(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 373 728**

(51) Int. Cl.:
H04S 3/02 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Número de solicitud europea: **05761091 .7**
(96) Fecha de presentación: **07.07.2005**
(97) Número de publicación de la solicitud: **1769655**
(97) Fecha de publicación de la solicitud: **04.04.2007**

(54) Título: **MÉTODO, DISPOSITIVO, APARATO CODIFICADOR, APARATO DECODIFICADOR Y SISTEMA DE AUDIO.**

(30) Prioridad:
14.07.2004 EP 04103365

(45) Fecha de publicación de la mención BOPI:
08.02.2012

(45) Fecha de la publicación del folleto de la patente:
08.02.2012

(73) Titular/es:
Koninklijke Philips Electronics N.V.
Groenewoudseweg 1
5621 BA Eindhoven, NL y
Dolby International AB

(72) Inventor/es:
VAN LOON, Machiel, W.;
BREEBAART, Dirk, J.;
HOTHO, Gerard, H.;
SCHUIJERS, Erik, G., P.;
PURNHAGEN, Heiko y
RÖDÉN, Karl, J.

(74) Agente: **Arizti Acha, Monica**

ES 2 373 728 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).

DESCRIPCION

Método, dispositivo, aparato codificador, aparato decodificador y sistema de audio.

La invención se refiere a un método y un dispositivo para procesar una señal de mezcla descendente estéreo que comprende señales estéreo primera y segunda, codificando la señal de mezcla descendente estéreo y parámetros espaciales asociados una señal de audio de N canales. La invención también se refiere a un aparato codificador que comprende un codificador y un dispositivo de este tipo.

La invención también se refiere a un método y un dispositivo para procesar una señal de mezcla descendente estéreo obtenida por un método y dispositivo de este tipo. La invención también se refiere a un aparato decodificador que comprende un dispositivo de este tipo para procesar una señal de mezcla descendente estéreo.

La invención también se refiere a un sistema de audio que comprende un aparato codificador de este tipo y un aparato decodificador de este tipo.

Durante mucho tiempo, ha prevalecido la reproducción en estéreo de la música, por ejemplo, en el entorno doméstico. Durante los años 70, se realizaron algunos experimentos con reproducción de cuatro canales en un equipo de música doméstico.

En salas más grandes, tales como salas de cine, durante mucho tiempo ha estado presente una reproducción multicanal del sonido. Se desarrollaron Dolby Digital® y otros sistemas para proporcionar una reproducción realista e impresionante del sonido en una sala grande.

Tales sistemas multicanal se han introducido en el cine en casa y están adquiriendo un gran interés. Por tanto, los sistemas que tienen cinco canales de gama completa y un canal de gama parcial o canal de efectos de baja frecuencia (LFE), denominados sistemas 5.1, son comunes en el mercado actual. También existen otros sistemas, tales como 2.1, 4.1, 7.1 e incluso 8.1.

Con la introducción de SACD y DVD, la reproducción de audio multicanal va ganando terreno. Muchos consumidores ya tienen la posibilidad de reproducción multicanal en sus hogares, y el material de fuente multicanal se está volviendo muy popular. Sin embargo, muchas personas aún tienen sistemas de reproducción de sólo 2 canales, y la transmisión habitualmente tiene lugar por medio de 2 canales. Por este motivo, se desarrollaron técnicas de matrizado como, por ejemplo, Dolby Surround®, para hacer posible una transmisión de audio multicanal por medio de 2 canales. La señal transmitida puede reproducirse directamente con un sistema de reproducción de 2 canales. Cuando un decodificador apropiado está disponible, es posible la reproducción multicanal. Los decodificadores ampliamente conocidos para este fin son Dolby Pro Logic® (I y II), (Kenneth Gundry, “A new active matrix decoder for surround sound”, en Proc. AES 19ª International Conference on Surround Sound, junio de 2001) y Circle Surround® (I y II) (patente estadounidense n.º 6.198.827: sistema matricial 5-2-5).

Debido al aumento en la popularidad del material multicanal, la codificación eficaz de material multicanal se está volviendo más importante. El matrizado reduce el número de canales de audio requerido para la transmisión y por tanto se reduce el ancho de banda o tasa de transmisión de bits requeridos. Una ventaja adicional de la técnica matricial es que es retrocompatible con sistemas de reproducción en estéreo. Para una reducción adicional de la tasa de transmisión de bits, puede aplicarse un codificador de audio convencional para codificar la señal estéreo matrizada.

Otra posibilidad para reducir la tasa de transmisión de bits es codificando todos los canales individuales sin matrizado. Este método da como resultado una tasa de transmisión de bits superior, ya que tienen que codificarse cinco canales en lugar de dos, pero la reconstrucción espacial puede estar mucho más cerca de la original que aplicando matrizado.

En principio, el proceso de matrizado es una operación con pérdidas. Por tanto, la reconstrucción perfecta de los 5 canales a partir de sólo una mezcla de 2 canales es, en general, imposible. Esta propiedad limita la calidad de percepción máxima de la reconstrucción de 5 canales.

Recientemente se ha desarrollado un sistema que codifica audio multicanal como una señal de audio estéreo de 2 canales y un pequeño número de parámetros espaciales o parámetros de información de codificador *P*. Por consiguiente, este sistema es retrocompatible para la reproducción en estéreo. Los parámetros espaciales o parámetros de información de codificador P transmitidos determinan cómo el decodificador debe reconstruir cinco canales a partir de la señal de mezcla descendente estéreo de dos canales disponible. Debido al hecho de que el proceso de mezcla ascendente se controla por los parámetros transmitidos, la calidad de percepción de la reconstrucción de 5 canales mejora de manera considerable en comparación con los algoritmos de mezcla ascendente sin parámetros de control (por ejemplo, Dolby Pro Logic).

En resumen, pueden aplicarse tres métodos diferentes para generar una reconstrucción de 5 canales a partir de una mezcla de dos canales proporcionada:

1) Reconstrucción ciega. Este método intenta estimar la matriz de mezcla ascendente basándose en sólo las propiedades de la señal, sin ninguna información proporcionada.

2) Técnicas de matrizado, por ejemplo, Dolby Pro Logic. Aplicando una determinada matriz de mezcla descendente, puede mejorarse la reconstrucción desde 2 hasta 5 canales debido a determinadas propiedades de la señal que están determinadas por la matriz de mezcla descendente aplicada.

3) Mezcla ascendente controlada por parámetros. En este método, los parámetros de información de codificador *P* se almacenan normalmente en partes auxiliares de un flujo de bits, garantizando la retrocompatibilidad con sistemas de reproducción en estéreo normales. Sin embargo, estos sistemas son generalmente no retrocompatibles con sistemas de matrizado.

