

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 499 640**

51 Int. Cl.:

**G10L 19/008** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **14.08.2012 E 12745880 (0)**

97 Fecha y número de publicación de la concesión europea: **23.07.2014 EP 2617031**

54 Título: **Matrices óptimas de mezcla y uso de descorreladores en el procesamiento de audio espacial**

30 Prioridad:

**17.08.2011 US 201161524647 P**  
**21.02.2012 EP 12156351**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**29.09.2014**

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR  
FÖRDERUNG DER ANGEWANDTEN  
FORSCHUNG E.V. (100.0%)**  
**Hansastraße 27c**  
**80686 München, DE**

72 Inventor/es:

**VILKAMO, JUHA;**  
**BÄCKSTRÖM, TOM;**  
**KÜCH, FABIAN y**  
**KUNTZ, ACHIM**

74 Agente/Representante:

**PONTI SALES, Adelaida**

**ES 2 499 640 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

Matrices óptimas de mezcla y uso de descorreladores en el procesamiento de audio espacial

- 5 **[0001]** La presente invención se refiere al procesamiento de señales de audio y, en particular, a un aparato y un procedimiento que emplea matrices óptimas de mezcla y, además, al uso de descorreladores en el procesamiento de audio espacial.
- 10 **[0002]** El procesamiento de audio es cada vez más importante. En el procesamiento perceptual de audio espacial, un supuesto típico es que el aspecto espacial de un sonido reproducido por un parlante es determinado especialmente por las potencias y las dependencias alineadas en tiempo entre los canales de audio en bandas de frecuencia perceptuales. Esto se basa en la noción de que estas características, cuando son reproducidas por parlantes, se transfieren a diferencias de nivel interaural, diferencias de tiempo interaural y coherencias interaurales, las cuales son las señales binaurales de percepción espacial. A partir de este concepto han surgido diversos procedimientos de procesamiento espacial, incluyendo la mezcla ascendente, véase
- 15 [1] C. Faller, "Multiple-Loudspeaker Playback of Stereo Signals", Journal of the Audio Engineering Society, Vol. 54, Nº 11, pp. 1051-1064, junio de 2006, microfonía espacial, véase, por ejemplo,
- 20 [2] V. Pulkki, "Spatial Sound Reproduction with Directional Audio Coding", Journal of the Audio Engineering Society, Vol. 55, Nº 6, pp. 503-516, junio de 2007, y
- [3] C. Tournery, C. Faller, F. Küch, J. Herre, "Converting Stereo Microphone Signals Directly to MPEG Surround", 128th AES Convention, mayo de 2010; y transmisión estéreo y multicanal eficiente, véase, por ejemplo,
- 25 [4] J. Breebaart, S. van de Par, A. Kohlrausch and E. Schuijers, "Parametric Coding of Stereo Audio", EURASIP Journal on Applied Signal Processing, Vol. 2005, Nº 9, pp. 1305-1322, 2005; y
- 30 [5] J. Herre, K. Kjörling, J. Breebaart, C. Faller, S. Disch, H. Purnhagen, J. Koppens, J. Hilpert, J. Rödén, W. Oomen, K. Linzmeier and K. S. Chong, "MPEG Surround – The ISO/MPEG Standard for Efficient and Compatible Multichannel Audio Coding", Journal of the Audio Engineering Society, Vol. 56, Nº 11, pp. 932-955, noviembre de 2008.
- Las pruebas de escucha han confirmado el beneficio del concepto en cada aplicación, véase, por ejemplo, [1, 4, 5] y,
- 35 [6] J. Vilkamo, V. Pulkki, "Directional Audio Coding: Virtual Microphone-Based Synthesis and Subjective Evaluation", Journal of the Audio Engineering Society, Vol. 57, Nº 9, pp. 709-724, septiembre de 2009.
- 40 **[0003]** Todas estas tecnologías, si bien tienen distinta aplicación, tienen la misma tarea central, que consiste en generar a partir de un conjunto de canales de entrada un conjunto de canales de salida con potencias y dependencias definidas en función del tiempo y la frecuencia, que se puede suponer que es la tarea común subyacente en el procesamiento de audio perceptual y espacial. Por ejemplo, en el contexto de la Codificación de Audio Direccional (DirAC) véase, por ejemplo, [2], los canales de origen son típicamente señales de micrófono de primer orden, que se procesan por medio de mezcla, panoramización de amplitud y descorrelación, para aproximarse perceptualmente a un campo de sonido medido. En la mezcla ascendente (véase [1]), los canales de entrada estéreo, nuevamente en función del tiempo y la frecuencia, son distribuidos adaptativamente a una configuración de sonido envolvente.
- 45 **[0004]** Un objetivo de la presente invención consiste en proporcionar conceptos mejorados para generar a partir de un conjunto de canales de entrada un conjunto de canales de salida con propiedades definidas. El objetivo de la presente invención se resuelve por medio de un aparato de acuerdo con la reivindicación 1, un procedimiento de acuerdo con la reivindicación 25 y un programa informático de acuerdo con la reivindicación 26.
- 50 **[0005]** Se proporciona un aparato para generar una señal de salida de audio que tiene dos o más canales de salida de audio a partir de una señal de entrada de audio que tiene dos o más canales de entrada de audio. El aparato comprende un proveedor y un procesador de señales. El proveedor está adaptado para proporcionar primeras propiedades de covarianza de la señal de entrada de audio. El procesador de señales está adaptado para generar la señal de salida de audio mediante la aplicación de una regla de mezcla en por lo menos dos de los dos o más canales de entrada de audio. El procesador de señales está configurado para determinar la regla de mezcla sobre la base de las primeras propiedades de covarianza de la señal de entrada de audio y sobre la base de las segundas propiedades de covarianza de la señal de salida de audio, en donde las segundas propiedades de covarianza son distintas de las primeras propiedades de covarianza.
- 55
- 60

5 [0006] Por ejemplo, las potencias de los canales y las dependencias alineadas en tiempo pueden ser expresadas por la parte real de una matriz de covarianza de la señal, por ejemplo, en bandas de frecuencia perceptuales. A continuación se expone un concepto generalmente aplicable para procesar un sonido espacial en este dominio. El concepto comprende una solución de mezcla adaptativa para alcanzar propiedades de covarianza de destino dadas (las segundas propiedades de covarianza), por ejemplo, una matriz de covarianza de destino dada, por el mejor uso de los componentes independientes en los canales de entrada. En una forma de realización se pueden prever medios para inyectar la cantidad necesaria de potencia de sonido descorrelacionado, cuando el objetivo no se logra de otro modo. Dicho concepto es robusto en su función y puede ser aplicado en numerosos casos de uso. Las propiedades de covarianza de destino pueden ser proporcionadas, por ejemplo, por un usuario. Por ejemplo, un aparato de acuerdo con una forma de realización puede tener medios de modo que un usuario pueda introducir las propiedades de covarianza.

15 [0007] De acuerdo con una forma de realización, el proveedor puede adaptarse para proporcionar las primeras propiedades de covarianza, en donde las primeras propiedades de covarianza tienen un primer estado para un primer intervalo de tiempo-frecuencia, y en donde las primeras propiedades de covarianza tienen un segundo estado, diferente del primer estado, para un segundo intervalo de tiempo-frecuencia, diferente del primer intervalo de tiempo-frecuencia. El proveedor no tiene que llevar a cabo necesariamente el análisis para obtener las propiedades de covarianza, pero puede proporcionar estos datos de un almacenamiento, una entrada de usuario o fuentes similares.

20 [0008] En otra forma de realización, el procesador de señales puede adaptarse para determinar la regla de mezcla sobre la base de las segundas propiedades de covarianza, en donde las segundas propiedades de covarianza tienen un tercer estado para un tercer intervalo de tiempo-frecuencia, y en donde las segundas propiedades de covarianza tienen un cuarto estado, que es diferente del tercer estado para un cuarto intervalo de tiempo-frecuencia, que es diferente del tercer intervalo de tiempo-frecuencia.

25 [0009] De acuerdo con otra forma de realización, el procesador de señales está adaptado para generar la señal de salida de audio mediante la aplicación de la regla de mezcla de modo que cada uno de los dos o más canales de salida de audio depende de cada uno de los dos o más canales de entrada de audio.

30 [0010] En otra forma de realización, el procesador de señales puede adaptarse para determinar la regla de mezcla de modo que una medida de error se reduce al mínimo. Una medida de error puede ser, por ejemplo, una señal de diferencia absoluta entre una señal de salida de referencia y una señal de salida real.

35 [0011] En una forma de realización, una medida de error puede ser, por ejemplo, una media que depende de

$$\|y_{\text{ref}} - y\|^2$$

40 en donde  $y$  es la señal de salida de audio, en donde

$$y_{\text{ref}} = Qx,$$

45 en donde  $x$  especifica la señal de entrada de audio y en donde  $Q$  es una matriz de mapeo, que puede ser específica de la aplicación, de modo que  $y_{\text{ref}}$  especifica una señal de salida de audio de destino de referencia.

[0012] De acuerdo con una forma de realización adicional, el procesador de señales puede adaptarse para determinar la regla de mezcla de modo que

$$e = E [\|y_{\text{ref}} - y\|^2]$$

50 se reduce al mínimo, en donde  $E$  es un operador de expectativa, en donde  $y_{\text{ref}}$  es un punto de referencia definido, y en donde  $y$  es la señal de salida de audio.

55 [0013] De acuerdo con otra forma de realización, el procesador de señales puede estar configurado para determinar la regla de mezcla mediante la determinación de las segundas propiedades de covarianza, en donde el procesador de señales puede estar configurado para determinar las segundas propiedades de covarianza sobre la base de las primeras propiedades de covarianza.

60 [0014] De acuerdo con una forma de realización adicional, el procesador de señales puede adaptarse para determinar una matriz de mezcla como la regla de mezcla, en donde el procesador de señales puede adaptarse para determinar la matriz de mezcla sobre la base de las primeras propiedades de covarianza y sobre la base de las segundas propiedades de covarianza.

5 [0015] En otra forma de realización, el proveedor puede adaptarse para analizar las primeras propiedades de covarianza mediante la determinación de una primera matriz de covarianza de la señal de entrada de audio y en donde el procesador de señales puede estar configurado para determinar la regla de mezcla sobre la base de una segunda matriz de covarianza de la señal de salida de audio como las segundas propiedades de covarianza.

10 [0016] De acuerdo con otra forma de realización, el proveedor puede adaptarse para determinar la primera matriz de covarianza de modo que cada valor de la diagonal de la primera matriz de covarianza puede indicar una potencia de uno de los canales de entrada de audio y de modo que cada valor de la primera matriz de covarianza que no es un valor de la diagonal puede indicar una correlación entre canales entre un primer canal de entrada de audio y un segundo canal de entrada de audio diferente.

15 [0017] De acuerdo con una forma de realización adicional, el procesador de señales puede estar configurado para determinar la regla de mezcla sobre la base de la segunda matriz de covarianza, en donde cada valor de la diagonal de la segunda matriz de covarianza puede indicar una potencia de uno de los canales de salida de audio y en donde cada valor de la segunda matriz de covarianza que no es un valor de la diagonal puede indicar una correlación entre canales entre un primer canal de salida de audio y un segundo canal de salida de audio.

20 [0018] De acuerdo con otra forma de realización, el procesador de señales puede adaptarse para determinar la regla de mezcla de modo que:

$$\mathbf{M} = \mathbf{K}_y \mathbf{P} \mathbf{K}_x^{-1}$$

25 de modo que

$$\begin{aligned} \mathbf{K}_x \mathbf{K}_x^T &= \mathbf{C}_x \\ \mathbf{K}_y \mathbf{K}_y^T &= \mathbf{C}_y \end{aligned}$$

30 en donde  $\mathbf{M}$  es la matriz de mezcla, en donde  $\mathbf{C}_x$  es la primera matriz de covarianza, en donde  $\mathbf{C}_y$  es la segunda matriz de covarianza, en donde  $\mathbf{K}_x^T$  es una primera matriz transpuesta de una primera matriz descompuesta  $\mathbf{K}_x$ , en donde  $\mathbf{K}_y^T$  es una segunda matriz transpuesta de una segunda matriz descompuesta  $\mathbf{K}_y$ , en donde  $\mathbf{K}_x^{-1}$  es una matriz inversa de la primera matriz descompuesta  $\mathbf{K}_x$  y en donde  $\mathbf{P}$  es una primera matriz unitaria.

35 [0019] En una forma de realización adicional, el procesador de señales puede adaptarse para determinar la regla de mezcla de modo que

$$\mathbf{M} = \mathbf{K}_y \mathbf{P} \mathbf{K}_x^{-1}$$

en donde

40 
$$\mathbf{P} = \mathbf{V} \mathbf{U}^T$$

en donde  $\mathbf{U}^T$  es una tercera matriz transpuesta de una segunda matriz unitaria  $\mathbf{U}$ , en donde  $\mathbf{V}$  es una tercera matriz unitaria, en donde

45 
$$\mathbf{U} \mathbf{S} \mathbf{V}^T = \mathbf{K}_x^T \mathbf{Q}^T \mathbf{K}_y$$

en donde  $\mathbf{Q}^T$  es una cuarta matriz transpuesta de la matriz de mezcla descendente  $\mathbf{Q}$ , en donde  $\mathbf{V}^T$  es una quinta matriz transpuesta de la tercera matriz unitaria  $\mathbf{V}$ , y en donde  $\mathbf{S}$  es una matriz de la diagonal.

50 [0020] De acuerdo con otra forma de realización, el procesador de señales está adaptado para determinar una matriz de mezcla como regla de mezcla, en donde el procesador de señales está adaptado para determinar la matriz de mezcla sobre la base de las primeras propiedades de covarianza y sobre la base de las segundas propiedades de covarianza, en donde el proveedor está adaptado para proporcionar o analizar las primeras propiedades de covarianza mediante la determinación de una primera matriz de covarianza de la señal de entrada de audio, y en donde el procesador de señales está configurado para determinar la regla de mezcla sobre la base de una segunda matriz de covarianza de la señal de salida de audio como las segundas propiedades de covarianza, en donde el procesador de señales está configurado para modificar por lo menos algunos valores de diagonal de una matriz de

la diagonal  $\mathbf{S}_x$  cuando los valores de la matriz de la diagonal  $\mathbf{S}_x$  son cero o menores que un valor umbral predeterminado, de modo que los valores son mayores que o iguales al valor umbral, en donde el procesador de señales está adaptado para determinar la matriz de mezcla sobre la base de la matriz de la diagonal. Sin embargo, el valor umbral no necesariamente está predeterminado pero también puede depender de una función.

[0021] En una forma de realización adicional, el procesador de señales está configurado para modificar dichos por lo menos algunos valores de diagonal de la matriz de la diagonal  $\mathbf{S}_x$ , en donde  $\mathbf{K}_x = \mathbf{U}_x \mathbf{S}_x \mathbf{V}_x^T$ , y en donde

$\mathbf{C}_x = \mathbf{K}_x \mathbf{K}_x^T$ , en donde  $\mathbf{C}_x$  es la primera matriz de covarianza, en donde  $\mathbf{S}_x$  es la matriz de la diagonal, en donde  $\mathbf{U}_x$  es una segunda matriz,  $\mathbf{V}_x^T$  es una tercera matriz transpuesta, y en donde  $\mathbf{K}_x^T$  es una cuarta matriz transpuesta de la quinta matriz  $\mathbf{K}_x$ . Las matrices  $\mathbf{V}_x$  y  $\mathbf{U}_x$  pueden ser matrices unitarias.

[0022] De acuerdo con otra forma de realización, el procesador de señales está adaptado para generar la señal de salida de audio mediante la aplicación de la regla de mezcla en por lo menos dos de los dos o más canales de entrada de audio para obtener una señal intermedia  $\mathbf{y}' = \hat{\mathbf{M}}\mathbf{x}$  y mediante la adición de una señal residual  $r$  a la señal intermedia para obtener la señal de salida de audio.

[0023] En otra forma de realización, el procesador de señales está adaptado para determinar la matriz de mezcla sobre la base de una matriz de ganancia de la diagonal  $\mathbf{G}$  y una matriz intermedia  $\hat{\mathbf{M}}$ , de modo que  $\mathbf{M}' = \mathbf{G}\hat{\mathbf{M}}$ , en donde la matriz de ganancia de la diagonal tiene el valor

$$\mathbf{G}(i, i) = \sqrt{\frac{\mathbf{C}_y(i, i)}{\hat{\mathbf{C}}_y(i, i)}}$$

en donde  $\hat{\mathbf{C}}_y = \hat{\mathbf{M}}\mathbf{C}_x\hat{\mathbf{M}}^T$ ,

en donde  $\mathbf{M}'$  es la matriz de mezcla, en donde  $\mathbf{G}$  es la matriz de ganancia de la diagonal y en donde  $\hat{\mathbf{M}}$  es la matriz intermedia, en donde  $\mathbf{C}_y$  es la segunda matriz de covarianza y en donde  $\hat{\mathbf{M}}^T$  es una quinta matriz transpuesta de la matriz  $\hat{\mathbf{M}}$ .

[0024] Las formas de realización preferidas de la presente invención se explicarán con referencia a las figuras en las que:

La Fig. 1 ilustra un aparato para generar una señal de salida de audio que tiene dos o más canales de salida de audio a partir de una señal de entrada de audio que tiene dos o más canales de entrada de audio de acuerdo con una forma de realización.

La Fig. 2 representa un procesador de señales de acuerdo con una forma de realización.

La Fig. 3 muestra un ejemplo de aplicación de una combinación lineal de vectores  $\mathbf{L}$  y  $\mathbf{R}$  para lograr un conjunto nuevo de vectores  $\mathbf{R}'$  y  $\mathbf{L}'$ .

