

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 755 675**

51 Int. Cl.:

H04S 7/00 (2006.01)

H04S 3/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **07.04.2014 PCT/EP2014/056917**

87 Fecha y número de publicación internacional: **16.10.2014 WO14166863**

96 Fecha de presentación y número de la solicitud europea: **07.04.2014 E 14716549 (2)**

97 Fecha y número de publicación de la concesión europea: **11.09.2019 EP 2984857**

54 Título: **Aparato y método para escalado de señales centrales y mejora estereofónica basada en una relación señal a mezcla a dos canales**

30 Prioridad:

12.04.2013 EP 13163621
28.08.2013 EP 13182103

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
23.04.2020

73 Titular/es:

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

72 Inventor/es:

UHLE, CHRISTIAN;
PROKEIN, PETER;
HELLMUTH, OLIVER;
SCHARRER, SEBASTIAN y
HABETS, EMANUEL

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 755 675 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

Aparato y método para escalado de señales centrales y mejora estereofónica basada en una relación señal a mezcla a dos canales

5 La presente invención se refiere a procesamiento de señales de audio, y en particular, a un escalado de señales centrales y mejora estereofónica basada en la relación señal a mezcla a dos canales.

10 Las señales de audio en general son una mezcla de sonidos directos y sonidos ambientales (o difusos). Las señales directas se emiten por fuentes de sonido, por ejemplo, un instrumento musical, un vocalista o un altavoz, y llegan por la ruta más corta posible al receptor, por ejemplo el oído del oyente o un micrófono. Cuando se escucha un sonido directo, se percibe como que proviene de la dirección de la fuente de sonido. Las pistas auditivas relevantes para la ubicación y para otras propiedades sonoras espaciales son la diferencia del nivel interaural (ILD), diferencia de tiempo interaural (ITD) y la coherencia interaural. Las ondas sonoras directas que provocan ILD e ITD idénticas se perciben como que provienen de la misma dirección. En la ausencia de sonido ambiental, las señales que alcanzan el oído izquierdo y derecho o cualquier otro conjunto de sensores separados son coherentes.

15 Los sonidos ambientales, en contraste, se emiten mediante muchas fuentes de sonido separadas o límites reflectantes de sonido que contribuyen al mismo sonido. Cuando una onda sonora alcanza una pared en una habitación, una porción de ésta se refleja, y la superposición de todas las reflexiones en una habitación, la reverberación, son un ejemplo destacado para sonidos ambientales. Otros ejemplos son aplausos, ruido de murmullo y ruido del viento. Los sonidos ambientales se perciben como que son difusos, no localizables, y provocan una impresión de envolvimiento (de estar "inmerso en sonido") por el oyente. Cuando se captura un campo de sonido ambiental utilizando un conjunto de sensores separados, las señales grabadas son al menos parcialmente incoherentes.

20 La técnica anterior relacionada acerca de la separación, descomposición o escalado se basa ya sea en información panorámica, es decir diferencias de nivel inter-canal (ICLD) y diferencias de tiempo inter-canal (ICTD), o se basa en características de señal de sonidos directos y sonidos ambientales. Los métodos que aprovechan las ICLD en grabaciones estereofónicas de dos canales son el método de mezcla multicanal descrito en [7], el algoritmo de resíntesis y discriminación en azimut (ADRes) [8], la mezcla multicanal a partir de señales de entrada de dos canales a tres canales propuesta por Vickers [9], y la extracción de señal central descrita en [10].

30 La técnica de estimación de desmezcla degenerada (DUET) [11, 12], se basa en la agrupación de los intervalos de tiempo-frecuencia en conjuntos con ICLD e ICTD similares. Una restricción del método original es que la frecuencia máxima que se puede procesar es igual a la mitad de la velocidad del sonido a través de la separación máxima de micrófono (debido a las ambigüedades en la estimación ICTD) que se ha abordado en [13]. El desempeño del método disminuye cuando las fuentes se traslapan en el dominio del tiempo-frecuencia y cuando incrementan la reverberación. Otros métodos basados en ICLD e ICTD son el algoritmo ADRes modificado [14], que extiende el algoritmo ADRes [8] para el procesamiento de grabaciones de micrófono separados, el método se basa en la correlación de tiempo-frecuencia (AD-TIFCORR) [15] para mezclas retardadas en tiempo, la estimación de dirección de la matriz de mezcla (DEMIX) para mezclas anecoicas [16], que incluye una medición de confianza que solamente una fuente está activa en un periodo de tiempo-frecuencia particular, la ubicación y separación de fuentes de esperanza-maximización basadas en modelo (MESSL) [17], y métodos que imitan el mecanismo de audición humano binaural como en, por ejemplo [18, 19].

40 A pesar de los métodos para la separación ciega de fuentes (BSS) que utilizan pistas espaciales de componentes de señales directas mencionadas anteriormente, también la extracción y atenuación de señales ambientales están relacionadas con el método presentado. Los métodos basados en la coherencia inter-canal (ICC) en señales de dos canales se describen en [22, 7, 23]. La aplicación del filtrado adaptativo se ha propuesto en [24], con la razón fundamental que las señales directas se pueden predecir a través de canales mientras los sonidos difusos se obtienen a partir del error de predicción.

50 Un método para la mezcla multicanal de señales estereofónicas de dos canales basado en filtrado Wiener multicanal estima tanto, las ICLD de los sonidos directos como las densidades espectrales de potencia (PSD) de las componentes de señales directas y ambientales [25].

60 Los enfoques para la extracción de señales ambientales a partir de grabaciones de un solo canal incluyen el uso de factorización no negativa de matrices de una representación en tiempo-frecuencia de la señal de entrada, donde la señal ambiental se obtiene del residuo de esa aproximación [26], aprendizaje supervisado y extracción de características de bajo nivel [27], y la estimación de la respuesta al impulso de un sistema reverberante y filtrado inverso en el dominio de la frecuencia [28].

El documento US 2010/296672 A1 describe un procedimiento de mezcla multicanal de dominio de frecuencia, que

5 usa descomposición de señal basada en vectores y métodos para mejorar la selectividad de la extracción del canal central. Los procedimientos de mezcla multicanal descritos no realizan una descomposición primaria /ambiental explícita. Esto reduce la complejidad y mejora la calidad de la derivación del canal central. Se describe un método de mezclado multicanal de una señal estéreo de tres canales a una señal de tres canales. Se añaden un vector de entrada izquierdo y un vector de entrada derecho para llegar a una magnitud de suma. De manera similar, la diferencia entre el vector de entrada izquierdo y el vector de entrada derecho se determina para llegar a una magnitud de diferencia. La diferencia entre la magnitud de suma y la magnitud de diferencia se escala para calcular una estimación de la magnitud del canal central, y esta estimación se usa para calcular un vector de salida central. Se calculan un vector de salida izquierdo y un vector de salida derecho. Se completa el método generando el vector de salida izquierdo, el vector de salida central, y el vector de salida derecho.

15 El documento EP 2 464 145 A1 muestra un aparato para descomponer una señal de entrada que tiene un número de al menos tres canales de entrada, que comprende un mezclador a dos canales para mezclar a dos canales la señal de entrada para obtener una señal mezclada a dos canales que tiene un número menor de canales. Además, se proporciona un analizador para analizar la señal mezclada a dos canales para derivar un resultado de análisis, y el resultado del análisis se reenvía a un procesador de señal para procesar la señal de entrada o una señal derivada de la señal de entrada para obtener la señal descompuesta.

20 El objeto de la presente invención es proporcionar conceptos mejorados para procesamiento de señales de audio. El objeto de la presente invención se resuelve mediante un aparato según la reivindicación 1, mediante un sistema según la reivindicación 8, mediante un método según la reivindicación 9 y mediante un programa informático según la reivindicación 10.

25 Se proporciona un aparato según la reivindicación 1 para generar una señal de audio modificada que comprende dos o más canales de audio modificados de una señal de entrada de audio que comprende dos o más canales de entrada de audio. El aparato comprende un generador de información para generar información de señal a mezcla a dos canales. El generador de información se adapta para generar información de señal al combinar un valor espectral de cada uno de los dos o más canales de entrada de audio de una primera forma. Además, el generador de información se adapta para generar información de mezcla a dos canales al combinar el valor espectral de cada uno de los dos o más canales de entrada de audio de una segunda forma que es diferente de la primera forma. Además, el generador de información se adapta para combinar la información de señal y la información de mezcla a dos canales para obtener información de señal a mezcla a dos canales. Además, el aparato comprende un atenuador de señal para atenuar los dos o más canales de entrada de audio dependiendo de la información de señal a mezcla a dos canales para obtener los dos más canales de audio modificados.

35 En una realización particular, el aparato, por ejemplo, se puede adaptar para generar una señal de audio modificada que comprende tres o más canales de audio modificados de una señal de entrada de audio que comprende tres o más canales de entrada de audio.

40 En una realización, el número de canales de audio modificados es igual a o menor que el número de canales de entrada de audio, o en donde el número de los canales de audio modificados es menor que el número de los canales de entrada de audio. Por ejemplo, según una realización particular, el aparato se puede adaptar para generar una señal de audio modificada que comprende dos o más canales de audio modificados de una señal de entrada de audio que comprende dos o más canales de entrada de audio, en donde el número de los canales de audio modificados es igual al número de los canales de entrada de audio.

50 Se propone que las realizaciones proporcionen nuevos conceptos para el escalado del nivel del centro virtual en señales de audio. Las señales de entrada se procesan en el dominio del tiempo-frecuencia de tal forma que se amplifican o atenúan las componentes de sonido directo que tienen energía aproximadamente igual en todos los canales. Las ponderaciones espectrales de valores reales se obtienen de la relación de la suma de las densidades espectrales de potencia de todas las señales de canal de entrada y la densidad espectral de potencia de la señal de suma. Las aplicaciones de los conceptos presentados son la mezcla multicanal de grabaciones estereofónicas de dos canales para su reproducción utilizando configuraciones de sonido envolvente, mejora estereofónica, mejora del dialogo, y como pre-procesamiento para análisis de audio semántico.

55 Las realizaciones proporcionan nuevos conceptos para amplificar o atenuar la señal central en una señal de audio. En contraste con los conceptos anteriores, se toman en cuenta tanto el desplazamiento lateral como el grado de difusión de las componentes de señal. Además, el uso de parámetros semánticamente significativos se analiza a fin de apoyar al usuario cuando se emplean implementaciones de los conceptos.

60 Algunas realizaciones se centran en el escalado de señal central, es decir la amplificación o atenuación de las señales centrales en las grabaciones de audio. La señal central, por ejemplo, se define aquí como la suma de todas las componentes de señal directa que tienen intensidad aproximadamente igual en todos los canales y diferencias en tiempo despreciables entre los canales.