Puede ser de interés combinar los métodos 2 y 3 mencionados anteriormente en un único sistema. Esto garantiza una calidad máxima, dependiendo del decodificador disponible. Para los consumidores que tienen un decodificador de envolvente de matriz, tal como Dolby Pro Logic o Circle Surround, se obtiene una reconstrucción según el proceso matricial. Si se dispone de un decodificador que puede interpretar los parámetros transmitidos, puede obtenerse una reconstrucción de calidad superior. Los consumidores que no tienen un decodificador de envolvente de matriz o un decodificador que pueda interpretar los parámetros espaciales todavía pueden disfrutar de la retrocompatibilidad estéreo. Sin embargo, un problema de combinar los métodos 2 y 3 es que se modificará la mezcla descendente estéreo transmitida real. Esto, a su vez, puede tener un efecto negativo sobre la reconstrucción de 5 canales usando los parámetros espaciales.

El documento US 5 818 941 A (EMBREE ET AL, 06-10-1998) y el documento US 6 697 491 B1 (GRIESINGER DAVID H, 24-02-2004) dan a conocer decodificadores de envolvente en los que se aplica la decodificación paramétrica direccional a los N canales en lugar de a los 2 canales de mezcla descendente.

Es un objeto de la invención proporcionar un método que permita la combinación de codificación paramétrica de audio multicanal con técnicas de matrizado, método con el que pueda realizarse una reconstrucción multicanal de calidad completa, independientemente del decodificador disponible.

Según la invención, este objeto se logra por medio de un método de procesamiento de una señal de mezcla descendente estéreo que comprende señales estéreo primera y segunda, codificando la señal de mezcla descendente estéreo y parámetros espaciales asociados una señal de audio de N canales, comprendiendo el método las etapas de:

- sumar una primera señal y una tercera señal para obtener una primera señal de salida, comprendiendo dicha primera señal dicha primera señal estéreo modificada por una primera función compleja, y comprendiendo dicha tercera señal dicha segunda señal estéreo modificada por una tercera función compleja; y

- sumar una segunda señal y una cuarta señal para obtener una segunda señal de salida, comprendiendo dicha cuarta señal dicha segunda señal estéreo modificada por una cuarta función compleja y comprendiendo dicha segunda señal dicha primera señal estéreo modificada por una segunda función compleja;

- en el que dichas funciones complejas son funciones de dichos parámetros espaciales y se eligen de modo que un valor de energía de la diferencia entre la primera señal y la segunda señal sea mayor que o igual al valor de energía de la suma de la primera y la segunda señal, y de manera que el valor de energía de la diferencia entre la cuarta señal y la tercera señal sea mayor que o igual al valor de energía de la suma de la cuarta señal y la tercera señal. Por consiguiente, se permite la orientación hacia delante/atrás en el decodificador.

El valor de energía de estas señales de diferencia y suma puede basarse en la norma a 2 (es decir, suma de cuadrados sobre varias muestras) o el valor absoluto de estas señales. Además, en este caso pueden aplicarse otras medidas de energía convencionales.

En una realización de la invención, la señal de audio de N canales comprende señales de canal delantero y señales de canal trasero, y dichos parámetros espaciales comprenden una medida de la contribución relativa de los canales traseros en la mezcla descendente estéreo en comparación con la contribución de los canales delanteros en la misma. Esto se debe a que es necesaria la selección de la contribución del canal trasero.

La magnitud de dicha segunda función compleja puede ser menor que la magnitud de dicha primera función compleja para permitir la orientación hacia atrás izquierda/derecha y/o la magnitud de dicha tercera función compleja es menor que la magnitud de dicha cuarta función compleja.

La segunda función compleja y/o la tercera función compleja pueden comprender un desplazamiento de fase, que es sustancialmente igual a más o menos 90 grados con el fin de evitar la cancelación de señal con contribución de canal delantero.

En otra realización de la invención, dicha primera función comprende partes de función primera y segunda, en la que la salida de dicha parte de segunda función aumenta cuando dichos parámetros espaciales indican que una contribución de los canales traseros en dicha primera señal estéreo aumenta en comparación con la contribución de los canales delanteros, y dicha parte de segunda función comprende un desplazamiento de fase que es sustancialmente igual a más o menos 90 grados. Esto es para evitar la cancelación de señal con canales delanteros. Además, dicha cuarta función puede comprender partes de función tercera y cuarta, en la que la salida de dicha parte de cuarta función

aumenta cuando dichos parámetros espaciales indican que la contribución de los canales traseros en dicha segunda señal estéreo aumenta en comparación con la contribución de los canales delanteros, y dicha parte de cuarta función comprende un desplazamiento de fase que es sustancialmente igual a más o menos 90 grados.

La parte de primera función puede tener un signo opuesto en comparación con dicha parte de cuarta función. La segunda función puede tener un signo opuesto en comparación con dicha tercera función. Las partes de segunda función y cuarta función pueden tener el mismo signo, y la parte de tercera función y segunda función puede tener el mismo signo.

En otro aspecto de la invención, se proporciona un dispositivo para procesar una señal estéreo según los métodos mencionados anteriormente, y un aparato codificador que comprende un dispositivo de este tipo.

En otro aspecto de la invención, se proporciona un método para procesar una señal de mezcla descendente estéreo que comprende señales estéreo primera y segunda, comprendiendo el método la etapa de invertir la operación de procesamiento según los métodos mencionados anteriormente.

En otro aspecto de la invención, se proporciona un dispositivo para procesar una señal de mezcla descendente estéreo según el método mencionado anteriormente de procesamiento de una señal de mezcla descendente estéreo, y un aparato decodificador que comprende un dispositivo de este tipo.

En aún otro aspecto de la invención, se proporciona un sistema de audio, que comprende un aparato codificador de este tipo y un aparato decodificador de este tipo.

Los objetos, características y ventajas adicionales de la invención serán evidentes a partir de la siguiente descripción detallada de la invención con referencia a realizaciones de la misma y a los dibujos adjuntos, en los que:

La figura 1 es un diagrama de bloques de un sistema de audio codificador/decodificador que incluye postprocesamiento y postprocesamiento inverso según la invención.

La figura 2 es un diagrama de bloques de una realización de un dispositivo para procesar una señal estéreo según la invención.

La figura 3 es un diagrama de bloques detallado similar a la figura 2, que muestra detalles adicionales de la invención.

La figura 4 es un diagrama de bloques detallado similar a la figura 3, que muestra detalles todavía adicionales de la invención.

La figura 5 es un diagrama de bloques detallado similar a la figura 3, que muestra detalles aún adicionales de la invención.

La figura 6 es un diagrama de bloques de una realización de un dispositivo para procesar una señal de mezcla descendente estéreo según la invención.

