(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 387 869**

(51) Int. CI.:
G10L 19/00 (2006.01)
H04S 3/00 (2006.01)
H04S 1/00 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Número de solicitud europea: **09802568 .7**
(96) Fecha de presentación: **03.07.2009**
(97) Número de publicación de la solicitud: **2319037**
(97) Fecha de publicación de la solicitud: **11.05.2011**

(54) Título: **Reconstrucción de datos de audio multicanal**

(30) Prioridad:
30.07.2008 FR 0855249

(45) Fecha de publicación de la mención BOPI:
03.10.2012

(45) Fecha de la publicación del folleto de la patente:
03.10.2012

(73) Titular/es:
France Télécom
6 place d'Alleray
75015 Paris, FR

(72) Inventor/es:
VIRETTE, David y
PHILIPPE, Pierrick

(74) Agente/Representante:
Pérez Barquín, Eliana

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

Reconstrucción de datos de audio multicanal

La invención se refiere a la ocultación de datos de espacialización defectuosos, para la reconstrucción de datos de audio multicanal. Los datos de audio multicanal son típicamente reconstruidos a partir de al menos datos de espacialización y de datos de audio en un número de canales restringido, por ejemplo datos monocanal.

Los datos de audio multicanal están destinados típicamente a varias pistas de audio respectivas. Pueden utilizarse varias fuentes sonoras respectivas para contribuir a proporcionar al oyente la ilusión de un sonido envolvente.

Los datos de audio multicanal pueden comprender por ejemplo datos en estéreo en dos canales, o bien incluso datos 5.1 en seis canales, en particular para aplicaciones de Home Cinema. La invención también puede aplicarse en el campo de las conferencias de audio espacializadas, en el que los datos que corresponden a un locutor experimentan un tratamiento de espacialización para proporcionar al oyente la ilusión de que la voz de este locutor proviene de una posición particular del espacio.

Los datos de espacialización se utilizan para obtener datos multicanal a partir de los datos en un número inferior de canales, por ejemplo datos monocanal. Estos datos de espacialización pueden comprender por ejemplo diferentes diferencias de nivel inter-vías o ILD (del inglés “Interchannel Level Difference”), correlaciones inter-vías o ICC (del inglés “Interchannel Cross Correlation”), retardos entre vías o ITD (del inglés “Interchannel Time Difference”), diferencias de fase entre vías o IPD (del inglés “Interchannel Phase Difference”), u otras.

Puede ocurrir que los datos de audio recibidos, que comprenden al menos los datos monocanal y los datos de espacialización, sean defectuosos, es decir que algunos datos falten o que sean erróneos.

La detección de esta transmisión defectuosa puede realizarse por medio de un código de tipo CRC (del inglés “Cyclic Redundancy Check”).

Se conoce el alivio de estos defectos sustituyendo valores defectuosos por valores predichos. Estos valores predichos pueden determinarse siguiendo un modelo de predicción conocido.

Se conocen varios modelos de predicción. Por ejemplo, se selecciona como valor predicho un valor arbitrario, un valor precedente, un valor determinado a partir de los datos de audio recibidos anteriormente siguiendo, por ejemplo, procedimientos de predicción lineal, u otros.

Cuando se reciben de forma defectuosa datos monocanal, la sustitución de los valores defectuosos por valores predichos de datos monocanal se revela en general relativamente satisfactoria.

Sin embargo, cuando se reciben de forma defectuosa datos de espacialización, la sustitución de los valores defectuosos por valores predichos puede revelarse insatisfactoria.

Las variaciones fuertes de los datos de espacialización a lo largo del tiempo se traducen para el oyente en la sensación de desplazamientos bruscos de las fuentes sonoras.

Por ejemplo, si se sustituyen los valores defectuosos por un valor arbitrario correspondiente a una ausencia de espacialización, la sensación de un retorno a un sonido monocanal puede ser molesta para el oyente, en particular en el caso de señales binaurales. En efecto, las señales binaurales, es decir que permiten una restitución fiel del espacio 3D a nivel de las orejas, corresponden a menudo a fuentes sonoras virtuales relativamente fijas en el espacio.

Existe, por lo tanto, una necesidad de una mejor ocultación de los defectos de los datos de espacialización durante la reconstrucción de datos de audio multicanal. Por ejemplo, el documento US 2005/182996 (BRUHN STEFAN), 18 de agosto de 2005, describe la construcción de un modelo de predicción sobre la base de una combinación apropiada de canales.

De acuerdo con un primer aspecto, la invención tiene por objeto un procedimiento de procesamiento de datos sonoros, para la reconstrucción de datos de audio multicanal a partir de, al menos, datos en un número de canales restringido y de datos de espacialización, comprendiendo este procedimiento, tal como se define en la reivindicación 1, una etapa de test de la validez de datos de espacialización de una trama recibida. Si este test muestra que estos datos de espacialización son válidos:

a/ mediante modelo respectivo de una pluralidad de modelos de predicción, se predice, según este modelo, un valor de espacialización,

b/ se selecciona un modelo de predicción, a partir de los valores de espacialización predichos de este modo y, a

partir de los datos de espacialización recibidos efectivamente, para poder, en caso de recepción posterior de datos de espacialización defectuosos, predecir según este modelo seleccionado un valor de espacialización, y utilizar este valor de espacialización predicho para la reconstrucción de los datos de audio multicanal.

Durante la etapa b/, se confronta cada valor de espacialización predicho con un valor estimado a partir de los datos de espacialización recibidos. En particular, puede preverse calcular, por modelo, un valor de semejanza a partir, por un lado, del valor de espacialización predicho siguiendo este modelo y, por otro lado, de un valor estimado a partir de los datos de espacialización recibidos. Se selecciona entonces el modelo de predicción para el cual el valor de semejanza indica una mayor adecuación entre el valor predicho y el valor estimado.

De este modo, se utilizan datos de espacialización considerados como válidos para seleccionar entre una pluralidad de modelos de predicción un modelo de predicción a adoptar en caso de recepción de datos de espacialización considerados como defectuosos. Dicho procedimiento adaptativo según el contenido permite paliar los defectos de los datos de espacialización de forma más satisfactoria que en la técnica anterior en la que se utiliza un solo modelo de predicción.

Por “un número de canales restringido”, se entiende un número de canales inferior al número de canales de los datos multicanal. Por ejemplo, los datos en un número de canales restringido pueden comprender datos monocanal.

Los datos de espacialización y, de forma más general, los datos de audio recibidos, pueden provenir de un canal de transmisión. Por ejemplo, estos datos pueden recibirse por Internet. Como alternativa, los datos de audio recibidos pueden leerse en un soporte de almacenamiento, por ejemplo un DVD (del inglés “Digital Versatile Disk”), u otro. La invención no está en absoluto limitada por la procedencia de los datos de audio recibidos.

Los datos de audio recibidos pueden comprender una señal codificada, una señal demultiplexada y/o decodificada, valores numéricos, u otros.