Diversas aplicaciones de reproducción y procesamiento de señales de audio se benefician a partir del escalado de señales centrales, por ejemplo la mezcla multicanal, mejora de dialogo, y análisis de audio semántico.

5 La mezcla multicanal se refiere al procedimiento de crear una señal de salida dada una señal de entrada con menos canales. Su aplicación principal es la reproducción de señales de dos canales que utilizan configuraciones de sonido envolvente como se especifica en [1], a manera de ejemplo. La investigación sobre la calidad subjetiva del audio espacial [2] indica que la ubicabilidad [3], la localización y el ancho son atributos descriptivos destacados del sonido. Los resultados de una evaluación subjetiva de algoritmos de mezcla de 2 a 5 canales [4] mostraron que el uso de un altavoz central adicional puede reducir la imagen estereofónica. El trabajo presentado está motivado por la suposición de que la ubicabilidad, la localización y el ancho se pueden preservar o incluso mejorar cuando el altavoz central adicional reproduce principalmente componentes de señal directa que se dirigen al centro, y cuando estos componentes de señal se atenúan en las señales de los altavoces fuera del centro.

10
15 La mejora del dialogo se refiere a la mejora de la inteligibilidad del habla, por ejemplo en el audio de películas y radiodifusión, y a menudo se desea cuando los sonidos de fondo son muy fuertes con respecto al dialogo [5]. Esto aplica en particular a personas que tienen problemas de audición, oyentes no nativos, en entornos ruidosos o cuando la diferencia de nivel de enmascaramiento binaural se reduce debido a la colocación estrecha de altavoces. El método del concepto se puede aplicar para procesamiento de señales de entrada donde el dialogo se dirige al centro a fin de atenuar los sonidos de fondo y permitiendo por lo tanto una mejor inteligibilidad del habla.

20 El análisis de audio semántico (o análisis de contenido de audio) comprende procedimientos para deducir descriptores significativos de las señales de audio, por ejemplo transcripción o rastreo de ritmo de la melodía principal. El desempeño de los métodos computacionales a menudo se deteriora cuando los sonidos de interés están incorporados en los sonidos de fondo, ver por ejemplo [6]. Ya que es una práctica común en la producción de audio que las fuentes de sonido de interés (por ejemplo cantantes e instrumentos principales) se dirijan al centro, se puede aplicar extracción central como una etapa de pre-procesamiento para la atenuación de los sonidos de fondo y la reverberación.

25 30 Según una realización, el generador de información se puede configurar para combinar la información de señal y la información de mezcla a dos canales de tal forma que la información de señal a mezcla a dos canales indica una relación de la información de señal a la información de mezcla a dos canales.

35 En una realización, el generador de información se puede configurar para procesar el valor espectral de cada uno de los dos o más canales de entrada de audio para obtener dos o más valores procesados, y en donde el generador de información se puede configurar para combinar los dos o más valores procesados para obtener la información de señal. Además, el generador de información se puede configurar para combinar el valor espectral de cada uno de los dos o más canales de entrada de audio para obtener un valor combinado, y en donde el generador de información se puede configurar para procesar el valor combinado para obtener la información de mezcla a dos canales.

40 Según la invención, el generador de información se configura para procesar el valor espectral de cada uno de los dos o más canales de entrada de audio al multiplicar dicho valor espectral por el complejo conjugado de dicho valor espectral para obtener una auto densidad espectral de potencia del valor espectral de cada uno de los dos o más canales de entrada de audio.

En la invención, el generador de información se configura para procesar el valor combinado al determinar una densidad espectral de potencia del valor combinado.

45 50 Según la invención, el generador de información se configura para generar la información de señal $s(m, k, \beta)$ según la fórmula:

$$s(m, k, \beta) = \sum_{i=1}^N \Phi_{i,i}(m, k)^\beta,$$

55 en donde N indica el número de canales de entrada de audio de la señal de entrada de audio, en donde $\Phi_{i,i}(m, k)$ indica la auto espacio densidad espectral de potencia del valor espectral del i-ésimo canal de señal de audio, en donde β es un número real con $\beta > 0$, en donde m indica un índice de tiempo, y en donde k indica un índice de frecuencia. Por ejemplo, según una realización particular $\beta \geq 1$.

60 En una realización, el generador de información se puede configurar para determinar la relación de señal a mezcla de dos canales como la información de señal a mezcla a dos canales según la fórmula $R(m, k, \beta)$

$$R(m, k, \beta) = \left(\frac{\sum_{i=1}^N \Phi_{i,i}(m, k)^\beta}{\Phi_d(m, k)^\beta} \right)^{\frac{1}{2\beta-1}}$$

5 en donde $\Phi_d(m, k)$ indica la densidad espectral de potencia del valor combinado, y en donde $\Phi_d(m, k)^\beta$ es la información de mezcla a dos canales.

Según un ejemplo, el generador de información se puede configurar para generar la información de señal $\Phi_1(m, k)$ según la fórmula.

$$10 \quad \Phi_1(m, k) = \mathcal{E}\{\mathbf{W}\mathbf{X}(m, k)(\mathbf{W}\mathbf{X}(m, k))^H\}$$

en donde el generador de información se configura para generar la información de mezcla a dos canales $\Phi_2(m, k)$ según la fórmula

$$15 \quad \Phi_2(m, k) = \mathcal{E}\{\mathbf{V}\mathbf{X}(m, k)(\mathbf{V}\mathbf{X}(m, k))^H\},$$

y

20 en donde el generador de información se configura para generar la relación de señal a mezcla a dos canales como la información de señal a mezcla a dos canales $R_g(m, k, \beta)$ según la fórmula

$$R_g(m, k, \beta) = \left(\frac{\text{tr}\{\Phi_1(m, k)^\beta\}}{\text{tr}\{\Phi_2(m, k)^\beta\}} \right)^{\frac{1}{2\beta-1}}$$

25 en donde $\mathbf{X}(m, k)$ indica la señal de entrada de audio, en donde

$$\mathbf{X}(m, k) = [X_1(m, k) \cdots X_N(m, k)]^T$$

30 en donde N indica el número de canales de entrada de audio de la señal de entrada de audio, en donde m indica un índice de tiempo, y en donde k indica un índice de frecuencia, en donde $X_1(m, k)$ indica el primer canal de entrada de audio, en donde $X_N(m, k)$ indica el N-ésimo canal de entrada de audio, en donde V indica una matriz o un vector, en donde W indica una matriz o un vector, en donde H indica la transpuesta conjugada de una matriz o un vector, en donde $\mathcal{E}\{\cdot\}$ es una operación esperanza, en donde β es un número real con $\beta > 0$, y en donde $\text{tr}\{\cdot\}$ es la traza de una matriz. Por ejemplo, según una realización particular $\beta \geq 1$.

35 En un ejemplo, V puede ser un vector fila de longitud N cuyos elementos son iguales a uno y W puede ser la matriz identidad de tamaño N x N.

Según un ejemplo, $V = [1, 1]$, en donde $W = [1, -1]$ y en donde $N = 2$.

40 En una realización, el atenuador de señal se puede adaptar para atenuar los dos o más canales de entrada de audio dependiendo de una función de ganancia $G(m, k)$ según la fórmula

$$\mathbf{Y}(m, k) = G(m, k)\mathbf{X}(m, k)$$

45 en donde la función de ganancia $G(m, k)$ depende de la información de señal a mezcla a dos canales, y en donde la función de ganancia $G(m, k)$ es una función monótonamente creciente de la información de señal a mezcla a dos canales o una función monótonamente decreciente de la información de señal a mezcla a dos canales,

50 en donde $\mathbf{X}(m, k)$ indica la señal de entrada de audio, en donde $\mathbf{Y}(m, k)$ indica la señal de audio modificada, en donde m indica un índice de tiempo, y en donde k indica un índice de frecuencia.

Según una realización, la función de ganancia $G(m, k)$ puede ser una primera función $G_{c1}(m, k, \beta, \gamma)$, una segunda función $G_{c2}(m, k, \beta, \gamma)$, una tercera función $G_{s1}(m, k, \beta, \gamma)$ o una cuarta función $G_{s2}(m, k, \beta, \gamma)$.

5 en donde

$$G_{c1}(m, k, \beta, \gamma) = (1 + R_{\min} - R(m, k, \beta))^{\gamma}$$

en donde

10

$$G_{c2}(m, k, \beta, \gamma) = \left(\frac{R_{\min}}{R(m, k, \beta)}\right)^{\gamma}$$

en donde

15

$$G_{s1}(m, k, \beta, \gamma) = R(m, k, \beta)^{\gamma}$$

en donde

$$G_{s2}(m, k, \beta, \gamma) = \left(1 + R_{\min} - \frac{R_{\min}}{R(m, k, \beta)}\right)^{\gamma}$$

20

en donde β es un número real con $\beta > 0$,

en donde γ es un número real con $\gamma > 0$ y

25

en donde R_{\min} indica el mínimo de R .

Además, se proporciona un sistema según la reivindicación 8. El sistema comprende un compensador de fase para generar una señal de audio compensada en fase que comprende dos o más canales de audio compensados en fase de una señal de audio sin procesar que comprende dos o más canales de audio sin procesar. Además, el sistema comprende un aparato según una de las realizaciones anteriormente descritas para recibir la señal de audio compensada en fase como una señal de entrada de audio y para generar una señal de audio modificada que comprende dos o más canales de audio modificados de la señal de entrada de audio que comprende los dos o más canales de audio compensados en fase como dos o más canales de entrada de audio. Uno de los dos o más canales de audio sin procesar es un canal de referencia. El compensador de fase está adaptado para estimar para cada uno de los canales de audio sin procesar de los dos o más canales de audio sin procesar que no son el canal de referencia una función de transferencia de fase entre el canal de audio sin procesar y el canal de referencia. Además, el compensador de fase se adapta para generar la señal de audio compensada en fase al modificar cada canal de audio sin procesar de los canales de audio sin procesar que no son el canal de referencia dependiendo de la función de transferencia de fase del canal de audio sin procesar.

40

Además, se proporciona un método según la reivindicación 9 para generar una señal de audio modificada que comprende dos o más canales de audio modificados de una señal de entrada de audio que comprende dos o más canales de entrada de audio. El método comprende:

45

- generar información de señal al combinar un valor espectral de cada uno de los dos o más canales de entrada de audio de una primera forma.

50

- generar información de mezcla a dos canales al combinar el valor espectral de cada uno de los dos o más canales de entrada de audio de una segunda forma que es diferente de la primera forma.

55

- generar información de señal a mezcla a dos canales al combinar la información de señal y la información de mezcla a dos canales. Y:

- Atenuar los dos o más canales de entrada de audio dependiendo de la información de señal a mezcla a dos canales para obtener los dos o más canales de audio modificados.