El método de la invención puede hacer posible una decodificación de matriz sin distorsionar la reconstrucción multicanal paramétrica. Es posible porque las técnicas de matrizado se aplican en el codificador tras el mezclado descendente, en contradicción con el matrizado usual, que se realiza antes del mezclado descendente. El matrizado de la mezcla descendente se controla por los parámetros espaciales.

Si la matriz aplicada puede invertirse, el decodificador puede anular el matrizado basándose en los parámetros de información de codificador *P* transmitidos.

De manera convencional, el matrizado se aplica en la señal original de entrada de N canales. Sin embargo, este enfoque no es adecuado en este caso, puesto que la inversión de este matrizado, que es un prerrequisito para la correcta reconstrucción de N canales, es generalmente imposible, ya que sólo 2 canales están disponibles en el decodificador. Por tanto, una característica de esta invención es sustituir la técnica de matrizado, que se aplica normalmente en la mezcla de 5 canales, por una modificación controlada por parámetros de la mezcla de dos canales.

La figura 1 es un diagrama de bloques de un sistema de audio codificador/decodificador que incorpora la invención. En el sistema 1 de audio, se suministra una señal de audio de N canales a un codificador 2. El codificador 2 transforma la señal de audio de N canales en señales de canal estéreo L_0 y R_0 y parámetros de información de codificador *P*, por medio de lo que un decodificador 3 puede decodificar la información y aproximadamente reconstruir la señal original de N canales que va a emitirse desde el decodificador 3. Las señales de N canales pueden ser señales para un sistema 5.1, que comprende un canal central, dos canales delanteros, dos canales envolventes y un canal de efectos de baja frecuencia (LFE).

De manera convencional, las señales de canal estéreo codificadas L_0 y R_0 y los parámetros de información de codificador *P* se transmiten o distribuyen al usuario de una manera adecuada, tal como por CD, DVD, difusión, disco láser, DBS, cable digital, Internet o cualquier otro sistema de transmisión o distribución, indicado por el círculo 4 en la figura 1. Puesto que las señales estéreo izquierda y derecha L_0 y R_0 se transmiten o distribuyen, el sistema 1 es

compatible con el enorme número de equipos de recepción que sólo pueden reproducir señales estéreo. Si el equipo de recepción incluye un decodificador multicanal paramétrico, el decodificador puede decodificar las señales de N canales proporcionando una estimación de las mismas basándose en la información en los canales estéreo L_0 y R_0 así como los parámetros de información de codificador *P*.

Ahora supóngase una señal de audio de N canales, siendo N un número entero que es mayor que 2, y donde $z_1[n]$, $z_2[n]$,....., $z_N[n]$ describen las formas de onda discretas en el dominio de tiempo de los N canales. Estas N señales se segmentan usando una segmentación común, preferiblemente usando ventanas de análisis de solapamiento. Posteriormente, cada segmento se convierte al dominio de frecuencia, usando una transformación compleja (por ejemplo, FFT). Sin embargo, las estructuras complejas del banco de filtros también pueden ser apropiadas para obtener mosaicos de tiempo/frecuencia. Este proceso da como resultado representaciones de subbanda, segmentadas de las señales de entrada, lo que se indicará por $Z_1[k]$, $Z_2[k]$,...., $Z_N[k]$ indicando *k* el índice de frecuencia.

A partir de estos N canales, se crean 2 canales de mezcla descendente, concretamente $L_0[k]$ y $R_0[k]$. Cada canal de mezcla descendente es una combinación lineal de las N señales de entrada:

$$L_0[k] = \sum_{i=1}^{N} \alpha_i Z_i[k]$$

$$R_0[k] = \sum_{i=1}^{N} \beta_i Z_i[k]$$

Los parámetros α_i y β_i se eligen para que la señal estéreo constituida por $L_o[k]$ y $R_o[k]$ tenga una buena imagen estéreo.

En la señal estéreo resultante, un postprocesador 5 puede aplicar un procesamiento de tal manera que afecte principalmente a la contribución de un canal específico i en la mezcla estéreo. Como procesamiento puede elegirse una técnica de matrizado específica. Esto da como resultado las señales de matriz compatible izquierda y derecha $L_{Ow}[k]$ y $R_{Ow}[k]$. Éstas, junto con los parámetros espaciales se transmiten al decodificador tal como se ilustra por el círculo 6 en la figura 1. El dispositivo para procesar una señal estéreo obtenida de un codificador comprende el postprocesador 5. El aparato codificador según la invención comprende el codificador 2 y el postprocesador 5.

Las señales postprocesadas L_{0w} y R_{0w} pueden suministrarse a un receptor estéreo convencional (no mostrado) para su reproducción. Alternativamente, las señales postprocesadas L_{0w} y R_{0w} pueden suministrarse a un decodificador matricial (no mostrado), por ejemplo un decodificador Dolby Pro Logic® o un decodificador Circle Surround®. Aún otra posibilidad es suministrar las señales postprocesadas L_{0w} y R_{0w} a un postprocesador 7 inverso para anular el procesamiento del postprocesador 5. Las señales resultantes L_0 y R_0 pueden suministrarse por el postprocesador 7 a un decodificador 3 multicanal. El dispositivo para procesar una señal de mezcla descendente estéreo comprende el postprocesador 7 inverso. El aparato decodificador según la invención comprende el decodificador 3 y el postprocesador 7 inverso.

En el decodificador 3, los N canales de entrada se reconstruyen de la siguiente manera:

$$\hat{Z}_i[k] = C_{1,Z_i} L_O[k] + C_{2,Z_i} R_O[k],$$

donde $\hat{Z}_i[k]$ es una estimación de $Z_i[k]$. Los filtros $C_{1,Zi}$ y $C_{2,Zi}$ son preferiblemente dependientes del tiempo y la frecuencia, y sus funciones de transferencia se derivan a partir de los parámetros de información de codificador *P* transmitidos.

La figura 2 muestra cómo puede realizarse este bloque 5 de postprocesamiento para hacer posible la decodificación de matriz. La señal de entrada izquierda $L_o[k]$ se modifica por una primera función compleja g_1, lo que da como resultado una primera señal $L_{OwL}[k]$ que se alimenta a la salida izquierda $L_{Ow}[k]$. La señal de entrada izquierda $L_o[k]$ también se modifica por una segunda función compleja g_2, que da como resultado una segunda señal $R_{OwL}[k]$ que se alimenta a la salida derecha $R_{Ow}[k]$. Las funciones g_1 y g_2 se eligen para que la señal diferencia L_{OwL} - R_{OwL} tenga una energía igual a o mayor que la señal suma L_{OwL} + R_{OwL}. Esto se debe a que, en la decodificación de matriz, la relación de la señal suma y diferencia se usa para realizar la orientación hacia delante/atrás. Cuando la señal diferencia se hace más grande, se orienta más la señal de entrada hacia atrás. Debido a esto $R_{OwL}[k]$ debe aumentar cuando la contribución de la parte trasera izquierda en $L_o[k]$ aumenta. Este procedimiento de control se realiza mediante las funciones g_1 y g_2, siendo ambas funciones de los parámetros espaciales *P*. Estas funciones se eligen de manera que la cantidad de procesamiento del canal de entrada izquierdo aumenta cuando la contribución de la parte trasera izquierda en $L_o[k]$ aumenta.

