



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 347 274**

51 Int. Cl.:
G10L 19/00 (2006.01)
G10L 19/14 (2006.01)
H04S 3/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06711111 .2**
96 Fecha de presentación : **16.03.2006**
97 Número de publicación de la solicitud: **1866911**
97 Fecha de publicación de la solicitud: **19.12.2007**

54 Título: **Codificación de audio multicanal ajustable a escala.**

30 Prioridad: **30.03.2005 EP 05102506**
18.04.2005 EP 05103077

45 Fecha de publicación de la mención BOPI:
27.10.2010

45 Fecha de la publicación del folleto de la patente:
27.10.2010

73 Titular/es: **Koninklijke Philips Electronics N.V.**
Groenewoudseweg 1
5621 BA Eindhoven, NL

72 Inventor/es: **Myburg, Francois, P. y**
Schuijers, Erik, G., P.

74 Agente: **Zuazo Araluze, Alexander**

ES 2 347 274 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 de audio multicanal ajustable a escala.

5 La invención se refiere al campo de la codificación de audio de alta calidad. Especialmente, la invención se refiere al campo de la codificación de alta calidad de datos de audio multicanal. Más específicamente, la invención define codificadores y decodificadores y procedimientos para codificar y decodificar datos de audio multicanal.

10 Aunque son posibles numerosas disposiciones/configuraciones multicanal, la disposición/configuración 5.1 es la más popular (véase también la figura 1). La configuración 5.1 multicanal típica consiste en cinco altavoces, concretamente altavoces frontal izquierdo (Lf), frontal derecho (Rf), central (C), envolvente izquierdo (Ls) y envolvente derecho (Rs) complementados por un altavoz LFE (mejora de baja frecuencia) adicional que se colocará en un ángulo arbitrario. En el pasado se consideraron diversos enfoques para comprimir datos de audio multicanal, tal como los datos de audio multicanal 5.1. A continuación se da una breve visión general.

15 En la norma de audio MPEG-2, ISO/IEC 13818-3:1998 *Information technology -- Generic coding of moving pictures and associated audio information -- Part 3: Audio*, se realiza una provisión para codificación de audio multicanal mientras se mantiene la retrocompatibilidad con la norma de audio MPEG-1, ISO/IEC 11172-3:1993 *Information technology Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s -- Part 3: Audio*, que se centra únicamente en la codificación de audio mono y estéreo. La retrocompatibilidad se logra formando una señal estéreo básica, derivada a partir del contenido multicanal, que se coloca en la parte de datos del flujo de bits de MPEG-1. Tres señales adicionales se colocan entonces en la parte de datos auxiliares del flujo de bits de MPEG-1. Esta técnica se denomina matrización. Un decodificador de audio MPEG-1 puede generar una señal estéreo significativa (Lo, Ro) a partir del flujo de bits, mientras que un decodificador de audio MPEG-2 puede extraer los canales adicionales y reconstruir una versión decodificada de los cinco canales de entrada. La retrocompatibilidad se consigue a costa de una alta tasa de transmisión de bits. Normalmente, se requiere una tasa de transmisión de bits de 640 kbit/s para obtener una alta calidad de audio para material de cinco canales con MPEG-2 Capa II.

30 En la codificación de audio avanzada (AAC) de MPEG-2, norma ISO/IEC TR 13818-5:1997/Amd 1:1999 *Advanced Audio Coding (AAC)*, se codifica audio multicanal en un formato no retrocompatible. Esto permite más libertad al codificador y tiene la ventaja de que puede conseguirse una calidad de audio superior (transparente) a una tasa de transmisión de bits de 320 kbit/s, en comparación con MPEG-2 Capa II a 640 kbit/s. En una configuración de 5(1) canales, AAC puede codificar los pares de canales que son simétricos respecto al oyente empleando la herramienta *Mid-Side (MS) stereo*: (Lf, Rf) y (Ls, Rs). Los canales central (C) y (opcionalmente) LFE se codifican por separado. Alternativamente, la codificación estéreo de intensidad (IS) puede emplearse para combinar varios canales de audio en un canal y, adicionalmente, proporcionar información de ajuste a escala para cada canal.

40 En la codificación de audio multicanal paramétrica, indicaciones (o parámetros espaciales) perceptivamente relevantes, tales como diferencias de intensidad entre canales (IID), diferencias de tiempo entre canales (ITD) y coherencia entre canales (ICC), se miden entre canales en una señal multicanal. Una descripción más detallada de los parámetros espaciales puede encontrarse en Christof Faller: "*Coding of Spatial Audio Compatible with Different Playback Formats*", AES Convention Paper, AES 117th Convention, San Francisco, EE.UU., 28-31 de octubre de 2004. Además, la representación multicanal se mezcla de manera descendente para dar una señal estéreo o mono que puede codificarse con un codificador convencional mono o estéreo. Un requisito importante es que la mezcla descendente estéreo o mono debe tener una calidad de audio suficiente, por ejemplo al menos comparable con la Recomendación de la ITU-R BS.775-1 sobre mezcla descendente. La información transmitida comprende por tanto una versión codificada de la señal mono o estéreo y de los parámetros espaciales. La mezcla descendente mono o estéreo se codifica a una tasa de transmisión de bits sustancialmente inferior que la requerida para codificar la señal de audio multicanal original, y los parámetros espaciales requieren un ancho de banda de transmisión muy pequeño. Por tanto, la mezcla descendente y los parámetros espaciales pueden codificarse a una tasa de transmisión de bits total que es tan sólo una fracción de la tasa de transmisión de bits requerida cuando se codifican todos los canales. El decodificador paramétrico genera una aproximación de alta calidad de la señal de audio multicanal original a partir de la mezcla descendente mono o estéreo transmitida y los parámetros espaciales.

55 Un enfoque de codificación/decodificación de audio estéreo ajustable a escala se describe en la solicitud de patente europea publicada con el número EP 0918407 A2.

60 Puede considerarse que un objeto de la presente invención es proporcionar un codificador de señal de audio multicanal ajustable a escala que proporcione una gran eficacia, que proporcione una alta calidad de señal y, al mismo tiempo, que proporcione una señal codificada que sea retrocompatible.

Según un primer aspecto, la invención proporciona un codificador de audio adaptado para codificar una señal de audio multicanal, comprendiendo el codificador:

65 - un módulo de combinación de codificador para generar una parte de señal dominante y una parte de señal residual que son una representación combinada de señales de audio primera y segunda, obteniéndose las partes de señal dominante y residual aplicando un procedimiento matemático a las señales de audio primera y segunda, en el que

el procedimiento matemático implica un primer parámetro espacial que comprende una descripción de propiedades espaciales de las señales de audio primera y segunda,

- un generador de parámetros para generar

- un primer conjunto de parámetros que comprende un segundo parámetro espacial, y

- un segundo conjunto de parámetros que comprende un tercer parámetro espacial, y

- un generador de salida para generar una señal de salida codificada que comprende

- una primera parte de salida que comprende la parte de señal dominante y el primer conjunto de parámetros, y

- una segunda parte de salida que comprende la parte de señal residual y el segundo conjunto de parámetros.

En el módulo de combinación de codificador, se combinan señales de audio primera y segunda en partes de señal dominante y residual. Por “partes de señal dominante y residual” se entienden dos señales de audio en las que la señal dominante contiene las partes dominantes o principales de las señales de audio primera y segunda, mientras que la señal residual contiene una parte residual o menos importante de las señales de audio primera y segunda. Por “parámetro espacial” se entiende un parámetro que puede expresarse matemáticamente y estar basado en o derivarse de una o más propiedades espaciales de un par de señales. Una lista no exhaustiva de tales propiedades espaciales que es posible calcular es: diferencias de intensidad entre canales (IID), diferencias de tiempo entre canales (ITD) y coherencia entre canales (ICC). El módulo de combinación de codificador preferiblemente genera las partes de señal dominante y residual de tal manera que estas partes de señal están menos correlacionadas que las señales de audio primera y segunda. Preferiblemente, las partes de señal dominante y residual se generan de modo que no están correlacionadas, es decir son ortogonales, o al menos deben estar lo menos correlacionadas posible.

La parte de señal residual puede filtrarse paso bajo antes de convertirse en un flujo de bits de salida, con el fin de representarse en un flujo de bits que requiera de este modo sólo una cantidad muy limitada de tasa de transmisión de bits. Una frecuencia de corte para tal filtrado paso bajo puede estar en el intervalo de 500 Hz a 10 kHz, por ejemplo 2 kHz.