La Fig. 4 ilustra un diagrama de bloques de un aparato de acuerdo con otra forma de realización.

La Fig. 5 muestra un diagrama que representa una señal de micrófono de coincidencia estéreo a codificador MPEG de Sonido Envoltante de acuerdo con una forma de realización.

La Fig. 6 representa un aparato de acuerdo con otra forma de realización relativa a la mezcla descendente ICC/corrección de nivel para un codificador SAM-a-MPS.

La Fig. 7 representa un aparato de acuerdo con una forma de realización de un mejoramiento para arreglos de micrófonos de espacio reducido.

La Fig. 8 ilustra un aparato de acuerdo con otra forma de realización para el mejoramiento ciego de la calidad de sonido espacial en la reproducción estéreo o multicanal.

La Fig. 9 ilustra el mejoramiento de las configuraciones estrechas de parlantes.

La Fig. 10 representa una forma de realización que proporciona una representación mejorada de la Codificación de Audio Direccional sobre la base de una señal de micrófono de formato B.

5 Las Figs. 11A, 11B y 11C ilustran la tabla 1 que muestra ejemplos numéricos de una forma de realización, y

Las Figs. 12A y 12B representan el listado 1 que muestra una aplicación de Matlab de un procedimiento de acuerdo con una forma de realización.

10 **[0025]** La Fig. 1 ilustra un aparato para generar una señal de salida de audio que tiene dos o más canales de salida de audio a partir de una señal de entrada de audio que tiene dos o más canales de entrada de audio de acuerdo con una forma de realización. El aparato comprende un proveedor 110 y un procesador de señales 120. El proveedor 110 está adaptado para recibir la señal de entrada de audio que tiene dos o más canales de entrada de audio. Asimismo, el proveedor 110 es una adaptado para analizar las primeras propiedades de covarianza de la señal de entrada de audio. El proveedor 110 está adaptado además para proporcionar las primeras propiedades de covarianza al procesador de señales. El procesador de señales 120 está adaptado además para recibir la señal de entrada de audio. El procesador de señales 120 está adaptado además para generar la señal de salida de audio mediante la aplicación de una regla de mezcla en por lo menos dos de los dos o más canales de entrada de la señal de entrada de audio. El procesador de señales 120 está configurado para determinar la regla de mezcla sobre la base de las primeras propiedades de covarianza de la señal de entrada de audio y sobre la base de las segundas propiedades de covarianza de la señal de salida de audio, en donde las segundas propiedades de covarianza son distintas de las primeras propiedades de covarianza.

25 **[0026]** La Fig. 2 ilustra un procesador de señales de acuerdo con una forma de realización. El procesador de señales comprende una unidad de formulación de matriz de mezcla óptima 210 y una unidad de mezcla 220. La unidad de formulación de matriz de mezcla óptima 210 formula una matriz de mezcla óptima. Para ello, la unidad de formulación de matriz de mezcla óptima 210 utiliza las primeras propiedades de covarianza 230 (por ejemplo, propiedades de covarianza de entrada) de una señal de entrada de audio de banda de frecuencia estéreo o multicanal tal como se recibe, por ejemplo, por un proveedor 110 de la forma de realización de la Fig. 1. Asimismo, la unidad de formulación de matriz de mezcla óptima 210 determina la matriz de mezcla sobre la base de las segundas propiedades de covarianza 240, por ejemplo, una matriz de covarianza de destino, que puede ser dependiente de la aplicación. La matriz de mezcla óptima formulada por la unidad de formulación de matriz de mezcla óptima 210 se puede utilizar como una matriz de mapeo de canales. La matriz de mezcla óptima entonces puede ser proporcionada a la unidad de mezcla 220. La unidad de mezcla 220 aplica la matriz de mezcla óptima a la entrada de banca de frecuencia estéreo o multicanal para obtener una salida de banda de frecuencia estéreo o multicanal de la señal de salida de audio. La señal de salida de audio tiene las segundas propiedades de covarianza deseadas (propiedades de covarianza de destino).

40 **[0027]** Con el fin de explicar las formas de realización de la presente invención con más detalle se incluyen definiciones. Ahora se definen las señales complejas de entrada y salida de media cero  $\mathbf{x}_i(t, f)$  y  $\mathbf{y}_j(t, f)$ , en donde  $t$  es el índice de tiempo, en donde  $f$  es el índice de frecuencia, en donde  $i$  es el índice de canal de entrada, y en donde  $j$  es el índice de canal de salida. Asimismo se definen los vectores de señales de la señal de entrada de audio  $\mathbf{x}$  y la señal de salida de audio  $\mathbf{y}$ :

$$\mathbf{x}_{N_x}(t, f) = \begin{bmatrix} x_1(t, f) \\ x_2(t, f) \\ \vdots \\ x_{N_x}(t, f) \end{bmatrix} \quad \mathbf{y}_{N_y}(t, f) = \begin{bmatrix} y_1(t, f) \\ y_2(t, f) \\ \vdots \\ y_{N_y}(t, f) \end{bmatrix} \quad (1)$$

45

en donde  $N_x$  y  $N_y$  son el número total de canales de entrada y de salida. Asimismo,

se definen  $N = \max(N_y, N_x)$  y las señales de igual dimensión completadas con 0:

50

$$\begin{aligned} \mathbf{x}(t, f) &= \begin{bmatrix} \mathbf{x}_{N_x}(t, f) \\ \mathbf{0}_{(N-N_x) \times 1} \end{bmatrix} \\ \mathbf{y}(t, f) &= \begin{bmatrix} \mathbf{y}_{N_y}(t, f) \\ \mathbf{0}_{(N-N_y) \times 1} \end{bmatrix}. \end{aligned} \quad (2)$$

[0028] Las señales completadas con cero se pueden utilizar en la formulación hasta cuando la solución derivada se extiende hasta distintas longitudes de vectores.

[0029] Como se ha explicado anteriormente, la medida ampliamente utilizada para describir el aspecto espacial de un sonido multicanal es la combinación de las potencias de canales y las dependencias alineadas en tiempo. Estas propiedades están incluidas en la parte real de las matrices de covarianza, definidas como:

$$\begin{aligned} \mathbf{C}_x &= E [\text{Re}\{\mathbf{x}\mathbf{x}^H\}] \\ \mathbf{C}_y &= E [\text{Re}\{\mathbf{y}\mathbf{y}^H\}] \end{aligned} \quad (3)$$

[0030] En la ecuación (3) y a continuación,  $E[\ ]$  es el operador de expectativa,  $\text{Re}\{\}$  es el operador de parte real, y  $\mathbf{x}^H$  y  $\mathbf{y}^H$  son los transpuestos conjugados de  $\mathbf{x}$  y  $\mathbf{y}$ . El operador de expectativa  $E[\ ]$  es un operador matemático. En las aplicaciones prácticas se sustituye por una estimación tal como una media de un determinado intervalo de tiempo. En las secciones siguientes, el uso del término matriz de covarianza se refiere a esta definición de valor real.  $\mathbf{C}_x$  y  $\mathbf{C}_y$  son simétricas y positivas semi-definidas y, por lo tanto, se pueden definir las matrices reales  $\mathbf{K}_x$  y  $\mathbf{K}_y$ , de modo que:

$$\begin{aligned} \mathbf{C}_x &= \mathbf{K}_x \mathbf{K}_x^T \\ \mathbf{C}_y &= \mathbf{K}_y \mathbf{K}_y^T. \end{aligned} \quad (4)$$

Dichas descomposiciones se pueden obtener, por ejemplo, mediante el uso de la descomposición de Cholesky o la descomposición Eigen, véase, por ejemplo,

[7] Golub, G.H. y Van Loan, C.F., "Matrix computations", Johns Hopkins Univ Press, 1996.

[0031] Cabe señalar que existe un número infinito de descomposiciones que cumplen la ecuación (4). Para cualquier matriz ortogonal  $\mathbf{P}_x$  y  $\mathbf{P}_y$ , las matrices  $\mathbf{K}_x \mathbf{P}_x$  y  $\mathbf{K}_y \mathbf{P}_y$  también cumplen la condición ya que

$$\begin{aligned} \mathbf{K}_x \mathbf{P}_x \mathbf{P}_x^T \mathbf{K}_x^T &= \mathbf{K}_x \mathbf{K}_x^T = \mathbf{C}_x \\ \mathbf{K}_y \mathbf{P}_y \mathbf{P}_y^T \mathbf{K}_y^T &= \mathbf{K}_y \mathbf{K}_y^T = \mathbf{C}_y. \end{aligned} \quad (5)$$

en casos que utilizan estéreo, la matriz de covarianza se da a menudo en forma de potencias de canales y correlación entre canales (ICC, por sus siglas en inglés), por ejemplo, en [1, 3, 4]. Los valores de diagonal de  $\mathbf{C}_x$  son las potencias de canales y la ICC entre los dos canales es

$$\text{ICC}_x = \frac{\mathbf{C}_x(1,2)}{\sqrt{\mathbf{C}_x(1,1)\mathbf{C}_x(2,2)}} \quad (6)$$

y correspondientemente para  $\mathbf{C}_y$ . Los índices en los paréntesis indican fila y columna de la matriz.

[0032] La definición que queda es la matriz de mapeo determinada por la aplicación  $\mathbf{Q}$ , que comprende la información sobre cuáles canales de entrada se van a utilizar en la composición de cada canal de salida. Con  $\mathbf{Q}$  se puede definir una señal de referencia

$$\mathbf{y}_{\text{ref}} = \mathbf{Q}\mathbf{x}. \quad (7)$$

[0033] La matriz de mapeo  $\mathbf{Q}$  puede comprender cambios en la dimensionalidad y, el cambio de escala, la combinación y reordenación de los canales. Debido a la definición de las señales completadas con cero,  $\mathbf{Q}$  es aquí una matriz cuadrada  $N \times N$  que puede comprender cero filas o columnas. Algunos ejemplos de  $\mathbf{Q}$  son:

- 5 - Mejoramiento espacial:  $\mathbf{Q} = \mathbf{I}$ , en aplicaciones en donde la salida debería ser similar a la entrada.
- Mezcla descendente:  $\mathbf{Q}$  es una matriz de mezcla descendente.
- Síntesis espacial a partir de señales de micrófonos de primer orden:  $\mathbf{Q}$  puede ser, por ejemplo, una matriz de mezcla de micrófonos de Ambisonic, que significa que  $\mathbf{y}_{ref}$  es un conjunto de señales de micrófonos virtuales.

10 [0034] A continuación se formula el modo de generar una señal  $\mathbf{y}$  a partir de una señal  $\mathbf{x}$ , con una limitación de que  $\mathbf{y}$  tiene la matriz de covarianza definida por la aplicación  $\mathbf{C}_y$ . La solicitud también define una matriz de mapeo  $\mathbf{Q}$  que da un punto de referencia para la optimización. La señal de entrada  $\mathbf{x}$  tiene la matriz de covarianza medida  $\mathbf{C}_x$ . Como se ha indicado, los conceptos propuestos para llevar a cabo esta transformada utilizan, en primer lugar, un concepto que implica sólo una mezcla óptima de los canales, ya que el uso de descorreladores comprende generalmente la calidad de la señal, y en segundo lugar, la inyección de potencia descorrelacionada cuando el objetivo no se logra de otro modo.

15 [0035] La relación de entrada-salida de acuerdo con estos conceptos se puede expresar del siguiente modo

$$\mathbf{y} = \mathbf{M}\mathbf{x} + \mathbf{r} \quad (8)$$

20 en donde  $\mathbf{M}$  es una matriz de mezcla real de acuerdo con el concepto primario y  $\mathbf{r}$  es una señal residual de acuerdo con el concepto secundario.

25 [0036] A continuación se sugieren conceptos para la modificación de la matriz de covarianza.

[0037] En primer lugar, la tarea de acuerdo con el concepto primario se resuelve sólo por la mezcla cruzada de los canales de entrada. La ecuación (8) luego se simplifica a

$$\mathbf{y} = \mathbf{M}\mathbf{x}. \quad (9)$$

30 A partir de las ecuaciones (3) y (9), se obtiene

$$\begin{aligned} \mathbf{C}_y &= E [\text{Re}\{\mathbf{y}\mathbf{y}^H\}] \\ &= E [\text{Re}\{\mathbf{M}\mathbf{x}\mathbf{x}^H\mathbf{M}^T\}] = \mathbf{M}\mathbf{C}_x\mathbf{M}^T. \end{aligned} \quad (10)$$

35 [0038] A partir de las ecuaciones (5) y (10) se deduce que

$$\mathbf{K}_y\mathbf{P}_y\mathbf{P}_y^T\mathbf{K}_y^T = \mathbf{M}\mathbf{K}_x\mathbf{P}_x\mathbf{P}_x^T\mathbf{K}_x^T\mathbf{M}^T \quad (11)$$

40 [0039] a partir del cual proviene un conjunto de soluciones para  $\mathbf{M}$  que cumplen la ecuación (10)

$$\mathbf{M} = \mathbf{K}_y\mathbf{P}_y\mathbf{P}_x^T\mathbf{K}_x^{-1} = \mathbf{K}_y\mathbf{P}\mathbf{K}_x^{-1} \quad (12)$$

45 [0040] La condición para estas soluciones es que exista  $\mathbf{K}_x^{-1}$ . La matriz ortogonal  $\mathbf{P} = \mathbf{P}_y\mathbf{P}_x^T$  es el parámetro libre restante.

[0041] A continuación se describe cómo se obtiene una matriz  $\mathbf{P}$  que proporciona una matriz óptima  $\mathbf{M}$ . De todos los  $\mathbf{M}$  en la ecuación (12) se ha examinado uno que produce una salida más cercana al punto de referencia definido  $\mathbf{y}_{ref}$ , es decir, que reduce al mínimo

$$e = E [\|\mathbf{y}_{ref} - \mathbf{y}\|^2] \quad (13a)$$

es decir, que reduce al mínimo

$$e = E [\|\mathbf{y}_{\text{ref}} - \mathbf{y}\|^2] = E [\|\mathbf{Q}\mathbf{x} - \mathbf{M}\mathbf{x}\|^2]. \quad (13)$$

**[0042]** Ahora se define una señal  $\mathbf{w}$ , de modo que  $E[\text{Re}\{\mathbf{w}\mathbf{w}^H\}] = \mathbf{I}$ . Se puede elegir  $\mathbf{w}$  de modo que  $\mathbf{x} = \mathbf{K}_x\mathbf{w}$ , ya que

$$\begin{aligned} E[\text{Re}\{\mathbf{x}\mathbf{x}^H\}] &= E[\text{Re}\{\mathbf{K}_x\mathbf{w}\mathbf{w}^H\mathbf{K}_x^T\}] \\ &= \mathbf{K}_xE[\text{Re}\{\mathbf{w}\mathbf{w}^H\}]\mathbf{K}_x^T \\ &= \mathbf{K}_x\mathbf{K}_x^T = \mathbf{C}_x. \end{aligned} \quad (14)$$

5

**[0043]** A continuación, se deduce que

$$\mathbf{M}\mathbf{x} = \mathbf{M}\mathbf{K}_x\mathbf{w} = \mathbf{K}_y\mathbf{P}\mathbf{w}. \quad (15)$$

10 **[0044]** La ecuación (13) se puede expresar como

$$\begin{aligned} e &= E [\|\mathbf{Q}\mathbf{x} - \mathbf{M}\mathbf{x}\|^2] \\ &= E [\|\mathbf{Q}\mathbf{K}_x\mathbf{w} - \mathbf{K}_y\mathbf{P}\mathbf{w}\|^2] \\ &= E [\|(\mathbf{Q}\mathbf{K}_x - \mathbf{K}_y\mathbf{P})\mathbf{w}\|^2] \\ &= E [\mathbf{w}^H(\mathbf{Q}\mathbf{K}_x - \mathbf{K}_y\mathbf{P})^T(\mathbf{Q}\mathbf{K}_x - \mathbf{K}_y\mathbf{P})\mathbf{w}]. \end{aligned} \quad (16)$$

**[0045]** A partir de  $E[\text{Re}\{\mathbf{w}\mathbf{w}^H\}] = \mathbf{I}$ , se puede demostrar fácilmente para una matriz simétrica real  $\mathbf{A}$  que  $E[\mathbf{w}^H\mathbf{A}\mathbf{w}] = \text{tr}(\mathbf{A})$ , que es la traza de la matriz. De ello se deduce que la ecuación (16) adopta la forma

15

$$e = \text{tr} [(\mathbf{Q}\mathbf{K}_x - \mathbf{K}_y\mathbf{P})^T(\mathbf{Q}\mathbf{K}_x - \mathbf{K}_y\mathbf{P})]. \quad (17)$$

20

**[0046]** Para trazas de matriz se puede confirmar fácilmente que

$$\begin{aligned} \text{tr}(\mathbf{A} + \mathbf{B}) &= \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B}) \\ \text{tr}(\mathbf{A}) &= \text{tr}(\mathbf{A}^T) \\ \text{tr}(\mathbf{P}^T\mathbf{A}\mathbf{P}) &= \text{tr}(\mathbf{A}). \end{aligned} \quad (18)$$

25

**[0047]** Al utilizar estas propiedades, la ecuación (17) adopta la forma

$$\begin{aligned} e &= \text{tr}(\mathbf{K}_x^T\mathbf{Q}^T\mathbf{Q}\mathbf{K}_x) + \text{tr}(\mathbf{K}_y^T\mathbf{K}_y) \\ &\quad - 2\text{tr}(\mathbf{K}_x^T\mathbf{Q}^T\mathbf{K}_y\mathbf{P}). \end{aligned} \quad (19)$$

30

**[0048]** Sólo el último término depende de  $\mathbf{P}$ . El problema de optimización es así

$$\mathbf{P} = \arg \min_{\mathbf{P}} e = \arg \max_{\mathbf{P}} [\text{tr}(\mathbf{K}_x^T\mathbf{Q}^T\mathbf{K}_y\mathbf{P})]. \quad (20)$$

**[0049]** Se puede demostrar fácilmente para una matriz de la diagonal no negativa  $\mathbf{S}$  y cualquier matriz ortogonal  $\mathbf{P}_s$  que

35

$$\text{tr}(\mathbf{S}) \geq \text{tr}(\mathbf{S}\mathbf{P}_s). \quad (21)$$

**[0050]** De ese modo, mediante la definición de la descomposición de valor singular  $\mathbf{USV}^T = \mathbf{K}_x^T \mathbf{Q}^T \mathbf{K}_y$ , en donde  $\mathbf{S}$  no es negativa y diagonal y  $\mathbf{U}$  y  $\mathbf{V}$  son ortogonales, se deduce que

$$\begin{aligned} \text{tr}(\mathbf{S}) &\geq \text{tr}(\mathbf{SV}^T \mathbf{PU}) = \text{tr}(\mathbf{USV}^T \mathbf{PUU}^T) \\ &= \text{tr}(\mathbf{K}_x^T \mathbf{Q}^T \mathbf{K}_y \mathbf{P}) \end{aligned} \quad (22)$$

5 para cualquier  $\mathbf{P}$  ortogonal. La igualdad se cumple para

$$\boxed{\mathbf{P} = \mathbf{VU}^T} \quad (23)$$

por lo que  $\mathbf{P}$  produce el máximo de  $\text{tr}(\mathbf{K}_x^T \mathbf{Q}^T \mathbf{K}_y \mathbf{P})$  y el mínimo de la medida de error en la ecuación (13).