La magnitud de g_2 es preferiblemente menor que la magnitud de g_1. Esto permite la orientación hacia atrás izquierda/derecha en el decodificador.

La señal de entrada derecha $R_o[k]$ se modifica por una cuarta función g_4, que da como resultado una cuarta señal $R_{OwR}[k]$, que se alimenta a la salida derecha R_{Ow}[k]. La señal de entrada derecha R_o[k] también se modifica por una tercera función g_3, que da como resultado una tercera señal L_{OwR}[k], que se alimenta a la salida izquierda L_{Ow}[k]. Las funciones g_3 y g_4 se eligen de manera que la cantidad de procesamiento del canal de entrada derecho aumenta cuando la contribución de la parte trasera derecha en R_o[k] aumenta, y también de manera que restar L_{0wR} de R_{0wR} da como resultado una señal mayor que al sumarlas.

La magnitud de g_3 es preferiblemente menor que la magnitud de g_4. Esto permite la orientación hacia atrás izquierda/derecha en el decodificador.

La salida puede describirse por medio de la siguiente ecuación matricial:

$$\begin{bmatrix} L_{ow} \\ R_{ow} \end{bmatrix} = H \begin{bmatrix} L_0 \\ R_0 \end{bmatrix} = \begin{bmatrix} g_1 & g_3 \\ g_2 & g_4 \end{bmatrix} \begin{bmatrix} L_0 \\ R_0 \end{bmatrix}$$

A continuación se describe un codificador multicanal paramétrico. Se aplican las siguientes ecuaciones:

$$L_0[k] = L[k] + C_s[k]$$

$$R_0[k] = R[k] + C_s[k]$$

en las que C_s[k] es la señal monofónica que resulta tras combinar el canal LFE y el canal central. Las siguientes ecuaciones son válidas para L[k] y R[k]:

$$L[k] = (c_1 \;\; c_2) \begin{pmatrix} L_f[k] \\ L_s[k] \end{pmatrix}$$

$$R[k] = (c_3 \;\; c_4) \begin{pmatrix} R_f[k] \\ R_s[k] \end{pmatrix}$$

donde L_f es el canal delantero izquierdo, L_s el envolvente izquierdo, R_f el delantero derecho y R_s el envolvente derecho. Las constantes c_1 a c_4 controlan el proceso de mezcla descendente y pueden tener un valor complejo y/o ser dependientes del tiempo y la frecuencia. Una mezcla descendente de estilo ITU se obtiene para (c_1, c_3 = sqrt(2); c_2, c_4=1).

En el decodificador, se realiza la siguiente reconstrucción:

$$\hat{L}[k] = \beta L_0[k] + (\gamma - 1) R_0[k]$$

$$\hat{R}[k] = (\beta - 1) L_0[k] + \gamma R_0[k]$$

$$\hat{C}[k] = (1 - \beta) L_0[k] + (1 - \gamma) R_0[k]$$

donde $\hat{L}$[k] es una estimación de $L[k]$, $\hat{R}[k]$ una estimación de $R[k]$ y $\hat{C}[k]$ una estimación de $C_s[k]$. Los parámetros β y γ se determinan en el codificador y se transmiten al decodificador, es decir, son un subconjunto de los parámetros de información de codificador P. Adicionalmente, la señal de información P puede incluir niveles de señal (relativos) entre canales delanteros y envolventes correspondientes, es decir, una diferencia de intensidad entre canales (IID) entre L_f, L_s, y R_f, R_s, respectivamente. Una expresión conveniente para la IID_1, que describe la relación de energía entre L_f y L_s

viene dada por

$$IID_L = \frac{\sum_k L_f[k]L_f^*[k]}{\sum_k L_s[k]L_s^*[k]}$$

Cuando se usan estos parámetros, el esquema en la figura 2 puede reemplazarse por el esquema en la figura 3. Para el procesamiento del canal izquierdo $L_o[k]$, sólo son necesarios los parámetros que determinan la contribución delantera/trasera en el canal de entrada izquierdo, que son los parámetros IID_L y β. Para el procesamiento del canal de entrada derecho, sólo son necesarios los parámetros IID_R y γ. La función g_2 puede reemplazarse ahora por la función g_3, pero con un signo opuesto.

En la figura 4, las funciones g_1 y g_4 se dividen en dos partes de función paralelas. La función g_1 se divide en g_{11} y g_{12}. La función g_4 se divide en g_{11} y $-g_{12}$. Las señales de salida de la parte de función g_{12} y la función g_3 son las contribuciones de los canales traseros. La parte de función g_{12} y la función g_3 necesitan sumarse con el mismo signo en una salida de modo que eviten la cancelación de señal y con signo opuesto en las diferentes salidas.

La parte de función g_{12} y la función g_3 contienen un desplazamiento de fase de más o menos 90 grados. Esto es para evitar la cancelación de la contribución de canal delantero (salida de la parte de función g_{11}).

La figura 5 da una descripción más detallada de este bloque. El parámetro w_l determina la cantidad de procesamiento de $L_o[k]$ y w_r de $R_o[k]$. Cuando w_l es igual a 0, $L_o[k]$ no se procesa, y cuando w_l es igual a 1, $L_o[k]$ se procesa al máximo. Lo mismo se aplica para w_r con respecto a R_o[k].

Las siguientes ecuaciones generalizadas se aplican para los parámetros de postprocesamiento w_l y w_r:

$$w_l = f_1(P)$$

$$w_r = f_r(p)$$

Los bloques Φ^{-90} son filtros pasa todo que realiza un desplazamiento de fase de 90 grados. Los bloques G_1 y G_2 en la figura 5 son ganancias. Las salidas resultantes son:

$$\begin{bmatrix} L_{0w} \\ R_{0w} \end{bmatrix} = H \begin{bmatrix} L_0 \\ R_0 \end{bmatrix}, \text{ con:}, \qquad H = \begin{bmatrix} 1 - w_l + w_l\Phi^{-90} & w_r\Phi^{-90}G_2 \\ -w_l\Phi^{-90}G_1 & 1 - w_r - w_r\Phi^{-90} \end{bmatrix}$$

donde:

$$G_1 = f_1(w_l, w_r)$$

$$G_2 = f_2(w_l, w_r)$$

Así las funciones g_1.....g_4 se reemplazan por funciones más específicas:

$$g_1 = 1 - w_l + w_l\Phi^{-90}$$

$$g_2 = -w_l\Phi^{-90}G_1$$

$$g_3 = w_r\Phi^{-90}G_2$$

$$g_4 = 1 - w_r = w_r\Phi^{-90}$$

La inversa de la matriz H viene dada por (si det(H)≠0):

$$H^{-1} = \frac{1}{1 - w_l - w_r + w_l w_r + (w_l - w_r)\Phi^{-90} + (G_1 G_2 - 1) w_l w_r \Phi^{-180}} \begin{bmatrix} 1 - w_r - w_r \Phi^{-90} & -w_r \Phi^{-90} G_2 \\ w_l \Phi^{-90} G_1 & 1 - w_l + w_l \Phi^{-90} \end{bmatrix}$$

Por tanto, la utilización de funciones adecuadas en la matriz H permite invertir el proceso de matrizado.