El módulo de combinación de codificador puede estar adaptado para combinar señales de audio primera, segunda y tercera, para dar partes de señal dominante primera y segunda en lugar de combinar dos señales de audio en una señal dominante, tal como se ha descrito anteriormente.

El codificador según el primer aspecto proporciona una representación codificada ajustable a escala de las señales de audio primera y segunda. Usando la primera parte de salida, o parte de capa base, es posible decodificar las señales de audio primera y segunda con una calidad de sonido resultante aceptable usando decodificadores existentes. Sin embargo, usando un decodificador que pueda utilizar la segunda parte de salida, o parte de capa de refinamiento, es posible obtener una calidad de señal superior. Por tanto, el uso de la segunda parte de salida en un decodificador puede considerarse opcional y sólo es necesario en caso de que se desee la mejor calidad de sonido posible.

En una realización preferida, la parte de señal residual comprende una diferencia entre las señales de audio primera y segunda. La parte de señal residual puede definirse precisamente como una diferencia entre las señales de audio primera y segunda.

En realizaciones preferidas, el procedimiento matemático comprende una rotación en un espacio de señal bidimensional.

El tercer parámetro espacial puede comprender una diferencia entre el segundo parámetro espacial y el primer parámetro espacial. El tercer parámetro espacial puede implicar codificación diferencial.

El segundo parámetro espacial puede comprender un parámetro ICC basado en coherencia. El tercer parámetro espacial puede comprender una diferencia entre un parámetro ICC basado en coherencia y un parámetro ICC basado en correlación. En una realización preferida, el segundo parámetro espacial comprende un parámetro ICC basado en coherencia, mientras que el tercer parámetro espacial comprende una diferencia entre el segundo parámetro espacial y un parámetro ICC basado en correlación.

El codificador puede estar adaptado además para codificar una señal de audio tercera, cuarta, quinta y sexta, o incluso más, según los principios del primer aspecto combinando estas señales de audio junto con las señales de audio primera y segunda y generar las partes de salida primera y segunda en respuesta a ello. Preferiblemente, un codificador de este tipo está adaptado para codificar una señal de audio 5.1 usando una configuración que comprende una pluralidad de módulos de combinación de codificador. En principio, el principio del codificador según el primer aspecto puede usarse para codificar cualquier dato de audio con formato multicanal.

En un segundo aspecto, la invención proporciona un decodificador de audio para generar una señal de audio mul-

ticanal basándose en una señal codificada, comprendiendo el decodificador:

- un módulo de combinación de decodificador para generar señales de audio primera y segunda basándose en una parte de señal dominante, una parte de señal residual y conjuntos de parámetros espaciales primero y segundo, comprendiendo los parámetros espaciales una descripción de propiedades espaciales de las señales de audio primera y segunda, en el que la parte de señal residual y los segundos parámetros espaciales están implicados en la determinación de una matriz de mezclado que se usa para generar las señales de audio primera y segunda.

Tal como se describe en conexión con el primer aspecto, pueden usarse decodificadores existentes para decodificar la señal de salida codificada a partir de un codificador según la invención utilizando únicamente la parte de señal dominante y los primeros parámetros espaciales. Sin embargo, el decodificador según el segundo aspecto podrá utilizar la segunda parte de salida codificada, es decir la parte de señal residual y un parámetro espacial, para determinar una matriz de mezclado que es inversa de manera idéntica a la combinación de codificador implicada en el proceso de codificación, y por tanto puede obtenerse una regeneración perfecta de las señales de audio primera y segunda.

En realizaciones preferidas, el decodificador comprende un decorrelacionador para recibir la parte de señal dominante y generar una parte de señal dominante decorrelacionada en respuesta a ello. Preferiblemente, una suma de la parte de señal residual y la parte de señal dominante decorrelacionada está implicada en la determinación de la matriz de mezclado. El decodificador puede comprender un atenuador para atenuar la parte de señal dominante decorrelacionada antes de sumarla a la parte de señal residual.

En realizaciones preferidas, la matriz de mezclado aplica una rotación en un espacio de señal bidimensional a las partes de señal dominante y residual.

El decodificador puede estar adaptado para recibir una pluralidad de conjuntos de parámetros primero y segundo y una pluralidad de partes de señal residual para generar una pluralidad de conjuntos de señales de audio primera y segunda en respuesta a ello. En una realización preferida, el decodificador está adaptado para recibir tres conjuntos de parámetros primero y segundo y tres partes de señal residual para generar tres conjuntos de señales de audio primera y segunda en respuesta a ello, en esta realización, el decodificador puede generar seis canales de audio independientes, tales como según el formato 5.1 u otro formato multicanal.

En realizaciones preferidas, el decodificador comprende una pluralidad de matrices de mezclado de canal uno a dos dispuestas en una configuración adecuada para permitir al decodificador decodificar una señal codificada que represente más de dos señales de audio. Por ejemplo, el decodificador puede comprender una configuración de cinco matrices de mezclado dispuestas para generar seis señales de audio y de este modo decodificar, por ejemplo, una señal de audio 5.1 codificada.

En un tercer aspecto, la invención proporciona un procedimiento de codificación de una señal de audio multicanal que comprende las etapas de

- 1) generar una parte de señal dominante y una parte de señal residual que son una representación combinada de señales de audio primera y segunda, obteniéndose la partes de señal dominante y residual aplicando un procedimiento matemático a las señales de audio primera y segunda, en el que el procedimiento matemático implica un primer parámetro espacial que comprende una descripción de propiedades espaciales de las señales de audio primera y segunda,

- 2) generar un primer conjunto de parámetros que comprende un segundo parámetro espacial,

- 3) generar un segundo conjunto de parámetros que comprende un tercer parámetro espacial, y

- 4) generar una señal de salida codificada que comprende una primera parte de salida que comprende la parte de señal dominante y el primer conjunto de parámetros, y una segunda parte de salida que comprende la parte de señal residual y el segundo conjunto de parámetros.

Las mismas ventajas y comentarios descritos en conexión con el primer aspecto se aplican al tercer aspecto.

En un cuarto aspecto, la invención proporciona un procedimiento para generar una señal de audio multicanal basándose en una señal codificada, comprendiendo el procedimiento las etapas de:

- 1) recibir la señal codificada que comprende una parte de señal dominante, una parte de señal residual, y conjuntos de parámetros espaciales primero y segundo que comprenden una descripción de propiedades espaciales de señales de audio primera y segunda,

- 2) determinar una matriz de mezclado basándose en la parte de señal residual y el segundo conjunto de parámetros espaciales,

- 3) generar las señales de audio primera y segunda basándose en la matriz de mezclado determinada.

El procedimiento puede comprender la etapa de decorrelacionar la parte de señal dominante y generar una parte de señal dominante decorrelacionada en respuesta a ello. El procedimiento puede comprender además la etapa de sumar la parte de señal residual y la parte de señal dominante decorrelacionada. La determinación de la matriz de mezclado puede basarse en la parte de señal residual y la parte de señal dominante decorrelacionada sumadas.

Preferiblemente, el procedimiento comprende recibir una pluralidad de conjuntos de parámetros primero y segundo y una pluralidad de partes de señal residual para generar una pluralidad de conjuntos de señales de audio primera y segunda en respuesta a ello. En una realización preferida, el procedimiento comprende recibir tres conjuntos de parámetros primero y segundo y tres partes de señal residual para generar tres conjuntos de señales de audio primera y segunda en respuesta a ello. En esta realización, el procedimiento puede generar seis canales de audio independientes tal como en un formato multicanal 5.1 o equivalente.

Las mismas ventajas y comentarios descritos para el segundo aspecto se aplican para el cuarto aspecto.

En un quinto aspecto, la invención proporciona una señal de audio multicanal codificada que comprende

- una primera parte de señal que comprende una parte de señal dominante y un primer conjunto de parámetros que comprenden una descripción de propiedades espaciales de señales de audio primera y segunda, y

- una segunda parte de señal que comprende una parte de señal residual y un segundo conjunto de parámetros que comprenden una descripción de propiedades espaciales de las señales de audio primera y segunda.

La señal de audio según el quinto aspecto proporciona las mismas ventajas que las expuestas en conexión con el primer aspecto, puesto que esta señal es idéntica a una señal de salida codificada a partir del codificador según el primer aspecto.