Las etapas a/ y b/ pueden realizarse sistemáticamente después de la recepción de una trama considerada como válida. De este modo, los procesamientos se distribuyen en el tiempo.

Puede preverse, en particular cuando las etapas a/ y b/ se realizan para cada trama válida, escribir en la memoria un identificador del modelo de predicción seleccionado y esto para poder, en caso de recepción posterior de datos de espacialización defectuosos, recuperar rápidamente el modelo de predicción a aplicar.

Como alternativa, la ejecución de las etapas a/ y/o b/ puede estar sometida a la realización de ciertas condiciones, lo que puede permitir evitar realizar cálculos inútiles.

Por ejemplo, cuando una trama se considera como válida, los datos de espacialización se almacenan en una memoria, al menos de forma temporal. Las etapas a/ y b/ se realizan (a partir de los datos almacenados de este modo), solamente en caso de recepción posterior de datos de espacialización considerados como defectuosos. De este modo, se evita realizar en particular las predicciones de la etapa a/ cuando esto no es necesario.

De acuerdo con otro ejemplo, puede preverse realizar las predicciones de la etapa a/ sistemáticamente después de la recepción de una trama considerada como válida, mientras que la etapa b/ solamente se realiza (a partir de los datos de espacialización de la trama o las tramas precedentes, conservadas en la memoria) en caso de recepción de una trama defectuosa.

El valor estimado puede ser uno de los datos de espacialización, por ejemplo el valor estimado puede comprender un ILD. En este caso, puede preverse, durante la etapa b/ comparar los valores de espacialización predichos directamente con datos de espacialización recibidos.

Como alternativa, el valor estimado puede derivar solamente de los datos de espacialización. Por ejemplo el valor estimado puede comprender una ganancia procedente de los ILD para una trama y una banda de frecuencias dadas, un retardo u otro. En este caso, puede preverse, durante la etapa b/, comparar los valores de espacialización predichos con valores obtenidos a partir de datos de espacialización recibidos.

Ventajosamente, para al menos un modelo, se confrontan además valores de espacialización predichos anteriormente con valores estimados correspondientes. De este modo, la elección del modelo de predicción más adecuado para el contenido puede realizarse de forma más precisa.

Por ejemplo, pueden utilizarse los datos de espacialización recibidos en varias tramas y confrontar, para varias tramas, los valores predichos y los valores estimados.

En particular, por trama de una secuencia de tramas recibidas, y para al menos un modelo, puede predecirse, siguiendo este modelo, un valor de espacialización, de modo que se predice una secuencia de valores de espacialización. Para este modelo, el valor de semejanza puede calcularse a partir, por un lado, de esta secuencia

de valores de espacialización predichos y, por otro lado, de una secuencia de valores estimados a partir de los datos de la secuencia de tramas.

Ventajosamente, no se utilizarán datos de espacialización defectuosos durante la etapa de elección del modelo de predicción, para evitar falsear esta elección.

Como alternativa, puede bastar con los datos de espacialización corrientes recibidos, por ejemplo, en una misma trama, para la elección del modelo de predicción.

Los datos pueden ser defectuosos, debido a degradaciones introducidas durante la transmisión, o a degradaciones de un soporte de almacenamiento de los datos. La invención no está limitada a este origen de defectos. Por ejemplo, en el caso de una transmisión jerarquizada en capas ("scalable coding" en inglés) para la cual un emisor u otro elemento de una red de transmisión puede elegir no transmitir un conjunto de datos, pueden faltar datos entre los datos de espacialización recibidos.

El carácter defectuoso de los datos de espacialización puede detectarse según procedimientos conocidos, por ejemplo por medio de un código de tipo CRC.

La invención no está en absoluto limitada por la forma de la escritura en memoria del identificador del modelo de predicción elegido. Se puede, por ejemplo, copiar en una memoria programa todas las instrucciones de un programa correspondiente a este modelo, o bien simplemente memorizar un nombre de modelo en una memoria eventualmente volátil.

Durante la etapa a/, la predicción del valor de espacialización se realiza siguiendo un modelo de predicción, es decir en particular que los datos utilizados para la predicción pueden variar siguiendo el modelo. Por ejemplo, para un modelo que consiste en asignar un valor arbitrario al valor de espacialización, no es necesario ningún dato para la predicción. Para un modelo que consiste en retomar un valor de espacialización precedente y/o en ponderar un valor de espacialización precedente, este valor de espacialización precedente se utiliza durante la predicción.

Ventajosamente, la etapa a/ se realiza para datos de espacialización correspondientes a una banda de frecuencias dada. De este modo, pueden realizarse varias predicciones en paralelo, en diferentes bandas de frecuencias. En efecto, en el caso de una señal en estéreo, la elección del modelo de predicción más preciso puede estar vinculada a la frecuencia: según la banda de frecuencias considerada, esto puede conducir a elegir modelos de predicción diferentes.

De acuerdo con otro aspecto (véase la reivindicación 9), la invención tiene por objeto un programa informático que comprende instrucciones para la implementación del procedimiento expuesto anteriormente, cuando estas instrucciones son ejecutadas por un procesador.

De acuerdo con otro aspecto más, la invención tiene como aspecto un dispositivo de ocultación de datos de espacialización defectuosos tal como se define en la reivindicación 10. Este dispositivo comprende una unidad de memoria, la cual puede comprender una o varias memorias, para almacenar una pluralidad de juegos de instrucciones, correspondiendo cada juego de instrucciones a un modelo de predicción. Este dispositivo comprende además medios de recepción para recibir datos de espacialización. Un módulo de test permite testar la validez de los datos de espacialización recibidos por los medios de recepción. En caso de recepción de datos de espacialización detectados como válidos por el módulo de test, un módulo de estimación permite, mediante el juego de instrucciones almacenado en la unidad de memoria, ejecutar este juego de instrucciones para predecir un valor de espacialización. Un módulo de selección permite elegir un modelo de predicción, a partir de los valores de espacialización predichos por el módulo de estimación y a partir de los datos de espacialización recibidos por los medios de recepción. El dispositivo de ocultación comprende además un módulo de predicción dispuesto para, en caso de recepción de datos de espacialización considerados como defectuosos por el módulo de detección, predecir, de acuerdo con el modelo elegido por el módulo de selección, un valor de espacialización.

De acuerdo con otro aspecto más (véase la reivindicación 11), la invención tiene por objeto un aparato de reconstrucción de datos de audio multicanal. Este aparato comprende medios de reconstrucción multicanal, para reconstruir datos de audio multicanal a partir de al menos datos en un número de canales restringido, por ejemplo datos monocanal. Este aparato comprende además el dispositivo de ocultación descrito anteriormente. El módulo de predicción se dispone para, en caso de recepción de datos de espacialización considerados como defectuosos por el módulo de detección, proporcionar el valor de espacialización predicho a los medios de reconstrucción multicanal para la reconstrucción de los datos de audio multicanal.

El aparato de reconstrucción de datos de audio multicanal puede estar integrado en un procesador, o bien incluso comprender un aparato de tipo ordenador, cadena Hi-Fi u otro.