La inversión puede realizarse en el decodificador sin la necesidad de transmitir información adicional, ya que los parámetros w_l y w_r pueden calcularse a partir de los parámetros transmitidos. Por tanto, la señal original estéreo estará disponible nuevamente lo que es necesario para la decodificación paramétrica de la mezcla multicanal.

Pueden lograrse incluso mejores resultados si las ganancias G_1 y G_2 son una función de la diferencia de intensidad entre canales (IID) entre los canales envolventes. En ese caso, esta IID también debe transmitirse al decodificador.

Dada la descripción de parámetros mencionada anteriormente, se usan las siguientes funciones para la operación de postprocesamiento:

$$w_l = f_1(\alpha_l) f_2(\beta)$$

$$w_r = f_3(\alpha_r) f_4(\gamma)$$

En este caso f_1.....f_4 pueden ser funciones arbitrarias. Por ejemplo:

$$f_1(IID) = f_3(IID) = \frac{IID}{1 + IDD}$$

$$f_2(\beta) = f_4(\beta) = \begin{Bmatrix} 2\beta - 1 \text{ si } 0{,}5 < \beta < 1 \\ 1 \text{ si } \beta \geq 1 \\ 0 \text{ si } \beta \leq 0{,}5 \end{Bmatrix}$$

El filtro pasa todo Φ^{-90} puede realizarse de manera eficaz realizando una multiplicación en el dominio de frecuencia (de valor complejo) con el operador complejo j(j^2 = -1). Para las ganancias G_1 y G_2 puede tomarse una función de w_l, w_r tal como se realiza en Circle Surround, pero también es adecuada una constante con el valor $1/\sqrt{2}$. Esto da como resultado la matriz:

$$H = \begin{pmatrix} 1 - w_l + w_l j & \frac{1}{2}\sqrt{2 w_r j} \\ -\frac{1}{2}\sqrt{2 w_l j} & 1 - w_r - w_r j \end{pmatrix}$$

La determinante de esta matriz es igual a:

$$\det(H) = \left(1 - w_l - w_r + \frac{3}{2} w_l w_r\right) + j(w_l - w_r)$$

La parte imaginaria de esta determinante sólo será igual a cero cuando $w_l = w_r$. En ese caso, se aplica lo siguiente para la determinante:

$$\det(H) = 1 - 2 w_l + \frac{3}{2} w_l^2$$

Esta función tiene un mínimo de $\det(H) = \frac{1}{3}$ para $w_l = \frac{2}{3}$.

Por consiguiente, también para $w_l = w_r$ esta matriz puede invertirse. Por tanto, para las ganancias $G_1 = G_2 = 1/\sqrt{2}$ la matriz H siempre se puede invertir, independientemente de los valores w_l y w_r.

La figura 6 es un diagrama de bloques de una realización del postprocesador 7 inverso. Como el postprocesamiento, la inversión se realiza mediante una multiplicación matricial para cada banda de frecuencia:

$$\begin{bmatrix} L_0 \\ R_0 \end{bmatrix} = H^{-1} \begin{bmatrix} L_{0w} \\ R_{0w} \end{bmatrix} = \begin{bmatrix} k_1 & k_3 \\ k_2 & k_4 \end{bmatrix} \begin{bmatrix} L_{0w} \\ R_{0w} \end{bmatrix} \qquad \text{con} \qquad \begin{aligned} k_1 &= \frac{1}{g_1 g_4 - g_2 g_3} g_4 \\ k_2 &= \frac{-1}{g_1 g_4 - g_2 g_3} g_2 \\ k_3 &= \frac{-1}{g_1 g_4 - g_2 g_3} g_3 \\ k_4 &= \frac{1}{g_1 g_4 - g_2 g_3} g_1 \end{aligned}$$

Por consiguiente, cuando las funciones g_1.....g_4 pueden determinarse en el decodificador, pueden determinarse las funciones k_1..... k_4. Las funciones k_1..... k_4 son funciones del conjunto de parámetros P, como las funciones g_1.....g_4. Para la inversión, por tanto, necesitan conocerse las funciones g_1.....g_4 y el conjunto de parámetros P.

La matriz H puede invertirse cuando la determinante de la matriz H es diferente de cero, es decir:

$$\det(H) = g_1 g_4 - g_2 g_3 \neq 0$$

Esto puede lograrse mediante una elección apropiada de las funciones g_1.....g_4.

Otra aplicación de la invención es realizar la operación de postprocesamiento en la señal estéreo en el lado de decodificador sólo (es decir, sin postprocesamiento en el lado de codificador). Usando este enfoque, el decodificador puede generar una señal estéreo mejorada a partir de una señal estéreo no mejorada. Esta operación de postprocesamiento en el lado de decodificador sólo puede realizarse adicionalmente en una situación en la que, en el codificador, la señal de entrada multicanal se decodifica en una única señal (monofónica) y parámetros espaciales asociados. En el decodificador, la señal monofónica puede convertirse en primer lugar en una señal estéreo (usando los parámetros espaciales) y a continuación esta señal estéreo puede postprocesarse tal como se describió anteriormente. Alternativamente, la señal monofónica puede decodificarse directamente mediante un decodificador multicanal.

Debe observarse que el uso del verbo “comprender” y sus conjugaciones no excluye otros elementos o etapas y que el uso del artículo indefinido “un” o “una” no excluye una pluralidad de elementos o etapas. Además, los signos de referencia en las reivindicaciones no deben interpretarse como limitativos del alcance de las reivindicaciones.

La invención se ha descrito con referencia a realizaciones específicas. Sin embargo, la invención no se limita a las diversas realizaciones descritas sino que pueden modificarse y combinarse de diferentes maneras como es evidente para un experto que lea la presente memoria descriptiva.