En un sexto aspecto, la invención proporciona un medio de almacenamiento que tiene almacenada en el mismo una señal tal como en el quinto aspecto. El medio de almacenamiento puede ser un disco duro, un disco flexible, un CD, un DVD, una tarjeta SD, una tarjeta Memory Stick, un chip de memoria, etc.

En un séptimo aspecto, la invención proporciona un código de programa ejecutable por ordenador adaptado para realizar el procedimiento según el tercer aspecto cuando se ejecuta en un ordenador.

En un octavo aspecto, la invención proporciona un medio de almacenamiento legible por ordenador que comprende un código de programa ejecutable por ordenador según el séptimo aspecto. El medio de almacenamiento puede ser un disco duro, un disco flexible, un CD, un DVD, una tarjeta SD, una tarjeta Memory Stick, un chip de memoria, etc.

En un noveno aspecto, la invención proporciona un código de programa ejecutable por ordenador adaptado para realizar el procedimiento según el cuarto aspecto cuando se ejecuta en un ordenador.

En un décimo aspecto, la invención proporciona un medio de almacenamiento legible por ordenador que comprende un código de programa ejecutable por ordenador según el noveno aspecto. El medio de almacenamiento puede ser un disco duro, un disco flexible, un CD, un DVD, una tarjeta SD, una tarjeta Memory Stick, un chip de memoria, etc.

En un undécimo aspecto, la invención proporciona un dispositivo que comprende un codificador según el primer aspecto. El dispositivo puede ser tal como un equipo de audio de entretenimiento en casa tal como amplificadores de sonido envolvente, receptores de sonido envolvente, reproductores/grabadores de DVD, etc. En principio, el dispositivo puede ser cualquier dispositivo de audio que pueda tratar datos de audio multicanal, por ejemplo en el formato 5.1.

En un duodécimo aspecto, la invención proporciona un dispositivo que comprende un decodificador según el segundo aspecto. El dispositivo puede ser tal como un equipo de audio de entretenimiento en casa tal como amplificadores de sonido envolvente, receptores de sonido envolvente, receptores de A/V, módulos decodificadores, reproductores/grabadores de DVD, etc.

La señal según el quinto aspecto es adecuada para su transmisión a través de una cadena de transmisión. Una cadena de transmisión de este tipo puede comprender un servidor que almacene las señales, una red para la distribución de las señales y clientes que reciban las señales. El lado del cliente puede comprender hardware tal como por ejemplo ordenadores, receptores de A/V, módulos decodificadores, etc. Por tanto, la señal según el quinto aspecto es adecuada para la transmisión de difusión de vídeo digital, difusión de audio digital o radio por Internet, etc.

Se apreciará que en todos los aspectos anteriores, las señales de audio primera y segunda pueden ser señales de ancho de banda completo. Opcionalmente, las señales de audio primera y segunda representan representaciones de subbandas de señales de audio de ancho de banda completo respectivas. En otras palabras, el procesamiento de señales según la invención puede aplicarse a señales de ancho de banda completo o aplicarse en subbandas.

A continuación se describe la invención con más detalle con referencia a las figuras adjuntas, de las que

la figura 1 muestra un esquema de una disposición de altavoces multicanal 5.1,

ES 2 347 274 T3

la figura 2 muestra una unidad de combinación de codificador según la invención,

la figura 3 muestra un codificador preferido para codificar una señal de audio 5.1 basándose en una combinación de codificador para dar una señal mono,

la figura 4 muestra un decodificador preferido correspondiente al codificador de la figura 3,

la figura 5 muestra un codificador preferido para codificar una señal de audio 5.1 basándose en una combinación de codificador para dar una señal estéreo,

la figura 6 muestra un decodificador preferido correspondiente al codificador de la figura 5, y

la figura 7 muestra una gráfica que ilustra los resultados de una prueba de audición realizada con el principio de codificación según la invención.

Aunque la invención es susceptible de diversas modificaciones y formas alternativas, se han mostrado realizaciones específicas a modo de ejemplo en los dibujos y se describirán más detalladamente a continuación en el presente documento. Ha de entenderse, sin embargo, que la invención no pretende limitarse a las formas particulares dadas a conocer. Más bien, la invención cubrirá todas las modificaciones, equivalentes y alternativas que se encuentren dentro del alcance de la invención según se define en las reivindicaciones adjuntas.

La figura 1 muestra un esquema de una disposición de audio multicanal 5.1 típica con un oyente LP situado en el centro de cinco altavoces C, Lf, Ls, Rf y Rs que reciben señales de audio independientes. Éstas se proporcionan para producir en el oyente LP una impresión de audio espacial. La disposición 5.1 proporciona además una señal LFE de *subwoofer* separada. Por tanto, una representación de señal completa para tal disposición multicanal requiere en total de seis canales de audio independientes, y por tanto es necesaria una gran tasa de transmisión de bits para representar una señal de audio para un sistema de este tipo con una calidad de audio total. A continuación se describirán realizaciones de la invención que pueden proporcionar una alta calidad de audio en un sistema 5.1 con una baja tasa de transmisión de bits.

La figura 2 muestra una unidad de combinación de codificador 2-1 EU según la invención. Se introducen señales de audio primera y segunda x_1 , x_2 en un módulo de combinación de codificador ECM en el que se realiza un procedimiento matemático sobre las señales de audio primera y segunda x_1 , x_2 , que comprende preferiblemente una rotación de señal, con el fin de combinar las señales de audio primera y segunda x_1 , x_2 y generar una representación paramétrica de las mismas que comprende una parte de señal dominante m y una parte de señal residual s . Un primer parámetro espacial SP1, es decir un parámetro que describe propiedades de señal espaciales de las señales de audio primera y segunda x_1 , x_2 , está implicado en el procedimiento matemático de combinación de codificador.

Un generador de parámetros PG genera conjuntos de parámetros primero y segundo PS1, PS2 basándose en las señales de audio primera y segunda x_1 , x_2 . El primer conjunto de parámetros PS1 comprende un segundo parámetro espacial SP2, y el segundo conjunto de parámetros PS2 comprende un tercer parámetro espacial SP3. La señal de salida codificada comprende una primera parte de salida OP1 que comprende la parte de señal dominante m y el primer conjunto de parámetros PS1, mientras que una segunda parte de salida OP2 comprende la parte de señal residual s y el segundo conjunto de parámetros PS2.

Mediante una elección apropiada de los parámetros espaciales segundo y tercero SP2, SP3 en relación con el primer parámetro espacial SP1 es posible obtener la inversa de la combinación de codificador o procedimiento de rotación en el lado del decodificador, y por tanto pueden decodificarse las señales de audio primera y segunda x_1 , x_2 de manera transparente.

Preferiblemente, el codificador pone la primera parte de salida en una capa base de su flujo de bits de salida, mientras que la segunda parte de salida se pone en una capa de refinamiento del flujo de bits de salida. Durante la decodificación es posible usar sólo la capa base, si es aceptable una calidad de señal reducida, mientras que puede obtenerse la mejor calidad de señal posible si también se incluye la capa de refinamiento en el proceso de decodificación.

El principio de codificación descrito proporciona un codificador de audio multicanal híbrido ajustable a escala con retrocompatibilidad total. El decodificador puede usarse para los siguientes escenarios: 1) señal mono o estéreo decodificada sólo, 2) salida multicanal decodificada sin usar las señales residuales, y 3) salida multicanal decodificada con señales residuales.

A continuación se describen realizaciones preferidas de módulos de combinación de codificador y parámetros espaciales. Un módulo de combinación de codificador preferido combina señales de audio primera y segunda x_1 , x_2 para dar una parte de señal dominante m y una parte de señal residual s maximizando la amplitud de la suma de las señales rotadas según:

$$\begin{pmatrix} m[k] \\ s[k] \end{pmatrix} = \begin{pmatrix} sc_{corr} & sc_{corr} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1[k] \\ x_2[k] \end{pmatrix}, \text{ donde}$$

$$sc_{corr} = \min \left\{ sc_{corr,max}, \frac{1}{c_l \cos(\chi + \beta) + c_r \cos(-\chi + \beta)} \right\}.$$

(Ec 1)

Los coeficientes de rotación de amplitud implicados en sc_{corr} se derivan a partir de ICC e IID, es decir se basan en propiedades espaciales de las señales de audio primera y segunda x_1 , x_2 . Estos coeficientes de rotación de amplitud se calculan preferiblemente según:

$$\chi = \frac{1}{2} \cos^{-1}(ICC), \beta = \tan^{-1} \left(\tan(\chi) \frac{c_r - c_l}{c_r + c_l} \right), c_l = \sqrt{\frac{IID}{1 + IID}}, c_r = \sqrt{\frac{1}{1 + IID}}.$$

La señal residual s se selecciona para que sea la diferencia entre x_1 y x_2 . Obsérvese que esta matriz es siempre invertible, ya que sc_{corr} nunca puede ser cero, lo que significa que puede conseguirse una reconstrucción perfecta siempre que se conozca sc_{corr} . Un valor adecuado para la constante de corte $sc_{corr,max}$ es 1,2.