Los diversos componentes del aparato de reconstrucción, por ejemplo los medios de reconstrucción, el dispositivo de ocultación, el módulo de detección u otro, pueden ser distintos o estar fusionados.

Otras particularidades y ventajas de la presente invención quedarán claras en la descripción detallada a continuación, realizada en referencia a los dibujos adjuntos, en los que:

la figura 1 muestra un ejemplo de dispositivo de codificación conversacional,

la figura 2 muestra un ejemplo de dispositivo de decodificación que comprende un ejemplo de aparato de reconstrucción de acuerdo con una realización de la invención,

la figura 3 es un ejemplo de algoritmo de un procedimiento de acuerdo con una realización de la invención,

la figura 4 es un gráfico que muestra un ejemplo de evolución posible de la ganancia, y

la figura 5 muestra un dispositivo adecuado para ejecutar un programa informático de acuerdo con un aspecto de la invención.

Referencias idénticas designan objetos idénticos o similares de una figura a otra.

En los ejemplos ilustrados por las figuras, el número de canales de los datos de audio multicanal es de dos exactamente pero, por supuesto, pueden preverse más canales. Los datos de audio multicanal pueden comprender por ejemplo datos 5.1 en seis canales. La invención también puede aplicarse en el campo de las conferencias de audio espacializadas.

En particular, se puede hacer remisión a la norma MPEG Surround, es decir que se puede utilizar o simular una estructura en árbol para generar más de de dos vías.

En los ejemplos representados, los datos de audio se agrupan por tramas o paquetes, con el índice n.

La figura 1 muestra un ejemplo de codificador, para el cual informaciones en estéreo son transmitidas por bandas de frecuencias y se aplican en el ámbito frecuencial.

Para ello, el codificador integra medios de transformación tiempo-frecuencia 10, por ejemplo un DSP (del inglés “Digital Signal Processor”) adecuados para realizar una transformada, por ejemplo una transformada de Fourier discreta o DFT (del inglés “Discrete Fourier Transform”), una transformada MDCT (del inglés “Modified Discrete Cosine Transform”), una transformada MCLT (del inglés “Modulated Complex Lapped Transform”).

De este modo se obtiene, a partir de los valores $S_L(n)$, $S_R(n)$ correspondientes a las señales temporales izquierda y derecha, valores de señales frecuenciales $S_L(k)$ y derecha $S_R(k)$.

A continuación se aplica una matrización a las señales de la vía izquierda $S_L(k)$ y derecha $S_R(k)$, mediante medios de matrización 11.

Estos medios 11 permiten determinar a partir de la señal en estéreo $S_L(k)$, $S_L(k)$, una señal monocanal M(k) y una señal residual E(k). La señal monocanal M(k) es típicamente la media-suma de las señales izquierda $S_L(k)$ y derecha $S_R(k)$. La señal residual E(k) puede ser igual a la mitad de la diferencia entre las señales izquierda $S_L(k)$ y derecha $S_R(k)$.

Puede preverse que la matrización sea adaptativa para que la señal monocanal M(k) transporte más información. Para ello el procedimiento implementado por los medios de matrización 11 puede evolucionar con el tiempo, para evitar la anulación de componentes que estarían en oposición de fase entre las vías izquierdas y derechas.

Medios de estimación de datos de espacialización 12 permiten estimar a partir de la señal monocanal M(k) y de la señal residual E(k) datos de espacialización, por ejemplo parámetros en estéreo. Estos parámetros en estéreo pueden ser conocidos por el especialista en la técnica y comprender, por ejemplo, diferencias de nivel inter-vías (ILD), correlaciones inter-vías (ICC) y retardos o diferencias de fase entre vías (IPD / ITD).

Estos parámetros en estéreo $ILD^{(b)}$ puede determinarse mediante bandas de frecuencias, indexadas por la variable b. Estas bandas pueden estar constituidas según una escala frecuencial próxima a la percepción humana. Por ejemplo, pueden utilizarse entre 8 y 20 bandas frecuenciales, según la precisión deseada y la riqueza del espectro considerado.

Medios de cuantificación, codificación y multiplexado 13 permiten cuantificar y codificar los parámetros en estéreo $ILD^{(b)}$ para permitir una transmisión a un caudal reducido.

La señal monocanal M(k) también es cuantificada y codificada por los medios 13, en el ámbito transformado como se presenta en la figura 1 o, como alternativa, en el ámbito temporal. Pueden utilizarse algoritmos normalizados para

procesar esta señal monocanal M(k), por ejemplo un codificador de voz de tipo ITU G.729.1 o G.718. También podrá tratarse de un codificador de audio genérico de tipo MPEG-4 AAC o HE-AAC.

La señal residual E(k) es transmitida opcionalmente, recurriendo también a una codificación normalizada o una técnica de transmisión apropiada para esta señal en el ámbito frecuencial o temporal.

La señal codificada S_{cod} obtenida a la salida de los medios de cuantificación, codificación y multiplexado 13 es transmitida, por ejemplo por vía de radio.

Como alternativa, se podrá prever que el codificador conduzca a obtener datos en más de un canal monofónico, siempre que el número de canales de los datos obtenidos a la salida del codificador sea inferior al número de canales de los datos a la entrada del codificador.

La figura 2 muestra un ejemplo de decodificador susceptible de recibir una señal S'_{cod} correspondiente a la señal S_{cod} transmitida.

Medios de decodificación y de demultiplexado 29 permiten extraer de la señal S'_{cod} recibida datos monocanal M'(k), datos de espacialización $ILD'^{(b)}$, así como, eventualmente, datos residuales E'(k).

El decodificador comprende, además, un aparato de reconstrucción 26 para reconstruir datos de audio multicanal $S'_L(k)$, $S'_R(k)$, a partir de datos monocanal M'(k), datos de espacialización $ILD'^{(b)}$ y eventuales datos residuales E'(k).

La figura 3 muestra un algoritmo ejecutable por el aparato de reconstrucción 26 de la figura 2. Estas dos figuras serán, por lo tanto, comentadas simultáneamente.

El aparato de reconstrucción 26 comprende un dispositivo de ocultación 20 para proporcionar valores de sustitución en caso de datos de espacialización $ILD'^{(b)}$ defectuosos y medios de reconstrucción multicanal 27 para la reconstrucción propiamente dicha.

Los medios de reconstrucción multicanal 27 pueden realizar, por ejemplo, durante una etapa 300, combinaciones de tipo:

$$\begin{cases} S'_L(k) = E'_L(k) + W_L(b,n).M_L(k) \\ S'_R(k) = E'_R(k) + W_R(b,n).M_R(k) \end{cases}$$

Donde:

k designa el índice frecuencial en cuestión,

b designa la banda asignada por los parámetros en estéreo transmitidos,

$M_L(k)$, una señal en el ámbito frecuencial, obtenida durante una etapa 301 a partir de los datos monocanal M'(k), aplicando de forma conocida por el especialista en la técnica un desfase o un retardo correspondiente a la vía izquierda, obteniéndose este desfase o este retardo de datos de espacialización no representados, y

$M_R(k)$, una señal en el ámbito frecuencial, obtenida de forma equivalente durante la etapa 301, para la vía derecha.