Además, se proporciona un programa informático según la reivindicación 10 para implementar el método descrito anteriormente cuando se ejecuta en un ordenador o atenuador de señal.

5 Breve Descripción de las Figuras

En lo siguiente, se describen las realizaciones de la presente invención en más detalles con referencia a las figuras, en las cuales:

10 La figura 1 ilustra un aparato según una realización;

La figura 2 ilustra la relación de señal a mezcla a dos canales como una función de las diferencias de nivel inter-canal y como una función de la coherencia inter-canal según una realización;

15 La figura 3 ilustra ponderaciones espectrales como una función de la coherencia inter-canal y de las diferencias de nivel inter-canal según una realización;

La figura 4 ilustra ponderaciones espectrales como una función de la coherencia inter-canal y de las diferencias de nivel inter-canal según otra realización;

20 La figura 5 ilustra ponderaciones espectrales como una función de la coherencia inter-canal y de las diferencias de nivel inter-canal según una realización adicional;

25 Las figuras 6a-e ilustran espectrogramas de las señales de fuente directa y las señales de canal izquierdo y derecho de la señal de mezcla;

La figura 7 ilustra la señal de entrada y la señal de salida de la extracción de señal central según una realización;

30 La figura 8 ilustra los espectrogramas de la señal de salida según una realización;

La figura 9 ilustra la señal de entrada y la señal de salida para la atenuación de señal central según otra realización;

La figura 10 ilustra los espectrogramas de la señal de salida según una realización;

35 Las figuras 11a-d ilustran dos señales de voz que se han mezclado para obtener señales de entrada con y sin diferencias de tiempo inter-canal;

Las figuras 12a-c ilustran las ponderaciones espectrales calculadas a partir de una función de ganancia según una realización; y

40 La figura 13 ilustra un sistema según una realización.

45 La figura 1 ilustra un aparato para generar una señal de audio modificada que comprende dos o más canales de audio modificados de una señal de entrada de audio que comprende dos o más canales de entrada de audio según una realización.

El aparato comprende un generador 110 de información para generar información de señal a mezcla a dos canales.

50 El generador 110 de información se adapta para generar información de señal al combinar un valor espectral de cada uno de los dos o más canales de entrada de audio de una primera forma. Además, el generador 110 de información se adapta para generar información de mezcla a dos canales al combinar el valor espectral de cada uno de los dos o más canales de entrada de audio de una segunda forma que es diferente de la primera forma.

55 Además, el generador 110 de información se adapta para combinar la información de señal y la información de mezcla a dos canales para obtener información de señal a mezcla a dos canales. Por ejemplo, la información de señal a mezcla a dos canales puede ser una relación de señal a mezcla a dos canales, por ejemplo, un valor de señal a mezcla a dos canales.

60 Además, el aparato comprende un atenuador 120 de señal para atenuar los dos o más canales de entrada de audio dependiendo de la información de señal a mezcla a dos canales para obtener los dos o más canales de audio modificados.

Según una realización, el generador de información se puede configurar para combinar la información de señal y la información de mezcla a dos canales de tal forma que la información de señal a mezcla a dos canales indica una

relación de la información de señal a la información de mezcla a dos canales. Por ejemplo, la información de señal puede ser un primer valor y la información de mezcla a dos canales puede ser un segundo valor y la información de señal a mezcla a dos canales indica una relación del valor de señal al valor de mezcla a dos canales. Por ejemplo, la información de señal a mezcla a dos canales puede ser el primer valor dividido por el segundo valor. O, por ejemplo, si el primer valor y el segundo valor son valores logarítmicos, la información de señal a mezcla a dos canales puede ser la diferencia entre el primer valor y el segundo valor.

En lo siguiente, se describen los conceptos y el modelo de señal subyacente y se analizan para el caso de señal de entrada que ofrece estereofonía de diferencia de amplitud.

La razón fundamental es calcular y aplicar las ponderaciones espectrales de valores reales como una función del grado de difusión y la posición lateral de las fuentes directas. El procesamiento como se demuestra aquí se aplica en el dominio STFT, sin embargo, no se limita a un banco de filtros particular. La N señal de entrada de canal se denota por

$$\mathbf{x}[n] = [x_1[n] \cdots x_N[n]]^T \quad (1)$$

en donde n denota el índice de tiempo discreto. La señal de entrada se asume que es una mezcla aditiva de señales directas $s_l[n]$ y sonidos ambientales $a_l[n]$,

$$x_l[n] = \sum_{i=1}^K d_{i,l}[n] * s_i[n] + a_l[n], \quad l = 1, \dots, N \quad (2)$$

donde P es el número de fuentes de sonido, $d_{i,l}[n]$ denota las respuestas al impulso de las rutas directas de la i-ésima fuente en el l-ésimo canal de longitud de $L_{i,l}$ muestras, y las componentes de señal ambiente no se correlacionan mutuamente o se correlacionan débilmente. En la siguiente descripción se asume que el modelo de señal corresponde a la estereofonía de diferencia de amplitud, es decir $L_{i,l} = 1, \forall i, l$.

La representación en el dominio del tiempo-frecuencia de $x[n]$ está dada por

$$\mathbf{X}(m, k) = [X_1(m, k) \cdots X_N(m, k)]^T \quad (3)$$

con el índice de tiempo m y el índice de frecuencia k. Las señales de salida se denotan por

$$\mathbf{Y}(m, k) = [Y_1(m, k) \cdots Y_N(m, k)]^T \quad (4)$$

y se obtienen por medio de la ponderación espectral

$$\mathbf{Y}(m, k) = G(m, k)\mathbf{X}(m, k) \quad (5)$$

con las ponderaciones de valores reales $G(m, k)$. Las señales de salida en el dominio del tiempo se calculan al aplicar el procesamiento inverso del banco de filtros. Para el cálculo de las ponderaciones espectrales, la señal de suma, en lo sucesivo se denota como la señal de mezcla a dos canales, se calcula como

$$X_d(m, k) = \sum_{i=1}^N X_i(m, k) \quad (6)$$

La matriz de las PSD de la señal de entrada, que comprende estimaciones de la (auto-)PSD en la diagonal principal, en tanto que los elementos fuera de la diagonal son estimaciones de la PSD transversal, está dada por

$$\Phi_{i,l}(m, k) = \mathcal{E}\{X_i(m, k)X_l^*(m, k)\}, \quad i, l = 1 \dots N \quad (7)$$

donde X^* denota el complejo conjugado de X , y $\mathcal{E}\{\cdot\}$ es la operación de esperanza con respecto a la dimensión temporal. En las simulaciones presentadas los valores de esperanza se estiman utilizando cálculo de promedio recursivo de un solo polo,

$$5 \quad \Phi_{i,l}(m, k) = \alpha X_i(m, k) X_l^*(m, k) + (1 - \alpha) \Phi_{i,l}(m - 1, k) \quad (8)$$

donde el coeficiente de filtro α determina el tiempo de integración. Además, la cantidad $R(m, k, \beta)$ se define como

$$R(m, k, \beta) = \left(\frac{\sum_{i=1}^N \Phi_{i,i}(m, k)^\beta}{\Phi_d(m, k)^\beta} \right)^{\frac{1}{2\beta-1}} \quad (9)$$

10 donde $\Phi_d(m, k)$ es la PSD de la señal de mezcla a dos canales y β es un parámetro que se abordará en lo siguiente. La cantidad $R(m, k; 1)$ es la relación de señal a mezcla a dos canales (SDR), es decir la relación de la PSD total y la PSD de la señal de mezcla a dos canales. La potenciación a $\frac{1}{2\beta-1}$ asegura que el intervalo de $R(m, k; \beta)$ sea independiente de β .

15 El generador 110 de información se puede configurar para determinar la relación de señal a mezcla a dos canales según la ecuación (9).

20 Según la ecuación (9) la información de señal $s(m, k, \beta)$ que se puede determinar por el generador 110 de información se define como

$$s(m, k, \beta) = \sum_{i=1}^N \Phi_{i,i}(m, k)^\beta$$

25 Como se puede ver de lo anterior, $\Phi_{i,i}(m, k)$ se define como $\Phi_{i,i}(m, k) = \mathcal{E}\{X_i(m, k) X_i^*(m, k)\}$. Por lo tanto, para determinar la información de señal $s(m, k, \beta)$, el valor espectral $X_i(m, k)$ de cada uno de los dos o más canales de entrada de audio se procesa para obtener el valor $\Phi_{i,i}(m, k)^\beta$ procesado para cada uno de los dos o más canales de entrada de audio, y los valores procesados obtenidos $\Phi_{i,i}(m, k)^\beta$ entonces se combinan, por ejemplo, como en la ecuación (9) al sumar los valores procesados obtenidos $\Phi_{i,i}(m, k)^\beta$.

30 Por lo tanto, el generador 110 de información se puede configurar para procesar el valor espectral $X_i(m, k)$ de cada uno de los dos o más canales de entrada de audio para obtener dos o más valores procesados $\Phi_{i,i}(m, k)^\beta$, y el generador 110 de información se puede configurar para combinar los dos o más valores procesados para obtener la señal de información $s(m, k, \beta)$. De forma más general, el generador 110 de información se adapta para generar información de señal $s(m, k, \beta)$ al combinar un valor espectral $X_i(m, k)$ de cada uno de los dos o más canales de entrada de audio de una primera forma.

Además, según la ecuación (9) la información de mezcla a dos canales $d(m, k, \beta)$ que se puede determinar por el generador 110 de información se define como

$$40 \quad d(m, k, \beta) = \Phi_d(m, k)^\beta$$

Para formar $\Phi_d(m, k)^\beta$, primeramente se forma $X_d(m, k)$ según la ecuación (6) anterior:

$$X_d(m, k) = \sum_{i=1}^N X_i(m, k)$$

45 Como se puede ver, primeramente, el valor espectral $X_i(m, k)$ de cada uno de los dos o más canales de entrada de audio se combina para obtener un valor combinado $X_d(m, k)$, procesamiento, como en la ecuación (6), al sumar el valor espectral $X_i(m, k)$ de cada uno de los dos o más canales de entrada de audio.

50 Entonces, para obtener $\Phi_d(m, k)$ se forma la densidad espectral de potencia de $X_d(m, k)$, por ejemplo, según

$$\Phi_d(m, k) = \mathcal{E}\{X_d(m, k) X_d^*(m, k)\},$$

5 y entonces, $\Phi_d(m, k)^\beta$ se puede determinar. En términos más generales, el valor combinado obtenido $X_d(m, k)$ se ha procesado para obtener la información de mezcla a dos canales $d(m, k, \beta) = \Phi_d(m, k)^\beta$.