10 **[0051]** Un aparato de acuerdo con una forma de realización determina una matriz de mezcla óptima  $\mathbf{M}$ , de modo que se reduce al mínimo un error  $e$ . Cabe señalar que las propiedades de covarianza de la señal de entrada de audio y la señal de salida de audio pueden variar para distintos intervalos de tiempo-frecuencia. Para ello, un proveedor de un aparato de acuerdo con una forma de realización está adaptado para analizar las propiedades de covarianza del canal de entrada de audio que pueden ser diferentes para los distintos intervalos de tiempo-frecuencia. Asimismo, el procesador de señales de un aparato de acuerdo con una forma de realización está adaptado para determinar una regla de mezcla, por ejemplo, una matriz de mezcla  $\mathbf{M}$  sobre la base de las segundas propiedades de covarianza de la señal de salida de audio, en donde las segundas propiedades de covarianza pueden tener distintos valores para distintos intervalos de tiempo-frecuencia.

20 **[0052]** Como la matriz de mezcla determinada  $\mathbf{M}$  se aplica sobre cada uno de los canales de entrada de audio de la señal de entrada de audio, y como cada uno de los canales de salida de audio resultantes de la señal de salida de audio pueden depender, por lo tanto, de cada uno de los canales de entrada de audio, un procesador de señales de un aparato de acuerdo con una forma de realización está adaptado, por lo tanto, para generar la señal de salida de audio mediante la aplicación de la regla de mezcla de modo que cada uno de los dos o más canales de salida de audio depende de cada uno de los dos o más canales de entrada de audio de la señal de entrada de audio.

25 **[0053]** De acuerdo con otra forma de realización se sugiere utilizar la descorrelación cuando  $\mathbf{K}_x^{-1}$  no existe o es inestable. En las formas de realización descritas anteriormente se proporciona una solución para la determinación de una matriz de mezcla óptima en donde se supone que  $\mathbf{K}_x^{-1}$  existe. Sin embargo,  $\mathbf{K}_x^{-1}$  no siempre existe o su inversa puede implicar multiplicadores muy grandes si algunos de los componentes principales en  $\mathbf{x}$  son muy pequeños. Un modo efectivo de regularizar la inversa consiste en emplear la descomposición de valor singular  $\mathbf{K}_x = \mathbf{U}_x \mathbf{S}_x \mathbf{V}_x^T$ . En consecuencia, la inversa es

$$\mathbf{K}_x^{-1} = \mathbf{V}_x \mathbf{S}_x^{-1} \mathbf{U}_x^T. \quad (24)$$

35 **[0054]** Los problemas surgen cuando algunos de los valores de diagonal de la matriz de la diagonal no negativa  $\mathbf{S}_x$  son cero o muy pequeños. Un concepto que regulariza fuertemente la inversa es entonces sustituir estos valores con valores más grandes. El resultado de este procedimiento es  $\mathbf{S}_x$ , y la inversa correspondiente  $\hat{\mathbf{K}}_x^{-1} = \mathbf{V}_x \hat{\mathbf{S}}_x^{-1} \mathbf{U}_x^T$ , y la matriz de mezcla correspondiente  $\hat{\mathbf{M}} = \mathbf{K}_y \mathbf{P} \hat{\mathbf{K}}_x^{-1}$ .

40 **[0055]** Esta regularización significa efectivamente que dentro del proceso de mezcla, la amplificación de algunos de los componentes pequeños principales en  $\mathbf{x}$  se reduce y, por consiguiente, su intacta (sic) a la señal de salida  $\mathbf{y}$  también se reduce y la covarianza de destino  $\mathbf{C}_y$  generalmente no se alcanza.

45 **[0056]** Por esto, de acuerdo con una forma de realización, el procesador de señales puede estar configurado para modificar por lo menos algunos valores de diagonal de una matriz de la diagonal  $\mathbf{S}_x$ , en donde los valores de la matriz de la diagonal  $\mathbf{S}_x$  son cero o menores que un valor umbral (el valor de umbral puede estar predeterminado o puede depender de una función), de modo que los valores son mayores que o iguales al valor umbral, en donde el procesador de señales puede estar adaptado para determinar la matriz de mezcla sobre la base de la matriz de la diagonal.

50

[0057] De acuerdo con una forma de realización, el procesador de señales puede estar configurado para modificar dichos por lo menos algunos valores de diagonal de la matriz de la diagonal  $\mathbf{S}_x$ , en donde  $\mathbf{K}_x = \mathbf{U}_x \mathbf{S}_x \mathbf{V}_x^T$ , y en donde  $\mathbf{C}_x = \mathbf{K}_x \mathbf{K}_x^T$  en donde  $\mathbf{C}_x$  es la primera matriz de covarianza, en donde  $\mathbf{S}_x$  es la matriz de la diagonal, en donde  $\mathbf{U}_x$  es una segunda matriz,  $\mathbf{V}_x^T$  es una tercera matriz transpuesta y en donde  $\mathbf{K}_x^T$  es una cuarta matriz transpuesta de la quinta matriz  $\mathbf{K}_x$ .

[0058] La pérdida anterior de un componente de la señal puede compensarse totalmente con una señal residual  $\mathbf{r}$ . La relación de entrada-salida original se elaborará con la inversa regularizada.

$$\begin{aligned} \mathbf{y} &= \hat{\mathbf{M}}\mathbf{x} + \mathbf{r} = \mathbf{K}_y \mathbf{P} \hat{\mathbf{K}}_x^{-1} \mathbf{x} + \mathbf{r} \\ &= \mathbf{K}_y \mathbf{P} \mathbf{V}_x \hat{\mathbf{S}}_x^{-1} \mathbf{U}_x^T \mathbf{x} + \mathbf{r} \end{aligned} \quad (25)$$

[0059] Ahora se define un componente aditivo  $\mathbf{c}$  de modo que en lugar de  $\hat{\mathbf{S}}_x^{-1} \mathbf{U}_x^T \mathbf{x}$ , se obtiene  $\hat{\mathbf{S}}_x^{-1} \mathbf{U}_x^T \mathbf{x} + \mathbf{c}$ . Adicionalmente se define una señal independiente  $\mathbf{w}'$ , de modo que  $E [\text{Re}\{\mathbf{w}' \mathbf{w}'^H\}] = \mathbf{I}_y$

$$\mathbf{c} = \sqrt{\mathbf{I} - (\hat{\mathbf{S}}_x^{-1} \mathbf{S}_x)^2} \mathbf{w}'. \quad (26)$$

[0060] Puede demostrarse fácilmente que una señal

$$\begin{aligned} \mathbf{y}' &= \mathbf{K}_y \mathbf{P} \mathbf{V}_x (\hat{\mathbf{S}}_x^{-1} \mathbf{U}_x^T \mathbf{x} + \mathbf{c}) \\ &= \hat{\mathbf{M}}\mathbf{x} + \mathbf{K}_y \mathbf{P} \mathbf{V}_x \mathbf{c} \end{aligned} \quad (27)$$

tiene covarianza  $\mathbf{C}_y$ . La señal residual para compensar la regularización es entonces

$$\mathbf{r} = \mathbf{K}_y \mathbf{P} \mathbf{V}_x \mathbf{c}. \quad (28)$$

[0061] A partir de las ecuaciones (27) y (28) se deduce que

$$\mathbf{C}_r = E [\text{Re}\{\mathbf{r} \mathbf{r}^H\}] = \mathbf{C}_y - \hat{\mathbf{M}} \mathbf{C}_x \hat{\mathbf{M}}^T. \quad (29)$$

[0062] Como  $\mathbf{c}$  ha sido definida como una señal estocástica, se deduce que la propiedad relevante de  $\mathbf{r}$  es su matriz de covarianza. De este modo, cualquier señal que es independiente con respecto a  $\mathbf{x}$  que se procesa para tener la covarianza  $\mathbf{C}_r$  sirve como una señal residual que idealmente, reconstruye la matriz de covarianza de destino  $\mathbf{C}_y$  en situaciones en las que se utilizó la regularización descripta. Dicha señal residual puede generarse fácilmente utilizando descorreladores y el procedimiento propuesto de mezcla de canales.

[0063] No resulta sencillo encontrar analíticamente el equilibrio óptimo entre la cantidad de potencia descorrelacionada y la amplificación de pequeños componentes de la señal. Esto es debido a que depende de factores específicos de la aplicación, tales como la estabilidad de las propiedades estadísticas de la señal de entrada, la ventana de análisis aplicado y SNR de la señal de entrada. Sin embargo, es bastante sencillo ajustar una función heurística para llevar a cabo este equilibrio sin desventajas obvias, como se realizó en el código de ejemplo proporcionado a continuación.

[0064] De acuerdo con esto, el procesador de señales de un aparato de acuerdo con una forma de realización puede estar adaptado para generar la señal de salida de audio mediante la aplicación de la regla de mezcla en dichas por lo menos dos de las dos o más señales de entrada de audio para obtener una señal intermedia  $\mathbf{y}' = \hat{\mathbf{M}}\mathbf{x}$  y mediante la adición de una señal residual  $\mathbf{r}$  a la señal intermedia para obtener la señal de salida de audio.

**[0065]** Se ha demostrado que cuando se aplica la regularización de la inversa de  $\mathbf{K}_x$ , los componentes de la señal faltantes en la salida total pueden complementarse totalmente con una señal residual  $\mathbf{r}$  con la covarianza  $\mathbf{C}_r$ . Por estos medios, se puede garantizar que la covarianza de destino  $\mathbf{C}_y$  siempre se alcanza. A continuación se presenta una forma de generar una señal residual correspondiente  $\mathbf{r}$ . Dicha forma comprende los pasos siguientes:

- 5 1. Generar un conjunto de señales que abarque la misma cantidad de canales de salida. La señal  $\mathbf{y}_{ref} = \mathbf{Q}\mathbf{x}$  se puede emplear porque tiene tantos canales como la señal de salida, y cada una de las señales de salida contiene una señal apropiada para ese canal en particular.
- 10 2. Descorrelacionar esta señal. Hay muchas maneras de descorrelacionar, incluyendo los filtros de todo paso, convoluciones con ráfagas de ruido, y retardos pseudo-aleatorios en bandas de frecuencia.
- 15 3. Medir (o suponer) la matriz de covarianza de la señal descorrelacionada. La medición es más simple y más robusta, pero debido a que las señales provienen de descorreladores, podrían suponerse incoherentes. Entonces, sólo la medición de la potencia sería suficiente.
- 20 4. Aplicar el procedimiento propuesto para generar una matriz de mezcla que, cuando se aplica a la señal descorrelacionada, genera una señal de salida con la matriz de covarianza  $\mathbf{C}_r$ . Utilizar aquí una matriz de mapeo  $\mathbf{Q} = \mathbf{I}$ , debido a que se desea afectar mínimamente el contenido de la señal.
- 25 5. Procesar la señal proveniente de los descorreladores con esta matriz de mezcla e introducirla en la señal de salida para complementar la falta de los componentes de la señal. Por esto se logra la  $\mathbf{C}_y$  de destino.

**[0066]** En una forma de realización alternativa, los canales descorrelacionados se añaden a la (por lo menos una) señal de entrada antes de la formulación de la matriz de mezcla óptima. En este caso, la entrada y la salida es de igual dimensión, y siempre que la señal de entrada tenga tantos componentes independientes de la señal como canales de entrada, no es necesario utilizar una señal residual  $\mathbf{r}$ . Cuando los descorreladores se utilizan de este modo, el uso de descorreladores es "invisible" para el concepto propuesto, porque los canales descorrelacionados son canales de entrada como cualquier otro.

**[0067]** Si el uso de descorreladores no es aconsejable, por lo menos las potencias de canales de destino pueden lograrse mediante la multiplicación de las filas de  $\hat{\mathbf{M}}$  de modo que

$$\mathbf{M}' = \mathbf{G}\hat{\mathbf{M}} \quad (30)$$

35 en donde  $\mathbf{G}$  es una matriz de ganancia de la diagonal con los valores

$$\mathbf{G}(i, i) = \sqrt{\frac{\mathbf{C}_y(i, i)}{\hat{\mathbf{C}}_y(i, i)}} \quad (31)$$

40 en donde  $\hat{\mathbf{C}}_y = \hat{\mathbf{M}}\mathbf{C}_x\hat{\mathbf{M}}^T$ .

**[0068]** En muchas aplicaciones, el número de canales de entrada y de salida es diferente. Como se describe en la Ecuación (2), la adición de ceros de la señal con una dimensión más pequeña se aplica para tener la misma dimensión que la superior. La adición de ceros implica una sobrecarga computacional debido a que algunas filas o columnas en  $\mathbf{M}$  resultante corresponden a los canales con potencia cero definida. Desde el punto de vista matemático, primero utilizar la adición de ceros y finalmente recortar la  $\mathbf{M}$  hasta la dimensión  $N_y \times N_x$  relevante equivale a reducir la sobrecarga mediante la introducción de la matriz  $\mathbf{\Lambda}$  que es una matriz de identidad añadida con ceros hasta la dimensión  $N_y \times N_x$ , por ejemplo,

$$\mathbf{\Lambda}_{3 \times 2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}. \quad (32)$$

50 Cuando  $\mathbf{P}$  se redefine de modo que

$$\mathbf{P} = \mathbf{\Lambda}\mathbf{U}^T \quad (33)$$

la **M** resultante es una matriz de mezcla  $N_y \times N_x$  que es la misma que la parte relevante de la **M** del caso con adición de ceros. En consecuencia, **C<sub>x</sub>**, **C<sub>y</sub>**, **K<sub>x</sub>** y **K<sub>y</sub>** pueden ser su dimensión natural y la matriz de mapeo **Q** es de dimensión  $N_y \times N_x$ .

5  
 [0069] La matriz de covarianza de entrada es siempre descomponible a  $\mathbf{C}_x = \mathbf{K}_x \mathbf{K}_x^T$  debido a que es una medida positiva semi-definida de una señal real. Sin embargo, es posible definir dichas matrices de covarianza de destino que no son descomponibles por la razón de que representan las dependencias imposibles de canal. Hay conceptos para garantizar la descomposición, tales como el ajuste de los valores Eigen negativos a ceros y la normalización de la potencia, véase, por ejemplo,

[8] R. Rebonato, P. Jäckel, "The most general methodology to create a valid correlation matrix for risk management and option pricing purposes", Journal of Risk, Vol. 2, N° 2, pp. 17-28, 2000.

15 [0070] Sin embargo, el uso más significativo del concepto propuesto consiste en solicitar sólo las matrices de covarianza posibles.

[0071] Para resumir lo anterior, la tarea común puede reformularse del siguiente modo. En primer lugar, una señal de entrada tiene una matriz de covarianza determinada. En segundo lugar, la aplicación define dos parámetros: la matriz de covarianza de destino y una regla, cuyos canales de entrada se van a utilizar en la composición de cada uno de los canales de salida. Para llevar a cabo esta transformada se sugiere utilizar los siguientes conceptos: El concepto primario, ilustrado por la Fig. 2, es que la covarianza de destino se logra utilizando una solución de mezcla óptima de los canales de entrada. Este concepto se considera primario ya que evita el uso del descorrelador, que a menudo compromete la calidad de la señal. El concepto secundario tiene lugar cuando no hay suficientes componentes independientes de potencia disponible razonable. La potencia descorrelacionada se inyecta para compensar la falta de estos componentes. Juntos, estos dos conceptos proporcionan los medios para llevar a cabo el ajuste robusto de matriz de covarianza en cualquier escenario dado.

[0072] La principal aplicación prevista del concepto propuesto está en el campo de la microfónica espacial [2,3], que es el campo en donde los problemas relacionados con covarianza de señales son particularmente evidentes debido a las limitaciones físicas de los micrófonos direccionales. Otros casos previstos de uso incluyen el mejoramiento estéreo y multicanal, la extracción ambiental, y la mezcla ascendente y descendente.