En particular, si no se aplica ningún desfase, entonces $M_R(k)=M_L(k)=M'(k)$.

E'_L es una señal específica para la vía izquierda, obtenida de forma conocida por el especialista en la técnica de los datos residuales E'(k) opcionalmente transmitidos; y E'_R, una señal específica para la vía derecha, obtenida de forma conocida por el especialista en la técnica de los datos residuales E'(k) opcionalmente transmitidos. La etapa de obtención de los datos E'_L, E'_R no se representa en la figura 3.

En caso de no transmisión de datos residuales:

$E'(k), E'_L=E'_R=0$.

W_L y W_R son las ganancias procedentes de datos de espacialización ILD'(b,n) para la banda b considerada y la trama n.

Las ganancias W_L y W_R pueden determinarse, por ejemplo, de la siguiente manera, por medio de valores W'_L y W'_R, durante una etapa 302:

$$W'_L(b,n) = \frac{2.ILD'(b,n)}{1+ILD'(b,n)}$$

$$W'_R(b,n) = \frac{2}{1+ILD'(b,n)}$$

donde ILD'(b,n) es el dato de espacialización ILD'(b) recibido para la trama n.

Entonces se realiza, durante una etapa 304, un alisado con una constante de tiempo α entre 0 y 1, por ejemplo α = 0,8, según:

$W_L(b,n) = \alpha.W'_L(b,n)+(1-\alpha).W_L(b,n-1)$, donde $W_L(b,n-1)$ designa el valor obtenido para la trama precedente.

Para la vía derecha, puede realizarse durante la etapa 304 el mismo alisado:

$W_R(b,n) = \alpha.W'R(b,n) + (1-\alpha).WR(b,n-1)$, donde $W_R(b,n-1)$ designa el valor obtenido para la trama precedente.

Como alternativa, puede utilizarse el valor obtenido para la vía izquierda, según por ejemplo:

$W_R(b,n)=2-W_R(b,n)$

El dispositivo de ocultación 20 permite evitar pérdidas eventuales de datos ILD'(b,n), de modo que los datos W_R y W_L puedan determinarse a pesar de todo.

El dispositivo de ocultación 20 comprende medios de recepción no representados para recibir, durante una etapa 305, los datos de espacialización ILD'(b,n), así como eventualmente los datos monocanal M'(k), y los datos residuales E'(k).

Estos medios de recepción pueden comprender, por ejemplo, un puerto de entrada, clavijas de entrada u otro.

Un módulo de test 22 conectado a estos medios de recepción permite testar durante una etapa 306 la validez de los datos de espacialización $ILD'^{(b)}$. Este módulo de test puede implementar una verificación de una codificación de tipo CRC, para verificar, por ejemplo, que la transmisión no ha conllevado ninguna degradación de los datos de espacialización.

El módulo de test 22 también puede leer algunos valores (no representados) extraídos de la señal S'_{cod} recibida, indicando estos valores eventuales supresiones de capas de datos transmitidos. En efecto, puede preverse que algunos elementos de la red de transmisión se abstienen de transmitir, en particular en caso de atasco de la red, o de reducción de la banda pasante del canal de transmisión, tal o cual conjunto de datos. Los conjuntos de datos no transmitidos pueden corresponder, por ejemplo, a detalles sonoros. Cuando el módulo de test 22 lee un valor que indica una supresión de algunos datos, se considera que estos datos faltan.

El dispositivo de ocultación 20 comprende una unidad de memoria 21 que almacena varios juegos de instrucciones, correspondiendo cada juego de instrucciones a un modelo de predicción.

Por ejemplo, según un primer modelo de predicción, cuando los datos de espacialización ILD'(b,n) son defectuosos para una trama n y una banda de frecuencia b dada, se selecciona

$$W_L^{(1)}(b,n) = W_L(b,n-1)$$

$$W_R^{(1)}(b,n) = W_R(b,n-1)$$

Las instrucciones correspondientes consisten entonces en copiar los valores $W_R(b,n-1)$, $W_L(b,n-1)$ obtenidos para la trama precedente.

Por ejemplo, según un segundo modelo de predicción, se selecciona:

$W_L^{(2)}(b,n) = \beta + (1-\beta).W_L(b,n-1)$, y

$W_R^{(2)}(b,n) = \beta + (1-\beta).W_R(b,n-1)$, con β entre 0 y 1.

De este modo, en caso de una sucesión de tramas para las cuales los datos de espacialización son defectuosos, $W_L^{(2)}(b,n)$ y $W_R^{(2)}(b,n)$ tienden a 1 y, por consiguiente, los datos de audio multicanal $S'_L(k)$, $S'_R(k)$ se aproximan a los datos monocanal M'(k). Dicho de otro modo, se suprimen poco a poco los efectos de espacialización para volver a una señal monocanal.

De acuerdo con otro ejemplo de modelo de predicción, se selecciona

$W_L^{(3)}(b,n) = 2.W_L(b,n-1) - W_L(b,n-2)$, y

$W_R^{(3)}(b,n) = 2.W_R(b,n-1) - W_R(b,n-2)$.

O bien también:

$W_L^{(4)}(b,n) = \frac{1}{2}.W_L(b,n-1) + \frac{1}{2}W_L(b,n-2)$, y

$W_R^{(4)}(b,n) = \frac{1}{2}.W_R(b,n-1) + \frac{1}{2}W_R(b,n-2)$.

O bien también se utiliza un filtro mediano:

$W_L^{(5)}(b,n) = \textit{Mediana}(W_L(b,n-1), W_L(b,n-2),...)$, y

$W_R^{(5)}(b,n) = \textit{Mediana}(W_R(b,n-1), W_R(b,n-2),...)$.

Eventualmente, para asegurar una mejor estabilidad, se utilizará en lugar de $W_L(b,n-i)$ y $W_R(b,n-i)$ respectivamente valores atenuados, por ejemplo $0{,}9.W_L(b,n-i)$ y $0{,}9.W_R(b,n-i)$. Puede preverse conservar en la unidad de memoria valores atenuados, para utilizarlos directamente aplicando uno de los modelos expuestos anteriormente.

Otros modelos también son posibles, por ejemplo es posible una predicción más general de la forma $W_L^{(m)} = \sum_{i=1}^{P} a_i.W_L(b,n-i)$ y $W_R^{(m)} = \sum_{i=1}^{P} a_i.W_R(b,n-i)$, con un orden de predicción P. Los coeficientes a_i pueden evolucionar en el tiempo, y reactualizarse utilizando un método de tipo Levinson-Durbin.

Estos ejemplos de modelos conducen a predecir valores de W_L y W_R. Como alternativa, los modelos pueden permitir predecir valores de las variables ILD'(b,n), de W'_L y W'_R, u otro.

