación, la descripción de las etapas de decodificación efectuadas por el decodificador se realizará para una banda de frecuencia determinada. Por supuesto, las etapas se llevan a cabo para cada una de las bandas de frecuencia a procesar.

35 El módulo 520 efectúa una demasterización  $N$  de los coeficientes de frecuencia de la transformada de la señal de suma de la banda de frecuencia considerada para encontrar las fuentes principales de la escena sonora.

40 Más precisamente, la matriz  $S_{princ}$  de los coeficientes de frecuencia para la banda de frecuencia actual de las  $n_{princ}$  fuentes principales se obtiene según la relación:  $S_{princ} = BN$ , donde  $N$  es de dimensión  $n_f \times n_{princ}$  y  $B$  es una matriz de dimensión  $n_{bin} \times n_f$  donde  $n_{bin}$  es el número de componentes (o binarios) de frecuencia retenidos en la banda de frecuencia considerada.

45  $N$  se calcula para permitir la inversión de la matriz de mezcla  $M$  utilizada en el codificador. Así que tenemos la siguiente relación:  $MN = I$ .

50 El número de líneas de la matriz  $N$  corresponde al número de canales de la señal de suma y el número de columnas corresponde al número de fuentes principales transmitidas. Para la matriz  $M$ , las dimensiones se invierten, siendo  $I$  una matriz de identidad de dimensiones  $n_{princ} \times n_{princ}$ .

55 Las líneas de  $B$  son los componentes de frecuencia en la banda de frecuencia actual, las columnas corresponden a los canales de señal de suma. Las líneas de  $S_{princ}$  son los componentes de frecuencia en la banda de frecuencia actual y cada columna corresponde a una fuente principal.

Cabe señalar que el número de fuentes principales  $n_{princ}$  es preferentemente inferior o igual al número  $n_f$  de canales de la señal de suma para asegurar que la operación se pueda invertir y posiblemente puede ser diferente para cada banda de frecuencia.

60 Cuando la escena es compleja, puede suceder que el número de fuentes a reconstruir en la banda de frecuencia actual para obtener una reconstrucción satisfactoria de la escena sea superior al número de canales de la señal de suma.

65 En este caso, las fuentes adicionales o secundarias se codifican y luego se decodifican a partir del flujo binario para la banda actual por el módulo 550 de decodificación del flujo binario.

Este módulo de decodificación decodifica las informaciones contenidas en el flujo binario y, en particular, las informaciones de patrones y si es necesario las fuentes secundarias.

La decodificación de las fuentes secundarias se efectúa mediante las operaciones inversas a las que se efectuaron con la codificación.

5 Cualquiera que sea el método de codificación que se ha retenido para fuentes secundarias, si se han transmitido datos de reconstrucción o informaciones de codificación de las fuentes secundarias en el flujo binario para la banda actual, los datos correspondientes se decodifican para reconstruir la matriz  $S_{sec}$  de los coeficientes de frecuencia en la banda actual de las  $n_{sec}$  fuentes secundarias. La forma de la matriz  $S_{sec}$  es similar a la matriz  $S_{princ}$ , es decir, que las líneas son los componentes de frecuencia en la banda de frecuencia actual y cada columna corresponde a una fuente secundaria.

15 Por lo tanto, es posible construir la matriz completa  $S$  de los coeficientes de frecuencia del conjunto de las  $n_{tot}=n_{princ}+n_{sec}$  fuentes necesarias para la reconstrucción de la señal multicanal en la banda considerada, obtenida reagrupando las dos matrices  $S_{princ}$  y  $S_{supp}$  según la relación  $S = (S_{princ} \ S_{supp})$ .  $S$  es por lo tanto una matriz de dimensión  $n_{bin} \times n_{tot}$ . También, la forma es idéntica a las matrices  $S_{princ}$  y  $S_{supp}$ : las líneas son los componentes de frecuencia en la banda de frecuencia actual, cada columna es una fuente, con  $n_{tot}$  fuentes en total.

20 En paralelo con la reconstrucción de las fuentes que acaba de describirse, se realiza la reconstrucción de los patrones.

Las informaciones de patrones se extraen del flujo binario en la etapa Decod. Fb por el módulo 550.

25 Las posibles salidas de este módulo de decodificación del flujo binario dependen de los métodos de codificación de los patrones utilizados en la codificación. Pueden estar en la forma de vectores de patrones de base de  $D_B$  básicos y de coeficientes asociados  $G_D$  y/o de parámetros de modelado  $P$ .

Estos datos se transmiten luego a un módulo de reconstrucción de las informaciones de patrones 560 que efectúa la decodificación de las informaciones de patrones mediante operaciones inversas a las efectuadas en la codificación.

30 El número de patrones a reconstruir es igual al número  $n_{tot}$  de fuentes en la banda de frecuencia considerada, siendo cada fuente asociada a un vector de patrón.

35 En el caso de la representación de los patrones a partir del patrón de base, la matriz de patrones  $D_i$  se escribe como la combinación lineal de estos patrones de base. Así que se puede escribir  $D_i = G_d D_B$ , donde  $D_B$  es la matriz de los patrones de base para el conjunto de las bandas y  $G_d$  la matriz de las ganancias asociadas. Esta matriz de ganancia tiene un número de líneas igual al número total de fuentes  $n_{tot}$  y un número de columnas igual al número de vectores de patrón de base.

40 En una variante de este modo de realización, se decodifican patrones de base por grupo de bandas de frecuencia consideradas, para representar más fielmente los patrones. Como se explica para la codificación, se puede, por ejemplo, proporcionar dos grupos patrones de base: uno para bajas frecuencias y uno para las altas frecuencias. Un vector de ganancias asociado con los patrones de base luego se decodifica para cada banda.