[0073] En la descripción anterior se han proporcionado definiciones y, a continuación, la derivación del concepto propuesto. En primer lugar se ha proporcionado la solución de mezcla cruzada, luego se ha proporcionado el concepto de inyección de la potencia de sonido correlacionado. Posteriormente se ha proporcionado una descripción del concepto con un número diferente de canales de entrada y de salida y también consideraciones en cuanto a la descomponibilidad de la matriz de covarianza. A continuación se presentan casos prácticos de uso y un conjunto de ejemplos numéricos y la conclusión. Asimismo se proporciona un ejemplo de código Matlab con la funcionalidad completa de acuerdo con este documento.

[0074] La característica espacial percibida de un sonido estéreo o multicanal está ampliamente definida por la matriz de covarianza de la señal en bandas de frecuencia. Se ha proporcionado un concepto para llevar a cabo una mezcla cruzada óptima y adaptativa de un conjunto de canales de entrada con propiedades de covarianza dadas a un conjunto de canales de salida con propiedades de covarianza arbitrariamente definibles. Se ha proporcionado otro concepto para inyectar potencia descorrelacionada sólo cuando sea necesario, cuando no haya disponibilidad de componentes de sonido independientes de potencia razonable. El concepto tiene una amplia variedad de aplicaciones en el campo del procesamiento de señales de audio espacial.

[0075] Las potencias de canales y las dependencias entre los canales (o la matriz de covarianza) de una señal multicanal pueden controlarse sólo mediante la mezcla cruzada lineal y variable en tiempo dependiendo de las características de entrada y las características de destino deseadas. Este concepto puede ilustrarse con una representación de factor de la señal en donde el ángulo entre los vectores corresponde a la dependencia de canales y la amplitud del vector es igual al nivel de la señal.

[0076] La Fig. 3 ilustra un ejemplo para aplicar una combinación lineal de vectores **L** y **R** para lograr un nuevo conjunto de vectores **R'** y **L'**. De manera similar, los niveles de canales de audio y su dependencia pueden modificarse con una combinación lineal. La solución general no incluye vectores sino una formulación de matriz que es óptima para cualquier número de canales.

[0077] La matriz de mezcla para las señales estéreo se puede formular fácilmente también en forma trigonométrica, como puede verse en la Fig. 3. Los resultados son los mismos que con las matemáticas de la matriz, pero la formulación es diferente.

**[0078]** Si los canales de entrada son muy dependientes sólo es posible lograr la matriz de covarianza de destino con el uso de descorreladores. También se ha proporcionado un procedimiento para inyectar descorreladores sólo cuando es necesario, por ejemplo, de forma óptima.

**[0079]** La Fig. 4 ilustra un diagrama de bloques de un aparato de acuerdo con una forma de realización que aplica la técnica de mezcla. El aparato comprende un módulo de análisis de la matriz de covarianza 410, y un procesador de señales (sin ilustrar), en donde el procesador de señales comprende un módulo de formulación de la matriz de mezcla 420 y un módulo de aplicación de la matriz de mezcla 430. Las propiedades de covarianza de entrada de una entrada de banda de frecuencia estéreo o multicanal son analizadas por un módulo de análisis de la matriz de covarianza 410. El resultado del análisis de la matriz de covarianza se introduce en un módulo de formulación de la matriz de mezcla 420.

**[0080]** El módulo de formulación de la matriz de mezcla 420 formula una matriz de mezcla sobre la base del resultado del análisis de la matriz de covarianza, sobre la base de una matriz de covarianza de destino y posiblemente también sobre la base de un criterio de error.

**[0081]** El módulo de formulación de la matriz de mezcla 420 introduce la matriz de mezcla en un módulo de aplicación de la matriz de mezcla 430. El módulo de aplicación de la matriz de mezcla 430 aplica la matriz de mezcla en la entrada de banda de frecuencia estéreo o multicanal para obtener una salida de banda de frecuencia estéreo o multicanal que tiene, por ejemplo, las propiedades predefinidas de covarianza de destino que dependen de la matriz de covarianza de destino.

**[0082]** Resumiendo lo anterior, el objetivo general del concepto consiste en mejorar, corregir y/o sintetizar el sonido espacial con un grado extremo de optimización en términos de calidad de sonido. El de destino, por ejemplo, las segundas propiedades de covarianza, está definido por la aplicación.

**[0083]** El concepto es perceptualmente significativo y también aplicable en toda la banda, especialmente en el procesamiento de la banda de frecuencia.

**[0084]** Los descorreladores se utilizan con el fin de mejorar (reducir) la correlación entre canales. Lo hacen, pero tienden a poner en peligro la calidad del sonido en general, especialmente con un componente de sonido transitorio.

**[0085]** El concepto propuesto evita, o en alguna aplicación reduce al mínimo, el uso de descorreladores. El resultado es la misma característica espacial, pero sin dicha pérdida de calidad de sonido.

**[0086]** Entre otros usos, la tecnología se puede emplear en un codificador SAM-a-MPS.

**[0087]** El concepto propuesto se ha implementado para mejorar una técnica de micrófono que genera una corriente de bits de Sonido Envolvente de MPEG (Grupo de Expertos en Imágenes en Movimiento, MPEG según sus siglas en inglés) de una señal desde micrófonos estéreo coincidentes de primer orden, véase, por ejemplo, [3]. El proceso incluye la estimación a partir de la señal estéreo de la dirección y la difusión del campo de sonido en bandas de frecuencia y la creación de dicha corriente de bits de Sonido Envolvente de MPEG que, cuando se decodifica en el extremo receptor, produce un campo de sonido que se aproxima perceptualmente al campo de sonido original.

**[0088]** En la Fig. 5 se ilustra un diagrama que representa una señal de micrófono de coincidencia estéreo al codificador de Sonido Envolvente de MPEG de acuerdo con una forma de realización, que emplea el concepto propuesto para crear la señal de mezcla descendente de Sonido Envolvente de MPEG a partir de la señal de micrófono dada. Todo el procesamiento se lleva a cabo en bandas de frecuencia.

**[0089]** Un módulo de determinación de datos espaciales 520 está adaptado para formular los datos de información de configuración que comprenden datos de sonido envolvente espacial y mezcla descendente ICC y/o niveles sobre la base de la información de dirección y difusión que depende de un modelo de campo de sonido 510. El modelo de campo de sonido en sí se basa en un análisis de ICCs de micrófonos y en los niveles de una señal la CCI de micrófono y los niveles de una señal de micrófono estéreo. El módulo de determinación de datos espaciales 520 a continuación proporciona las ICCs de mezcla descendente de destino y los niveles a un módulo de formulación de la matriz de mezcla 530. Asimismo, el módulo de determinación de datos espaciales 520 puede estar adaptado para formular datos de sonido envolvente espacial e ICCs de mezcla descendente y niveles como información lateral espacial de Sonido Envolvente de MPEG. El módulo de formulación de la matriz de mezcla 530 luego formula una matriz de mezcla sobre la base de datos de información de configuración proporcionados, por ejemplo las ICCs y los niveles de la mezcla descendente de destino, e introduce la matriz en un módulo de mezcla 540. El módulo de mezcla 540 aplica la matriz de mezcla en la señal de micrófono estéreo. Por esto, se genera una señal que tiene las ICCs y los niveles de destino. La señal que tiene las ICCs y los niveles de destino luego se proporciona a un

codificador central 550. En una forma de realización, los módulos 520, 530 y 540 son submódulos de un procesador de señales.

5 **[0090]** Dentro del proceso llevado a cabo por un aparato de acuerdo con la Fig. 5 debe generarse una mezcla descendente estéreo de Sonido Envolvente de MPEG. Esto incluye la necesidad de ajustar los niveles y las ICCs de la señal estéreo dado con un impacto mínimo sobre la calidad del sonido. Para este fin se aplicó el concepto de mezcla cruzada sugerido y pudo observarse el beneficio perceptual de la técnica anterior en [3].

10 **[0091]** La Fig. 6 ilustra un aparato de acuerdo con otra forma de realización relativa a la ICC de mezcla descendente/corrección de nivel para un codificador SAM-a-MPS, Un análisis de ICC y nivel se lleva a cabo en el módulo 602 y el modelo de campo de sonido 610 depende del análisis de la ICC y el nivel por el módulo 602. El módulo 620 corresponde al módulo 520, el módulo 630 corresponde al módulo 530 y el módulo 640 corresponde a módulo 540 de la Fig. 5, respectivamente. Lo mismo se aplica para el codificador central 650 que corresponde al  
15 **[0091]** codificador central 550 de la Fig. 5. El concepto descrito anteriormente puede estar integrado en un codificador SAM-a-MPS para crear desde las señales del micrófono la mezcla descendente MPS con ICC y niveles exactamente correctos. El concepto descrito anteriormente es también aplicable en la representación directa SAM-a-multicanal sin MPS con el fin de proporcionar una síntesis espacial ideal mientras reduce al mínimo la cantidad de uso del descorrelador.

20 **[0092]** Se esperan mejoras con respecto a la distancia de origen, la localización de origen, la estabilidad, la comodidad de escucha y el sonido envolvente.

25 **[0093]** La Fig. 7 representa un aparato de acuerdo con una forma de realización de un mejoramiento para arreglos de micrófonos de espacio reducido. Un módulo 705 está adaptado para llevar a cabo un análisis de matriz de covarianza de una señal de entrada de micrófono para obtener una matriz de covarianza de micrófono. La matriz de covarianza de micrófono se introduce en un módulo de formulación de la matriz de mezcla 730. Asimismo, la matriz de covarianza de micrófono se utiliza para obtener un modelo de campo de sonido 710. El modelo de campo de sonido 710 puede estar basado sobre otras fuentes distintas de la matriz de covarianza.

30 **[0094]** La información de dirección y difuso basada en el modelo de campo de sonido luego se introduce en un módulo de formulación de matriz de covarianza de destino 720 para generar una matriz de covarianza de destino. El módulo de formulación de la matriz de covarianza de destino 720 luego introduce la matriz de covarianza de destino generada en el módulo de formulación de la matriz de mezcla 730.

35 **[0095]** El módulo de formulación de la matriz de mezcla 730 está adaptado para generar la matriz de mezcla e introduce la matriz de mezcla generada en un módulo de aplicación de la matriz de mezcla 740. El módulo de aplicación de la matriz de mezcla 740 está adaptado para aplicar la matriz de mezcla en la señal de entrada del micrófono para obtener una señal de salida del micrófono que tiene las propiedades de covarianza de destino. En una forma de realización, los módulos 720, 730 y 740 son submódulos de un procesador de señales.

40 **[0096]** Dicho aparato se ajusta al concepto de DirAC y SAM, que consiste en estimar la dirección y la difusión del campo de sonido original y crear dicha salida que reproduce mejor la dirección y la difusión estimadas. Este procedimiento de procesamiento de señales requiere grandes ajustes de la matriz de covarianza con el fin de proporcionar la imagen espacial correcta. El concepto procesado es la solución a lo anterior. El concepto propuesto  
45 **[0096]** implica la distancia de origen, la localización de origen y/o la separación de origen, confortabilidad de escucha y/o sonido envolvente.

50 **[0097]** La Fig. 8 ilustra un ejemplo que muestra una forma de realización para el mejoramiento ciego de la calidad de sonido espacial en la reproducción estéreo o multicanal. En el módulo 805 se lleva a cabo un análisis de la matriz de covarianza, por ejemplo, un análisis de ICC o nivel del contenido estéreo o multicanal. A continuación, una regla de mejoramiento se aplica en el módulo de mejoramiento 815, por ejemplo, para obtener las ICCs de salida de las ICCs de entrada. Un módulo de formulación de la matriz de mezcla 830 genera una matriz de mezcla sobre la base del análisis de la matriz de covarianza llevado a cabo por el módulo 805 y sobre la base de la información derivada de la aplicación de la regla de mejoramiento que se llevó a cabo en el módulo de mejoramiento 815. La matriz de mezcla  
55 **[0097]** se aplica entonces al contenido estéreo o multicanal en el módulo 840 para obtener el contenido estéreo o multicanal ajustado que tiene las propiedades de covarianza de destino.

60 **[0098]** En cuanto al sonido multicanal, por ejemplo, mezclas o grabaciones, es bastante común encontrar suboptimalidad en el sonido espacial, especialmente en términos de ICC demasiado elevada. Una consecuencia típica abarca calidad reducida con respecto a la anchura, sonido envolvente, distancia, separación de origen, localización de origen y/o estabilidad de origen y confortabilidad de escucha. Se ha probado de manera informal que el concepto es capaz de mejorar estas propiedades con los elementos que tienen innecesariamente ICCs elevadas. Las mejoras observadas son la anchura, la distancia de origen, la localización/separación de origen, sonido envolvente y confortabilidad de escucha.

**[0099]** La Fig. 9 ilustra otra forma de realización para el mejoramiento de las configuraciones estrechas de parlantes (por ejemplo, tablets, TV). El concepto propuesto es probablemente beneficioso como herramienta para mejorar la calidad estéreo en las configuraciones de reproducción en donde un ángulo del parlante es demasiado estrecho (por ejemplo, *tablets*). El concepto propuesto contempla lo siguiente:

- repanoramización de fuentes dentro del arco dado para que coincidan con una configuración más amplia de parlantes
- aumentar la ICC para adaptarse mejor a la de una configuración más amplia de parlantes
- proporcionar un mejor punto de partida para llevar a cabo la cancelación de diafonía, por ejemplo, utilizando cancelación de diafonía sólo cuando no hay forma directa de crear las señales binaurales deseadas.

**[0100]** Se esperan mejoras con respecto a la anchura y con respecto a la cancelación de diafonía regular, calidad de sonido y robustez.

**[0101]** En otro ejemplo de aplicación ilustrado por la Fig. 10, una forma de realización proporciona una reproducción óptima de Codificación de Audio Direccional (DirAC) sobre la base de una señal de micrófono de formato B.,

**[0102]** La forma de realización de la Fig. 10 se basa en el hallazgo de que las unidades de reproducción de DirAC del estado de las técnica basadas en señales de micrófonos coincidentes aplican la descorrelación en una medida innecesaria, por lo tanto, comprometen la calidad de audio. Por ejemplo, si el campo de sonido analizado es difuso se aplica correlación total en todos los canales, aunque un formato B proporcione ya tres componentes de sonido incoherentes en el caso de un campo de sonido horizontal (W, X, Y). Este efecto está presente en diversos grados excepto cuando la difusión es cero.

**[0103]** Asimismo, los sistemas descritos anteriormente que utilizan micrófonos virtuales no garantizan una matriz de covarianza de salida correcta (niveles y correlaciones entre canales), porque los micrófonos virtuales efectúan el sonido de modo diferente dependiendo del ángulo de origen, el posicionamiento del parlante y la difusión del campo de sonido.

**[0104]** El concepto propuesto resuelve ambos problemas. Existen dos alternativas: proporcionar canales decorrelacionados como canales de entrada adicionales (como en la figura siguiente); o utilizar un concepto de descorrelador-mezcla.

**[0105]** En la Fig. 10, un módulo 1005 lleva a cabo un análisis de la matriz de covarianza. Un módulo de formulación de la matriz de covarianza de destino 1018 no sólo toma en cuenta un modelo de campo de sonido, sino también una configuración de parlantes cuando formula una matriz de covarianza de destino. Asimismo, un módulo de formulación de la matriz de mezcla 1030 genera una matriz de mezcla no sólo sobre la base de un análisis de la matriz de covarianza y la matriz de covarianza de destino, sino también sobre la base de un criterio de optimización, por ejemplo, una matriz de mezcla de micrófonos de formato B a virtual proporcionada por un módulo 1032. El modelo de campo de sonido 1010 puede corresponder al modelo de campo de sonido 710 de la Fig. 7. El módulo de aplicación de la matriz de mezcla 1040 puede corresponder al módulo de aplicación de la matriz de mezcla 740 de la Fig. 7.

**[0106]** En un ejemplo de aplicación adicional se proporciona una forma de realización para el ajuste espacial en los procedimientos de conversión de canales, por ejemplo, mezcla descendente. La conversión de canales, por ejemplo, haciendo automática la mezcla descendente 5.1 de la pista de audio 22.2 incluye canales de colapso. Esto puede incluir una pérdida o cambio de la imagen espacial que puede ser abordado con el concepto propuesto. Una vez más, existen dos alternativas: La primera alternativa utiliza el concepto en el dominio del mayor número de canales pero definiendo canales con cero potencia para los canales que faltan del número inferior; la otra alternativa formula la disolución de la matriz directamente para los números de canales diferentes.

**[0107]** La Fig. 11 ilustra la Tabla 1, que proporciona ejemplos numéricos de los conceptos descritos anteriormente. Cuando una señal con covarianza  $\mathbf{C}_x$  se procesa con una matriz de mezcla  $\mathbf{M}$  y se complementa con una señal residual posible con  $\mathbf{C}_r$ , la señal de salida tiene la covarianza  $\mathbf{C}_y$ . A pesar de que estos ejemplos numéricos son estáticos, el caso de uso típico del procedimiento propuesto es dinámico. El orden de los canales se supone que es L, R, C, Ls, Rs, (Lr, Rr).

**[0108]** La Tabla 1 muestra un conjunto de ejemplos numéricos para ilustrar el comportamiento del concepto propuesto en algunos casos de uso previstos. Las matrices se formularon con el código Matlab proporcionado en el Listado 1. El Listado 1 se ilustra en la Fig. 12.

**[0109]** El Listado 1 de la Fig. 12 ilustra una aplicación de Matlab del concepto propuesto. El código Matlab se utilizó en los ejemplos numéricos y proporciona la funcionalidad general del concepto propuesto.