10 Por lo tanto, el generador 110 de información se puede configurar para combinar el valor espectral $X_i(m, k)$ de cada uno de los dos o más canales de entrada de audio para obtener un valor combinado, y el generador 110 de información se puede configurar para procesar el valor combinado para obtener la información de mezcla a dos canales $d(m, k, \beta)$. De forma más general, el generador 110 de información se adapta para generar información de mezcla a dos canales $d(m, k, \beta)$ al combinar el valor espectral $X_i(m, k)$ de cada uno de los dos o más canales de entrada de audio de una segunda forma. La forma, de qué manera se genera la información de mezcla a dos canales ("segunda forma") difiere de la forma, de qué manera se genera la información de señal ("primera forma") y por lo tanto, la segunda forma es diferente de la primera forma.

15 El generador 110 de información se adapta para generar información de señal al combinar un valor espectral de cada uno de los dos o más canales de entrada de audio de una primera forma. Además, el generador 110 de información se adapta para generar información de mezcla a dos canales al combinar el valor espectral de cada uno de los dos o más canales de entrada de audio de una segunda forma que es diferente de la primera forma.

20 El gráfica superior de la figura 2 ilustra la relación de señal a mezcla a dos canales $R(m, k; 1)$ para $N=2$ como función de la ICLD $\Theta(m, k)$, mostrada para $\Psi(m, k) \in \{0, 0,2, 0,4, 0,6, 0,8, 1\}$. En la gráfica inferior de la figura 2 se ilustra la relación de señal a mezcla a dos canales $R(m, k; 1)$ para $N=2$ como función de ICC $\Psi(m, k)$ e ICLD $\Theta(m, k)$ en la gráfica 2D codificada por color.

25 En particular, la figura 2 ilustra la SDR para $N=2$ como una función de ICC $\Psi(m, k)$ e ICLD $\Theta(m, k)$, con

$$\Psi(m, k) = \frac{|\Phi_{1,2}(m, k)|}{\sqrt{\Phi_{1,1}(m, k)\Phi_{2,2}(m, k)}}, \quad (10)$$

30 y

$$\Theta(m, k) = \frac{\Phi_{1,1}(m, k)}{\Phi_{2,2}(m, k)}. \quad (11)$$

35 La figura 2 muestra que la SDR tiene las siguientes propiedades:

1. Está monótonamente relacionada tanto con, $\Psi(m, k)$ como $|\log \Theta(m, k)|$.
2. Para señales de entrada difusas, es decir $\Psi(m, k) = 0$, la SDR asume su valor máximo, $R(m, k, 1) = 1$.
- 40 3. Para sonidos directos dirigidos al centro, es decir, $\Theta(m, k) = 1$, la SDR asume su valor mínimo R_{\min} , donde $R_{\min} = 0,5$ para $N=2$.

45 Debido a estas propiedades, las ponderaciones espectrales apropiadas para el escalado de señales centrales se pueden calcular de la SDR al utilizar funciones monótonamente decrecientes para la extracción de señales centrales y funciones monótonamente crecientes para la atenuación de señales centrales.

Para la extracción de una señal central, funciones apropiadas de $R(m, k; \beta)$ son, por ejemplo,

$$G_{c1}(m, k, \beta, \gamma) = (1 + R_{\min} - R(m, k, \beta))^\gamma, \quad (12)$$

50 y

$$G_{c2}(m, k, \beta, \gamma) = \left(\frac{R_{\min}}{R(m, k, \beta)}\right)^\gamma. \quad (13)$$

donde se introduce un parámetro para controlar la atenuación máxima.

Para la atenuación de la señal central, funciones apropiadas de $R(m, k; \beta)$ son, por ejemplo,

$$G_{s1}(m, k, \beta, \gamma) = R(m, k, \beta)^\gamma, \quad (14)$$

y

$$G_{s2}(m, k, \beta, \gamma) = \left(1 + R_{\min} - \frac{R_{\min}}{R(m, k, \beta)}\right)^\gamma, \quad (15)$$

Las figuras 3 y 4 ilustran las funciones de ganancia (13) y (15), respectivamente, para $\beta = 1, \gamma = 3$. Las ponderaciones espectrales son constantes para $\Psi(m, k) = 0$. La atenuación máxima es $\gamma \cdot 6\text{dB}$, que también se aplica a las funciones de ganancia (12) y (14).

En particular, la figura 3 ilustra las ponderaciones espectrales $G_{c2}(m, k; 1, 3)$ en dB como función ICC $\Psi(m, k)$ e ICLD $\Theta(m, k)$.

Además, la figura 4 ilustra las ponderaciones espectrales $G_{s2}(m, k; 1, 3)$ en dB como función de ICC $\Psi(m, k)$ e ICLD $\Theta(m, k)$.

Además, la figura 5 ilustra las ponderaciones espectrales $G_{c2}(m, k; 2, 3)$ en dB y como función de ICC $\Psi(m, k)$ e ICLD $\Theta(m, k)$.

El efecto del parámetro β se muestra en la figura 5 para la función de ganancia en la ecuación (13) con $\beta = 2, \gamma = 3$. Con valores más grandes para β , la influencia de Ψ en las ponderaciones espectrales disminuye mientras la influencia de Θ incrementa. Esto conduce a mayor pérdida de las componentes de señal difusa en la señal de salida, y a mayor atenuación de las componentes de señal directa dirigidas fuera del centro, cuando se compara con la función de ganancia en la figura 3.

Post-procesamiento de ponderaciones espectrales: Antes de la ponderación espectral, las ponderaciones $G(m, k; \beta, \gamma)$ se pueden procesar de forma adicional por medio de operaciones de suavizado. El filtrado de paso bajo de fase cero a lo largo del eje de frecuencia reduce las distorsiones de convolución circular que se pueden presentar, por ejemplo, cuando el relleno con ceros en el cálculo STFT es demasiado corto o se aplica una ventana de síntesis rectangular. El filtrado de paso bajo a lo largo del eje del tiempo puede reducir las distorsiones de procesamiento, especialmente cuando es bastante pequeña la constante de tiempo para la estimación PSD.

En lo siguiente, se proporcionan ponderaciones espectrales generalizadas.

Las ponderaciones espectrales más generales se obtienen cuando se reescribe la ecuación (9),

$$R_g(m, k, \beta) = \left(\frac{\text{tr}\{\Phi_1(m, k)^\beta\}}{\text{tr}\{\Phi_2(m, k)^\beta\}} \right)^{\frac{1}{2\beta-1}}. \quad (16)$$

con

$$\Phi_1(m, k) = \mathcal{E}\{\mathbf{W}\mathbf{X}(m, k)(\mathbf{W}\mathbf{X}(m, k))^H\} \quad (17)$$

$$\Phi_2(m, k) = \mathcal{E}\{\mathbf{V}\mathbf{X}(m, k)(\mathbf{V}\mathbf{X}(m, k))^H\} \quad (18)$$

donde el superíndice H denota la transpuesta conjugada de una matriz o un vector, y W y V son matrices de mezcla o vectores de mezcla (fila).

Aquí, $\Phi_1(m, k)$ se puede considerar como información de señal y $\Phi_2(m, k)$ se puede considerar como información de mezcla a dos canales.

Por ejemplo, $\Phi_2 = \Phi_d$ cuando V es un vector de longitud N cuyos elementos son iguales a uno. La ecuación (16) es igual a (9) cuando V es un vector fila de longitud N cuyos elementos son iguales a uno y W es la matriz identidad de tamaño $N \times N$.

- 5 La SDR $R_g(m, k, \beta, W, V)$ generalizada cubre, por ejemplo, la relación de la PSD de la señal lateral y la PSD de la señal de mezcla a dos canales, para $W = [1, -1]$, $V = [1, 1]$, y $N=2$.

$$R(m, k, \beta) = \left(\frac{\Phi_s(m, k)^\beta}{\Phi_d(m, k)^\beta} \right)^{\frac{1}{2\beta-1}} \quad (19)$$

- 10 donde $\Phi_s(m, k)$ es la PSD de la señal lateral.

Según una realización, el generador 110 de información se adapta para generar información de señal $\Phi_1(m, k)$ al combinar un valor espectral $X_i(m, k)$ de cada uno de los dos o más canales de entrada de audio de una primera forma. Además, el generador 110 de información se adapta para generar información de mezcla a dos canales $\Phi_2(m, k)$ al combinar el valor espectral $X_i(m, k)$ de cada uno de los dos o más canales de entrada de audio de una segunda forma que es diferente de la primera forma.

En lo siguiente, se describe un caso más general de modelos de mezcla que ofrecen estereofonía de tiempo de llegada.

20 La derivación de las ponderaciones espectrales descrita anteriormente depende de la suposición que $L_{i,l} = 1, \forall i, l$, es decir las fuentes de sonido directo se alinean en tiempo entre los canales de entrada. Cuando la mezcla de las señales de fuente directa no se limita a estereofonía de diferencia de amplitud ($L_{i,l} > 1$), por ejemplo cuando se graban con micrófonos separados, la mezcla a dos canales de la señal de entrada $X_d(m, k)$ se somete a cancelación de fase. La cancelación de fase en $X_d(m, k)$ conduce a un incremento de los valores SDR y en consecuencia a las distorsiones de filtro de peine convencionales cuando se aplica la ponderación espectral como se describió anteriormente.

Las muescas del filtro de peine corresponden a las frecuencias

$$f_n = \frac{of_s}{2d}$$

para las funciones de ganancia (12) y (13) y

$$f_n = \frac{ef_s}{2d}$$

para las funciones de ganancia (14) y (15), donde f_s es la frecuencia de muestreo, o son enteros impares, e son enteros pares, y d es el retardo en las muestras.

40 Un primer enfoque para resolver este problema es compensar las diferencias en fase que resultan de la ICTD antes del cálculo de $X_d(m, k)$. La compensación de diferencia en fase (PDC) se logra al estimar la función de transferencia de fase inter-canal variante en el tiempo $\hat{P}_i(m, k) \in [-\pi \pi]$ entre el i -ésimo canal y un canal de referencia denotado por el índice r ,

$$\hat{P}_i(m, k) = \arg X_r(m, k) - \arg X_i(m, k), \quad i \in [1, \dots, N] \setminus r \quad (20)$$

donde el operador $A \setminus B$ denota la diferencia teórica de conjunto del conjunto B y el conjunto A , y aplicando un filtro de compensación paso todo variante en el tiempo $H_{c,i}(m, k)$ a la i -ésima señal de canal.

$$\tilde{X}_i(m, k) = H_{c,i}(m, k) X_i(m, k) \quad (21)$$

donde la función de transferencia de fase de $H_{c,i}(m, k)$ es

$$\arg H_{C,i}(m, k) = -\mathcal{E}\{\hat{P}_i(m, k)\}. \quad (22)$$

5 El valor de esperanza se estima utilizando cálculo de promedio recursivo de un solo polo. Se debe señalar que la fase salta de 2π que se presenta en frecuencias cerca a las frecuencias de muesca se necesitan compensar antes del cálculo de promedio recursivo.