45 Al final se reconstruyen tantos patrones como fuentes. Estos patrones se reagrupan en una matriz  $D_i$  donde las líneas corresponden a los valores de ángulo (tanto valor de ángulo como de canales en la señal multicanal a reconstruir) y cada columna corresponde al patrón de la fuente correspondiente, es decir, que la columna  $r$  de  $D_i$  del patrón de la fuente que está en la columna  $r$  de  $S$ .

50 A partir de la matriz  $S$  de los coeficientes de las fuentes y de la matriz  $D$  de los patrones asociados, los coeficientes de frecuencia de la señal multicanal reconstruida en la banda se calculan en el módulo de espacialización 530 en la etapa SPAT., según la relación:

$Y=SD$ , donde  $Y$  es la señal reconstruida en la banda. Las líneas de la matriz  $Y$  son los componentes de frecuencia en la banda de frecuencia actual y cada columna corresponde a un canal de la señal multicanal a reconstruir.

55 Al reproducir el mismo procesamiento en cada una de las bandas de frecuencia, se reconstruyen las transformadas de Fourier completas de los canales de señal a reconstruir para la trama temporal actual. Las señales temporales correspondientes entonces se obtienen por transformada de Fourier inversa  $T^{-1}$ , utilizando un algoritmo rápido implementado por el módulo de transformada inversa 540.

60 Se obtiene de este modo la señal multicanal  $S_m$  en la trama temporal actual. Las diferentes tramas temporales se combinan luego mediante el método convencional de adición con recubrimiento (u "overlap-add" en inglés) para reconstruir la señal multicanal completa.

65 De manera general, los suavizados temporales o de frecuencia de los parámetros se podrán usar tanto para el análisis como para la síntesis para asegurar transiciones suaves en la escena sonora. Se podrá reservar una señalización de cambio repentino de la escena sonora en el flujo binario para evitar suavizar el decodificador en el caso de detección

de un cambio rápido en la composición de la escena sonora. Por otra parte, se pueden utilizar métodos convencionales de adaptación de la resolución del análisis de tiempo-frecuencia (cambio en el tamaño de las ventanas de análisis y de síntesis a lo largo del tiempo).

5 De la misma manera que el codificador, un módulo de cambio de base puede efectuar un procesamiento previo  $P^{-1}$  para obtener una descomposición en ondas planas de las señales, un módulo de cambio de base 570 efectúa la operación inversa a partir de las señales en ondas planas para encontrar la señal multicanal original.

10 La codificación del modo de realización descrita con referencia a la figura 2 permite obtener una compresión eficaz cuando la complejidad de la escena sigue siendo limitada. Cuando la complejidad de la escena es mayor, es decir, cuando la escena contiene un gran número de fuentes activas en una banda de frecuencia o componentes grandes difusos, se necesita un número importante de fuentes y de patrón asociado para obtener una buena calidad de restitución de la escena. La eficacia de la compresión se reduce entonces.

15 Una variante de realización del procedimiento de codificación y un codificador que implementa este procedimiento se describe con referencia a la figura 6. Esta variante de realización permite mejorar la eficacia de codificación para las escenas complejas.

20 Para ello, el codificador tal como se representa en la figura 6 consta de los módulos 215, 210, 220, 230, 240 tales como los descritos con referencia a la figura 2.

También consta de los módulos 260, 270 y 290 tal como se describe con referencia a la figura 2.

25 Este codificador, sin embargo, consta de un módulo de codificación de las fuentes secundarias 620, que difiere del módulo 280 de la figura 2 en el caso donde el número de fuentes secundarias es importante.

En el caso de esta figura, este módulo de codificación 620 implementa un método de codificación paramétrica de fuentes secundarias.

30 Para ello, se tienen en cuenta los límites de la percepción auditiva espacial. En bandas de frecuencia donde el número de fuentes secundarias es importante, el campo es asimilable perceptivamente a un campo difuso y la representación del campo por una o varias características estadísticas del campo es suficiente para reconstruir un campo perceptivamente equivalente.

35 Este principio es asimilable al principio más utilizado convencionalmente en la codificación de audio para la representación de componentes ruidosos. De hecho, estos componentes se codifican comúnmente en forma de ruido blanco filtrado con características de filtrado que varían con el tiempo. Para reconstruir estos componentes de manera perceptivamente satisfactoria, solo es necesario el conocimiento de las características del filtrado (la envoltura espectral), pudiendo cualquier ruido blanco utilizarse durante la reconstrucción.

40 En el ámbito de la presente invención, se utiliza el hecho de que los componentes espacialmente difusos de la escena sonora pueden reconstruirse perceptivamente a partir del simple conocimiento del patrón correspondiente y controlando la coherencia del campo creado. Esto se puede hacer utilizando pseudofuentes construidas por descorrelación, a partir de un número limitado de fuentes transmitidas y utilizando los patrones de los componentes difusos estimados en la señal multicanal original. El objetivo es entonces reconstruir un campo sonoro estadística y perceptivamente equivalente al original, incluso si está constituido por señales cuyas formas de ondas son diferentes.

45 De este modo, para implementar este método, cierto número de fuentes secundarias no se transmiten y se reemplazan por pseudofuentes obtenidas por descorrelación de las fuentes transmitidas o por cualquier otra fuente artificial no correlacionada transmitida. Se evita de este modo la transmisión de datos correspondientes a estas fuentes y se mejora significativamente la eficacia de la codificación.