Por ejemplo, según un modelo de predicción equivalente al primer modelo expuesto anteriormente, cuando faltan datos de espacialización ILD'(b,n) para una trama n y una banda de frecuencia b dada, se selecciona ILD'(b,n)=ILD'(b,n-1). La instrucción correspondiente consiste entonces en copiar este valor ILD'(b,n-1) obtenido para la trama precedente.

Un módulo de estimación 23 permite ejecutar las instrucciones de los diferentes juegos de instrucciones. Este módulo 23 está activado, por ejemplo, para cada trama tal que los datos de espacialización ILD'(b,n) correspondientes son considerados como válidos por el módulo de test 22, o bien también solamente para las tramas consideradas como válidas y que preceden a una trama considerada como defectuosa.

Cuando este módulo 23 está activado, todos los juegos de instrucciones almacenados son ejecutados, durante etapas 307 repetidas en un bucle que recorre los juegos de instrucciones, con las etapas clásicas de inicialización,

de test y de incremento, para obtener un conjunto de valores $\{W_L^{(m)}, W_R^{(m)}\}$, indicando m el modelo utilizado.

Un módulo de selección 24 permite seleccionar uno de estos modelos confrontando los valores de espacialización predichos $\{W_L^{(m)}, W_R^{(m)}\}$, con valores de espacialización estimados W_L, W_R a partir de los datos de espacialización efectivamente recibidos ILD'(b,n).

Por ejemplo, para cada modelo, pueden calcularse durante etapas 308 valores de semejanza $\sigma_{L,m}^2$, $\sigma_{R,m}^2$, a partir de valores predichos $W_L^{(m)}(b,n)$, $W_R^{(m)}(b,n)$ y a partir de valores estimados $W_L(b,n)$, $W_R(b,n)$. Los valores de semejanza pueden comprender, por ejemplo, la varianza de cada predicción:

$$\sigma_{L,m}^2 = E\left[(W_L(b,n) - W_L^{(m)}(b,n))^2\right],$$

representando E la expectativa matemática, según por ejemplo:

$$E\left[x^2\right] = \frac{1}{N}\sum_{n=0}^{N-1} x^2(n)$$

De este modo, se utiliza una secuencia de N tramas recibidas para determinar N valores $W_L^{(m)}(b,n)$ y compararlas con N valores estimados $W_L(b,n)$.

Una fórmula equivalente se aplica para la vía derecha.

Como alternativa, puede preverse calcular una varianza de forma recursiva, por ejemplo según, para cada vía:

$$\sigma_{m,n}^2 = \alpha \cdot \sigma_{m,n-1}^2\left[x^2\right] + (1-\alpha)\cdot x^2(n),$$

donde α es, en este caso, una constante de tiempo, por ejemplo, igual a 0,975, y $\sigma_{m,n}^2$ designa la estimación de la varianza en la trama n.

De acuerdo con una realización alternativa y no representada, en lugar de estimar la varianza, se estima una probabilidad de los datos $W_L^{(m)}$, $W_R^{(m)}$ frente a los datos W_L, W_R obtenidos a partir de los valores efectivamente recibidos. Puede utilizarse, por ejemplo, un conjunto de estimadores:

$$P_m^L = P(W_L^{(m)}(b,n)/W_L(b,n))$$ y

$$P_m^R = P(W_R^{(m)}(b,n)/W_R(b,n))$$.

Mediante comparación de los estimadores de tipo σ_m^2 o P_m, puede seleccionarse el modelo de predicción para el cual el valor de semejanza indica una mayor adecuación entre valores predichos y valores estimados. Por ejemplo, se determina el índice m* del modelo que proporciona la mejor ocultación: éste será el índice que minimizará σ_m^2 o maximizará P_m en otra realización.

En aras de la sencillez, puede preverse seleccionar el índice que minimizará σ_m^2 en una sola de las vías, por ejemplo la vía izquierda.

Este valor m* constituye un identificador del modelo de predicción seleccionado y se almacena en la unidad de memoria 21 durante una etapa 309.

Queda claro que las etapas 307 pueden ejecutarse antes que las etapas 302, 304, o bien también en paralelo. Cada etapa 308 emplea, en este caso, valores obtenidos durante la etapa 304, y es ejecutada, por lo tanto, posteriormente

a esta etapa 304.

El dispositivo de ocultación 20 comprende, además, un módulo de predicción 25 para, en caso de recepción de datos de espacialización considerados como defectuosos, predecir durante una etapa 310 de acuerdo con el modelo identificado por el valor m*, valores de espacialización $W_L^{(m^*)}(b,n)$ y $W_R^{(m^*)}(b,n)$.

Este valor es suministrado a los medios de reconstrucción multicanal 27, los cuales están ahora en condiciones de reconstruir durante la etapa 300 los datos multicanal $S'_L(k)$, $S'_R(k)$ a pesar de los defectos de los datos de espacialización.

Medios de transformación frecuencia-tiempo 28, por ejemplo DSP, permiten recuperar datos de audio temporales $S'_L(n)$, $S'_R(n)$ a partir de los datos multicanal $S'_L(k)$, $S'_R(k)$ reconstruidos.

La figura 4 muestra un diagrama que representa un ejemplo de evolución del valor $W_L(b,n)$ para la segunda sub-banda de frecuencias, es decir b=1. En abscisas figura el índice de trama n, y en ordenadas los valores $W_L(1,n)$.

Para la parte A que corresponde, de forma grosera, a las tramas entre la 500ª y la 810ª tramas, los valores de $W_L(1,n)$ son en su mayor parte iguales a 1, lo que corresponde a una señal sonora relativamente monofónica.

Para la parte B, los valores de $W_L(1,n)$ corresponden a una señal localizada a la izquierda, mientras que para la parte C, los valores de $W_L(1,n)$ corresponden a una señal localizada a la derecha.

Para la parte D, los valores de $W_L(1,n)$ corresponden a una pluralidad de fuentes sonoras localizadas en diversos emplazamientos.

El mejor modelo de predicción seleccionado puede variar de acuerdo con el tipo de las variaciones de la ganancia.

De este modo, para la parte A, el modelo que consiste en repetir el valor obtenido para la trama precedente conduciría a repetir de forma errónea los picos de valores de $W_L(1,n)$. Un modelo más juicioso consistiría en seleccionar un valor arbitrario correspondiente a una señal monocanal, o bien en ponderar la ganancia obtenida para la trama precedente para aproximarse poco a poco a una ganancia de 1.

Por el contrario, para las partes B y C, el enfoque más juicioso puede consistir en repetir el valor de ganancia obtenido para la trama precedente.

Para la parte D, cuando la evolución de la ganancia es relativamente lenta y, por lo tanto, relativamente predecible, un enfoque juicioso consistiría en realizar una media ponderada de las ganancias obtenidas para P tramas precedentes. Cuando los parámetros en estéreo evolucionan más rápidamente, el enfoque más juicioso consistiría en volver a una señal monocanal para evitar cualquier artefacto.

De este modo, el modelo más juicioso puede cambiar de acuerdo con el tipo de variaciones de la ganancia de una trama a otra. El procedimiento de la figura 3 permite seleccionar, sin intervención humana, el modelo de predicción más adecuado.