- 5 [0110] Aunque las matrices se ilustran estáticas, en aplicaciones típicas varían en tiempo y frecuencia. El criterio de diseño se cumple por definición ya que si una señal con la covarianza  $\mathbf{C}_x$  se procesa con una matriz de mezcla  $\mathbf{M}$  y se complementa con una señal residual posible con  $\mathbf{C}_r$ , la señal de salida tiene la covarianza  $\mathbf{C}_y$ .
- 10 [0111] La primera y la segunda fila de la tabla ilustran un caso de uso de mejoramiento estéreo por medio de la descorrelación de las señales. En la primera fila hay un componente incoherente pequeño pero razonable entre los dos canales y, por lo tanto, se logra una salida totalmente incoherente sólo con la mezcla de canales. En la segunda fila, la correlación de entrada es muy alta, por ejemplo, el componente principal más pequeño es muy pequeño. La
- 15 [0112] La tercera fila muestra un caso de estéreo a mezcla ascendente 5.0. En este ejemplo, la matriz de covarianza de destino se fija de modo que el componente incoherente de la mezcla estéreo es distribuido por igual e incoherentemente a los parlantes laterales y traseros y el componente coherente se coloca en el parlante central. La señal residual es nuevamente distinta de cero dado que la dimensión de la señal aumenta.
- 20 [0113] La cuarta fila muestra un caso de mezcla ascendente simple de 5.0 a 7.0 en donde los dos canales traseros originales se mezclan incoherentemente en forma ascendente a los cuatro canales traseros nuevos. Este ejemplo ilustra que el procesamiento se enfoca en aquellos canales en donde se requieren ajustes.
- 25 [0114] La quinta fila representa un caso de mezcla descendente de una señal 5.0 a estéreo. La mezcla descendente pasiva, tal como la aplicación de una matriz de mezcla descendente estática  $\mathbf{Q}$ , amplificaría los componentes coherentes en comparación con los componentes incoherentes. Aquí, la matriz de covarianza de destino se definió para conservar la potencia que se cumple por la  $\mathbf{M}$  resultante.
- 30 [0115] La sexta y séptima fila ilustran el caso de uso de microfónica espacial coincidente. Las matrices de covarianza de entrada  $\mathbf{C}_x$  son el resultado de la colocación de micrófonos coincidentes ideales de primer orden en un campo difuso ideal. En la sexta fila, los ángulos entre los micrófonos son iguales, y en la séptima fila los micrófonos están orientados hacia ángulos estándares de una configuración 5.0. En ambos casos, los valores grandes fuera de la diagonal de  $\mathbf{C}_x$  ilustran la desventaja inherente de las técnicas pasivas de los micrófonos coincidentes de primer orden en el caso ideal, en donde la matriz de covarianza que representa mejor un campo difuso es diagonal, y esto se fijó por lo tanto como el objetivo. En ambos casos, la relación de la potencia correlacionada resultante sobre toda la energía es exactamente 2/5. Esto se debe a que hay tres componentes de señales independientes disponibles en las señales de los micrófonos coincidentes horizontales de primer orden, y dos son para agregar con el fin de alcanzar la matriz de covarianza diagonal de destino de cinco canales.
- 35 [0116] Se ha identificado que la percepción espacial en la reproducción estéreo y multicanal depende especialmente de la matriz de covarianza de la señal en las bandas de frecuencia perceptualmente relevantes.
- 40 [0117] Se ha presentado un concepto para controlar la matriz de covarianza de una señal mediante la mezcla cruzada óptima de los canales. Se han presentado medios para inyectar la potencia descorrelacionada en casos en que sea necesario ya que no se dispone de componentes de señales suficientemente independientes de potencia razonable.
- 45 [0118] Se ha descubierto que el concepto es robusto en su propósito y se ha identificado una amplia variedad de aplicaciones probables.
- 50 [0119] A continuación se presentan formas de realización que describen cómo generar  $\mathbf{C}_y$  sobre la base de  $\mathbf{C}_x$ . A modo de primer ejemplo se considera la mezcla ascendente estéreo a 5.0. En cuanto a la mezcla ascendente estéreo a 5.0,  $\mathbf{C}_x$  es una matriz 2x2 y  $\mathbf{C}_y$  es una matriz 5x5 (en este ejemplo, el canal subwoofer (parlante para graves) no se considera). Los pasos para generar  $\mathbf{C}_y$  sobre la base de  $\mathbf{C}_x$  en cada baldosa de tiempo-frecuencia, en el contexto de mezcla ascendente pueden ser, por ejemplo, los siguientes:
- 55 1. Estimar la potencia ambiental y directa en el canal izquierdo y derecho. El ambiente se caracteriza por un componente incoherente entre los canales que tiene igual potencia en ambos canales. La potencia directa es el resto cuando la porción de potencia ambiental se extrae de la potencia total, por ejemplo, el componente de potencia coherente, posiblemente con distintas energías en los canales izquierdo y derecho.
- 60 2. Estimar un ángulo del componente directo. Esto se lleva a cabo utilizando una panoramización de amplitud de ley inversa. Hay una relación de panoramización de amplitud en el componente directo, y sólo hay un ángulo entre los parlantes delanteros que corresponde a la misma.
3. Generar una matriz de 5x5 de ceros como  $\mathbf{C}_y$ .

4. Colocar la cantidad de potencia directa en la diagonal de  $\mathbf{C}_y$  correspondiente a los dos parlantes más próximos de la dirección analizada. La distribución de la potencia entre éstos puede adquirirse por las leyes de panoramización de amplitud. La panoramización de amplitud es coherente, por lo tanto agregar a la no diagonal correspondiente la raíz cuadrada del producto de las potencias de los dos canales.
5. Agregar a la diagonal de  $\mathbf{C}_y$ , correspondiente a los canales L, R, Ls y Rs, la cantidad de potencia que corresponde a la potencia del componente ambiental. La distribución equitativa es una buena opción. Ahora se obtiene la  $\mathbf{C}_y$  de destino.
- [0120]** A modo de otro ejemplo se considera la mejora. El objetivo es aumentar las calidades perceptuales tales como anchura o sonido envolvente mediante el ajuste de la coherencia entre canales a cero. Aquí se proporcionan dos ejemplos diferentes en dos formas para llevar a cabo la mejora. En la primera forma se selecciona un caso de uso del mejoramiento de estéreo, por lo que  $\mathbf{C}_x$  y  $\mathbf{C}_y$  son matrices de  $2 \times 2$ . Los pasos son los siguientes:
1. Formular la ICC (el valor normalizado de covarianza entre -1 y 1, por ejemplo con la fórmula proporcionada).
  2. Ajustar la ICC por una función. Por ejemplo,  $ICC_{nueva} = \text{signo}(ICC) * ICC^2$ . Este es un ajuste muy leve. O  $ICC_{nueva} = \text{signo}(ICC) * \max(0, \text{abs}(ICC) * 10 - 9)$ . Este es un ajuste mayor.
  3. Formular  $\mathbf{C}_y$  de manera que los valores de la diagonal sean los mismos que en  $\mathbf{C}_x$ , pero el valor no diagonal se formula utilizando  $ICC_{nueva}$ , con la misma fórmula que en el paso 1, pero a la inversa.
- [0121]** En la perspectiva anterior, la señal residual no es necesaria, debido a que el ajuste de la ICC está diseñado para que el sistema no requiera una gran amplificación de pequeños componentes de la señal.
- [0122]** El segundo tipo de aplicación del procedimiento en este caso de uso es el siguiente. Uno tiene una señal de entrada del canal N, de modo que  $\mathbf{C}_x$  y  $\mathbf{C}_y$  son matrices  $N \times N$ .
1. Formular  $\mathbf{C}_y$  a partir de  $\mathbf{C}_x$  fijando simplemente los valores de la diagonal en  $\mathbf{C}_y$  igual que en  $\mathbf{C}_x$ , y los valores no-diagonales a cero.
  2. Habilitar el procedimiento de compensación de ganancia en el procedimiento propuesto, en lugar de utilizar los residuales. La regularización en la inversa de  $\mathbf{K}_x$  controla que el sistema sea estable. La compensación de la ganancia se encarga de que las potencias se conserven.
- [0123]** Las dos formas descritas para llevar a cabo el mejoramiento proporcionan resultados similares. Esta última es más fácil de aplicar en el caso de uso de canales múltiples.
- [0124]** Por último, como tercer ejemplo, se considera el modelo Directo/de difusión, por ejemplo la Codificación de Audio Direccional (DirAC).
- [0125]** La DirAC y también los Micrófonos de Audio Espacial (SAM) proporcionan una interpretación de un campo de sonido con dirección de parámetros y difusión. La dirección es el ángulo de llegada del componente de sonido directo. La difusión es un valor entre 0 y 1, el cual proporciona información sobre la magnitud de la difusión de la potencia de sonido total, por ejemplo, que se supone que llega incoherentemente desde todas las direcciones. Esta es una aproximación del campo de sonido, pero cuando se aplica en bandas de frecuencias perceptuales, proporciona una representación perceptualmente buena del campo de sonido. La dirección, difusión, y la potencia total del campo de sonido conocidos se suponen en una baldosa de tiempo-frecuencia. Éstos se formulan utilizando la información en la matriz de covarianza del micrófono  $\mathbf{C}_x$ . Se obtiene una configuración de parlantes de canal N. Los pasos para generar  $\mathbf{C}_y$  son similares a la mezcla ascendente, de acuerdo con lo siguiente:
1. Generar una matriz de  $N \times N$  de ceros como  $\mathbf{C}_y$ .
  2. Colocar la cantidad de potencia directa, que es  $(1 - \text{difusión}) * \text{potencia total}$  en la diagonal de  $\mathbf{C}_y$  correspondiente a los dos parlantes más próximos de la dirección analizada. La distribución de la potencia entre éstos puede adquirirse por las leyes de panoramización de amplitud. La panoramización de amplitud es coherente, por lo tanto agregar a la no diagonal correspondiente una raíz cuadrada de los productos de las potencias de los dos canales.
  3. Distribuir en la diagonal de  $\mathbf{C}_y$  la cantidad de potencia de difusión, que es la potencia de difusión total \*. La distribución puede llevarse a cabo, por ejemplo, de modo que se coloque más potencia en aquellas direcciones en donde los parlantes son escasos. Ahora se obtiene la  $\mathbf{C}_y$  de destino.

- 5 **[0126]** Aunque algunos aspectos han sido descritos en el contexto de un aparato, es evidente que estos aspectos representan también una descripción del procedimiento correspondiente, en donde un bloque o dispositivo corresponde a un paso del procedimiento o una característica de un paso del procedimiento. En forma análoga, los aspectos descritos en el contexto de un paso del procedimiento también representan una descripción de un bloque o elemento o característica correspondiente de un aparato respectivo.
- 10 **[0127]** Dependiendo de determinados requisitos de implementación, las formas de realización de la invención se pueden implementar en hardware o en software. La implementación se puede llevar a cabo utilizando un medio de almacenamiento digital, por ejemplo un disco flexible, un DVD, un CD, una memoria ROM, una memoria PROM, una memoria EPROM, una memoria EEPROM o una memoria FLASH, que tienen señales de control de lectura electrónica almacenadas en los mismos, cuyas señales cooperan (o son capaces de cooperar) con un sistema de computación programable de forma tal que el procedimiento respectivo se realice.
- 15 **[0128]** Algunas formas de realización de acuerdo con la invención comprenden un portador de datos que tiene señales de control de lectura electrónica, las cuales son capaces de cooperar con un sistema de computadora programable, de tal manera que uno de los procedimientos descritos aquí se realice.
- 20 **[0129]** En general, las formas de realización de la presente invención se pueden implementar como un producto de programa informático con un código de programa, cuyo código de programa es operativo para llevar a cabo uno de los procedimientos cuando el producto de programa informático se ejecuta en una computadora. El código del programa se puede almacenar, por ejemplo, en un portador legible por computadora.
- 25 **[0130]** Otras formas de realización comprenden el programa informático para llevar a cabo uno de los procedimientos descritos en la presente, almacenados en un portador legible por computadora o un medio de almacenamiento no transitorio.
- 30 **[0131]** En otras palabras, una forma de realización del procedimiento de la invención es, por lo tanto, un programa informático que tiene un código de programa para llevar a cabo uno de los procedimientos descritos en la presente, cuando el programa informático se ejecuta en una computadora.
- 35 **[0132]** Otra forma de realización de los procedimientos de la invención es, por lo tanto, un portador de datos (o un medio de almacenamiento digital, o un medio legible por computadora) que comprende, grabado en el mismo, el programa informático para llevar a cabo uno de los procedimientos descritos en la presente.
- 40 **[0133]** Otra forma de realización del procedimiento de la invención es, por lo tanto, una corriente de datos o una secuencia de señales que representan el programa informático para llevar a cabo uno de los procedimientos descritos en la presente. La corriente de datos o la secuencia de señales, por ejemplo, pueden estar configuradas para ser transferidas a través de una conexión de comunicación de datos, por ejemplo, a través de Internet.
- 45 **[0134]** Una forma de realización adicional comprende un medio de procesamiento, por ejemplo, una computadora, o un dispositivo lógico programable, configurado o adaptado para llevar a cabo uno de los procedimientos descritos en la presente.
- 50 **[0135]** Otra forma de realización comprende una computadora que tiene el programa informático instalado en la misma para llevar a cabo uno de los procedimientos descritos en la presente.
- 55 **[0136]** En algunas formas de realización, un dispositivo lógico programable (por ejemplo, un arreglo de puerta programable de campo) se puede utilizar para llevar a cabo algunas o todas las funcionalidades de los procedimientos descritos en la presente. En algunas formas de realización, un arreglo de puerta programable de campo puede cooperar con un microprocesador para llevar a cabo uno de los procedimientos descritos en la presente. En general, los procedimientos serán llevados a cabo, preferentemente, por cualquier aparato de hardware.
- 60 **[0137]** Las formas de realización anteriormente descritas son simplemente ilustrativas de los principios de la presente invención. Se entiende que las modificaciones y variaciones de los arreglos y los detalles descritos en la presente serán evidentes para otros expertos en la materia. Es la intención, por lo tanto, de que la invención esté limitada solamente por el alcance de las reivindicaciones inminentes de la patente y no por los detalles específicos presentados a modo de descripción y explicación de las formas de realización de la presente.
- 60 Bibliografía:
- [0138]**

- [1] C. Faller, "Multiple-Loudspeaker Playback of Stereo Signals", Journal of the Audio Engineering Society, Vol. 54, N° 11, pp. 1051-1064, junio de 2006.
- 5 [2] V. Pulkki, "Spatial Sound Reproduction with Directional Audio Coding", Journal of the Audio Engineering Society, Vol. 55, N° 6, pp. 503-516, junio de 2007.
- [3] C. Tournery, C. Faller, F. KÜch, J. Herre, "Converting Stereo Microphone Signals Directly to MPEG Surround", 128th AES Convention, mayo de 2010.
- 10 [4] J. Breebaart, S. van de Par, A. Kohlrausch and E. Schuijers, "Parametric Coding of Stereo Audio", EURASIP Journal on Applied Signal Processing, Vol. 2005, N° 9, pp. 1305-1322, 2005.
- [5] J. Herre, K. Kjörling, J. Breebaart, C. Faller, S. Disch, H. Purnhagen, J. Koppens, J. Hilpert, J. Rödén, W. Oomen, K. Linzmeier and K. S. Chong, "MPEG Surround – The ISO/MPEG Standard for Efficient and Compatible Multichannel Audio Coding", Journal of the Audio Engineering Society, Vol. 56, N° 11, pp. 932-955, noviembre de 15 2008.
- [6] J. Vilkamo, V. Pulkki, "Directional Audio Coding: Virtual Microphone-Based Synthesis and Subjective Evaluation", Journal of the Audio Engineering Society, Vol. 57, N° 9, pp. 709-724, septiembre de 2009.
- 20 [7] Golub, G.H. y Van Loan, C.F., "Matrix computations", Johns Hopkins Univ Press, 1996.
- [8] R. Rebonato, P. Jäckel, "The most general methodology to create a valid correlation matrix for risk management and option pricing purposes", Journal of Risk, Vol. 2, N° 2, pp. 17-28, 2000.
- 25

**REIVINDICACIONES**

1. Un aparato para generar una señal de salida de audio que tiene dos o más canales de salida de audio a partir de una señal de entrada de audio que tiene dos o más canales de entrada de audio, el cual comprende:  
 5 un proveedor (110) para proporcionar primeras propiedades de covarianza de la señal de entrada de audio, y un procesador de señales (120) para generar la señal de salida de audio mediante la aplicación de una regla de mezcla en por lo menos dos de los dos o más canales de entrada de audio,  
 10 en el que el procesador de señales (120) está configurado para determinar la regla de mezcla sobre la base de las primeras propiedades de covarianza de la señal de entrada de audio y sobre la base de las segundas propiedades de covarianza de la señal de salida de audio, en donde las segundas propiedades de covarianza son distintas de las primeras propiedades de covarianza.

2. Un aparato de acuerdo con la reivindicación 1, en el cual el proveedor (110) está adaptado para proporcionar las primeras propiedades de covarianza, en donde las primeras propiedades de covarianza tienen un primer estado para un primer intervalo de tiempo-frecuencia, y en donde las primeras propiedades de covarianza tienen un segundo estado, diferente del primer estado, para un segundo intervalo de tiempo-frecuencia, diferente del primer intervalo de tiempo-frecuencia.

3. Un aparato de acuerdo con la reivindicación 1 o 2, en el que el procesador de señales (120) está adaptado para determinar la regla de mezcla sobre la base de las segundas propiedades de covarianza, en donde las segundas propiedades de covarianza tienen un tercer estado para un tercer intervalo de tiempo-frecuencia, y en donde las segundas propiedades de covarianza tienen un cuarto estado, que es diferente del tercer estado para un cuarto intervalo de tiempo-frecuencia, que es diferente del tercer intervalo de tiempo-frecuencia.