Para derivar sc_{corr} en el decodificador, el segundo conjunto de parámetros PS2 preferiblemente comprende una diferencia entre parámetros de coherencia y correlación y por tanto se transmiten junto con la correspondiente señal residual s en una capa de refinamiento en el flujo de bits ajustable a escala. El primer conjunto de parámetros PS1 se selecciona para comprender o bien parámetros de coherencia o bien parámetros de correlación y por tanto para transmitirse en la capa base junto con la parte de señal dominante m .

Cuando la señal residual s está disponible para el decodificador, se derivan parámetros de correlación, lo que facilita el cálculo de sc_{corr} , y puede determinarse la inversa de la matriz de mezclado de Ec 1:

$$\begin{pmatrix} x_1[k] \\ x_2[k] \end{pmatrix} = \begin{pmatrix} \frac{1}{2sc_{corr}} & 1 \\ \frac{1}{2sc_{corr}} & -1 \end{pmatrix} \begin{pmatrix} m[k] \\ s[k] \end{pmatrix}.$$

En otra realización preferida, el módulo de combinación de codificador se basa en el análisis de componentes principales (PCA) y mezcla las señales de audio primera y segunda x_1 , x_2 según:

$$\begin{pmatrix} m[k] \\ s[k] \end{pmatrix} = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} x_1[k] \\ x_2[k] \end{pmatrix},$$

donde un coeficiente α preferido se basa en ICC e IID según:

$$\alpha = \frac{1}{2} \tan^{-1} \left(\frac{2ICC \cdot c}{c^2 - 1} \right), c = 10^{\frac{IID}{20}}.$$

Opciones preferidas para codificar el segundo conjunto de parámetros PS2 que ha de incluirse en la capa de refinamiento son los parámetros de correlación que incluyen lo siguiente:

1) Codificación diferencial en tiempo o frecuencia de los parámetros de correlación, independientemente de los parámetros de coherencia en la capa base.

2) Codificación diferencial de los parámetros de correlación con respecto a los parámetros de coherencia en la capa base (es decir, $\Delta ICC = ICC_{\text{correlación}} - ICC_{\text{coherencia}}$).

Una combinación de 1 y 2, dependiendo de cuál requiera la menor cantidad de bits.

3) Las figuras 3 y 4 ilustran configuraciones preferidas de un codificador en formato 5.1 y un decodificador 5.1 correspondiente, respectivamente, que se basan en una combinación de codificador para dar una señal mono codificada. Las figuras 5 y 6 ilustran un codificador en formato 5.1 alternativo y un decodificador correspondiente, respectivamente, que se basan en una combinación de codificador para dar una señal estéreo codificada.

La figura 3 muestra una configuración de codificador basada en una combinación de seis señales de audio independientes lf, ls, rf, rs, co, lfe para dar una señal mono m, por ejemplo las seis señales de audio representan las señales lf, ls, rf, rs, co, lfe en un formato 5.1. El codificador comprende cinco unidades de combinación de codificador EU, tal como se ha descrito anteriormente, estando dispuestas estas unidades EU para combinar de manera sucesiva las seis señales lf, ls, rf, rs, co, lfe en una única señal mono m. Una etapa de segmentación y transformación inicial ST se realiza para pares de señales antes de la combinación de codificador. Esta etapa ST comprende la segmentación de las señales de audio en el dominio del tiempo para dar segmentos superpuestos y entonces la transformación de estos segmentos superpuestos en representaciones en el dominio de la frecuencia (indicadas mediante letras mayúsculas).

Tras la segmentación y la transformación ST, los dos canales izquierdos Lf y Ls se combinan para dar una parte de señal dominante L, conjuntos de parámetros primero y segundo PS1a, PS1b y una señal residual ResL. Los dos canales derechos Rf, Rs se combinan para dar una parte de señal dominante R, conjuntos de parámetros primero y segundo PS2a, PS2b y una señal residual ResR. Las partes de señal dominante L y R resultantes se combinan entonces para dar una parte de señal dominante LR, una parte de señal residual ResLR y parámetros primero y segundo PS4a, PS4b. El canal central C0 y el canal LFE de *subwoofer* se combinan para dar una parte de señal dominante C, conjuntos de parámetros primero y segundo PS3a, PS3b y una señal residual ResC. Finalmente, las partes de señal dominante C y LR se combinan para dar una parte de señal dominante M, una parte de señal residual ResM y parámetros primero y segundo PS5a, PS5b.

Preferiblemente, los conjuntos de parámetros primero y segundo PS1a-PS5a, PS1b-PS5b se determinan independientemente para varias bandas de frecuencia (subbandas) en un segmento antes de la cuantificación, la codificación y la transmisión, sin embargo, si se prefiere, el procesamiento puede realizarse sobre señales de ancho de banda completo. Tras aplicar el análisis y el procesamiento de las señales, puede aplicarse un procesamiento opcional IT, OLA: pueden transformarse de manera inversa segmentos IT de vuelta al dominio del tiempo, y pueden superponerse segmentos y sumarse OLA para obtener la señal de audio mono m en el dominio del tiempo. En general, el codificador genera una primera parte de salida que comprende la parte de señal dominante m y cinco conjuntos de parámetros PS1a-PS5a, y una segunda parte de salida que comprende cinco partes de señal residual ResL, ResR, ResLR, ResM, ResC, y cinco conjuntos de parámetros PS1b, PS5b.

La figura 4 muestra un decodificador correspondiente al codificador de la figura 3, es decir que está adaptado para recibir la señal de salida desde el codificador de la figura 3. El decodificador esencialmente aplica la inversa del procesamiento descrito para la figura 3. El decodificador comprende una segmentación (opcional) inicial y se aplica transformación en frecuencia ST a la parte de señal dominante m. El decodificador comprende cinco unidades de combinación de decodificador DU similares, de las que se indica una con una línea discontinua. La unidad de combinación de decodificador DU comprende una matriz de mezclado MM que genera señales primera y segunda basándose en una parte de señal dominante. La matriz de mezclado MM, es decir la inversa de la matriz de mezclado aplicada en el módulo de combinación de codificador ECM, se determina basándose en la parte de señal dominante, la parte residual y conjuntos de parámetros primero y segundo recibidos.

En la primera unidad de combinación de decodificador DU indicada en la figura 4, la señal dominante M se decorrelaciona en primer lugar en un decorrelacionador Dec y después se atenúa en un atenuador Att. La parte de señal dominante decorrelacionada y atenuada se suma entonces a la parte de señal residual ResM. Esta señal sumada se usa entonces para determinar la matriz de mezclado MM. El atenuador Att se ajusta en respuesta a la parte de señal residual ResM y el primer conjunto de parámetros PS5a. Finalmente, la matriz de mezclado MM se determina usando los conjuntos de parámetros primero y segundo PS5a, PS5b. La matriz de mezclado MM determinada combina entonces la parte de señal dominante M para dar una primera señal de salida LR y una segunda señal de salida C. Estas señales de salida primera y segunda LR, C se aplican entonces a unidades de combinación de codificador respectivas y se combinan sucesivamente para producir L, R, y C0, LFE, respectivamente. Finalmente, se aplica a L una combinación de decodificador para dar Lf y Lr, mientras que se aplica a R una combinación de decodificador para producir Rf y Rr. Tras aplicar análisis y procesamiento de señales, los segmentos se transforman de manera inversa IT de vuelta al dominio del tiempo, y se superponen los segmentos y suman OLA para obtener las representaciones lf, lr, rf, rr, co, lfe en el dominio del tiempo. Esta transformación inversa y superposición-suma IT, OLA son opcionales.