La señal de mezcla a dos canales se calcula según

$$X_d(m, k) = \sum_{i=1}^N \tilde{X}_i(m, k) \quad (23)$$

10 de tal forma que la PDC sólo se aplica para el cálculo de X_d y no afecta la fase de la señal de salida.

La figura 13 ilustra un sistema según una realización.

15 El sistema comprende un compensador 210 de fase para generar una señal de audio compensada en fase que comprende dos o más canales de audio compensados en fase de una señal de audio sin procesar que comprende dos o más canales de audio sin procesar.

20 Además, el sistema comprende un aparato 220 según una de las realizaciones descritas anteriormente para recibir la señal de audio compensada en fase como una señal de entrada de audio y para generar una señal de audio modificada que comprende dos o más canales de audio modificadas de la señal de entrada de audio que comprende los dos o más canales de audio compensados en fase como dos o más canales de entrada de audio.

25 Uno de los dos o más canales de audio sin procesar es un canal de referencia. El compensador 210 de fase se adapta para estimar para cada uno de los canales de audio sin procesar de los dos o más canales de audio sin procesar que no son el canal de referencia una función de transferencia de fase entre el canal de audio sin procesar y el canal de referencia. Además, el compensador 210 de fase se adapta para generar la señal de audio compensada en fase al modificar cada canal de audio sin procesar de los canales de audio sin procesar que no son el canal de referencia dependiendo de la función de transferencia de fase del canal de audio sin procesar.

30 En lo siguiente, se proporcionan explicaciones intuitivas de los parámetros de control, por ejemplo, un significado semántico de los parámetros de control.

35 Para la operación de efectos de audio digital es ventajoso proporcionar controles con parámetros semánticamente significativos. Las funciones de ganancia (12)-(15) se controlan por los parámetros α , β y γ . Se utiliza ingeniería de sonido e ingeniería de audio para constantes de tiempo, y que especifican α como constante de tiempo es intuitivo y según la práctica común. El efecto del tiempo de integración se puede experimentar mejor por la experimentación. A fin de soportar la operación de los conceptos provistos, se proponen descriptores para los parámetros restantes, es decir impacto para γ y grado de difusión para β .

40 El impacto de parámetro se puede comparar mejor con el orden de un filtro. Por analogía con la atenuación progresiva en el filtrado, la atenuación máxima es igual a γ . 6dB para $N=2$.

45 El grado de difusión de marca se propone aquí para enfatizar el hecho de que después de la atenuación de sonidos difusos y dirigidos, valores más grandes de β dan por resultado mayor pérdida de sonidos difusos. Un mapeo no

lineal del parámetro de usuario β_u , por ejemplo $\beta = \sqrt{\beta_u + 1}$, con $0 \leq \beta_u \leq 10$, es ventajoso de una manera que permite un comportamiento más consistente del procesamiento en contraposición a cuando se modifica β directamente (cuando la consistencia se refiere al efecto de un cambio del parámetro en el resultado a lo largo del intervalo del valor espectral).

50 En lo siguiente, se analizan brevemente los requerimientos de memoria y complejidad computacional.

55 Los requerimientos de memoria y complejidad computacional cambian de escala con el número de bandas del banco de filtros y dependen de la implementación del pos-procesamiento adicional de las ponderaciones espectrales. Una implementación de bajo costo del método se puede lograr cuando se ajusta $\beta = 1$, $\gamma \in \mathbb{N}$, calculando ponderaciones espectrales según la ecuación (12) o (14), y cuando no se aplica el filtro PDC. El cálculo de la SDR utiliza solamente una función no lineal intensiva de costo por su banda cuando $\beta \in \mathbb{N}$. Para $\beta = 1$, sólo se necesitan dos memorias intermedias para la estimación PSD, mientras que los métodos que hacen uso

explícito de la ICC, por ejemplo [7, 10, 20, 21, 23], requieren al menos tres memorias intermedias.

En lo siguiente, se discute el desempeño de los conceptos presentados por medio de ejemplos.

- 5 Primero, el procesamiento se aplica a una mezcla de amplitud dirigida de 5 grabaciones de instrumentos (batería, bajo, claves, 2 guitarras) muestreada a 44100 Hz de las cuales se visualiza un extracto de 3 segundos de duración. La batería, bajo y claves se dirigen al centro, una guitarra se dirige al canal izquierdo y la segunda guitarra se dirige al canal derecho, ambos con $|ICLD| = 20$ dB. Una reverberación de convolución que tiene respuestas al impulso estéreo con un RT60 de aproximadamente 1.4 segundos por canal de entrada se utiliza para generar componentes de señal ambiental. La señal reverberada se suma con una relación directa a ambiental de aproximadamente 8 dB después de la ponderación K [29].

- 15 Las figuras 6a-e muestran espectrogramas de las señales de fuente directa y las señales de canal izquierdo y derecho de la señal de mezcla. Los espectrogramas se calculan utilizando una STFT con una longitud de 2048 muestras, 50% de traslape, un tamaño de trama de 1024 muestras y una ventana sinusoidal. Se debe tener en cuenta que a modo de claridad solamente se muestran las magnitudes de los coeficientes espectrales que corresponden a las frecuencias de hasta 4 kHz. En particular, las figuras 6a-3 ilustran señales de entrada para el ejemplo de música.

- 20 En particular, las figuras 6a-e ilustran en la figura 6a señales de fuente, en donde batería, bajo y claves se dirigen al centro; en la figura 6b las señales de fuente, en donde la guitarra 1, en la mezcla se dirige a la izquierda; en la figura 6c en las señales de fuente en donde la guitarra 2, en la mezcla se dirige a la derecha; en la figura 6d un canal izquierdo de una señal de mezcla; y en la figura 6e un canal derecho de una señal de mezcla.

- 25 La figura 7 muestra la señal de entrada y la señal de salida para la extracción de señal central obtenida al aplicar GC2 ($m, k; 1, 3$). En particular, la figura 7 es un ejemplo para extracción central, en donde se ilustran las señales de tiempo de entrada (negras) y las señales de tiempo de salida (superpuestas en gris), en donde la gráfica superior de la figura 7 ilustra un canal izquierdo, y en donde la gráfica inferior de la figura 7 ilustra un canal derecho.

- 30 La constante de tiempo para el cálculo de tiempo recursivo en la estimación PSD aquí y en lo siguiente se ajusta a 200 ms.

- 35 La figura 8 ilustra los espectrogramas de la señal de salida. La inspección visual revela que las señales de fuente dirigidas fuera del centro (mostradas en la figura 6b y 6c) se atenúan ampliamente en los espectrogramas de salida. En particular, la figura 8 ilustra un ejemplo para extracción central, de forma más particular espectrogramas de las señales de salida. Los espectrogramas de salida también muestran que se atenúan las componentes de señal ambiental.

- 40 La figura 9 muestra la señal de entrada y la señal de salida de la atenuación de señal central obtenida al aplicar Gs2 ($m, k; 1, 3$). Las señales de tiempo ilustran que los sonidos transitorios de la batería se atenúan por el procesamiento. En particular, la figura 9 ilustra un ejemplo para atenuación central, en donde se ilustran las señales de tiempo de entrada (negras) y señales de tiempo de salida (superpuestas en gris).

- 45 La figura 10 ilustra los espectrogramas de la señal de salida. Se puede observar que las señales dirigidas al centro se atenúan, por ejemplo cuando se observa las componentes de sonido transitorio y los tonos sostenidos en el intervalo de baja frecuencia por debajo de 600 Hz y en comparación con la figura 6a. Los sonidos prominentes en la señal de salida corresponden a los instrumentos dirigidos fuera del centro y la reverberación. En particular, la figura 10 ilustra un ejemplo para atenuación central, de forma más particular, espectrogramas de las señales de salida.

- 50 La escucha informal a través de auriculares revela que es efectiva la atenuación de las componentes de señal. Cuando se escucha la señal central extraída, se hacen audibles las distorsiones de procesamiento como modulaciones ligeras durante las notas de la guitarra 2, similar al bombeo en compresión del intervalo dinámico. Se puede señalar que la reverberación se reduce y que la atenuación es más efectiva para bajas frecuencias que para altas frecuencias. Si esto se produce por la relación directa a ambiental más grande en las bajas frecuencias, el contenido de frecuencia de las fuentes de sonido o percepción subjetiva debido a fenómenos de desenmascaramiento no se puede resolver sin un análisis más detallado.

- 60 Cuando se escucha la señal de salida donde se atenúa el centro, la calidad de sonido total es ligeramente mejor cuando se compara con el resultado de extracción central. Las distorsiones de procesamiento son audibles como movimientos ligeros de las fuentes dirigidas hacia el centro cuando están activas las fuentes centradas dominantes, de forma equivalente al bombeo cuando se extrae el centro. La señal de salida suena menos directa como resultado de la cantidad incrementada de sonido ambiental en la señal de salida.

Para ilustrar el filtrado PDC, las figuras 11a-d muestran dos señales de voz que se han mezclado para obtener

señales de entrada con y sin ICTD. En particular, las figuras 11a-d ilustran señales de fuente de entrada para ilustrar la PDC, en donde la figura 11a ilustra señal de fuente 1; en donde la figura 11b ilustra señal de fuente 2; en donde la figura 11c ilustra un canal izquierdo de una señal de mezcla; y en donde la figura 11d ilustra un canal derecho de una señal de mezcla.

5 La señal de mezcla a dos canales se genera al mezclar las señales de fuente de voz con ganancias iguales a cada canal y al sumar ruido blanco con una SNR de 10 dB (K-ponderada) a la señal.

10 Las figuras 12a-c muestran las ponderaciones espectrales calculadas de la función de ganancia (13). En particular, las figuras 12a-c ilustran ponderaciones espectrales $G_{c2}(m, k, 1, 3)$ para demostrar el filtrado PDC, en donde la figura 12a ilustra ponderaciones espectrales para señales de entrada sin ICTD, PDC deshabilitada; la figura 12c ilustra ponderaciones espectrales para señales de entrada con ICTD, PDC deshabilitada; y la figura 12c ilustra ponderaciones espectrales para señales de entrada con ICTD, PDC habilitada.

15 Las ponderaciones espectrales en la gráfica superior están cerca a 0 dB cuando se activa el habla y se asume el valor mínimo en regiones de tiempo-frecuencia con SNR baja. La segunda gráfica muestra las ponderaciones espectrales para una señal de entrada donde la primera señal de voz (Figura 11a) se mezcla con una ICTD de 26 muestras. Las características de filtro de peine se ilustran en la figura 12b. La figura 12c muestra las ponderaciones espectrales cuando se habilita la PDC. Las distorsiones del filtrado de peine se reducen ampliamente, aunque no es perfecta la compensación cerca de las frecuencias de muesca a 848 Hz y 2544 Hz.