50 En un primer modo de realización, se elige una fuente para ser transmitida al decodificador y un descorrelador predefinido conocido tanto por el codificador como por el decodificador, para aplicar a la fuente transmitida para construir, al decodificador de las pseudofuentes.

En este modo de realización, por lo tanto, no es necesario transmitir datos de descorrelación, sino al menos una fuente que sirve como base para esta descorrelación debe ser transmitida (de manera efectiva y no paramétrica).

60 En un segundo modo de realización, el módulo de codificación de las fuentes secundarias 620 obtiene una representación paramétrica de las fuentes secundarias y también se transmite al módulo de construcción del flujo binario.

65 Esta representación paramétrica de fuentes secundarias o de fuentes difusas se efectúa, por ejemplo, mediante una envoltura espectral. También se puede utilizar una envoltura temporal.

En una variante de este modo de realización, las pseudofuentes se calculan mediante un módulo 630 de descorrelación que calcula las fuentes descorrelacionadas a partir de al menos una fuente principal o con al menos una fuente secundaria codificada para ser transmitida.

5 Se pueden usar varios descorreladores y varias fuentes iniciales y se puede seleccionar la fuente inicial asociada con un tipo de descorrelador que dé el mejor resultado de reconstrucción. Estos datos de descorrelación como, por ejemplo, el índice del correlador utilizado y los datos de elección de la fuente inicial como el índice de la fuente, luego se transmiten al módulo de construcción del flujo binario para su inserción.

10 El número de fuentes a transmitir se reduce, por lo tanto, mientras se mantiene una buena calidad perceptiva de la señal reconstruida.

La figura 7 representa un decodificador y un procedimiento de decodificación adaptado a la codificación según la variante de realización descrita en la figura 6.

15 Este decodificador consta de los módulos 510, 520, 530, 540, 570, 560 tales como se describe con referencia a la figura 5. Este decodificador difiere del descrito en la figura 5 por las informaciones decodificadas por el módulo de decodificación del flujo binario 720 y por el bloque de cálculo de descorrelación 710.

20 En efecto, el módulo 720 obtiene, además de informaciones de patrones, las fuentes de la escena sonora y, si corresponde, fuentes secundarias decodificadas, datos de representación paramétricos de ciertas fuentes secundarias o fuentes difusas y, posiblemente, informaciones sobre el descorrelador y las fuentes transmitidas a usar para reconstruir las pseudofuentes.

25 Estas últimas informaciones se usan luego por el módulo de descorrelación 710 que permite reconstruir las pseudofuentes secundarias que se combinarán con las fuentes principales y las otras fuentes secundarias potenciales en el módulo de espacialización como se describe con referencia a la figura 5.

30 Los codificadores y decodificadores tales como los descritos con referencia a las figuras 2, 6 y 5, 7 pueden integrarse en un equipo multimedia del tipo de decodificador de salón, ordenador o incluso equipo de comunicación, tal como un teléfono móvil o una agenda electrónica personal.

35 La figura 8a representa un ejemplo de tal equipo multimedia o dispositivo de codificación que consta de un codificador según la invención. Este tipo de dispositivo consta de un procesador PROC que coopera con un bloque de memoria BM que consta de una memoria de almacenamiento y/o de trabajo MEM.

El bloque de memoria puede constar, ventajosamente, de un programa informático que consta de instrucciones de código para la implementación de las etapas del procedimiento de codificación en el sentido de la invención, cuando estas instrucciones se ejecutan por el procesador PROC y, en particular, las etapas de

- 40
- descomposición (T) de la señal multicanal en bandas de frecuencia y las siguientes etapas por banda de frecuencia:
  - obtención de informaciones de patrón por fuente sonora de la escena sonora, siendo las informaciones representativas de la distribución espacial de la fuente sonora en la escena sonora;
  - selección de un conjunto de fuentes sonoras de la escena sonora que constituyen unas fuentes principales;
  - 45 - masterización de las fuentes principales seleccionadas para obtener una señal de suma con un número reducido de canales;
  - codificación de las informaciones de patrón y formación de un flujo binario que consta de las informaciones de patrones codificadas, siendo el flujo binario adecuado para transmitirse en paralelo a la señal de suma.

50 Típicamente, la descripción de la figura 2 retoma las etapas de un algoritmo de tal programa informático. El programa informático puede estar almacenado, igualmente, sobre un soporte de memoria legible por un lector del dispositivo o descargable en el espacio de memoria del equipo.

55 El dispositivo consta de un módulo de entrada adecuado para recibir una señal multicanal que representa una escena sonora, ya sea por una red de comunicación, o por lectura de un contenido almacenado en un medio de almacenamiento. Este equipo multimedia también puede constar de medios de captura de tal señal multicanal.

El dispositivo consta de un módulo de salida adecuado para transmitir un flujo binario Fb y una señal de suma Ss procedente de la codificación de la señal multicanal.

60 De la misma forma, la figura 8b ilustra un ejemplo de equipo multimedia o dispositivo de decodificación que consta de un decodificador según la invención.

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

El bloque de memoria puede constar, ventajosamente, de un programa informático que consta de instrucciones de código para la implementación de las etapas del procedimiento de decodificación en el sentido de la invención, cuando estas instrucciones se ejecutan por el procesador PROC y, en particular, las etapas de:

- 5
- extracción en el flujo binario y decodificación de informaciones de patrones representativas de la distribución espacial de las fuentes en la escena sonora;
  - demasterización de la señal de suma para obtener un conjunto de fuentes principales;
  - reconstrucción de la señal de audio multicanal mediante espacialización de al menos las fuentes principales con las informaciones de patrones decodificadas.