4. Un aparato de acuerdo con una de las reivindicaciones anteriores, en el cual el procesador de señales (120) está adaptado para generar la señal de salida de audio mediante la aplicación de la regla de mezcla de modo que cada uno de los dos o más canales de salida de audio depende de cada uno de los dos o más canales de entrada de audio.

5. Un aparato de acuerdo con una de las reivindicaciones anteriores, en el cual el procesador de señales (120) está adaptado para determinar la regla de mezcla de modo que una medida de error se reduce al mínimo.

6. Un aparato de acuerdo con la reivindicación 5, en el cual el procesador de señales (120) está adaptado para determinar la regla de mezcla de modo que la regla de mezcla depende de

$$\|y_{ref} - y\|^2$$

en el cual

$$y_{ref} = Qx ,$$

en donde x es la señal de entrada de audio, en donde Q es una matriz de mapeo, y en donde y es la señal de salida de audio.

7. Un aparato de acuerdo con una de las reivindicaciones anteriores, en el cual el procesador de señales (120) está configurado para determinar la regla de mezcla mediante la determinación de las segundas propiedades de covarianza, en donde el procesador de señales (120) está configurado para determinar las segundas propiedades de covarianza sobre la base de las primeras propiedades de covarianza.

8. Un aparato de acuerdo con una de las reivindicaciones anteriores, en el cual el procesador de señales (120) está adaptado para determinar una matriz de mezcla como la regla de mezcla, en donde el procesador de señales (120) está adaptado para determinar la matriz de mezcla sobre la base de las primeras propiedades de covarianza y sobre la base de las segundas propiedades de covarianza.

9. Un aparato de acuerdo con una de las reivindicaciones anteriores, en el cual el proveedor (110) está adaptado para proporcionar las primeras propiedades de covarianza mediante la determinación de una primera matriz de covarianza de la señal de entrada de audio y en donde el procesador de señales (120) está configurado para determinar la regla de mezcla sobre la base de una segunda matriz de covarianza de la señal de salida de audio como las segundas propiedades de covarianza.

10. Un aparato de acuerdo con la reivindicación 9, en el cual el proveedor (110) está adaptado para determinar la primera matriz de covarianza, de modo que cada valor de la diagonal de la primera matriz de covarianza indica

una potencia de uno de los canales de entrada de audio, y de modo que cada valor de la primera matriz de covarianza, que no es un valor de diagonal indica una correlación entre canales entre un primer canal de entrada de audio y un segundo canal de entrada de audio diferente.

5 11. Un aparato de acuerdo con la reivindicación 9 o 10, en el cual el procesador de señales (120) está configurado para determinar la regla de mezcla sobre la base de la segunda matriz de covarianza, en donde cada valor de la diagonal de la segunda matriz de covarianza indica una potencia de uno de los canales de salida de audio, y en donde cada valor de la segunda matriz de covarianza, que no es un valor de la diagonal, indica una correlación entre canales entre un primer canal de salida de audio y un segundo canal de salida de audio.

10 12. Un aparato de acuerdo con una de las reivindicaciones anteriores, en el cual el procesador de señales (120) está adaptado para determinar una matriz de mezcla como regla de mezcla, en donde el procesador de señales (120) está adaptado para determinar la matriz de mezcla sobre la base de las primeras propiedades de covarianza y sobre la base de las segundas propiedades de covarianza, en donde el proveedor (110) está adaptado para proporcionar las primeras propiedades de covarianza mediante la determinación de una primera matriz de covarianza de la señal de entrada de audio, y en donde el procesador de señales (120) está configurado para determinar la regla de mezcla sobre la base de una segunda matriz de covarianza de la señal de salida de audio como las segundas propiedades de covarianza, en donde el procesador de señales (120) está adaptado para determinar la matriz de mezcla, de modo que:

$$\mathbf{M} = \mathbf{K}_y \mathbf{P} \mathbf{K}_x^{-1},$$

de modo que

$$\begin{aligned} \mathbf{K}_x \mathbf{K}_x^T &= \mathbf{C}_x, \\ \mathbf{K}_y \mathbf{K}_y^T &= \mathbf{C}_y \end{aligned}$$

25 en donde  $\mathbf{M}$  es la matriz de mezcla, en donde  $\mathbf{C}_x$  es la primera matriz de covarianza, en donde  $\mathbf{C}_y$  es la segunda matriz de covarianza, en donde  $\mathbf{K}_x^T$  es una primera matriz transpuesta de una primera matriz descompuesta  $\mathbf{K}_x$ , en donde  $\mathbf{K}_y^T$  es una segunda matriz transpuesta de una segunda matriz descompuesta  $\mathbf{K}_y$ , en donde  $\mathbf{K}_x^{-1}$  es una matriz inversa de la primera matriz descompuesta  $\mathbf{K}_x$ , y en donde  $\mathbf{P}$  es una primera matriz unitaria.

35 13. Un aparato de acuerdo con la reivindicación 12, en el cual el procesador de señales (120) está adaptado para determinar la regla de mezcla de modo que

$$\mathbf{M} = \mathbf{K}_y \mathbf{P} \mathbf{K}_x^{-1},$$

en el cual

$$\mathbf{P} = \mathbf{V} \mathbf{\Lambda} \mathbf{U}^T,$$

40 en donde  $\mathbf{U}^T$  es una tercera matriz transpuesta de una segunda matriz unitaria  $\mathbf{U}$ , en donde  $\mathbf{V}$  es una tercera matriz unitaria, en donde  $\mathbf{\Lambda}$  es una matriz de identidad añadida con ceros, en donde

$$\mathbf{U} \mathbf{S} \mathbf{V}^T = \mathbf{K}_x^T \mathbf{Q}^T \mathbf{K}_y,$$

en donde  $\mathbf{Q}^T$  es una cuarta matriz transpuesta de la matriz de mapeo  $\mathbf{Q}$ ,

50 en donde  $\mathbf{V}^T$  es una quinta matriz transpuesta de la tercera matriz unitaria  $\mathbf{V}$ , y en donde  $\mathbf{S}$  es una matriz de la diagonal.

55 14. Un aparato de acuerdo con la reivindicación 1, en el cual el procesador de señales (120) está adaptado para determinar una matriz de mezcla como la regla de mezcla, en donde el procesador de señales (120) está adaptado para determinar la matriz de mezcla sobre la base de las primeras propiedades de covarianza y sobre la base de las segundas propiedades de covarianza, en el cual el proveedor (110) está adaptado para proporcionar las primeras propiedades de covarianza mediante la determinación de la primera matriz de covarianza de la señal de entrada de audio, y

en el cual el procesador de señales (120) está configurado para determinar la regla de mezcla sobre la base de una segunda matriz de covarianza de la señal de salida de audio como las segundas propiedades de covarianza.

5 en el cual el procesador de señales (120) está adaptado para determinar la regla de mezcla mediante la modificación de por lo menos algunos valores de la diagonal de una matriz de la diagonal  $\mathbf{S}_x$  cuando los valores de la matriz de la diagonal  $\mathbf{S}_x$  son cero o menores que un valor umbral, de modo que los valores son mayores que o iguales al valor umbral, en el cual la matriz de la diagonal depende de la primera matriz de covarianza.

10 15. Un aparato de acuerdo con la reivindicación 14, en el cual el procesador de señales (120) está configurado para modificar dichos por lo menos algunos valores de la diagonal de la matriz de la diagonal  $\mathbf{S}_x$ , en donde

$\mathbf{K}_x = \mathbf{U}_x \mathbf{S}_x \mathbf{V}_x^T$ , y en donde  $\mathbf{C}_x = \mathbf{K}_x \mathbf{K}_x^T$ , en donde  $\mathbf{C}_x$  es la primera matriz de covarianza, en donde  $\mathbf{S}_x$  es la matriz de la diagonal, en donde  $\mathbf{U}_x$  es una segunda matriz,  $\mathbf{V}_x^T$  es una tercera matriz transpuesta, y en donde  $\mathbf{K}_x^T$  es una cuarta matriz transpuesta de la quinta matriz  $\mathbf{K}_x$ , y en donde  $\mathbf{V}_x$  y  $\mathbf{U}_x$  son matrices unitarias.

15 16. Un aparato de acuerdo con la reivindicación 14 o 15, en el cual el procesador de señales (120) está adaptado para generar la señal de salida de audio mediante la aplicación de la matriz de mezcla en por lo menos dos de los dos o más canales de entrada de audio para obtener una señal intermedia y mediante la adición de una señal residual  $r$  a la señal intermedia para obtener la señal de salida de audio.

20 17. Un aparato de acuerdo con la reivindicación 14 o 15, en el cual el procesador de señales (120) está adaptado para determinar la matriz de mezcla sobre la base de una matriz de ganancia de la diagonal  $\mathbf{G}$  y una matriz intermedia  $\hat{\mathbf{M}}$ , de modo que  $\mathbf{M}' = \mathbf{G}\hat{\mathbf{M}}$ , en donde la matriz de ganancia de la diagonal tiene el valor

$$25 \quad \mathbf{G}(i, i) = \sqrt{\frac{\mathbf{C}_y(i, i)}{\hat{\mathbf{C}}_y(i, i)}}$$

$$\text{en donde } \hat{\mathbf{C}}_y = \hat{\mathbf{M}}\mathbf{C}_x\hat{\mathbf{M}}^T,$$

en donde  $\mathbf{M}'$  es la matriz de mezcla, en donde  $\mathbf{G}$  es la matriz de ganancia de la diagonal, en donde  $\mathbf{C}_y$  es la segunda matriz de covarianza y en donde  $\hat{\mathbf{M}}^T$  es una quinta matriz transpuesta de la matriz intermedia  $\hat{\mathbf{M}}$ .

30 18. Un aparato de acuerdo con la reivindicación 1, en el cual el procesador de señales (120) comprende: un módulo de formulación de la matriz de mezcla (420; 530; 630; 730; 830; 1030) para generar una matriz de mezcla como la regla de mezcla sobre la base de las primeras propiedades de covarianza, y un módulo de aplicación de la matriz de mezcla (430; 540; 640; 740; 840; 1040) para aplicar la matriz de mezcla en la señal de entrada de audio para generar la señal de salida de audio.

35 19. Un aparato de acuerdo con la reivindicación 18, en el cual el proveedor (110) comprende un módulo de análisis de la matriz de covarianza (410; 705; 805; 1005) para proporcionar las propiedades de covarianza de entrada de la señal de entrada de audio para obtener un resultado de análisis como las primeras propiedades de covarianza, y en el cual el módulo de formulación de la matriz de mezcla (420; 530; 630; 730; 830; 1030) está adaptado para generar la matriz de mezcla sobre la base del resultado del análisis.

40 20. Un aparato de acuerdo con la reivindicación 18 o 19, en el cual el módulo de formulación de la matriz de mezcla (420; 530; 630; 730; 830; 1030) está adaptado para generar la matriz de mezcla sobre la base de un criterio de error.

45 21. Un aparato de acuerdo con una de las reivindicaciones 18 a 20, en el cual el procesador de señales (120) comprende además un módulo de determinación de datos espaciales (520; 620) para determinar los datos de información de configuración que comprenden los datos espaciales de sonido envolvente, los datos de la correlación entre canales o los datos del nivel de la señal de audio, y en el cual el módulo de formulación de la matriz de mezcla (420; 530; 630; 730; 830; 1030) está adaptado para generar la matriz de mezcla sobre la base de los datos de la información de configuración.

22. Un aparato de acuerdo con una de las reivindicaciones 18 a 20, en el cual el procesador de señales (120) comprende además un módulo de formulación de la matriz de covarianza de destino (730; 1018) para generar una matriz de covarianza de destino sobre la base del resultado del análisis, y  
5 en el cual el módulo de formulación de la matriz de mezcla (420; 530; 630; 730; 830; 1030) está adaptado para generar una matriz de mezcla sobre la base de la matriz de covarianza de destino.
23. Un aparato de acuerdo con la reivindicación 22, en el cual el módulo de formulación de la matriz de covarianza de destino (1018) está configurado para generar la matriz de covarianza de destino sobre la base de una  
10 configuración de parlante.
24. Un aparato de acuerdo con la reivindicación 18 a 19, en el cual el procesador de señales (120) comprende además un módulo de mejoramiento (815) para la obtención de datos de la correlación entre canales de salida sobre la base de los datos de la correlación entre canales de entrada, diferentes de los datos de la correlación entre  
15 canales de entrada, y en el cual el módulo de formulación de la matriz de mezcla (420; 530; 630; 730; 830; 1030) está adaptado para generar la matriz de mezcla sobre la base de los datos de la correlación entre canales de salida.
25. Un procedimiento para generar una señal de salida de audio que tiene dos o más canales de salida de audio a partir de una señal de entrada de audio que tiene dos o más canales de entrada de audio, el cual  
20 comprende:  
proporcionar primeras propiedades de covarianza de la señal de entrada de audio, y  
generar la señal de salida de audio mediante la aplicación de una regla de mezcla en por lo menos dos de los dos o  
25 más canales de entrada de audio,  
en el cual la regla de mezcla está determinada sobre la base de las primeras propiedades de covarianza de la señal de entrada de audio y sobre la base de las segundas propiedades de covarianza de la señal de salida de audio,  
distintas de las primeras propiedades de covarianza.
26. Un programa de computadora adaptado para aplicar el procedimiento de acuerdo con la reivindicación 25  
30 cuando se ejecuta en una computadora o procesador.