20 La escucha informal muestra que el ruido aditivo se atenúa ampliamente. Cuando se procesan las señales sin ICTD, las señales de salida tienen un poco de sonido ambiental característico que resulta probablemente de la incoherencia de fase introducida por el ruido aditivo.

25 Cuando se procesan señales con ICTD, la primera señal de voz (Figura 11a) se atenúa ampliamente y son audibles fuertes distorsiones de filtrado de peine cuando no se aplica el filtrado PDC. Con filtrado PDC adicional, las distorsiones de filtrado de peine siguen ligeramente audibles, pero mucho menos molestas. La escucha informal a otro material revela ligeras distorsiones, que se pueden reducir ya sea al disminuir γ , al incrementar β , o al añadir una versión escalada de la señal de entrada sin procesar a la salida. En general, las distorsiones son menos audibles cuando se atenúa la señal central y son más audibles cuando se extrae la señal central. Las distorsiones de la imagen espacial percibida son muy pequeñas. Esto se puede atribuir al hecho de que las ponderaciones espectrales son idénticas para todas las señales de canal y no afectan las ICLD. Las distorsiones del filtrado de peine apenas son audibles cuando se procesan grabaciones naturales que ofrecen estereofonía de tiempo de llegada para quienes no se somete una mezcla a un canal a fuertes distorsiones del filtrado de peine audibles. Para el filtrado PDC se puede señalar que pequeños valores de la constante de tiempo del cálculo de promedio recursivo (en particular la compensación instantánea de diferencias de fase cuando se calcula X_d) introducen coherencia en las señales utilizadas para la mezcla a dos canales. En consecuencia, el procesamiento es agonístico con respecto al grado de difusión de la señal de entrada. Cuando se incrementa la constante de tiempo, se puede observar que

30 (1) disminuye el efecto de la PDC para señales de entrada con estereofonía de diferencia de amplitud y (2) se hace más audible el efecto de filtrado de peine en los inicios de la nota cuando las fuentes de sonido directo no están alineadas en tiempo entre los canales de entrada.

45 Se han proporcionado los conceptos para el escalado de la señal central en grabaciones de audio al aplicar ponderaciones espectrales de valores reales que se calculan de funciones monótonas de la SDR. La razón fundamental es que el escalamiento de señales centrales necesita tomar en cuenta tanto, el desplazamiento lateral de las fuentes directas como la cantidad de grado de difusión, y que estas características se capturan de forma implícita por la SDR. El procesamiento se puede controlar mediante parámetros de usuario semánticamente significativos y es en comparación con otras técnicas en el dominio de la frecuencia de baja carga de memoria y

50 complejidad computacional. Los conceptos propuestos dan buenos resultados cuando se procesan señales de entrada que ofrecen estereofonía de diferencia de amplitud, pero se pueden someter a distorsiones de filtrado de peine cuando las fuentes de sonido directo no están alineadas en el tiempo entre los canales de entrada. Un primer enfoque para resolver esto es compensar la fase no cero en la función de transferencia inter-canal.

55 Hasta aquí, se han probado los conceptos de las realizaciones por medio de escucha informal. Para grabaciones comerciales convencionales, los resultados son de buena calidad de sonido pero también dependen de la fuerza de separación deseada.

60 Aunque se han descrito algunos aspectos en el contexto de un aparato, es claro que estos aspectos también representan una descripción del método correspondiente, donde un bloque o dispositivo corresponde a una etapa del método o una característica de una etapa del método. De forma análoga, los aspectos descritos en el contexto de una etapa del método también representan una descripción de un bloque o elemento o característica correspondiente de un aparato correspondiente.

La señal descompuesta inventiva también se puede almacenar en un medio de almacenamiento digital o se puede transmitir en un medio de transmisión tal como un medio de transmisión inalámbrica o un medio de transmisión alámbrica tal como internet.

5 Dependiendo de ciertos requerimientos de implementación, las realizaciones de la invención se pueden implementar en hardware o software. La implementación se puede realizar utilizando un medio de almacenamiento digital, por ejemplo un disco flexible, un DVD, un CD, una ROM, una PROM, una EPROM o una memoria flash, que tiene señales de control electrónicamente leíbles almacenadas en el mismo, que cooperan (o pueden cooperar) con un sistema informático programable de tal forma que se realiza el método respectivo.

10 Algunas realizaciones según la invención comprenden un portador de datos no transitorio que tiene señales de control electrónicamente leíbles, que pueden cooperar con un sistema informático programable, de tal forma que se realiza uno de los métodos descritos en el presente documento.

15 En general, las realizaciones de la presente invención se pueden implementar como un producto de programa informático con un código de programa, el código de programa que es operativo para realizar uno de los métodos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede, por ejemplo, almacenarse en un portador leíble por máquina.

20 Otras realizaciones comprenden el programa informático para realizar uno de los métodos descritos en la presente, almacenado en un portador leíble por máquina.

En otras palabras, una realización del método inventivo es, por lo tanto, un programa informático que tiene un código de programa para realizar uno de los métodos descritos en el presente documento, cuando el programa informático se ejecuta en un ordenador.

25 Una realización adicional del método inventivo es, por lo tanto, un portador de datos (o un medio de almacenamiento digital, o un medio leíble por ordenador) que comprende, grabado en el mismo, el programa informático para realizar uno de los métodos descritos en el presente documento.

30 Una realización adicional del método inventivo es, por lo tanto, un flujo de datos o una secuencia de señales que representan el programa informático para realizar uno de los métodos descritos en el presente documento. El flujo de datos o la secuencia de señales pueden, por ejemplo, configurarse para que se transfiera mediante una conexión de comunicación de datos, por ejemplo mediante internet.

35 Una realización adicional comprende un medio de procesamiento, por ejemplo un ordenador, o un dispositivo lógico programable, configurado para o adaptado para realizar uno de los métodos descritos en el presente documento.

40 Una realización adicional comprende un ordenador que tiene instalado en la misma el programa informático para realizar uno de los métodos descritos en el presente documento.

En algunas realizaciones, se puede utilizar un dispositivo lógico programable (por ejemplo, un arreglo de compuertas programables en el campo) para realizar algunas o todas las funcionalidades de los métodos descritos en el presente documento. En algunas realizaciones, un arreglo de compuertas programables en el campo puede cooperar con un microprocesador a fin de realizar uno de los métodos descritos en el presente documento. En general, los métodos se realizan de manera preferible por cualquier aparato de hardware.

45 Las realizaciones descritas anteriormente son simplemente ilustrativas para los principios de la presente invención. Se entiende que serán evidentes modificaciones y variaciones de los arreglos y los detalles descritos en el presente documento para otros expertos en la técnica. Se propone que, por lo tanto, se limite solamente por el alcance de las reivindicaciones de patente inminentes y no por los detalles específicos presentados a manera de descripción y explicación de las realizaciones en el presente documento.

Bibliografía:

55 [1] International Telecommunication Union, Radiocommunication Assembly, "Multichannel stereophonic sound system with and without accompanying picture", Recommendation ITU-R BS.775-2, 2006, Ginebra, Suiza).

60 [2] J. Berg y F. Rumsey, "Identification of quality attributes of spatial sound by repertory grid technique", J. Audio Eng. Soc., vol. 54, pág. 365–379, 2006.

[3] J. Blauert, Spatial Hearing, MIT Press, 1996.

[4] F. Rumsey, "Controlled subjective assessment of two-to-five channel surround sound processing algorithms", J.

Audio Eng. Soc., vol. 47, pág. 563–582, 1999.

[5] H. Fuchs, S. Tuff, y C. Bustad, “Dialogue enhancement - technology and experiments”, EBU Technical Review, vol. Q2, pág. 1–11, 2012.

5 [6] J.-H. Bach, J. Anemüller, y B. Kollmeier, “Robust speech detection in real acoustic backgrounds with perceptually motivated features”, Speech Communication, vol. 53, pág. 690–706, 2011.

10 [7] C. Avendano y J.-M. Jot, “A frequency-domain approach to multi-channel upmix”, J. Audio Eng. Soc., vol. 52, 2004.

[8] D. Barry, B. Lawlor, y E. Coyle, “Sound source separation: Azimuth discrimination and resynthesis”, en Proc. Int. Conf. Digital Audio Effects (DAFx), 2004.

15 [9] E. Vickers, “Two-to-three channel upmix for center channel derivation and speech enhancement”, en Proc. Audio Eng. Soc. 127th Conv., 2009.

[10] D. Jang, J. Hong, H. Jung, y K. Kang, “Center channel separation based on spatial analysis”, en Proc. Int. Conf. Digital Audio Effects (DAFx), 2008.

20 [11] A. Jourjine, S. Rickard, y O. Yilmaz, “Blind separation of disjoint orthogonal signals: Demixing N sources from 2 mixtures”, en Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP), 2000.

25 [12] O. Yilmaz y S. Rickard, “Blind separation of speech mixtures via time-frequency masking”, IEEE Trans. on Signal Proc., vol. 52, pág. 1830–1847, 2004.

[13] S. Rickard, “The DUET blind source separation algorithm”, en Blind Speech Separation, S: Makino, T.-W. Lee, y H. Sawada, Eds. Springer, 2007.

30 [14] N. Cahill, R. Cooney, K. Humphreys, y R. Lawlor, “Speech source enhancement using a modified ADRes algorithm for applications in mobile communications”, en Proc. Audio Eng. Soc. 121st Conv., 2006.

[15] M. Puigt e Y. Deville, “A time-frequency correlation-based blind source separation method for time-delay mixtures”, en Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP), 2006.

35 [16] Simon Arberet, Remi Gribonval, y Frederic Bimbot, “A robust method to count and locate audio sources in a stereophonic linear anechoic mixture”, en Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP), 2007.

[17] M.I. Mandel, R.J. Weiss, y D.P.W. Ellis, “Model-based expectation-maximization source separation and localization”, IEEE Trans. on Audio, Speech and Language Proc., vol. 18, pág. 382–394, 2010.

[18] H. Viste y G. Evangelista, “On the use of spatial cues to improve binaural source separation”, en Proc. Int. Conf. Digital Audio Effects (DAFx), 2003.

45 [19] A. Favrot, M. Erne, y C. Faller, “Improved cocktail-party processing”, en Proc. Int. Conf. Digital Audio Effects (DAFx), 2006.