10

Típicamente, la descripción de la figura 5 retoma las etapas de un algoritmo de tal programa informático. El programa informático puede estar almacenado, igualmente, sobre un soporte de memoria legible por un lector del dispositivo o descargable en el espacio de memoria del equipo.

15

El dispositivo consta de un módulo de entrada adecuado para recibir un flujo binario  $F_b$  y una señal de suma  $S_s$  que proviene, por ejemplo, de una red de comunicación. Estas señales de entrada pueden provenir de una lectura en un medio de almacenamiento.

20

El dispositivo consta de un módulo de salida adecuado para transmitir una señal multicanal decodificada por el procedimiento de decodificación implementado por el equipo.

Este equipo multimedia también puede constar de medios de restitución de tipo altavoz o medios de comunicación adecuados para transmitir esta señal multicanal.

25

Muy evidentemente, tal equipo multimedia puede constar tanto del codificador como del decodificador según la invención. Como la señal de entrada es entonces la señal multicanal original y la señal de salida, la señal multicanal decodificada.

## REIVINDICACIONES

- 5 1. Procedimiento de codificación de una señal de audio multicanal que representa una escena sonora que consta de una pluralidad de fuentes sonoras, caracterizado por que consta de una etapa de descomposición (T) de la señal multicanal en bandas de frecuencia y las siguientes etapas por banda de frecuencia:
- obtención (OBT) de informaciones de patrón por fuente sonora de la escena sonora, siendo las informaciones de patrón representativas de al menos una dirección y de un ancho angular de la fuente sonora en la escena sonora;
  - 10 - selección (Select) de un conjunto de fuentes sonoras de la escena sonora que constituyen unas fuentes principales;
  - masterización (M) de las fuentes principales seleccionadas para obtener una señal de suma con un número reducido de canales;
  - 15 - codificación (Cod.Di) de las informaciones de patrón y formación (Con.Fb) de un flujo binario que consta de las informaciones de patrones codificadas, siendo el flujo binario adecuado para transmitirse en paralelo a la señal de suma.
- 20 2. Procedimiento de codificación según la reivindicación 1, caracterizado por que consta, además, de una etapa de codificación (Cod  $S_{sec}$ ) de fuentes secundarias de entre las fuentes no seleccionadas de la escena sonora y de inserción de informaciones de codificación de las fuentes secundarias en el flujo binario.
3. Procedimiento según la reivindicación 2, caracterizado por que las informaciones de codificación de las fuentes secundarias son envolturas espectrales codificadas de las fuentes secundarias.
- 25 4. Procedimiento según la reivindicación 2, caracterizado por que la codificación de las fuentes secundarias consta de las siguientes etapas:
- construcción de pseudofuentes que representen al menos una parte de las fuentes secundarias, por descorrelación con al menos una fuente principal y/o al menos una fuente secundaria codificada;
  - 30 - codificación de las pseudofuentes construidas; e
  - inserción en el flujo binario de un índice de fuente utilizada y de un índice de descorrelador utilizado para la etapa de construcción.
- 35 5. Procedimiento según la reivindicación 1, caracterizado por que la codificación de las informaciones de patrones se efectúa mediante un método de representación paramétrica.
6. Procedimiento según la reivindicación 5, caracterizado por que la representación paramétrica consta de informaciones de dirección de llegada, para la reconstrucción de un patrón simulando una onda plana.
- 40 7. Procedimiento según la reivindicación 5, caracterizado por que la representación paramétrica consta de índices de selección de forma de patrón en un diccionario de formas de patrones.
8. Procedimiento según la reivindicación 1, caracterizado por que la codificación de las informaciones de patrón se efectúa mediante un método de análisis de componente principal que entrega vectores de patrón de base asociados con ganancias que permiten la reconstrucción de los patrones iniciales.
- 45 9. Procedimiento según la reivindicación 1, caracterizado por que la codificación de las informaciones de patrón se efectúa mediante una combinación de un método de análisis de componente principal y de un método de representación paramétrica.
- 50 10. Procedimiento de decodificación de una señal de audio multicanal que representa una escena sonora que consta de una pluralidad de fuentes sonoras, a partir de un flujo binario y de una señal de suma, caracterizado por que consta de las etapas siguientes:
- extracción (decod.Fb) en el flujo binario y decodificación de informaciones de patrones representativas de al menos una dirección y de un ancho angular de las fuentes en la escena sonora;
  - 55 - demasterización (N) de la señal de suma para obtener un conjunto de fuentes principales;
  - reconstrucción (SPAT) de la señal de audio multicanal mediante espacialización de al menos las fuentes principales con las informaciones de patrones decodificadas.
- 60 11. Procedimiento de decodificación según la reivindicación 10, caracterizado por que consta, además, de las etapas siguientes:
- extracción del flujo binario, de informaciones de codificación de fuentes secundarias codificadas;
  - 65 - decodificación de las fuentes secundarias a partir de las informaciones de codificación extraídas;
  - reagrupación de las fuentes secundarias con las fuentes principales para la espacialización.

12. Procedimiento de decodificación según la reivindicación 11, caracterizado por que consta, además, de la etapa siguiente:

- 5 - decodificación de las fuentes secundarias mediante uso de una fuente efectivamente transmitida y de un descorrelador predefinido para reconstruir unas pseudofuentes representativas de al menos una parte de las fuentes secundarias.

13. Procedimiento de decodificación según la reivindicación 11, caracterizado por que consta, además, de las etapas siguientes:

- 10 - extracción del flujo binario, de un índice de fuente principal y/o de al menos una fuente secundaria codificada y de un índice de un descorrelador a aplicar a esta fuente;  
- decodificación de las fuentes secundarias mediante uso de la fuente y del índice de descorrelador para reconstruir pseudofuentes representativas de al menos una parte de las fuentes secundarias.