Esta selección del modelo de predicción más adecuado permite obtener una ocultación de mejor calidad en caso de datos defectuosos.

La figura 5 muestra un ordenador que comprende una pantalla 502, un teclado y una unidad central. Esta unidad central comprende una memoria 500 para almacenar un programa informático que comprende instrucciones correspondientes a las etapas del procedimiento descrito anteriormente. Esta unidad central comprende, además, un procesador 501 conectado a la memoria 500, para ejecutar estas instrucciones.

REIVINDICACIONES

1. Procedimiento de procesamiento de datos sonoros, para la reconstrucción de datos de audio multicanal a partir de al menos datos en un número de canales restringido y de datos de espacialización, comprendiendo dicho procedimiento una etapa de test (306) de validez de datos de espacialización de una trama recibida y, si dicho test muestra que dichos datos de espacialización recibidos son válidos, etapas de:

a/ mediante un modelo respectivo de una pluralidad de modelos de predicción, predicción de acuerdo con dicho modelo de un valor de espacialización (307), y

b/ elección de un modelo de predicción, a partir de los valores de espacialización predichos de este modo y, a partir de los datos de espacialización recibidos, para poder, en caso de recepción posterior de datos de espacialización defectuosos, predecir de acuerdo con dicho modelo seleccionado un valor de espacialización y utilizar dicho valor de espacialización predicho para la reconstrucción de los datos de audio multicanal;

y en el que, durante la etapa b/:

mediante un modelo respectivo de la pluralidad de modelos, se calcula un valor de semejanza $(\sigma_{L,m}^2, \sigma_{R,m}^2)$, a partir, por un lado, del valor de espacialización predicho según dicho modelo $(W_L^{(m)}(b,n), W_R^{(m)}(b,n))$ y, por otro lado, de un valor estimado ($W_L(b,n)$, $W_R(b,n)$) a partir de los datos de espacialización recibidos, y

se selecciona el modelo de predicción para el cual dicho valor de semejanza indica una mayor adecuación entre el valor de espacialización predicho y dicho valor estimado.

2. Procedimiento de acuerdo con la reivindicación 1, que comprende además, si el test muestra que los datos de espacialización recibidos son válidos, y previamente a la etapa a/, una etapa de almacenamiento de dichos datos de espacialización válidos; y en el que la etapa b/ se realiza en caso de recepción posterior de datos de espacialización defectuosos, a partir de dichos datos de espacialización almacenados.

3. Procedimiento de acuerdo con la reivindicación 2, en el que la etapa a/ se realiza en caso de recepción posterior de datos de espacialización defectuosos, a partir de dichos datos de espacialización almacenados.

4. Procedimiento de acuerdo con la reivindicación 1, en el que las etapas a/ y b/ se realizan sistemáticamente después de la recepción de una trama válida, comprendiendo el procedimiento además, después de la etapa b/, una etapa de escritura en memoria de un identificador del modelo de predicción seleccionado.

5. Procedimiento de acuerdo con la reivindicación 1, en el que el valor de espacialización predicho comprende una ganancia $(W_L^{(m)})$.

6. Procedimiento de acuerdo con la reivindicación 1, en el que el valor de espacialización predicho comprende un retardo.

7. Procedimiento de acuerdo con la reivindicación 1, en el que durante las etapas a/ y b/:

por trama de una secuencia de tramas recibidas, y para al menos un modelo de la pluralidad de modelos, se predice de acuerdo con dicho modelo un valor de especialización $(W_L^{(m)}(b,n), W_R^{(m)}(b,n))$ y,

para dicho modelo, el valor de semejanza $(\sigma_{L,m}^2, \sigma_{R,m}^2)$ se calcula a partir, por un lado, de la secuencia de valores de espacialización predichos según dicho modelo y, por otro lado, de una secuencia de valores estimados ($W_L(b,n)$, $W_R(b,n)$) a partir de los datos de espacialización de la secuencia de tramas recibidas.

8. Procedimiento de acuerdo con la reivindicación 1, en el que, la etapa a/ se realiza para los datos de espacialización correspondientes a una banda de frecuencias (b) dada.

9. Programa informático que comprende instrucciones para la implementación del procedimiento de acuerdo con la reivindicación 1, cuando dichas instrucciones son ejecutadas por un procesador.

10. Dispositivo de ocultación (20) de datos de espacialización defectuosos para la reconstrucción de datos de audio multicanal, comprendiendo el dispositivo:

una unidad de memoria (21) para almacenar una pluralidad de juegos de instrucciones, correspondiendo cada juego de instrucciones a un modelo de predicción,

medios de recepción para recibir datos de espacialización,

un módulo de test (22) de la validez de los datos de espacialización recibidos por los medios de recepción,

un módulo de estimación (23) adecuado para, en caso de recepción de datos de espacialización detectados como válidos por el módulo de test (22), y por juego de instrucciones almacenado en la unidad de memoria, ejecutar dicho juego de instrucciones para predecir un valor de espacialización, y

un módulo de selección (24) para seleccionar un modelo de predicción, a partir de los valores de espacialización predichos por el módulo de estimación, siendo este módulo de selección adecuado:

• para calcular, mediante modelo respectivo de la pluralidad de modelos, un valor de semejanza $(\sigma_{L,m}^2, \sigma_{R,m}^2)$ a partir, por un lado, del valor de espacialización predicho según dicho modelo $(W_L^{(m)}(b,n), W_R^{(m)}(b,n))$ y, por otro lado, de un valor estimado ($W_L(b,n)$, $W_R(b,n)$) a partir de los datos de espacialización recibidos por los medios de recepción, y

• para seleccionar el modelo de predicción para el cual dicho valor de semejanza indica una mayor adecuación entre el valor de espacialización predicho y dicho valor estimado;

comprendiendo el dispositivo de ocultación además un módulo de predicción (25) dispuesto para, en caso de recepción posterior de datos de espacialización considerados como defectuosos por el módulo de test (22), predecir un valor de espacialización de acuerdo con dicho modelo seleccionado por el módulo de selección.

11. Aparato de reconstrucción (26) de datos de audio multicanal, comprendiendo dicho aparato:

medios de reconstrucción multicanal (27), para reconstruir datos de audio multicanal a partir al menos de datos monocanal,

el dispositivo de ocultación (20) de acuerdo con la reivindicación 10, en el que el módulo de predicción (25) está dispuesto para, en caso de recepción de datos de espacialización considerados como defectuosos por el módulo de test (22), proporcionar el valor de espacialización predicho a los medios de reconstrucción multicanal para la reconstrucción de los datos de audio multicanal.