La figura 5 muestra una realización de codificador en la que se usan tres unidades de combinación de codificador, que funcionan cada una según los principios descritos en conexión con el codificador de la figura 3, para combinar seis señales de audio Lf, Lr, Rf, Rr, C0, LFE en pares para dar tres partes de señal dominante L, R, C con primeros conjuntos de parámetros PS1a-PS3a, segundos conjuntos de parámetros PS1b-PS3b y partes de señal residual ResL, ResR, ResC asociados. Una unidad de combinación de codificador 3-2 se aplica entonces a las tres partes de señal dominante L, R

y C dando como resultado dos partes de señal dominante L0, R0 y una parte de señal residual ResEo y un conjunto de parámetros PS4. Opcionalmente, se aplica una segmentación y transformación al dominio de la frecuencia ST inicial, y se aplica (opcionalmente) una transformación inversa IT y una superposición-suma OLA final, tal como se describe también en conexión con la figura 3.

La figura 6 muestra una configuración de decodificador adaptada para decodificar una salida del codificador de la figura 5. Tras una segmentación y transformación al dominio de la frecuencia ST (opcional) inicial de señales de entrada lo, ro, un módulo de combinación de decodificador 2-3 genera partes de señal dominante L, R, C en respuesta a las partes de señal dominante Lo, Ro y la parte de señal residual ResEo junto con el conjunto de parámetros PS4. Estas tres partes de señal dominante L, R, C se procesan entonces en respectivas unidades de combinación de decodificador similares a las unidades de combinación de decodificador DU descritas en conexión con el decodificador de la figura 4. Una transformación inversa IT y superposición-suma OLA final se aplica (opcionalmente) tal como también se describió anteriormente.

La figura 7 ilustra los resultados de una prueba de audición realizada para cinco oyentes entrenados. Los elementos musicales A-K usados son los especificados en el trabajo MPEG "*Spatial Audio Coding*". Para cada elemento A-K, se incluyeron en la prueba resultados para tres versiones codificadas: 1) decodificador sin residuales - mostrado a la izquierda, 2) codificador espacial con residuales, es decir un decodificador según la invención - mostrado en el centro y 3) referencia (oculta) - mostrada a la derecha. Un promedio total de los elementos A-K se muestra como TOT. Para cada versión codificada se indica un grado promedio GRD con un asterisco (*), mientras que se indican +/- desviación estándar respecto al mismo para respuestas de los oyentes.

Para el escenario 2) y 3) se usó el principio de codificador/decodificador ilustrado en las figuras 5 y 6. En el escenario 2) se descartaron las partes de señal residual. Para el escenario 3) se usaron tres partes de señal residual, banda limitada a 2 kHz: parte de señal residual para el canal izquierdo ResL, parte de señal residual para el canal central ResR, y parte de señal residual ResEo para el módulo de combinación de decodificador 3-2. Cada una de las señales residuales ResL, ResR, ResEo se codificó a una tasa de transmisión de bits de 8 kbit/s, y los parámetros espaciales adicionales (que fueron diferencias entre parámetros de correlación (capa de refinamiento) y coherencia (capa base)) requirieron una tasa de transmisión de bits estimada de 700 bit/s. Por tanto, la tasa de transmisión de bits total relacionada con los residuales adicionales es entonces de aproximadamente 25 kbit/s. Los parámetros espaciales convencionales (que van a colocarse en la capa base), requirieron de forma estimada 10 kbit/s. La tasa de transmisión de datos espaciales total es por tanto de aproximadamente 35 kbit/s. No se aplicó ningún códec central a la señal estéreo lo, ro.

A partir de los resultados, queda claro que puede obtenerse una gran mejora de la calidad utilizando tres señales residuales codificadas a una baja tasa de transmisión de bits. Además, el grado de calidad promedio total es de +/- 92, muy próximo a lo que se considera calidad de audio "transparente".

El codificador y el decodificador según la invención pueden aplicarse en todas las aplicaciones que impliquen codificación de audio multicanal, incluyendo: difusión de vídeo digital (DVB), difusión de audio digital (DAB), radio por Internet, distribución de música electrónica.

Los símbolos de referencia en las reivindicaciones sirven meramente para aumentar la legibilidad. Estos símbolos de referencia no deben en ningún caso interpretarse como que limitan el alcance de las reivindicaciones, sino que son únicamente ejemplos ilustrativos incluidos.

REIVINDICACIONES

1. Codificador de audio adaptado para codificar una señal de audio multicanal, comprendiendo el codificador:

- un módulo de combinación de codificador (ECM) para generar una parte de señal dominante (m) y una parte de señal residual (s) que son una representación combinada de señales de audio primera y segunda (x1, x2), obteniéndose las partes de señal dominante y residual (m, s) aplicando un procedimiento matemático a las señales de audio primera y segunda (x1, x2), en el que el procedimiento matemático implica un primer parámetro espacial (SP1) que comprende una descripción de propiedades espaciales de las señales de audio primera y segunda (x1, x2),

- un generador de parámetros (PG) para generar

- un primer conjunto de parámetros (PS1) que comprende un segundo parámetro espacial (SP2), y

- un segundo conjunto de parámetros (PS2) que comprende un tercer parámetro espacial (SP3), y

caracterizado porque el codificador comprende

- un generador de salida para generar una señal de salida codificada que comprende

- una primera parte de salida (OP1) que comprende la parte de señal dominante (m) y el primer conjunto de parámetros (PS1), y

- una segunda parte de salida (OP2) que comprende la parte de señal residual (s) y el segundo conjunto de parámetros (PS2).

2. Codificador de audio según la reivindicación 1, en el que el tercer parámetro espacial (SP3) comprende una diferencia entre el segundo parámetro espacial (SP2) y el primer parámetro espacial (SP1).

3. Codificador de audio según la reivindicación 1, en el que el segundo parámetro espacial (SP2) comprende un parámetro basado en coherencia.

4. Codificador de audio según la reivindicación 1, en el que el tercer parámetro espacial (SP3) comprende una diferencia entre un parámetro basado en coherencia y un parámetro basado en correlación.

5. Codificador de audio según la reivindicación 1, en el que la parte de señal residual (s) comprende una diferencia entre las señales de audio primera y segunda (x1, x2).

6. Codificador de audio según la reivindicación 1, en el que el módulo de combinación de codificador (ECM) está adaptado para generar las partes de señal dominante y residual (m, s) de tal manera que estas partes de señal (m, s) estén menos correlacionadas que las señales de audio primera y segunda (x1, x2).

7. Codificador de audio según la reivindicación 1, que está adaptado además para recibir señales de audio tercera, cuarta, quinta y sexta y mezclar de manera descendente estas señales junto con las señales de audio primera y segunda (x1, x2) y para generar las partes de salida primera y segunda en respuesta a ello.

8. Decodificador de audio para generar una señal de audio multicanal basándose en una señal codificada, comprendiendo el decodificador:

- una unidad de combinación de decodificador (DU) para generar señales de audio primera y segunda basándose en una parte de señal dominante, una parte de señal residual y conjuntos de parámetros primero y segundo, comprendiendo los conjuntos de parámetros primero y segundo una descripción de propiedades espaciales de las señales de audio primera y segunda, en el que la parte de señal residual y el segundo conjunto de parámetros están implicados en la determinación de una matriz de mezclado (MM) que se usa para generar las señales de audio primera y segunda.

9. Decodificador de audio según la reivindicación 8, comprendiendo el decodificador un decorrelacionador (Dec) para recibir la parte de señal dominante y generar una parte de señal dominante decorrelacionada en respuesta a ello.

10. Decodificador de audio según reivindicación 9, en el que una suma de la parte de señal residual y la parte de señal dominante decorrelacionada está implicada en la determinación de la matriz de mezclado (MM).

11. Decodificador de audio según la reivindicación 10, comprendiendo el decodificador un atenuador (Att) para atenuar la parte de señal dominante decorrelacionada antes de sumarla a la parte de señal residual.

12. Decodificador de audio según la reivindicación 8, adaptado para recibir una pluralidad de conjuntos de conjuntos de parámetros primero y segundo y una pluralidad de partes de señal residual para generar una pluralidad de conjuntos de señales de audio primera y segunda en respuesta a ello.

5 13. Decodificador de audio según la reivindicación 12, estando adaptado el decodificador para recibir tres conjuntos de conjuntos de parámetros primero y segundo y tres partes de señal residual para generar tres conjuntos de señales de audio primera y segunda en respuesta a ello.