14. Codificador de una señal de audio multicanal que representa una escena sonora que consta de una pluralidad de fuentes sonoras, caracterizado por que incluye:

- 20 - un módulo (210) de descomposición de la señal multicanal en banda de frecuencia;  
- un módulo (220) de obtención de informaciones de patrón adecuado para obtener estas informaciones por fuente sonora de la escena sonora y por banda de frecuencia, siendo las informaciones de patrón representativas de al menos una dirección y de un ancho angular de la fuente sonora en la escena sonora;  
- un módulo (260) de selección de un conjunto de fuentes sonoras de la escena sonora que constituyen unas fuentes principales;  
25 - un módulo (270) de masterización de las fuentes principales procedentes del módulo de selección para obtener una señal de suma con un número reducido de canales;  
- un módulo (230) de codificación de las informaciones de patrón y un módulo de formación de un flujo binario que consta de las informaciones de patrones codificadas, siendo el flujo binario adecuado para transmitirse en paralelo a la señal de suma.

15. Decodificador de una señal de audio multicanal que representa una escena sonora que consta de una pluralidad de fuentes sonoras, que recibe a la entrada un flujo binario y una señal de suma, caracterizado por que incluye:

- 35 - un módulo (550) de extracción y de decodificación de informaciones de patrones representativas de al menos una dirección y de un ancho angular de las fuentes en la escena sonora;  
- un módulo (520) de demasterización de la señal de suma para obtener un conjunto de fuentes principales;  
- un módulo (530) de reconstrucción de la señal de audio multicanal mediante espacialización de al menos las fuentes principales con las informaciones de patrones decodificadas.

40 16. Programa informático que consta de unas instrucciones de código para la implementación de las etapas de un procedimiento de codificación según una de las reivindicaciones 1 a 9 y/o de un procedimiento de decodificación según una de las reivindicaciones 10 a 13, cuando estas instrucciones se ejecutan por un procesador.

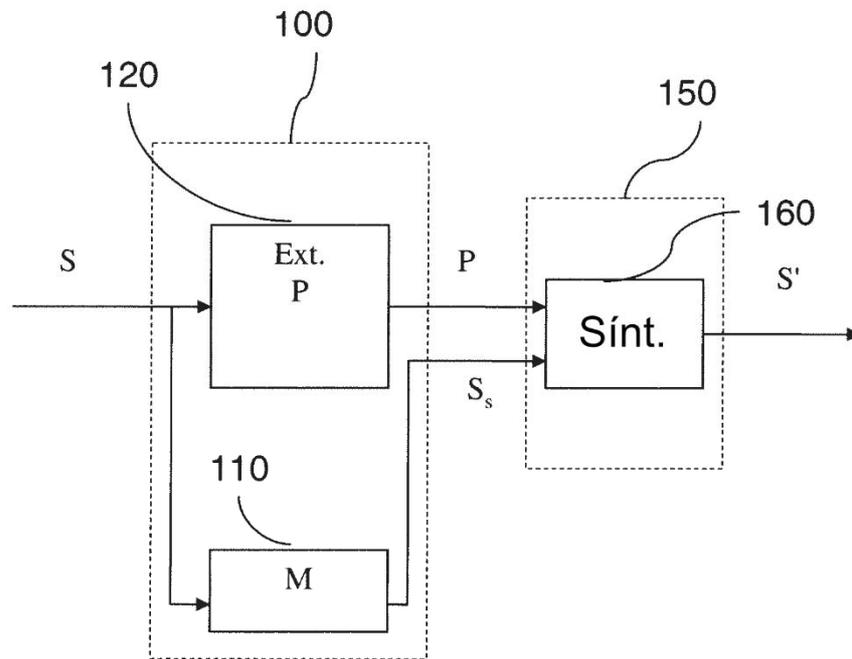


Fig.1 (Estado de la técnica)