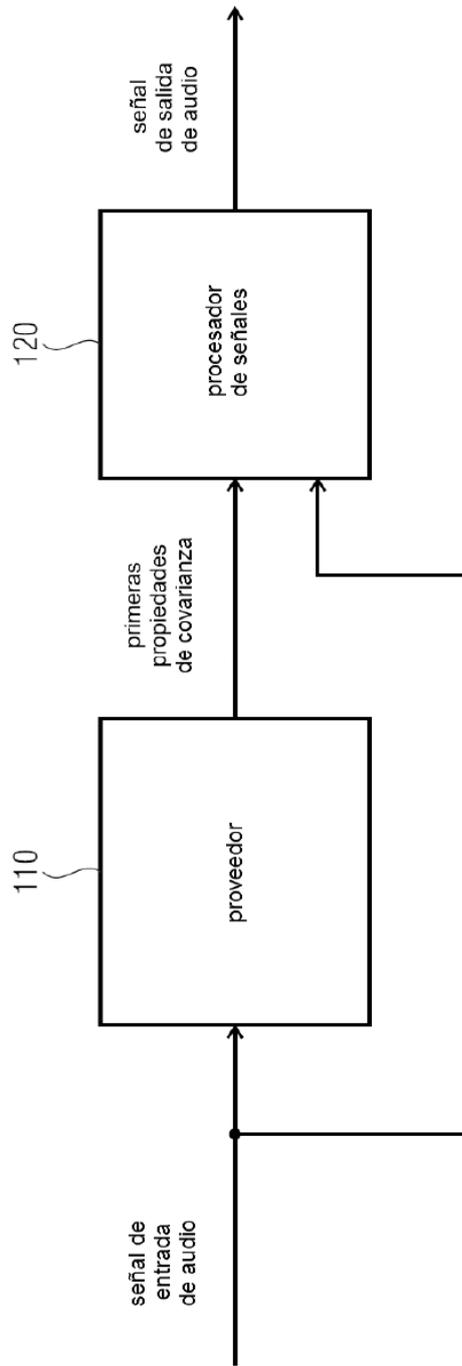


FIG 1

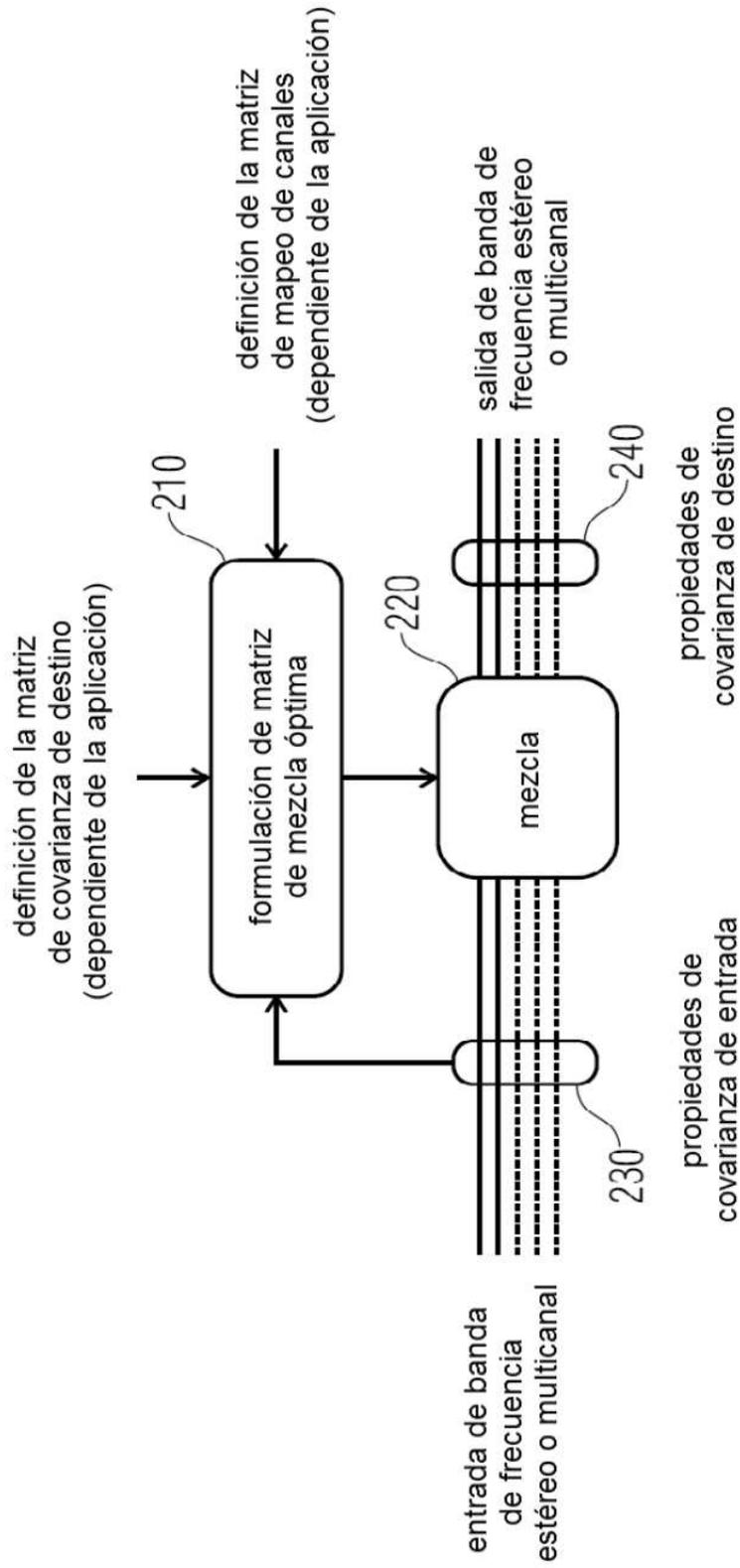


FIG 2

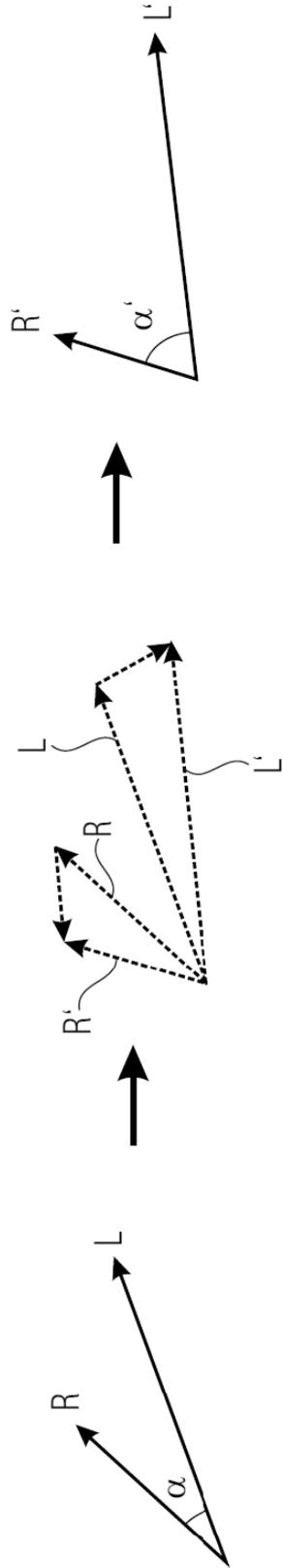


FIG 3

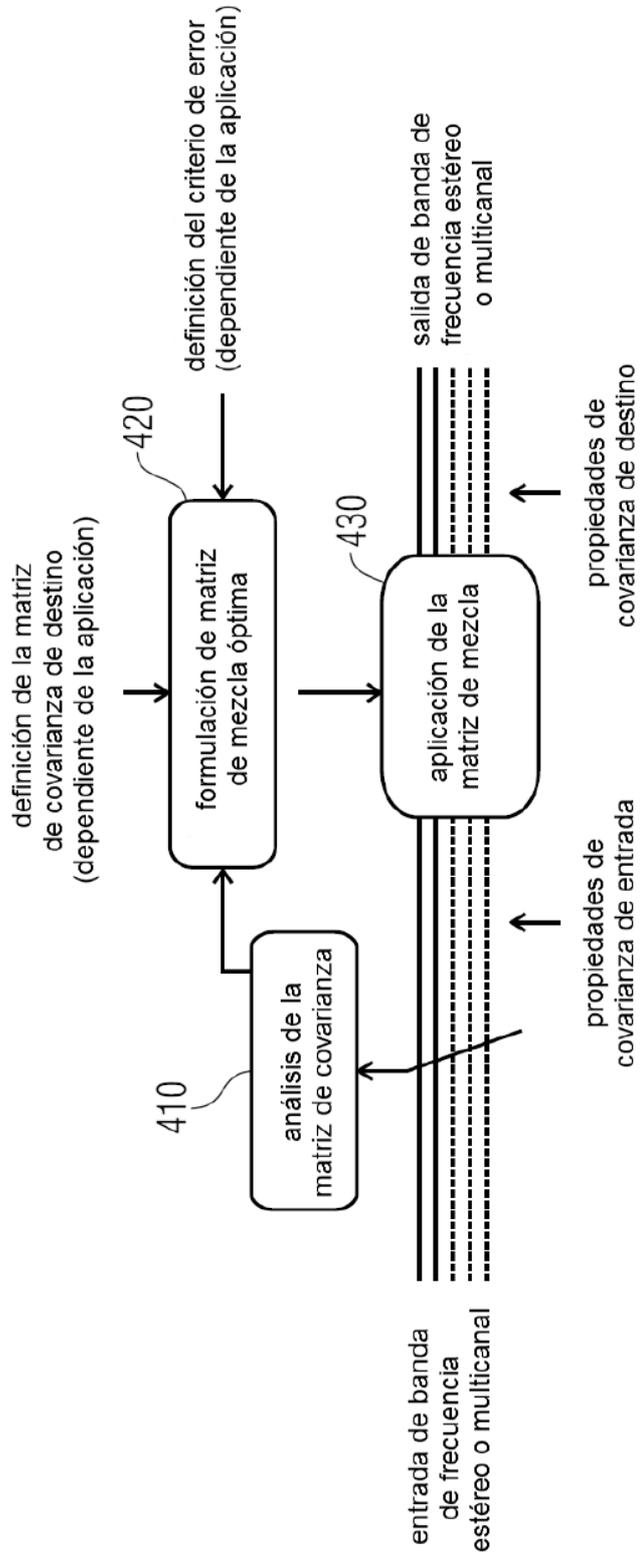


FIG 4

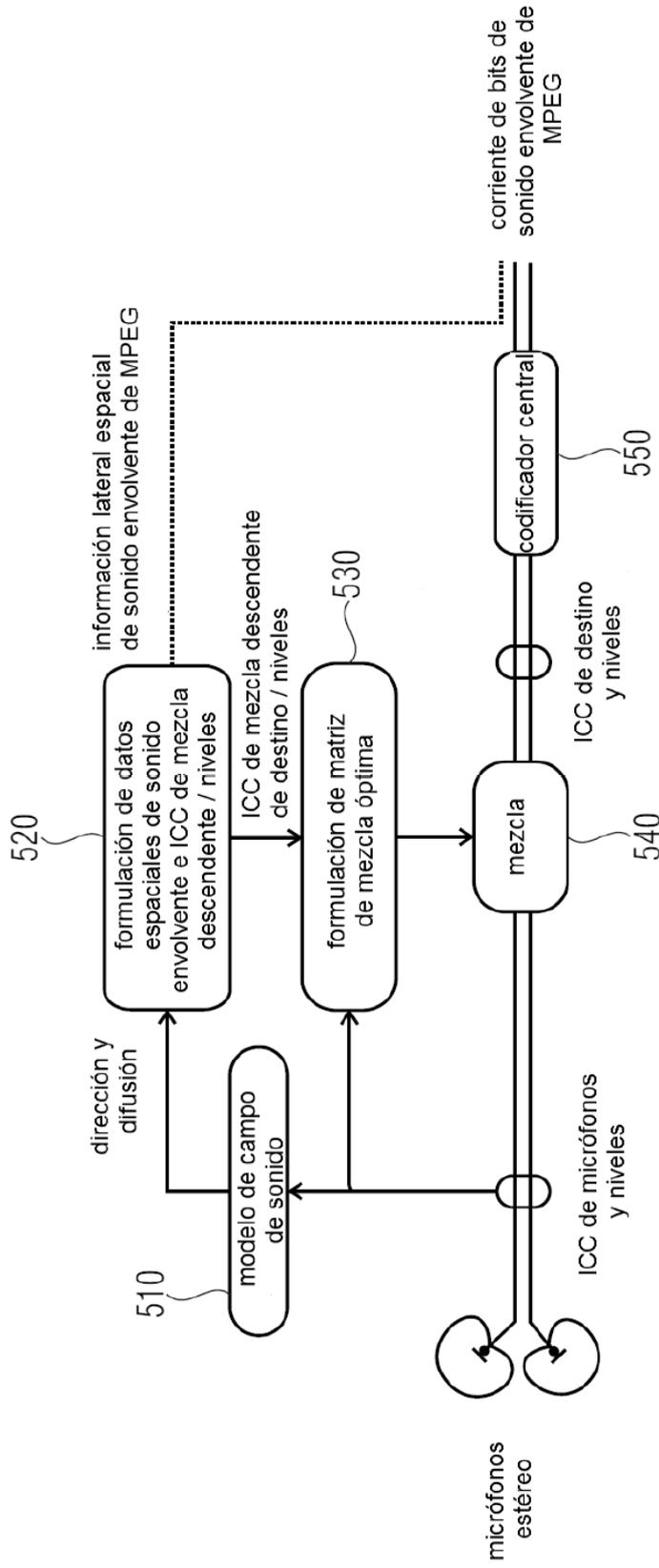


FIG 5

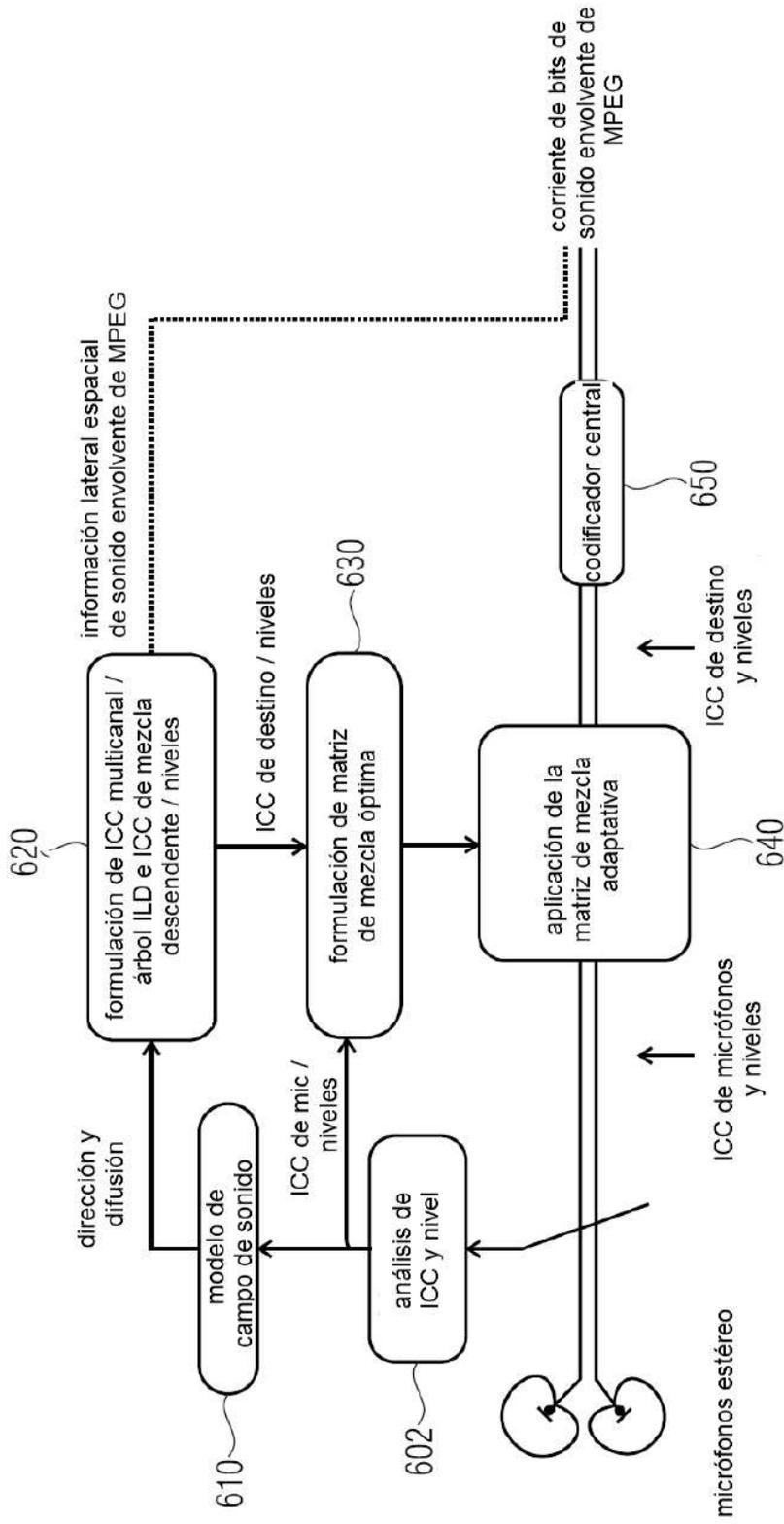


FIG 6

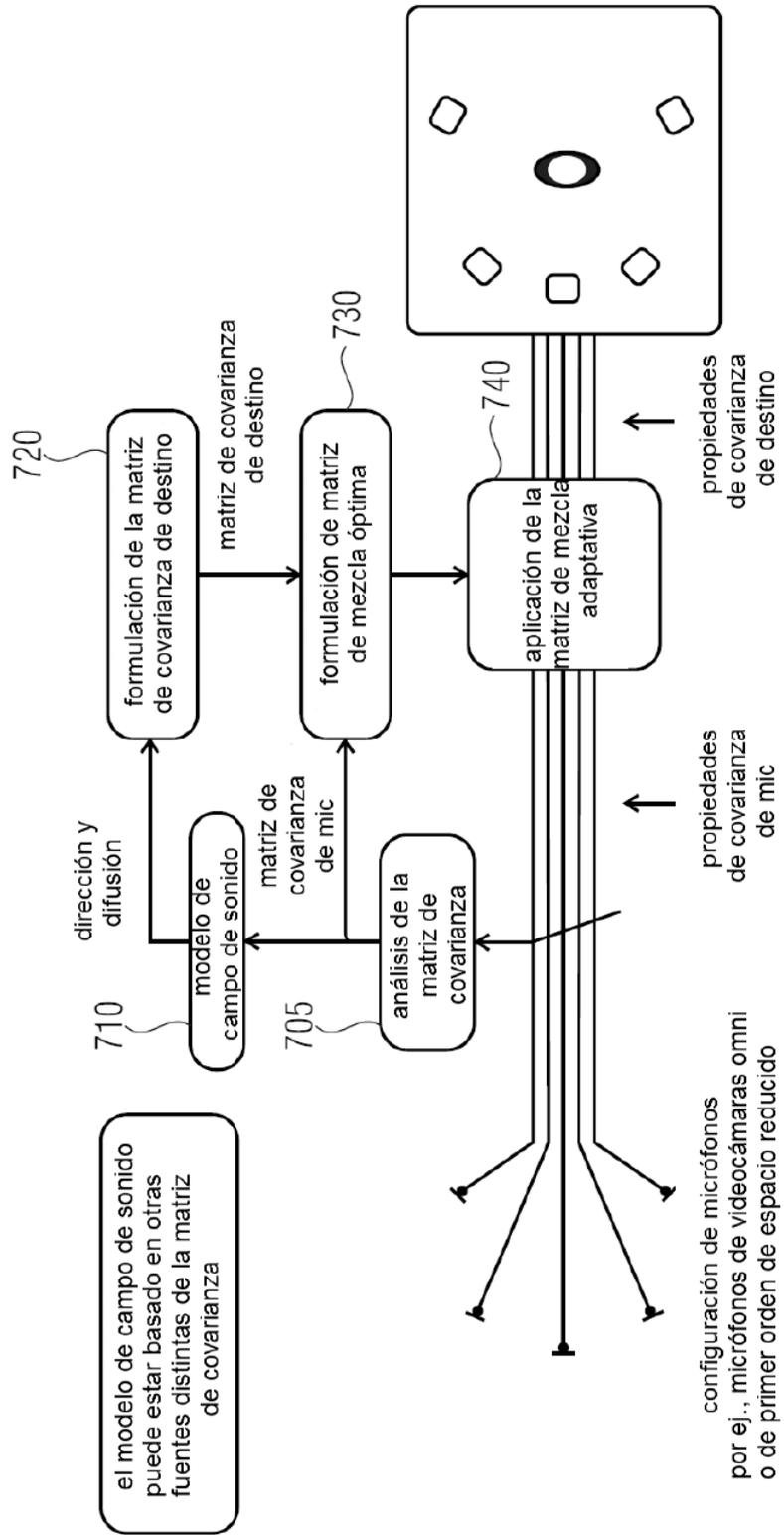


FIG 7

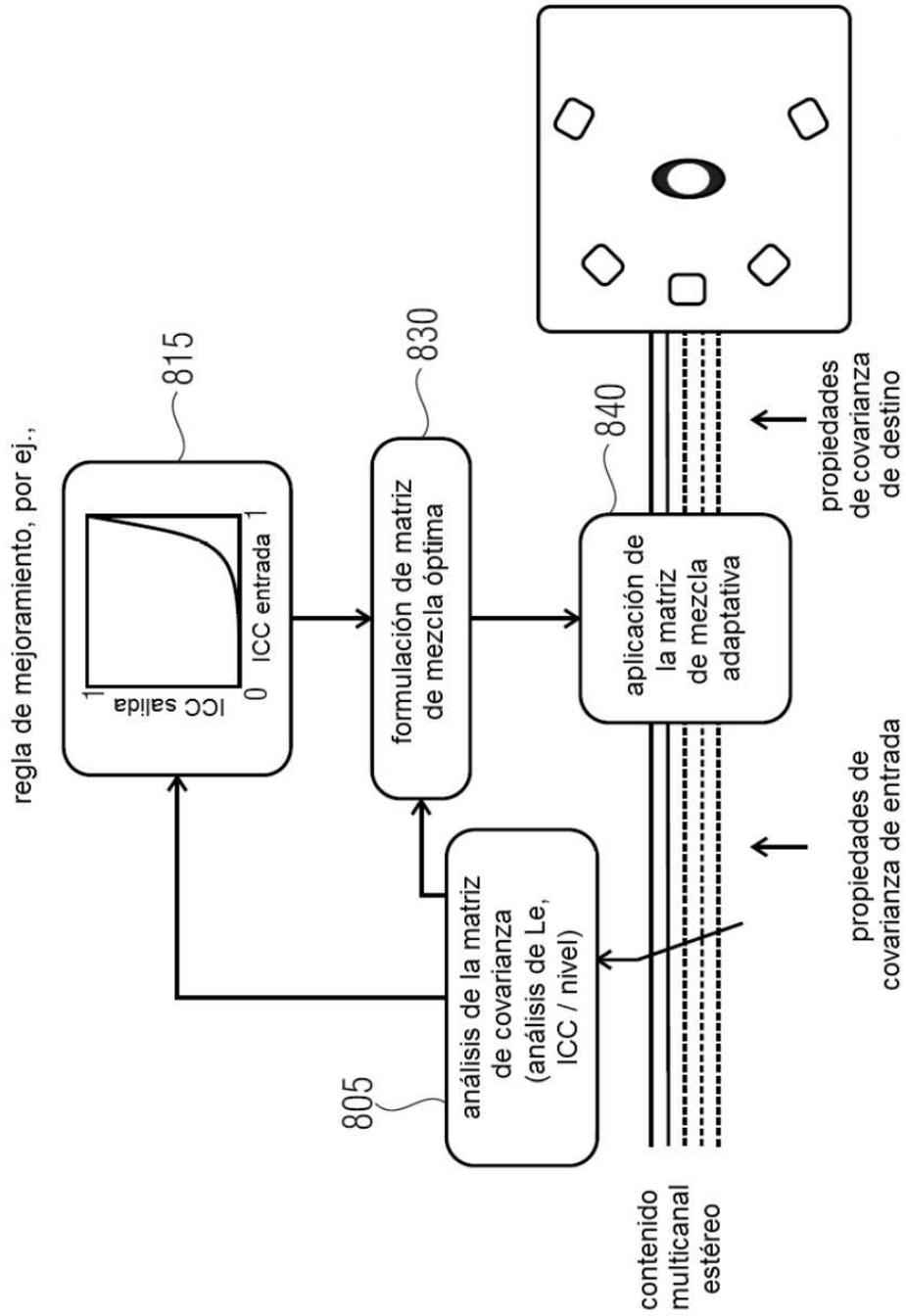


FIG 8

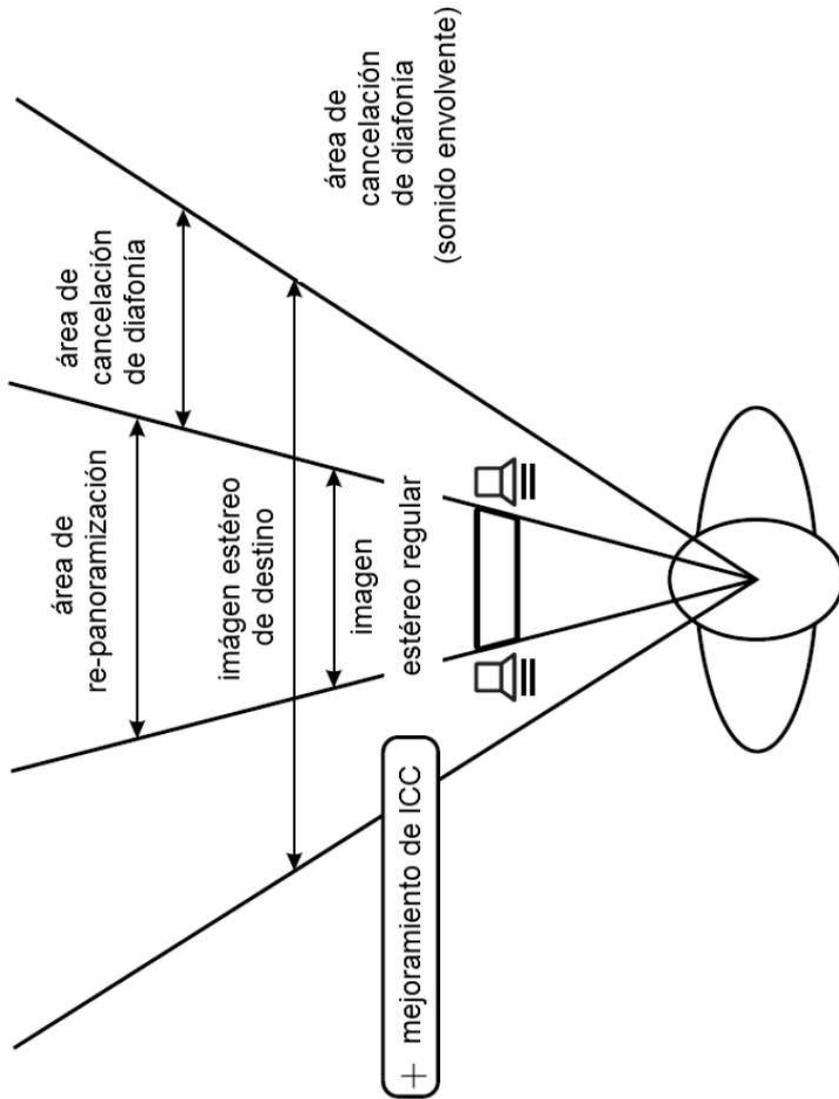


FIG 9

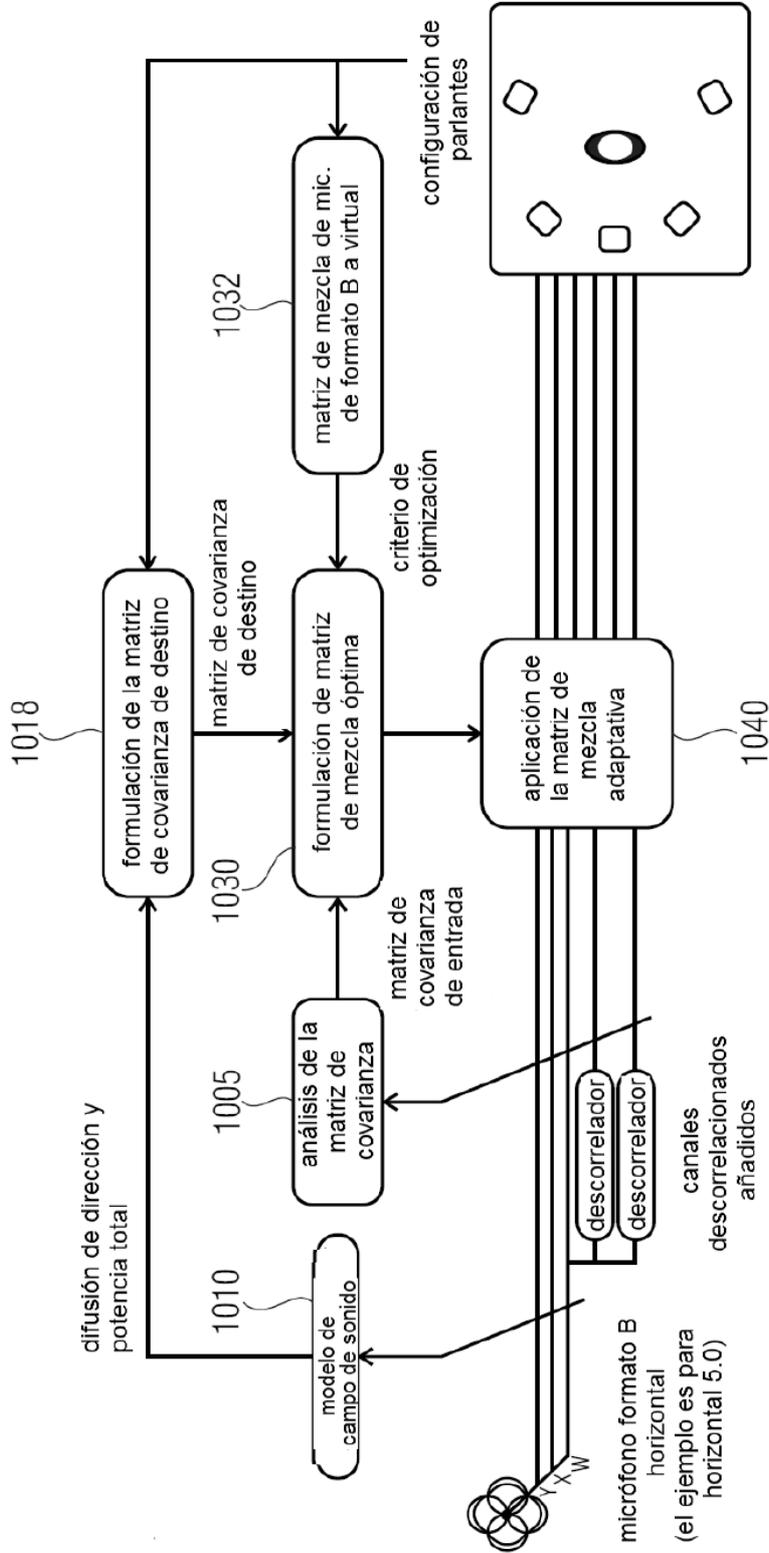


FIG 10

Contexto	$C_x$
Descorrelación: Alta ICC de entrada	$\begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}$
Descorrelación: Muy alta ICC de entrada	$\begin{bmatrix} 1 & 0.97 \\ 0.97 & 1 \end{bmatrix}$
Mezcla ascendente estéreo	$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$
Mezcla ascendente de 5.0 a 7.0	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$
Mezcla descendente: con igual coherencia no cero	$\begin{bmatrix} 1 & 0 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 0 & 0.5 \\ 0 & 0 & 1 & 0 & 0 \\ 0.5 & 0 & 0 & 1 & 0 \\ 0 & 0.5 & 0 & 0 & 1 \end{bmatrix}$
Descorrelación: cardioide coincidente de igual espacio 5.0	$\begin{bmatrix} 1 & 0.86 & 0.64 & 0.64 & 0.86 \\ 0.86 & 1 & 0.86 & 0.64 & 0.64 \\ 0.64 & 0.86 & 1 & 0.86 & 0.64 \\ 0.64 & 0.64 & 0.86 & 1 & 0.86 \\ 0.86 & 0.64 & 0.64 & 0.86 & 1 \end{bmatrix}$
Descorrelación: hipercardiode coincidente de distribución estándar 5.0	$\begin{bmatrix} 1 & 0.65 & 0.91 & 0.43 & -0.22 \\ 0.65 & 1 & 0.91 & -0.22 & 0.43 \\ 0.91 & 0.91 & 1 & 0.07 & 0.07 \\ 0.43 & -0.22 & 0.07 & 1 & -0.22 \\ -0.22 & 0.43 & 0.07 & -0.22 & 1 \end{bmatrix}$

Tabla 1



FIG 11A

	Q	$C_y$	
	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	
	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	
	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0.71 & 0.71 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0.5 \end{bmatrix}$	
	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0.71 & 0 \\ 0 & 0 & 0 & 0 & 0.71 \\ 0 & 0 & 0 & 0.71 & 0 \\ 0 & 0 & 0 & 0 & 0.71 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 \end{bmatrix}$	
	$\begin{bmatrix} 1 & 0 & 0.71 & 1 & 0 \\ 0 & 1 & 0.71 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 2.5 & 0.5 \\ 0.5 & 2.5 \end{bmatrix}$	
	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	
	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$	

Tabla 1

FIG 11	FIG 11A	FIG 11B	FIG 11C
--------	---------	---------	---------

FIG 11B

M	$C_r$
$\begin{bmatrix} 1.5 & -0.75 \\ -0.75 & 1.5 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$
$\begin{bmatrix} 2.1 & -1.4 \\ -1.4 & 2.1 \end{bmatrix}$	$\begin{bmatrix} 0.31 & -0.31 \\ -0.31 & 0.31 \end{bmatrix}$