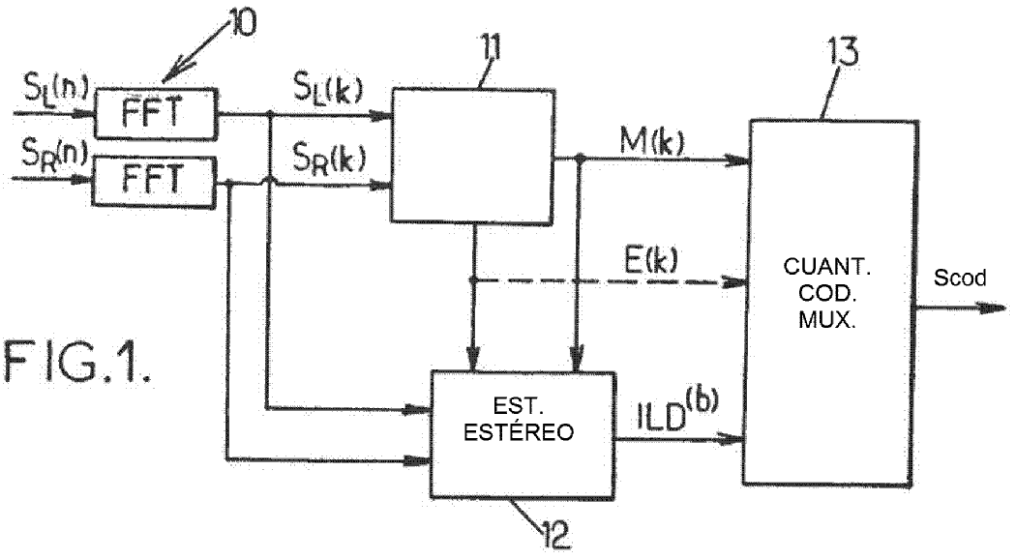
10
11
13
$S_L(n)$
FFT
$S_L(k)$
$S_R(n)$
FFT
$S_R(k)$
M(k)
E(k)
CUANT.
COD.
MUX.
Scod
EST.
ESTÉREO
ILD(b)
12

FIG.1.

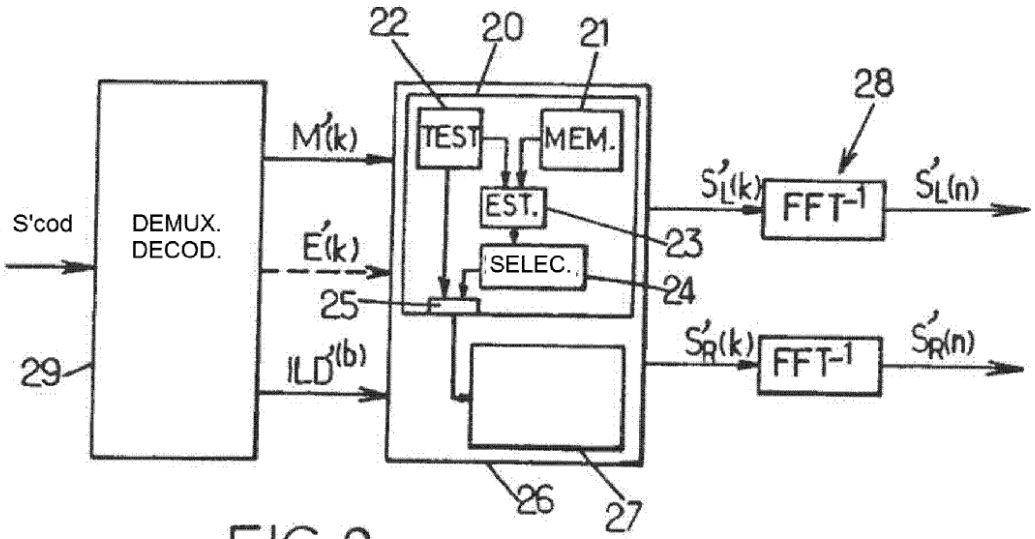
22
20
21
28
TEST
MEM.
M'(k)
EST.
$S'_L(k)$
FFT-1
$S'_L(n)$
S'cod
DEMUX.
DECOD.
E'(k)
23
SELEC.
24
25
$S'_R(k)$
FFT-1
$S'_R(n)$
29
ILD'(b)
26
27

FIG.2.

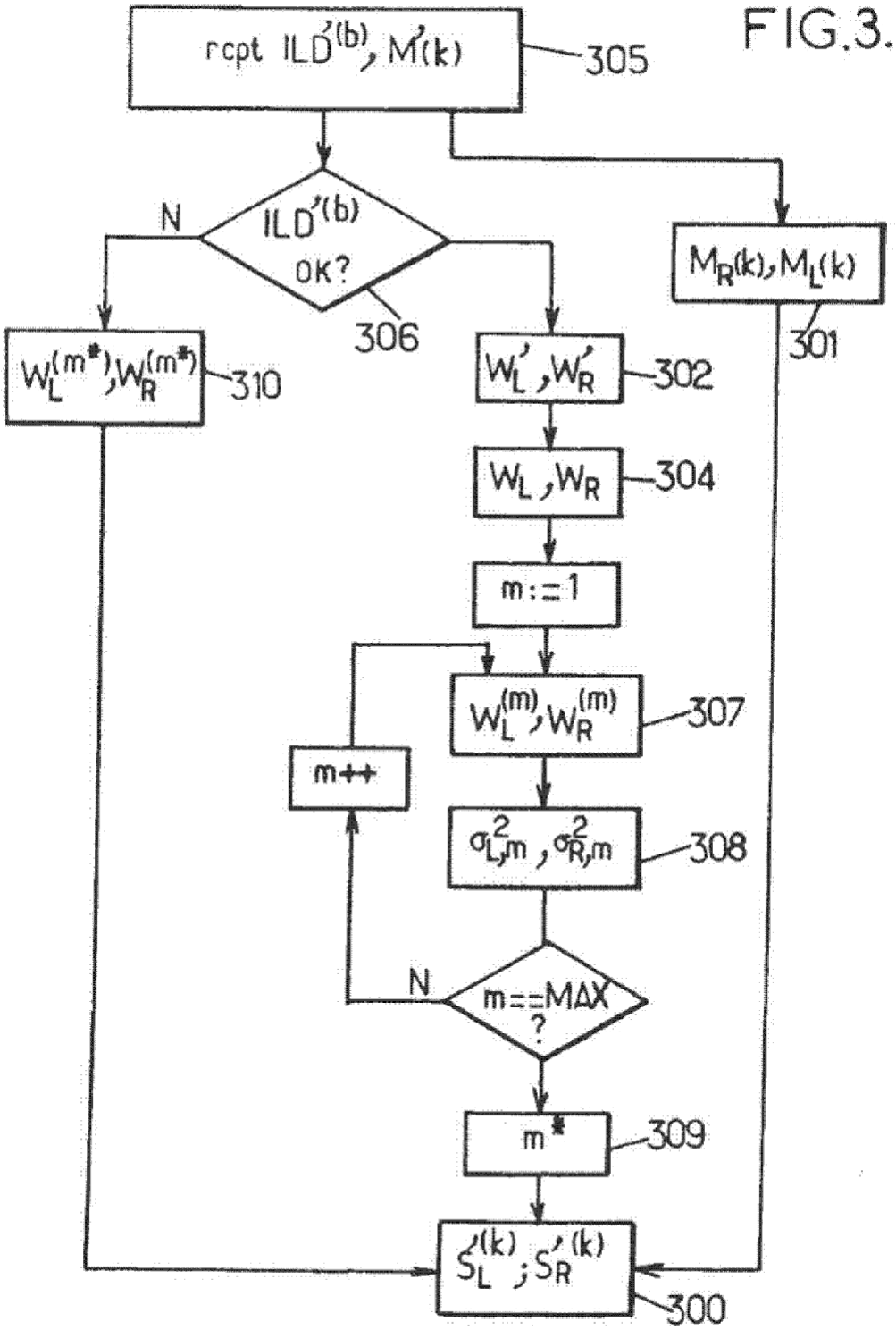
rcpt ILD'(b), M'(k)
305
ILD'(b) OK?
306
N
M_R(k), M_L(k)
301
W_L^(m*), W_R^(m*)
310
W'_L, W'_R
302
W_L, W_R
304
m:=1
W_L^(m), W_R^(m)
307
m++
σ²_L,m , σ²_R,m
308
N
m==MAX ?
m*
309
S'_L(k) ; S'_R(k)
300