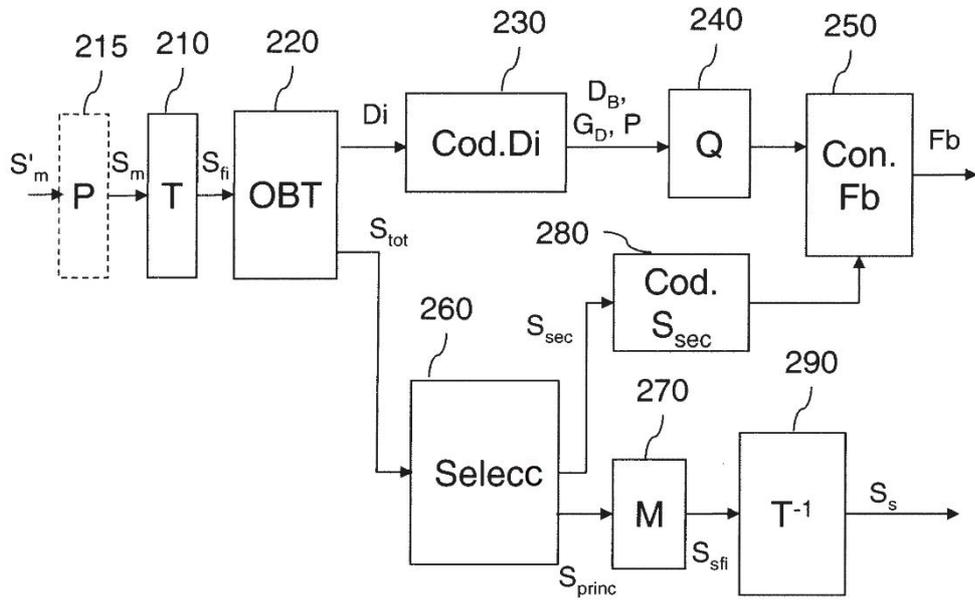


Fig.2

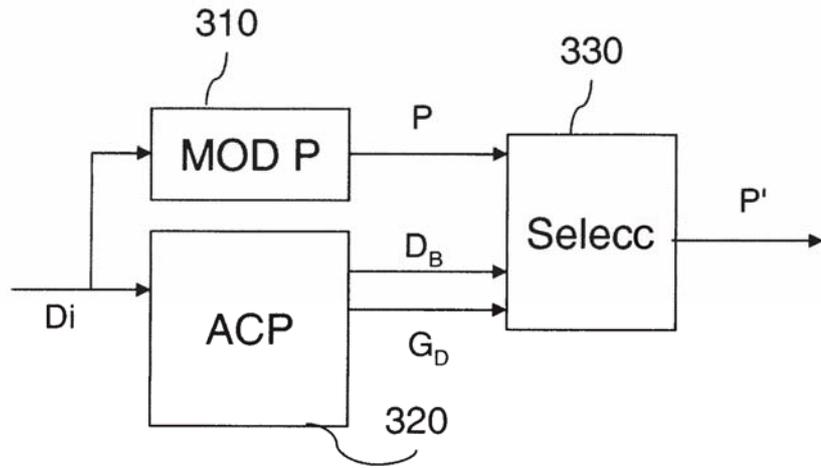


Fig.3a

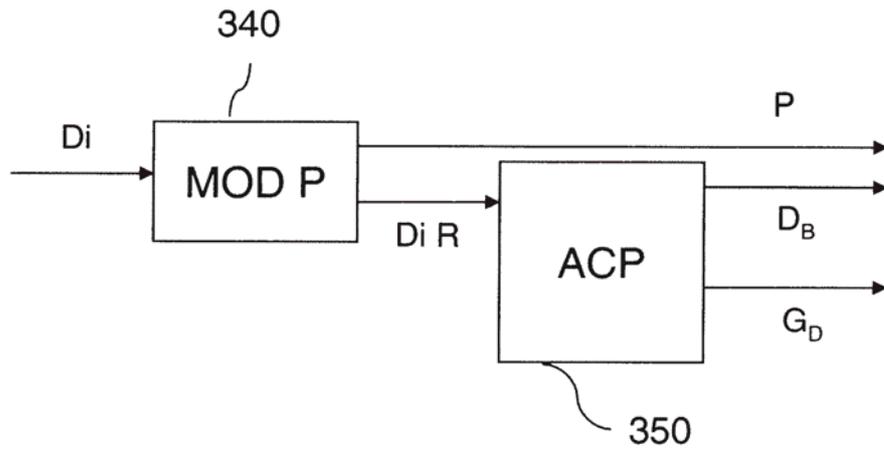


Fig.3b