10 14. Procedimiento de codificación de una señal de audio multicanal que comprende las etapas de:

1) generar una parte de señal dominante (m) y una parte de señal residual (s) que son una representación combinada de señales de audio primera y segunda (x1, x2), obteniéndose las partes de señal dominante y residual (m, s) aplicando un procedimiento matemático a las señales de audio primera y segunda (x1, x2), en el que el procedimiento matemático implica un primer parámetro espacial que comprende una descripción de propiedades espaciales de las señales de audio primera y segunda (x1, x2),

2) generar un primer conjunto de parámetros que comprende un segundo parámetro espacial,

3) generar un segundo conjunto de parámetros que comprende un tercer parámetro espacial, y
caracterizado por

4) generar una señal de salida codificada que comprende una primera parte de salida que comprende la parte de señal dominante (m) y el primer conjunto de parámetros, y una segunda parte de salida que comprende la parte de señal residual (s) y el segundo conjunto de parámetros.

15 15. Procedimiento de generación de una señal de audio multicanal basándose en una señal codificada, comprendiendo el procedimiento las etapas de:

30 1) recibir la señal codificada que comprende una parte de señal dominante, una parte de señal residual, y conjuntos de parámetros primero y segundo que comprenden una descripción de propiedades espaciales de señales de audio primera y segunda,

35 2) determinar una matriz de mezclado (MM) basándose en la parte de señal residual y el segundo conjunto de parámetros,

3) generar las señales de audio primera y segunda basándose en la matriz de mezclado determinada.

40 16. Procedimiento según la reivindicación 15, que comprende la etapa de decorrelacionar la parte de señal dominante y generar una parte de señal dominante decorrelacionada en respuesta a ello.

17. Procedimiento según la reivindicación 16, que comprende además la etapa de sumar la parte de señal residual y la parte de señal dominante decorrelacionada.

45 18. Procedimiento según la reivindicación 17, en el que la determinación de la matriz de mezclado (MM) se basa en la parte de señal residual y la parte de señal dominante decorrelacionada sumadas.

50 19. Procedimiento según la reivindicación 15, que comprende recibir una pluralidad de conjuntos de conjuntos de parámetros primero y segundo y una pluralidad de partes de señal residual para generar una pluralidad de conjuntos de señales de audio primera y segunda en respuesta a ello.

20. Procedimiento según la reivindicación 19, que comprende recibir tres conjuntos de conjuntos de parámetros primero y segundo y tres partes de señal residual para generar tres conjuntos de señales de audio primera y segunda en respuesta a ello.

55 21. Señal de audio multicanal codificada que comprende:

- una primera parte de señal (OP1) que comprende una parte de señal dominante (m) y un primer conjunto de parámetros (PS1) que comprende una descripción de propiedades espaciales de señales de audio primera y segunda (x1, x2), y

- una segunda parte de señal (OP2) que comprende una parte de señal residual (s) y un segundo conjunto de parámetros (PS2) que comprende una descripción de propiedades espaciales de las señales de audio primera y segunda (x1, x2).

65 22. Medio de almacenamiento que tiene almacenada en el mismo una señal según la reivindicación 21.

ES 2 347 274 T3

23. Código de programa ejecutable por ordenador adaptado para realizar el procedimiento según la reivindicación 14 cuando se ejecuta en un ordenador.

5 24. Medio de almacenamiento legible por ordenador que comprende un código de programa ejecutable por ordenador según la reivindicación 23.

25. Código de programa ejecutable por ordenador adaptado para realizar el procedimiento según la reivindicación 15 cuando se ejecuta en un ordenador.

10 26. Medio de almacenamiento legible por ordenador que comprende un código de programa ejecutable por ordenador según la reivindicación 25.

27. Dispositivo que comprende un codificador según la reivindicación 1.

15 28. Dispositivo que comprende un decodificador según la reivindicación 8.