REIVINDICACIONES

1. Método de procesamiento de una señal de mezcla descendente estéreo que comprende señales estéreo primera y segunda (L_0, R_0), codificando la señal de mezcla descendente estéreo y los parámetros espaciales asociados (*P*) una señal de audio de N canales, comprendiendo el método las etapas de:

- sumar una primera señal y una tercera señal para obtener una primera señal de salida (L_{0w}), comprendiendo dicha primera señal (L_{0wL}) dicha primera señal estéreo (L_0) modificada por una primera función compleja (g_1), y comprendiendo dicha tercera señal (L_{0wR}) dicha segunda señal estéreo (R_0) modificada por una tercera función compleja (g_3); y

- sumar una segunda señal y una cuarta señal para obtener una segunda señal de salida (R_{0w}), comprendiendo dicha cuarta señal (R_{0wR}) dicha segunda señal estéreo (R_0) modificada por una cuarta función compleja (g_4) y comprendiendo dicha segunda señal (R_{0wL}) dicha primera señal estéreo (L_0) modificada por una segunda función compleja (g_2);

- en el que dichas funciones complejas (g_1, g_3, g_3, g_4) son funciones de dichos parámetros espaciales (*P*) y se eligen de modo que un valor de energía de la diferencia (L_{0wL}-R_{0wL}) entre la primera señal y la segunda señal sea mayor que o igual al valor de energía de la suma (L_{0wL}+R_{0wL}) de la primera y la segunda señal, y de manera que el valor de energía de la diferencia (R_{0wR}-L_{0wR}) entre la cuarta señal y la tercera señal sea mayor que o igual al valor de energía de la suma (R_{0wR}+L_{0wR}) de la cuarta señal y la tercera señal.

2. Método según la reivindicación 1, en el que la señal de audio de N canales comprende señales de canal delantero y señales de canal trasero, y en el que dichos parámetros espaciales (*P*) comprenden una medida de la contribución relativa de los canales traseros en la mezcla descendente estéreo (L_0, R_0) en comparación con la contribución de los canales delanteros en la misma.

3. Método según la reivindicación 1 ó 2, en el que la magnitud de dicha segunda función compleja (g_2) es menor que la magnitud de dicha primera función compleja (g_1) y/o la magnitud de dicha tercera función compleja (g_3) es menor que la magnitud de dicha cuarta función compleja (g_4).

4. Método según la reivindicación 1, 2 ó 3, en el que dicha segunda función compleja (g_2) y/o dicha tercera función compleja (g_3) comprende un desplazamiento de fase que es sustancialmente igual a más o menos 90 grados.

5. Método según una cualquiera de las reivindicaciones anteriores, en el que dicha primera función (g_1) comprende partes de función primera y segunda (g_{11L}; g_{12L}), en el que la salida de dicha parte de segunda función (g_{12L}) aumenta cuando dichos parámetros espaciales (*P*) indican que una contribución de los canales traseros en dicha primera señal estéreo (L_0) aumenta en comparación con la contribución de los canales delanteros en dicha primera señal estéreo (L_0), y dicha parte de segunda función (g_{12L}) comprende un desplazamiento de fase que es sustancialmente igual a más o menos 90 grados.

6. Método según la reivindicación 5, en el que dicha cuarta función (g_4) comprende partes de función tercera y cuarta (g_{11R}; g_{12R}), en el que la salida de dicha parte de cuarta función (g_{12R}) aumenta cuando dichos parámetros espaciales (*P*) indican que la contribución de los canales traseros en dicha segunda señal estéreo (R_0) aumenta en comparación con la contribución de los canales delanteros en dicha segunda señal estéreo (R_0), y dicha parte de cuarta función (g_{12R}) comprende un desplazamiento de fase que es sustancialmente igual a más o menos 90 grados.

7. Método según la reivindicación 6, en el que dicha parte de primera función (g_{12L}) tiene un signo opuesto en comparación con dicha parte de cuarta función (g_{12R}).

8. Método según la reivindicación 6, en el que dicha segunda función (g_2) tiene un signo opuesto en comparación con dicha tercera función (g_3).

9. Método según la reivindicación 7 u 8, en el que dicha segunda función (g_2) y dicha parte de cuarta función (g_{12R}) tienen el mismo signo, y en el que dicha tercera función (g_3) y dicha parte de segunda función (g_{12L}) tienen el mismo signo.

10. Dispositivo (5) para procesar una señal de mezcla descendente estéreo que comprende señales estéreo primera y segunda (L_0, R_0), codificando la señal de mezcla descendente estéreo y parámetros espaciales asociados (*P*) una señal de audio de N canales, comprendiendo el dispositivo:

- primeros medios de suma para sumar una primera señal y una tercera señal para obtener una primera señal de salida (L_{0w}), comprendiendo dicha primera señal (L_{0wL}) dicha primera señal estéreo (L_0) modificada por una primera función compleja (g_1), y comprendiendo dicha tercera señal (L_{0wR}) dicha segunda señal estéreo (R_0) modificada por una tercera función compleja (g_3); y

- segundos medios de suma para sumar una segunda señal y una cuarta señal para obtener una segunda señal de salida (R_{0w}), comprendiendo dicha cuarta señal (R_{0wR}) dicha segunda señal estéreo (R_0) modificada por una cuarta función compleja (g_4), y comprendiendo dicha segunda señal (R_{0wL}) dicha primera señal estéreo (L_0) modificada por una segunda función compleja (g_2);

en el que dichas funciones complejas son funciones de dichos parámetros espaciales (*P*), de manera que un valor de energía de la diferencia (L_{0wL}-R_{0wL}) entre la primera señal y la segunda señal es mayor que o igual al valor de energía de la suma (L_{0wL}+R_{0wL}) de la primera y la segunda señal y de manera que el valor de energía de la diferencia (R_{0wR}-L_{0wR}) entre la cuarta señal y la tercera señal es mayor que o igual al valor de energía de la suma (R_{0wR}+L_{0wR}) de la cuarta señal y la tercera señal.

11. Aparato codificador que comprende:

- un codificador (2) para codificar una señal de audio de N canales en parámetros espaciales (*P*) y una señal de mezcla descendente estéreo que comprende señales estéreo primera y segunda (L_0,R_0), y

- un dispositivo (5) según la reivindicación 10 para procesar la señal de mezcla descendente estéreo.

12. Método de procesamiento de una señal de mezcla descendente estéreo que comprende señales estéreo primera y segunda (L_{0w}, R_{0w}), comprendiendo el método invertir la operación de procesamiento según el método según una cualquiera de las reivindicaciones 1 a 9.