[20] Patente de los Estados Unidos 7.630.500 B1, P.E. Beckmann, 2009

50 [21] Patente de los Estados Unidos 7.894.611 B2, P.E. Beckmann, 2011

[22] J.B. Allen, D.A. Berkeley, y J. Blauert, “Multimicrophone signal-processing technique to remove room reverberation from speech signals”, J. Acoust. Soc. Am., vol. 62, 1977.

55 [23] J. Merimaa, M. Goodwin, y J.-M. Jot, “Correlation-based ambience extraction from stereo recordings”, en Proc. Audio Eng. Soc. 123rd Conv., 2007.

[24] J. Usher y J. Benesty, “Enhancement of spatial sound quality: A new reverberation-extraction audio upmixer”, IEEE Trans. on Audio, Speech, and Language Processing, vol. 15, pág. 2141–2150, 2007.

60 [25] C. Faller, “Multiple-loudspeaker playback of stereo signals”, J. Audio Eng. Soc., vol. 54, 2006.

[26] C. Uhle, A. Walther, O. Hellmuth, y J. Herre, “Ambience separation from mono recordings using Non-negative Matrix Factorization”, en Proc. Audio Eng. Soc. 30th Int. Conf., 2007.

[27] C. Uhle y C. Paul, "A supervised learning approach to ambience extraction from mono recordings for blind upmixing", en Proc. Int. Conf. Digital Audio Effects (DAFx), 2008.

5 [28] G. Soulodre, "System for extracting and changing the reverberant content of an audio input signal", Patente de los Estados Unidos 8.036.767, Oct. 2011.

[29] International Telecommunication Union, Radiocommunication Assembly, "Algorithms to measure audio programme loudness and true-peak audio level", Recommendation ITUR BS.1770-2, Marzo 2011, Ginebra, Suiza.

10

REIVINDICACIONES

1. Aparato para generar una señal de audio modificada que comprende dos o más canales de audio modificados de una señal de entrada de audio que comprende dos o más canales de entrada de audio, en donde el aparato comprende:

un generador (110) de información para generar información de señal a mezcla a dos canales, en donde el generador (110) de información se adapta para generar información de señal al combinar un valor espectral de cada uno de los dos o más canales de entrada de audio de una primera forma, en donde el generador (110) de información se adapta para generar información de mezcla a dos canales al combinar el valor espectral de cada uno de los dos o más canales de entrada de audio de una segunda forma que es diferente de la primera forma, y en donde el generador (110) de información se adapta para combinar la información de señal y la información de mezcla a dos canales para obtener la información de señal a mezcla a dos canales, y

un atenuador (120) de señal para atenuar los dos o más canales de entrada de audio dependiendo de la información de señal a mezcla a dos canales para obtener los dos o más canales de audio modificados,

en donde el generador (110) de información se configura para procesar el valor espectral de cada uno de los dos o más canales de entrada de audio al multiplicar dicho valor espectral por el complejo conjugado de dicho valor espectral para obtener una auto densidad espectral de potencia de dicho valor espectral para cada uno de los dos o más canales de entrada de audio,

en donde el generador (110) de información se configura para combinar el valor espectral de cada uno de los dos o más canales de entrada de audio para obtener un valor combinado, y en donde el generador (110) de información se configura para procesar el valor combinado al determinar una densidad espectral de potencia del valor combinado,

caracterizado porque el generador (110) de información se configura para generar la información de señal $s(m, k, \beta)$ según la fórmula

$$s(m, k, \beta) = \sum_{i=1}^N \Phi_{i,i}(m, k)^\beta ,$$

en donde N indica el número de canales de entrada de audio de la señal de entrada de audio,

en donde $\Phi_{i,i}(m, k)$ indica la auto densidad espectral de potencia del valor espectral del i-ésimo canal de señal de audio,

en donde β es un número real con $\beta > 0$,

en donde m indica un índice de tiempo, y en donde k indica un índice de frecuencia.

2. Aparato según la reivindicación 1, en donde el generador (110) de información se configura para combinar la información de señales y la información de mezcla a dos canales de modo que la información de mezcla a dos canales indica una relación de la información de señales con respecto a la información de mezcla a dos canales.

3. Aparato según la reivindicación 1 ó 2, en donde el número de los canales de audio modificados es igual al número de los canales de entrada de audio, o en donde el número de los canales de audio modificados es más pequeño que el número de los canales de entrada de audio.

4. Aparato según una de las reivindicaciones anteriores,

en donde el generador (110) de información se configura para procesar el valor espectral de cada uno de los dos o más canales de entrada de audio para obtener dos o más valores procesados, y en donde el generador (110) de información se configura para combinar los dos o más valores procesados para obtener la información de señal, y

en donde el generador (110) de información se configura para procesar el valor combinado para obtener la información de mezcla a dos canales.

5. Aparato según una de las reivindicaciones anteriores,

en donde el generador (110) de información se configura para determinar una relación de señal a mezcla a dos canales como la información de señal a mezcla a dos canales según la fórmula $R(m, k, \beta)$

5

$$R(m, k, \beta) = \left(\frac{\sum_{i=1}^N \Phi_{i,i}(m, k)^\beta}{\Phi_d(m, k)^\beta} \right)^{\frac{1}{2\beta-1}}$$

en donde $\Phi_d(m, k)$ indica la densidad espectral de potencia del valor combinado, y

10

en donde $\Phi_d(m, k)^\beta$ es la información de mezcla a dos canales.

6. Aparato según la reivindicación 5, en donde el atenuador (120) de señal se adapta para atenuar los dos o más canales de entrada de audio dependiendo de una función de ganancia $G(m, k)$ según la fórmula

15

$$Y(m, k) = G(m, k)X(m, k),$$

en donde la función de ganancia $G(m, k)$ depende de la información de señal a mezcla a dos canales, y en donde la función de ganancia $G(m, k)$ es una función monótonamente creciente de la información de señal a mezcla a dos canales o una función monótonamente decreciente de la información de señal a mezcla a dos canales,

20

en donde $X(m, k)$ indica la señal de entrada de audio,

25

en donde $Y(m, k)$ indica la señal de audio modificada,

en donde m indica un índice de tiempo, y

en donde k indica un índice de frecuencia.

30

7. Aparato según la reivindicación 6,

en donde la función de ganancia $G(m, k)$ es una primera función $G_{c1}(m, k, \beta, \gamma)$, una segunda función $G_{c2}(m, k, \beta, \gamma)$, una tercera función $G_{s1}(m, k, \beta, \gamma)$, o una cuarta función $G_{s2}(m, k, \beta, \gamma)$,

35

en donde

$$G_{c1}(m, k, \beta, \gamma) = (1 + R_{\min} - R(m, k, \beta))^\gamma,$$

en donde

40

$$G_{c2}(m, k, \beta, \gamma) = \left(\frac{R_{\min}}{R(m, k, \beta)} \right)^\gamma$$

en donde

45

$$G_{s1}(m, k, \beta, \gamma) = R(m, k, \beta)^\gamma$$

en donde

$$G_{s_2}(m, k, \beta, \gamma) = \left(1 + R_{\min} - \frac{R_{\min}}{R(m, k, \beta)}\right)^\gamma$$

en donde β es un número real con $\beta > 0$,

5 en donde γ es un número real con $\gamma > 0$, y

en donde R_{\min} indica el mínimo de R .

8. Sistema que comprende:

10 un compensador (210) de fase para generar una señal de audio compensada en fase que comprende dos o más canales de audio compensados en fase de una señal de audio sin procesar que comprende dos o más canales de entrada de audio sin procesar, y

15 un aparato (220) según una de las reivindicaciones anteriores para recibir la señal de audio compensada en fase como una señal de entrada de audio y para generar una señal de audio modificada que comprende dos o más canales de audio modificados de la señal de entrada de audio que comprende los dos o más canales de audio compensados en fase como dos o más canales de entrada de audio,

20 en donde uno de los dos o más canales de entrada de audio sin procesar es un canal de referencia,

en donde el compensador (210) de fase se adapta para estimar cada canal de audio sin procesar de los dos o más canales de audio sin procesar que no son el canal de referencia una función de transferencia de fase entre dicho canal de audio sin procesar y el canal de referencia, y

25 en donde el compensador (210) de fase se adapta para generar la señal de audio compensada en fase al modificar cada canal de audio sin procesar de los canales de audio sin procesar que no son el canal de referencia dependiendo de la función de transferencia de fase de dicho canal de audio sin procesar.

30 9. Método para generar una señal de audio modificada que comprende dos o más canales de audio modificados de una señal de entrada de audio que comprende dos o más canales de entrada de audio, en donde el método comprende:

35 generar información de señal al combinar un valor espectral de cada uno de los dos o más canales de entrada de audio de una primera forma,

generar información de mezcla a dos canales al combinar el valor espectral de cada uno de los dos o más canales de entrada de audio de una segunda forma que es diferente de la primera forma,

40 generar información de señal a mezcla a dos canales al combinar la información de señal y la información de mezcla a dos canales, y

atenuar los dos o más canales de entrada de audio dependiendo de la información de señal a mezcla a dos canales para obtener los dos o más canales de audio modificados,

45 en donde el método comprende además:

50 procesar el valor espectral de cada uno de los dos o más canales de entrada de audio al multiplicar dicho valor espectral por el complejo conjugado de dicho valor espectral para obtener una auto densidad espectral de potencia de dicho valor espectral para cada uno de los dos o más canales de entrada de audio,

55 combinar el valor espectral de cada uno de los dos o más canales de entrada de audio para obtener un valor combinado, y procesar el valor combinado al determinar una densidad espectral de potencia del valor combinado,

estando el método caracterizado por generar la información de señal $s(m, k, \beta)$ según la fórmula

$$s(m, k, \beta) = \sum_{i=1}^N \Phi_{i,i}(m, k)^\beta,$$

en donde N indica el número de canales de entrada de audio de la señal de entrada de audio,

5 en donde $\Phi_i(m, k)$ indica la auto densidad espectral de potencia del valor espectral del i-ésimo canal de señal de audio,

en donde β es un número real con $\beta > 0$,

10 en donde m indica un índice de tiempo, y en donde k indica un índice de frecuencia.

10. Programa informático para implementar el método de la reivindicación 9, cuando se ejecuta en un ordenador o procesador de señales.