20

25

30

35

40

45

50

55

60

65

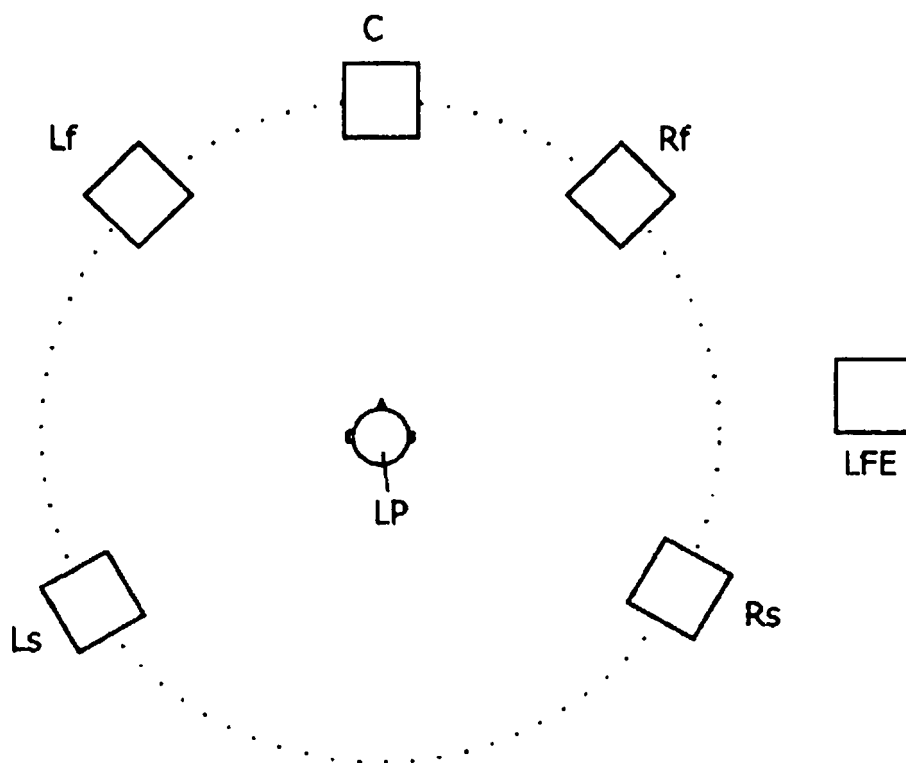


FIG.1

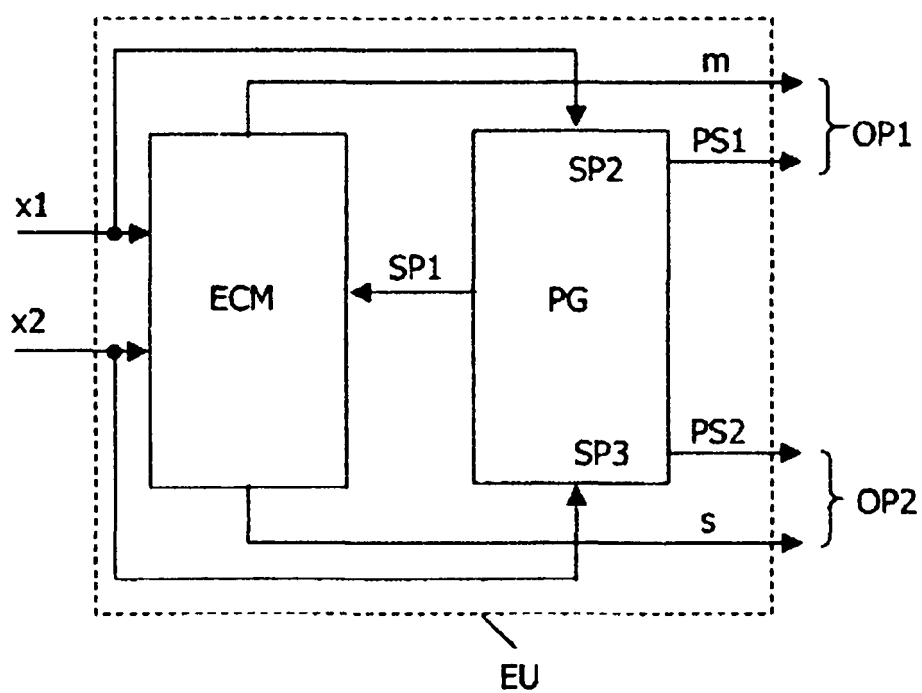


FIG.2

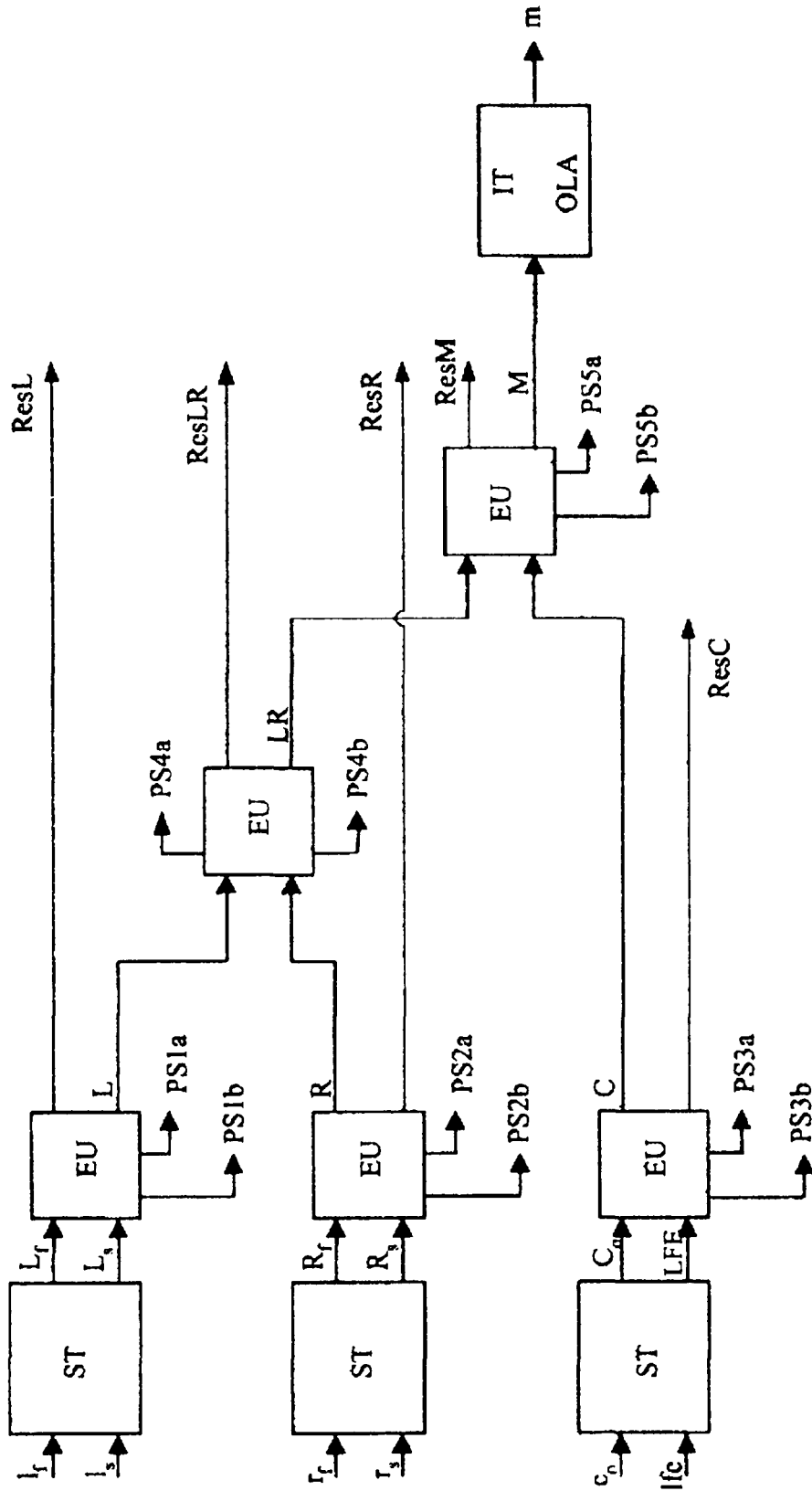
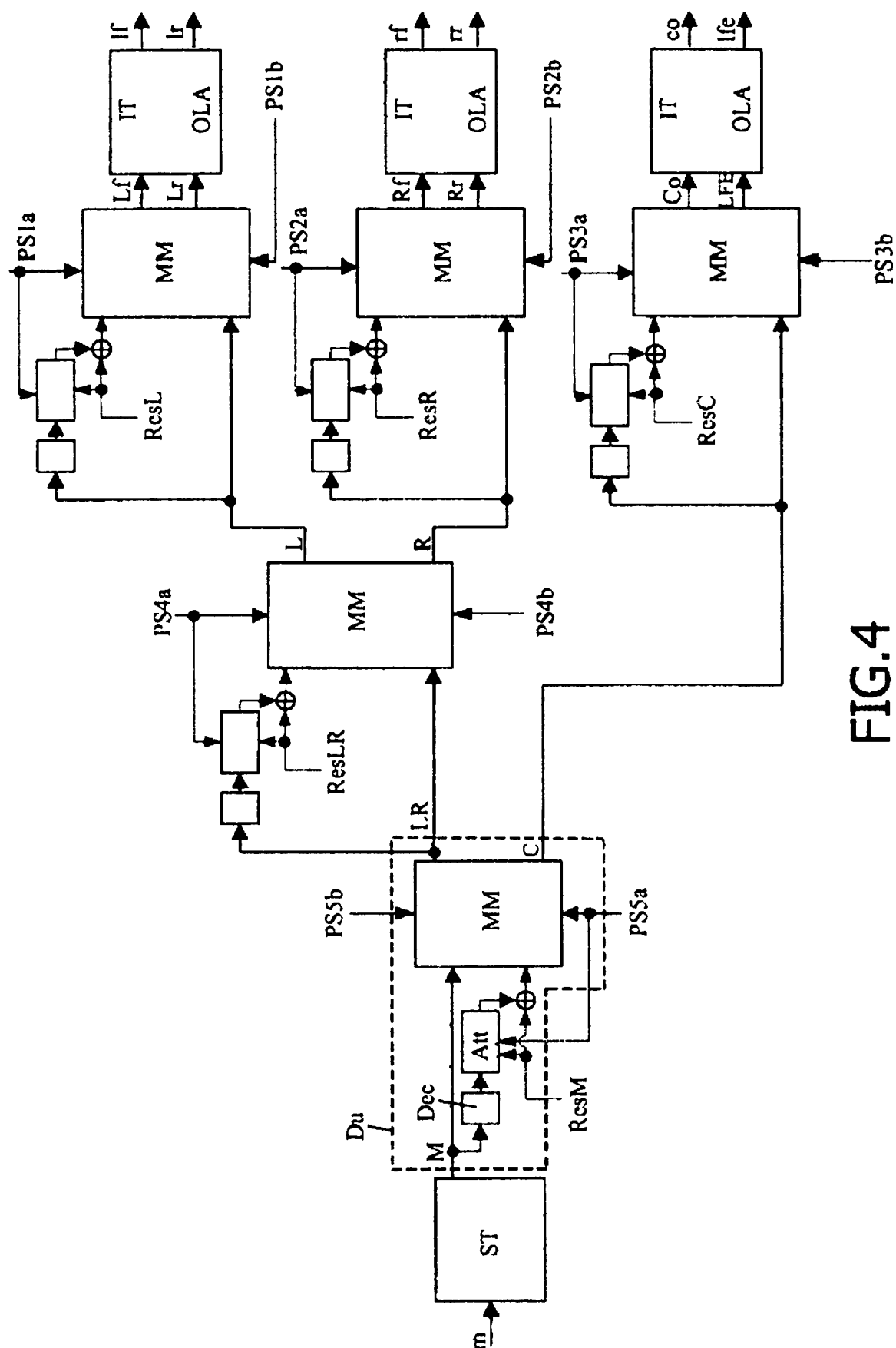