FIG.3.

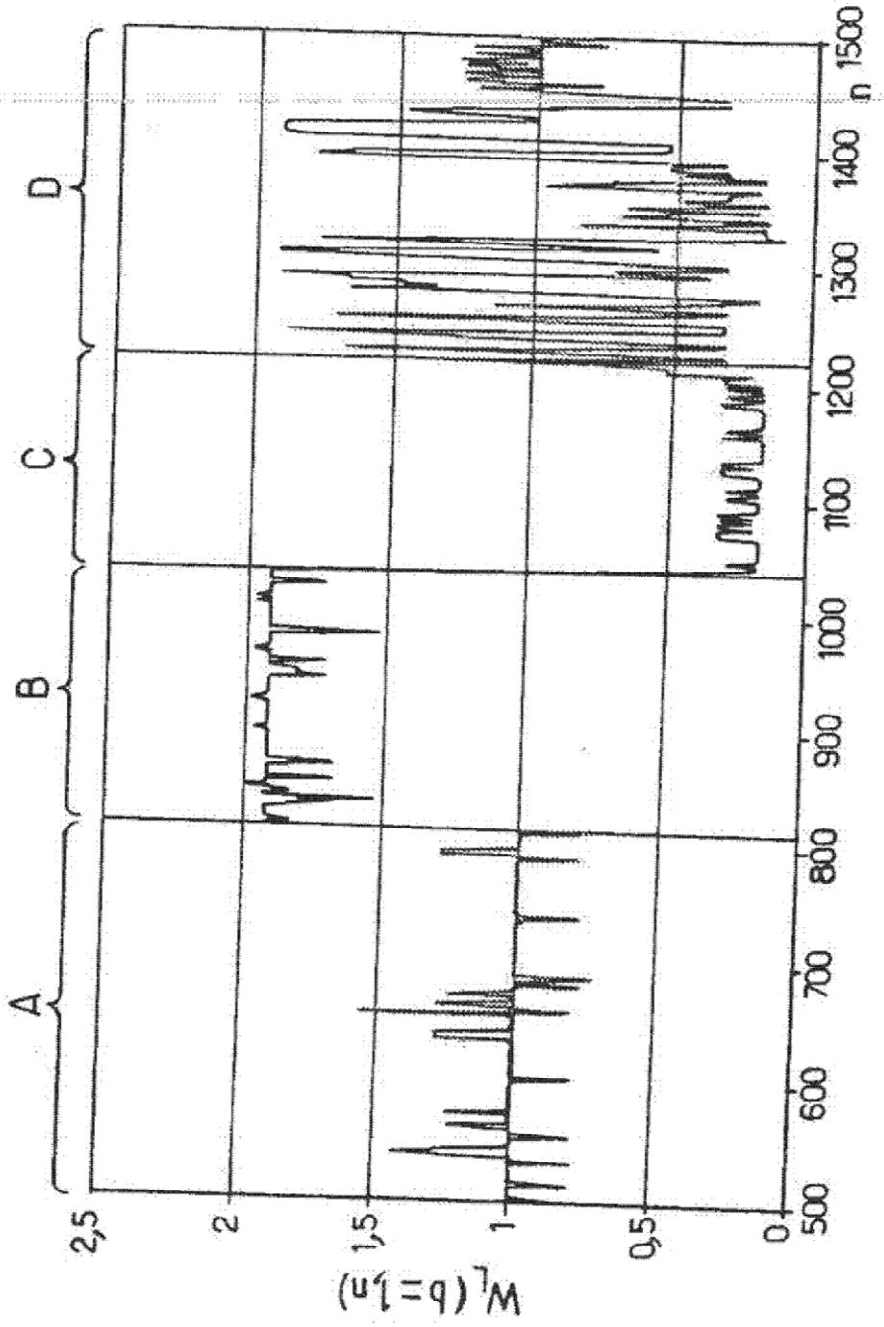
A
B
C
D
2,5
2
1,5
1
0,5
0
500
600
700
800
900
1000
1100
1200
1300
1400
1500
n
$W_L(b=1,n)$

FIG.4.

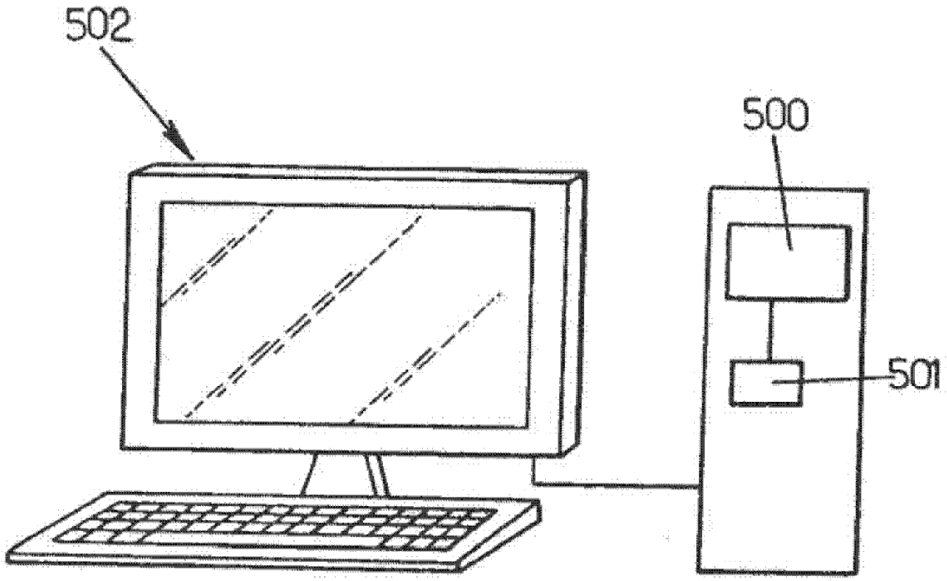

FIG.5.