$\begin{bmatrix} 0.33 & -0.17 \\ -0.17 & 0.33 \\ 0.47 & 0.47 \\ 0.33 & -0.17 \\ -0.17 & 0.33 \end{bmatrix}$	$\begin{bmatrix} 0.33 & 0.08 & -0.24 & -0.17 & 0.08 \\ 0.08 & 0.33 & -0.24 & 0.08 & -0.17 \\ -0.24 & -0.24 & 0.67 & -0.24 & -0.24 \\ -0.17 & 0.08 & -0.24 & 0.33 & 0.08 \\ 0.08 & -0.17 & -0.24 & 0.08 & 0.33 \end{bmatrix}$
$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0.5 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.25 & 0 & -0.25 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.25 & 0 & -0.25 & 0 \\ 0 & 0 & 0 & -0.25 & 0 & 0.25 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.25 & 0 & 0 & 0.25 \end{bmatrix}$
$\begin{bmatrix} 0.84 & 0.02 & 0.61 & 0.84 & 0.02 \\ 0.02 & 0.84 & 0.61 & 0.02 & 0.84 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$
$\begin{bmatrix} 1.7 & -0.53 & -0.05 & -0.05 & -0.53 \\ -0.53 & 1.7 & -0.53 & -0.05 & -0.05 \\ -0.05 & -0.53 & 1.7 & -0.53 & -0.05 \\ -0.05 & -0.05 & -0.53 & 1.7 & -0.53 \\ -0.53 & -0.05 & -0.05 & -0.53 & 1.7 \end{bmatrix}$	$\begin{bmatrix} 0.4 & -0.32 & 0.12 & 0.12 & -0.32 \\ -0.32 & 0.4 & -0.32 & 0.12 & 0.12 \\ 0.12 & -0.32 & 0.4 & -0.32 & 0.12 \\ 0.12 & 0.12 & -0.32 & 0.4 & -0.32 \\ -0.32 & 0.12 & 0.12 & -0.32 & 0.4 \end{bmatrix}$
$\begin{bmatrix} 2 & -0.51 & -0.83 & -0.53 & 0.41 \\ -0.51 & 2 & -0.83 & -0.41 & -0.53 \\ -0.83 & -0.83 & 2.1 & 0.04 & 0.04 \\ -0.53 & 0.41 & 0.04 & 1.2 & -0.07 \\ 0.41 & -0.53 & 0.04 & -0.07 & 1.2 \end{bmatrix}$	$\begin{bmatrix} 0.58 & -0.2 & -0.34 & -0.23 & 0.19 \\ -0.2 & 0.58 & -0.34 & 0.19 & -0.23 \\ -0.34 & -0.34 & 0.62 & 0.03 & 0.03 \\ -0.23 & 0.19 & 0.03 & 0.11 & -0.11 \\ 0.19 & -0.23 & 0.03 & -0.11 & 0.11 \end{bmatrix}$

Tabla 1

FIG 11	FIG 11A	FIG 11B	FIG 11C
--------	---------	---------	---------

FIG 11C

Listado 1: Aplicación de Matlab del método propuesto

```

1  función [M,Cr]=formulate_M_and_Cr (Cx, Cy, Q, flag)
2  % de indicador = 0: Uso esperado de residuales
3  % de indicador = 1: En cambio, fijar potencias
4  lambda= ojo (longitud(Cy), longitud(Cx));
5
6  % de Descomposición de Cy
7  [U_Cy,S_Cy]=svd(Cy);
8  Ky=U_Cy*sqrt(S_Cy);
9
10 % de Descomposición de Cx
11 [U_Cx,S_Cx]=svd(Cx);
12 Kx=U_Cx*sqrt(S_Cx);
13
14 % SVD de Kx
15 Ux=U_Cx;
16 Sx=sqrt(S_Cx);
17 % Vx = matriz de identidad
18
19 % A regularización simple de la inversa
20 Sx_diag=diag(Sx);
21 limit=max(Sx_diag)*0.2;
22 Sx_hat_diag=max(Sx_diag,limit);
23

```

FIG 12	FIG 12A
	FIG 12B

FIG 12A

```

24 % Formular  $Kx^{-1}$  regularizado
25 Kx_hat_inverse=diag(1./Sx_hat_diag)*Ux';
26
27 % Formular P óptimo
28 [U,S,V]=svd(Kx'*Q'*Ky);
29 P=V*lambda*U';
30
31 % Mediante el uso se obtiene M óptima
32 M=Ky*P*Kx_hat_inverse;
33
34 % Formular matriz de covarianza residual
35 Cy_hat = M*Cx*M';
36 Cr=Cy-Cy_hat;
37
38 % Usar compensación de potencia en vez de residuales
39 si el indicador == 1
40     ajuste =diag(Cy)./diag(Cy_hat+1e-20);
41     G=diag(sqrt(adjustment));
42     M=G*M;
43     Cr='innecesario';
44 fin

```

FIG 12	FIG 12A
	FIG 12B

FIG 12B