FIG.3



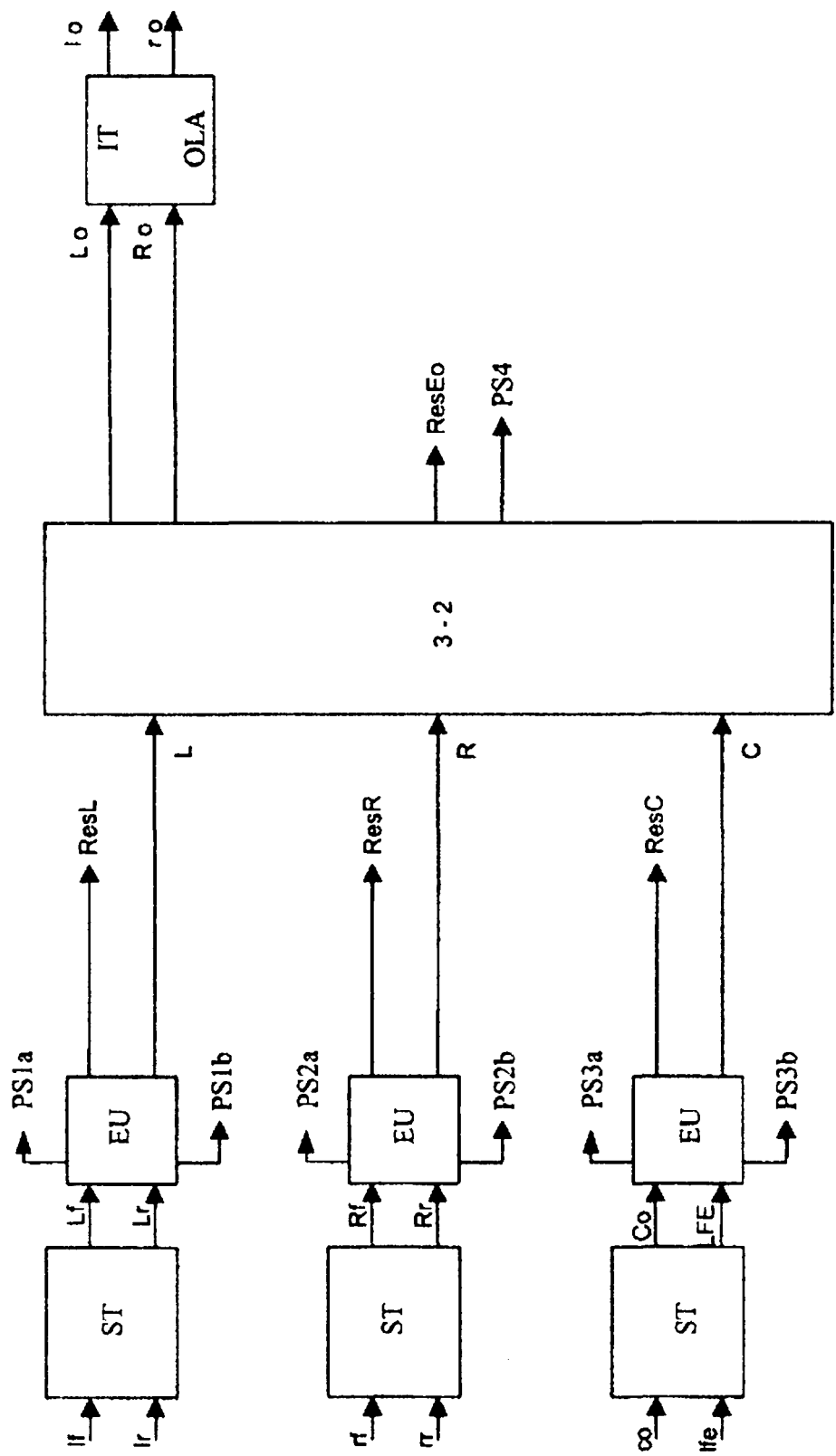


FIG.5

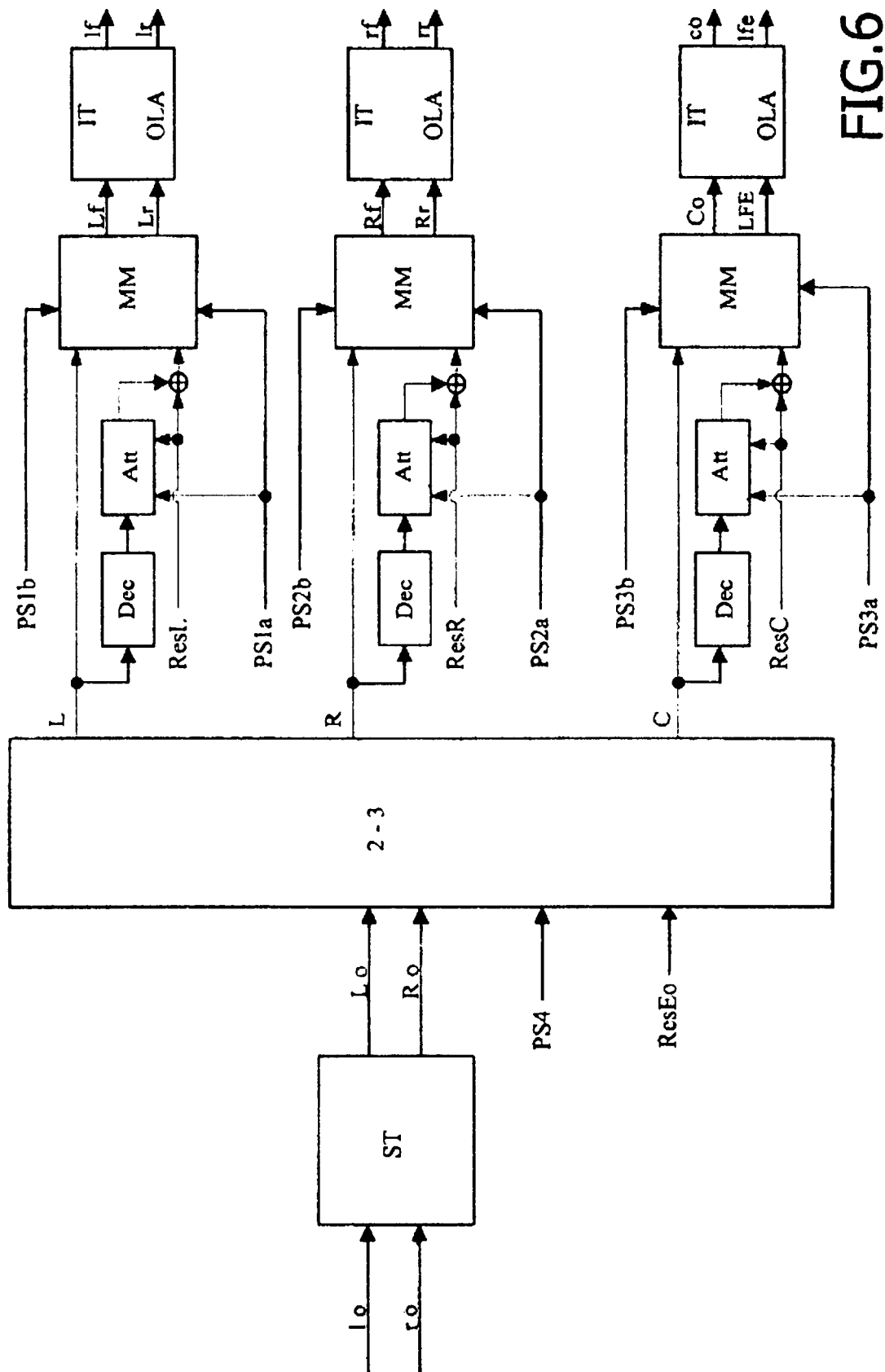


FIG.6

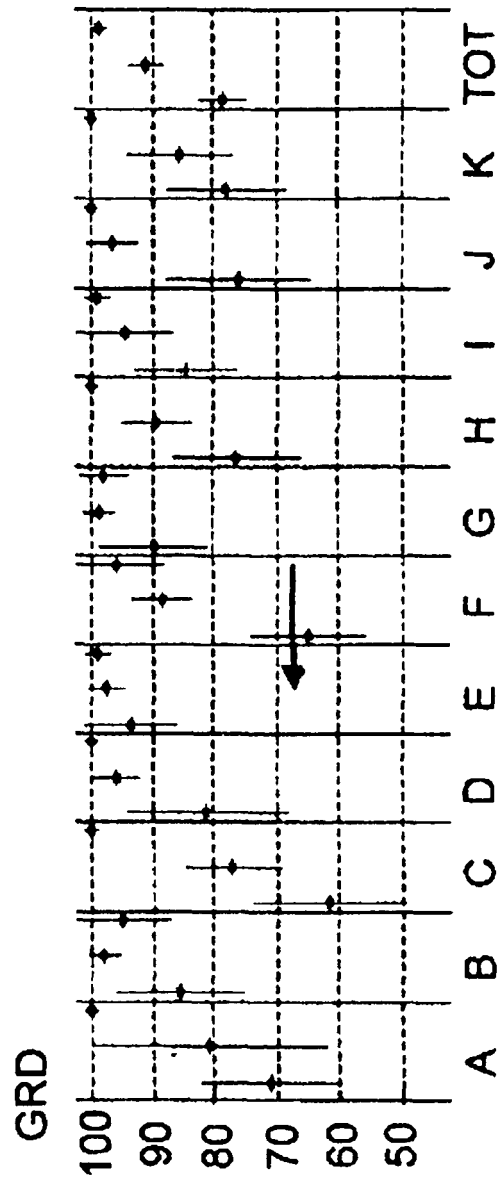


FIG.7