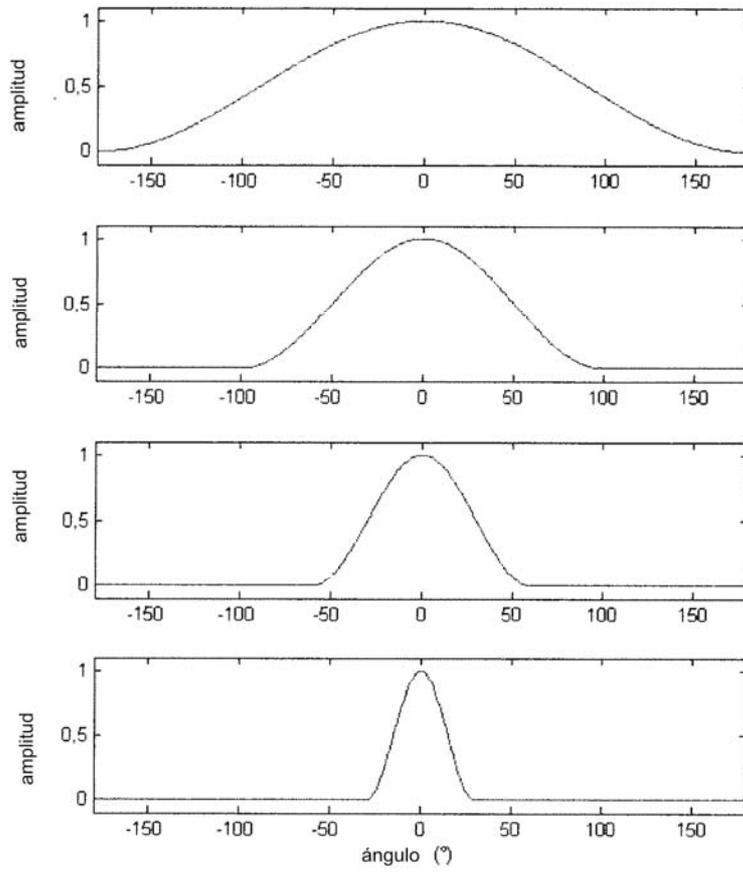


Fig.4

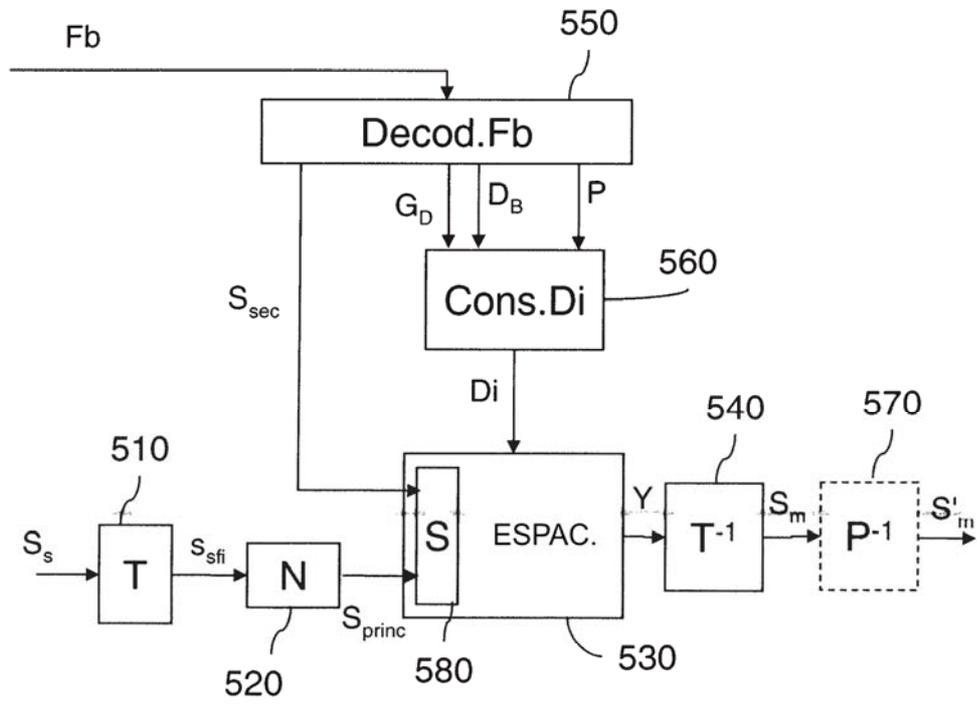


Fig.5

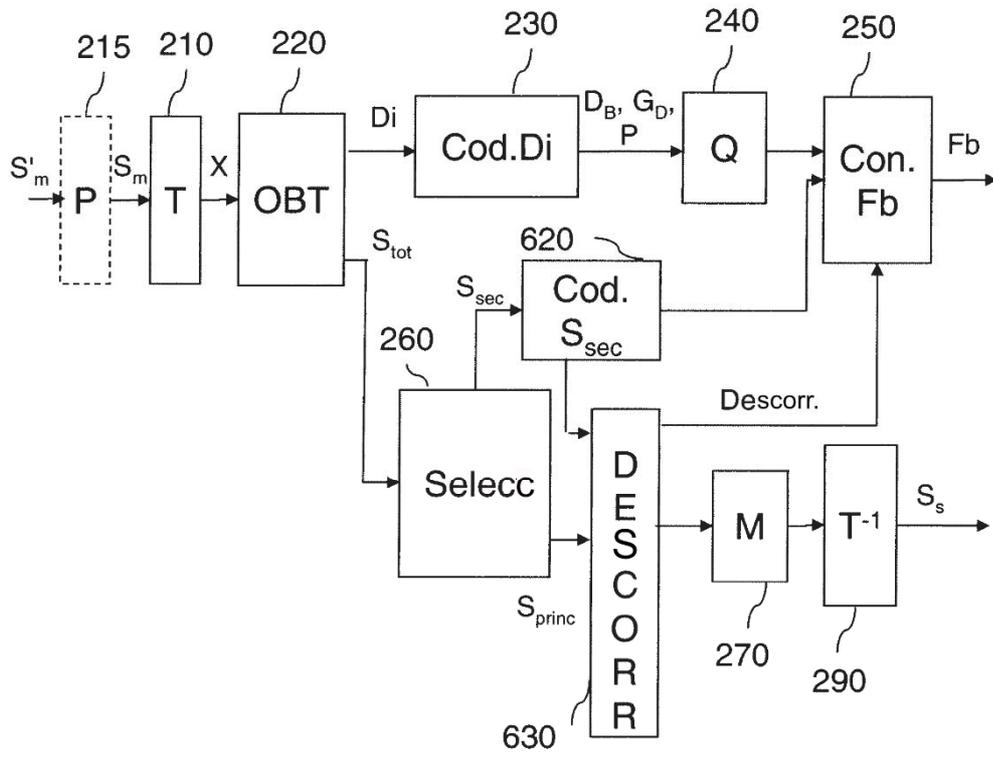


Fig.6