13. Método según la reivindicación 12, en el que la inversión comprende una multiplicación matricial

$$\begin{bmatrix} L_0 \\ R_0 \end{bmatrix} = \begin{bmatrix} k_1 & k_3 \\ k_2 & k_4 \end{bmatrix} \begin{bmatrix} L_{0w} \\ R_{0w} \end{bmatrix} \quad \text{con} \quad \begin{aligned} k_1 &= \frac{1}{g_1 g_4 - g_2 g_3} g_4 \\ k_2 &= \frac{-1}{g_1 g_4 - g_2 g_3} g_2 \\ k_3 &= \frac{-1}{g_1 g_4 - g_2 g_3} g_3 \\ k_4 &= \frac{1}{g_1 g_4 - g_2 g_3} g_1 \end{aligned},$$

donde L_0 y R_0 son señales de salida primera y segunda respectivas, y donde L_{0w} y R_{0w} son señales de entrada estéreo primera y segunda respectivas, y donde g_1; g_2, g_3 y g_4 son dichas funciones complejas primera, segunda, tercera y cuarta respectivas.

14. Dispositivo (7) para procesar una señal de mezcla descendente estéreo que comprende señales estéreo primera y segunda (L_{0w}, R_{0w}), comprendiendo el dispositivo medios para invertir la operación de procesamiento según el método según una cualquiera de las reivindicaciones 1 a 9.

15. Dispositivo (7) según la reivindicación 14, en el que los medios para invertir comprenden una multiplicación matricial

$$\begin{bmatrix} L_0 \\ R_0 \end{bmatrix} = \begin{bmatrix} k_1 & k_3 \\ k_2 & k_4 \end{bmatrix} \begin{bmatrix} L_{0w} \\ R_{0w} \end{bmatrix} \quad \text{con} \quad \begin{aligned} k_1 &= \frac{1}{g_1 g_4 - g_2 g_3} g_4 \\ k_2 &= \frac{-1}{g_1 g_4 - g_2 g_3} g_2 \\ k_3 &= \frac{-1}{g_1 g_4 - g_2 g_3} g_3 \\ k_4 &= \frac{1}{g_1 g_4 - g_2 g_3} g_1 \end{aligned},$$

donde L_0 y R_0 son señales de salida primera y segunda respectivas, y donde L_{0w} y R_{0w} son señales de entrada estéreo primera y segunda respectivas, y donde g_1, g_2, g_3 y g_4 son dichas funciones complejas primera, segunda, tercera y cuarta respectivas.

16. Aparato decodificador que comprende:

- un dispositivo (7) según la reivindicación 14 ó 15 para procesar una señal de mezcla descendente estéreo que

comprende señales estéreo primera y segunda (L_{0w}, R_{0w}), y

- un decodificador para decodificar las señales estéreo (L_0, R_0) procesadas en una señal de audio de N canales.

17. Sistema de audio que comprende un aparato codificador según la reivindicación 11 y un aparato decodificador según la reivindicación 16.

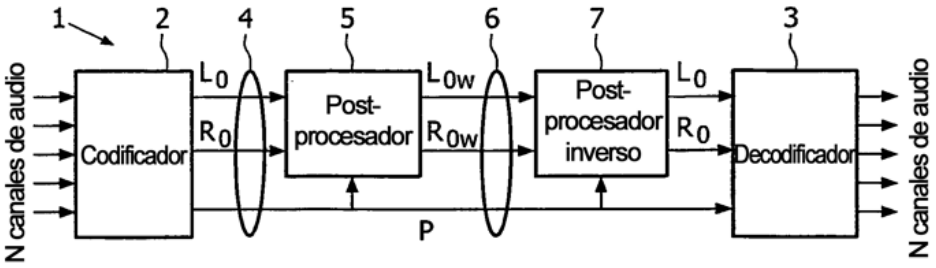
1
2
4
5
6
7
3
N canales de audio
Codificador
L0
R0
Post-procesador
L0w
R0w
Post-procesador inverso
L0
R0
Decodificador
P
N canales de audio

FIG. 1

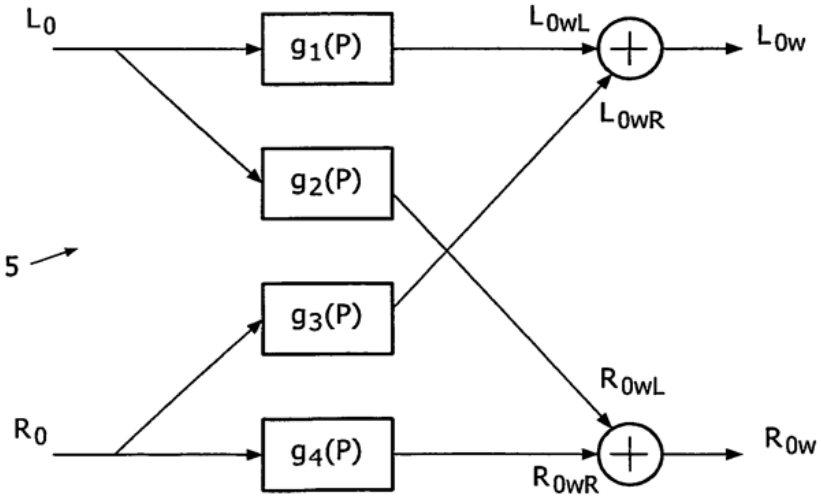
L0
g1(P)
L0wL
L0w
L0wR
g2(P)
5
g3(P)
R0wL
R0
g4(P)
R0w
R0wR

FIG. 2

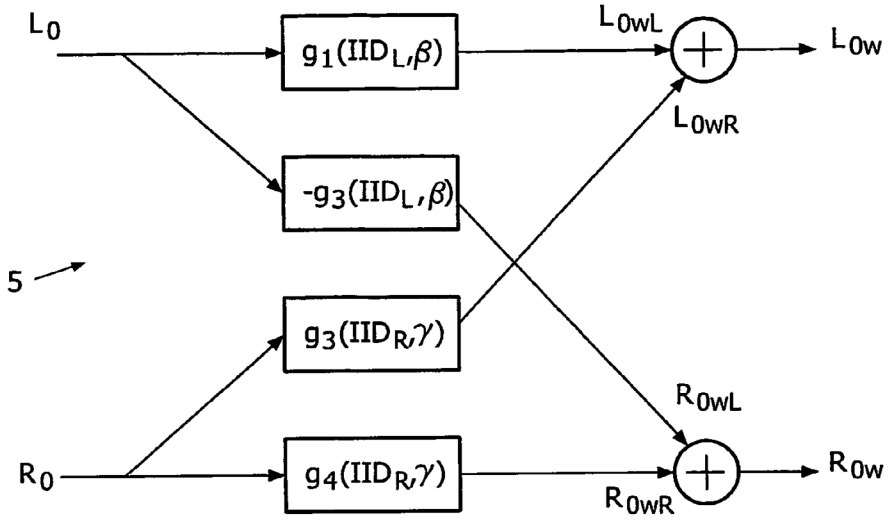
L0
g1(IIDL,β)
LOwL
LOw
LOwR
-g3(IIDL,β)
5
g3(IIDR,γ)
ROwL
R0
g4(IIDR,γ)
ROw
ROwR