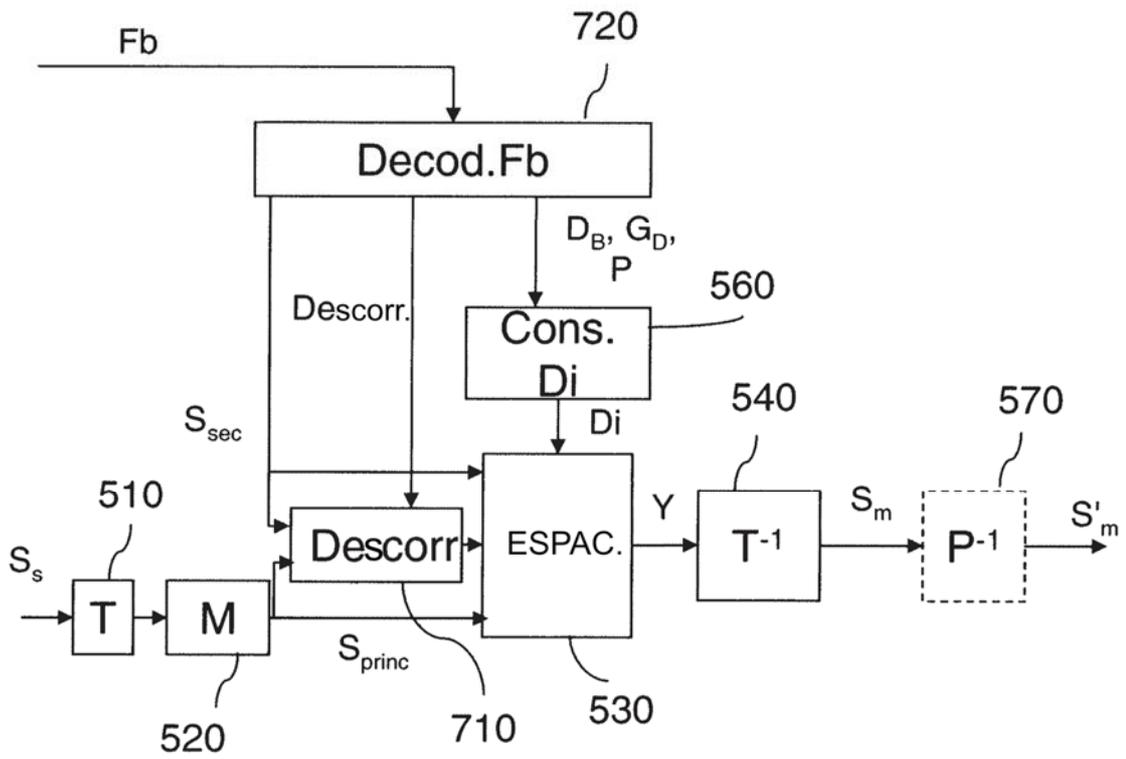


Fig.7

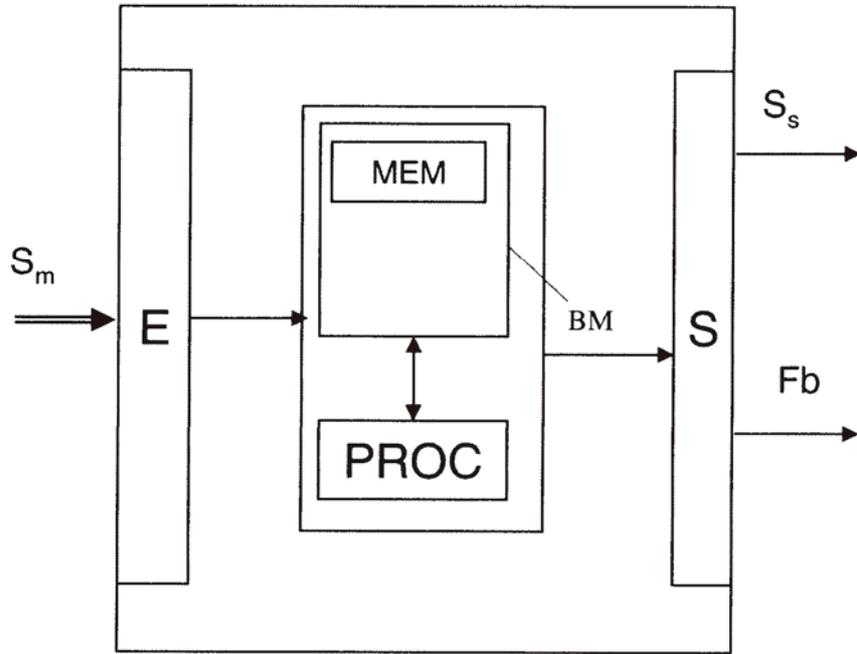


Fig.8a

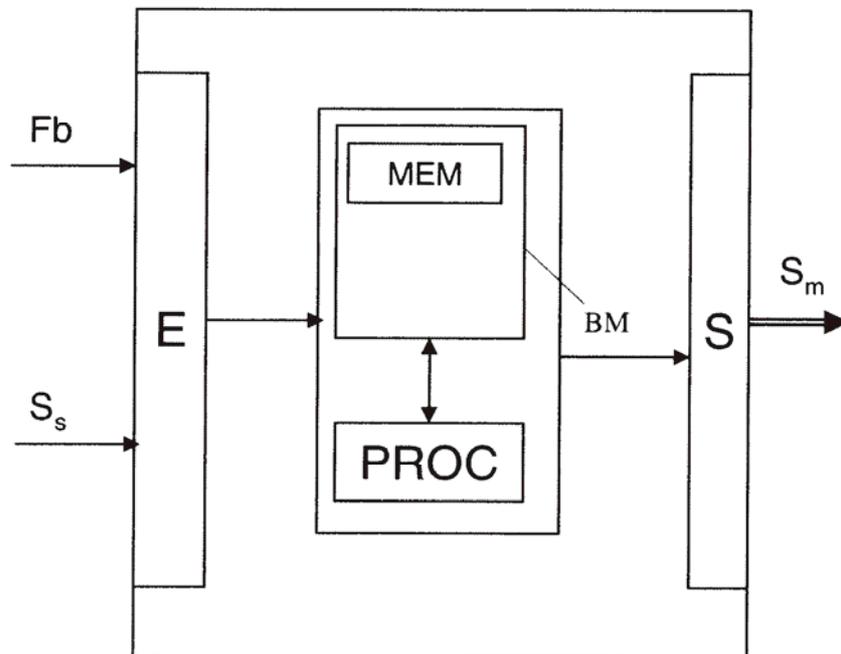


Fig.8b