FIG. 3

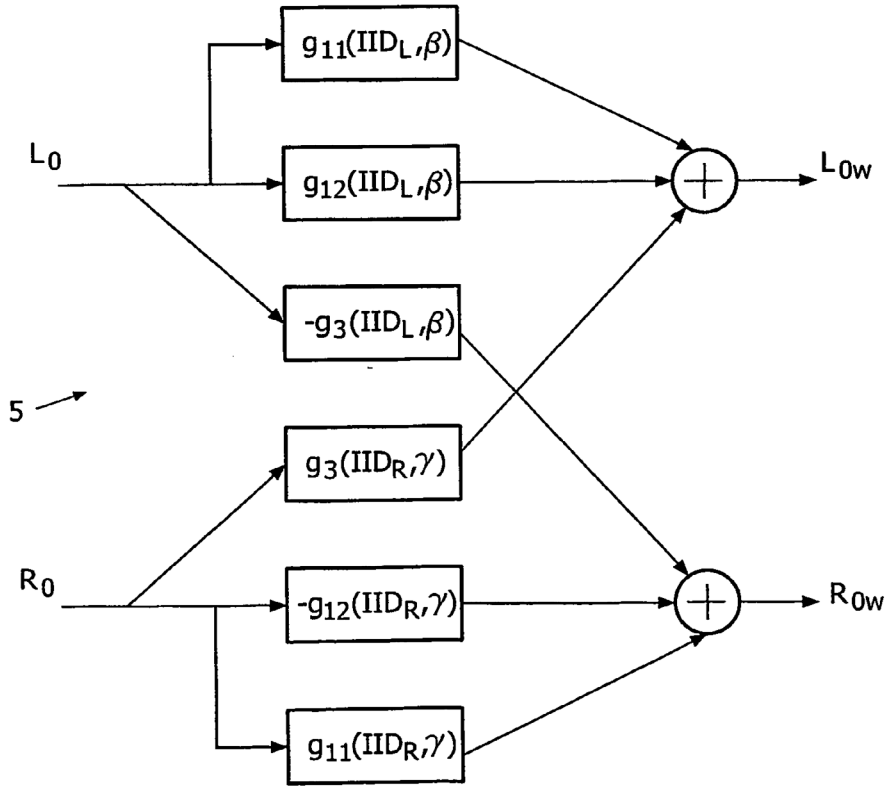
g11(IIDL,β)
L0
g12(IIDL,β)
LOw
-g3(IIDL,β)
5
g3(IIDR,γ)
R0
-g12(IIDR,γ)
ROw
g11(IIDR,γ)

FIG. 4

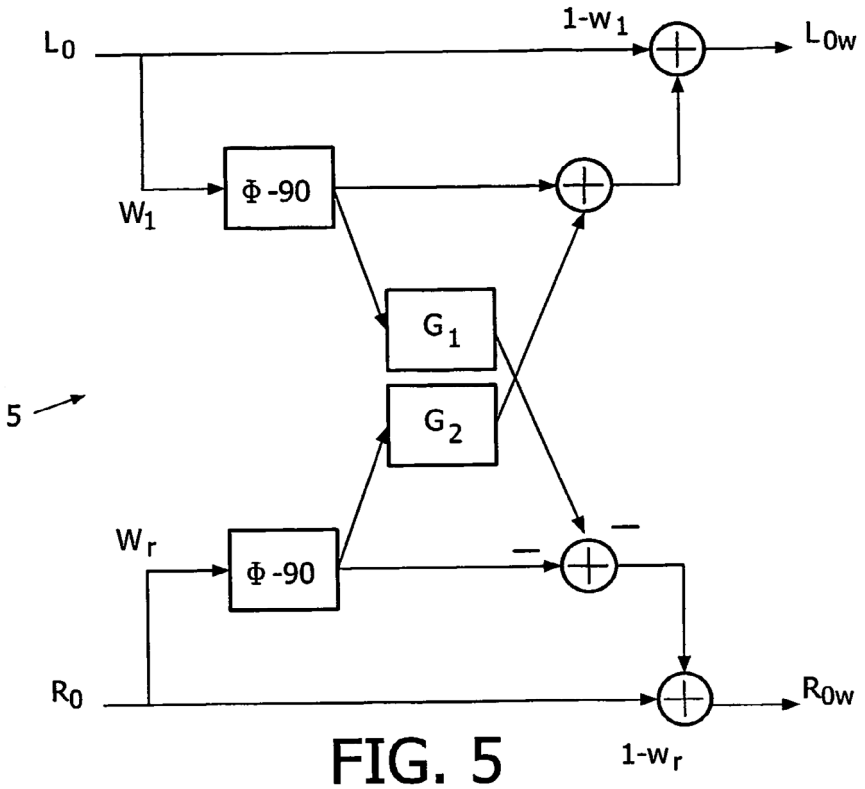
L0
1-w1
L0w
Φ -90
W1
G1
G2
5
Wr
Φ -90
R0
R0w
1-wr

FIG. 5

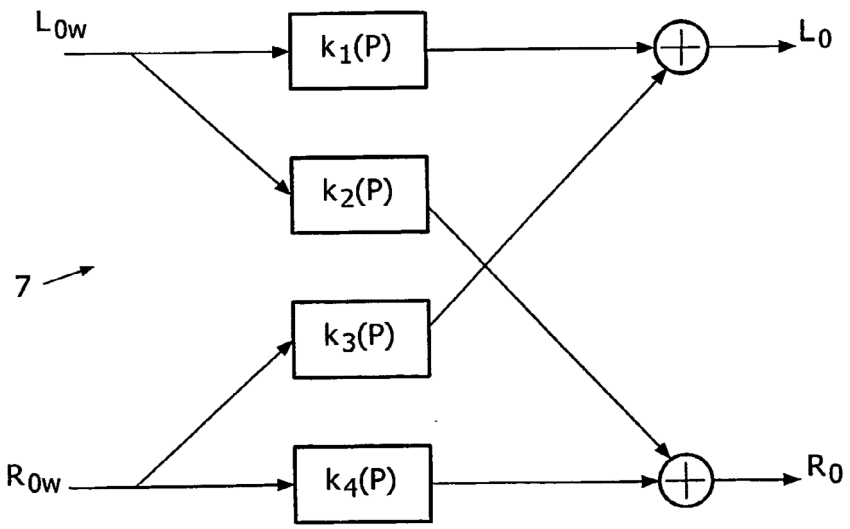
L0w
k1(P)
L0
k2(P)
7
k3(P)
R0w
k4(P)
R0

FIG. 6