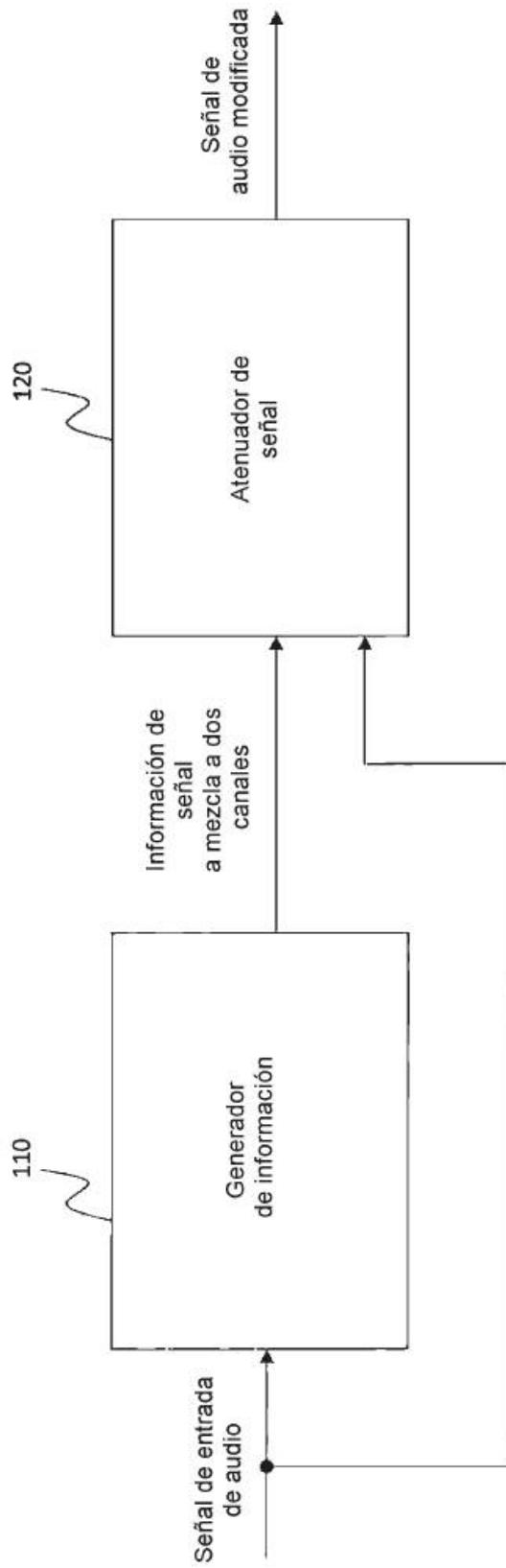


FIG 1

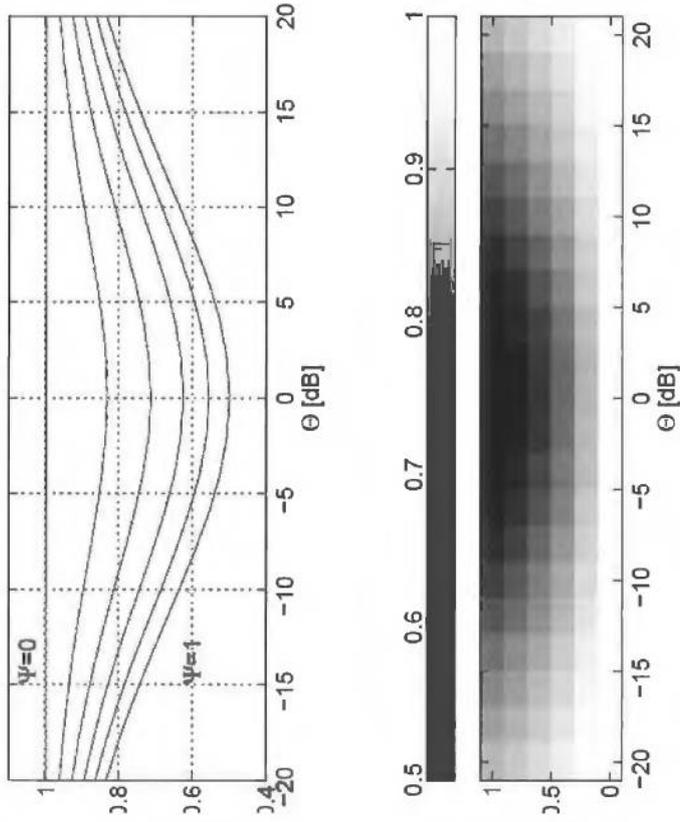


FIG 2

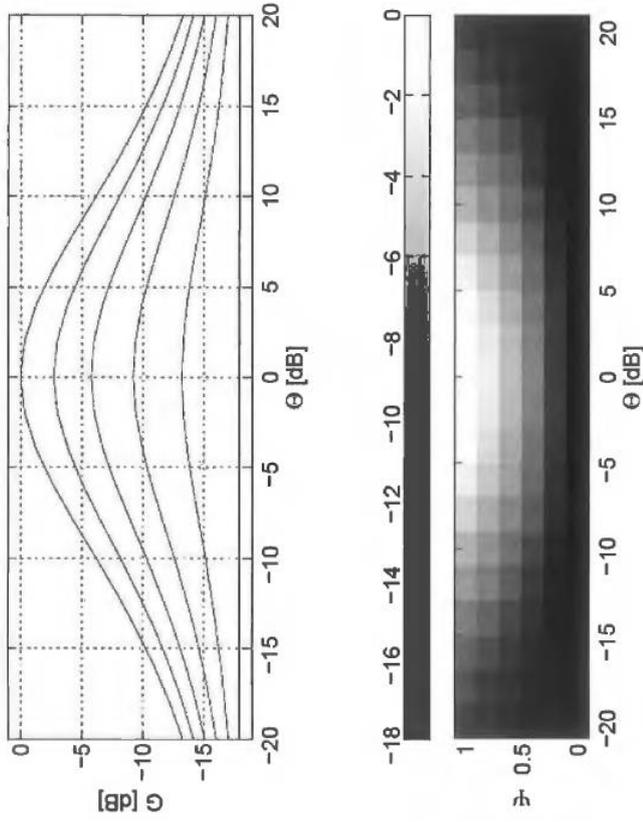


FIG 3

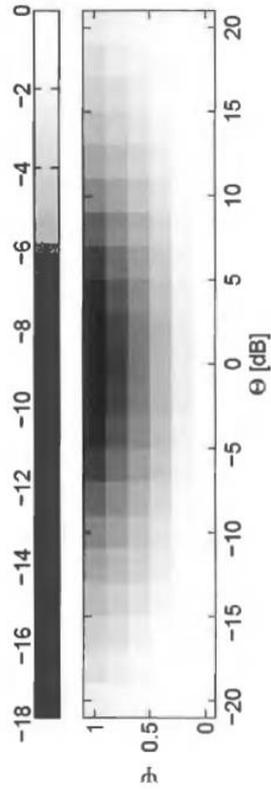
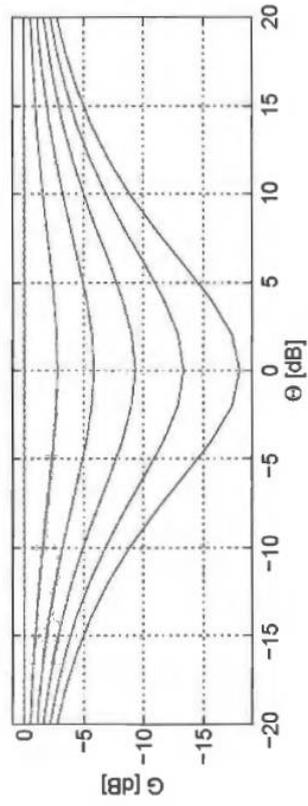


FIG 4

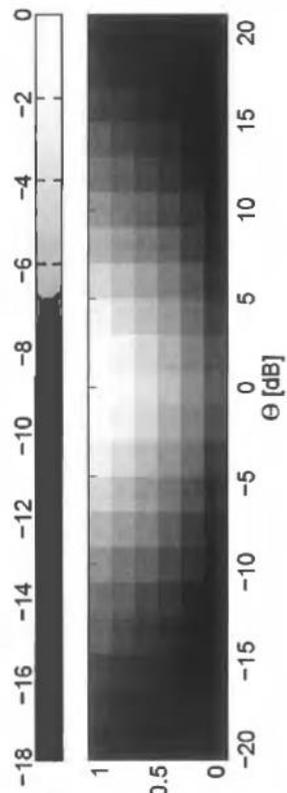
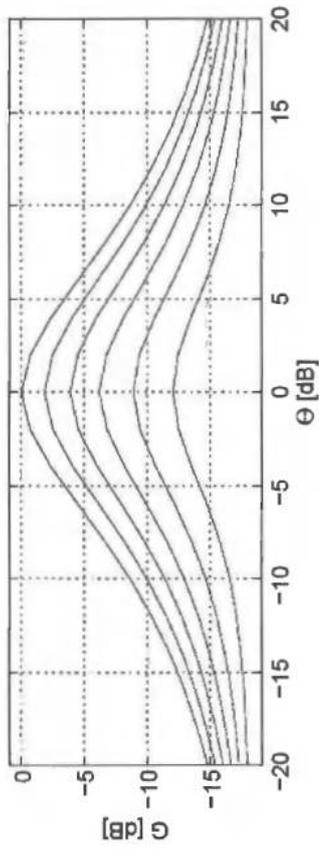


FIG 5

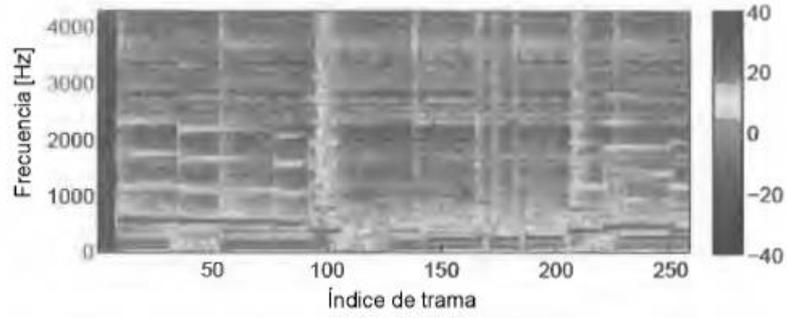


FIG 6A

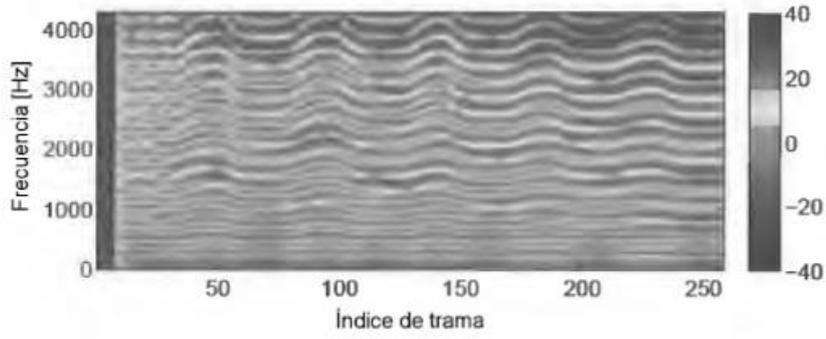


FIG 6B

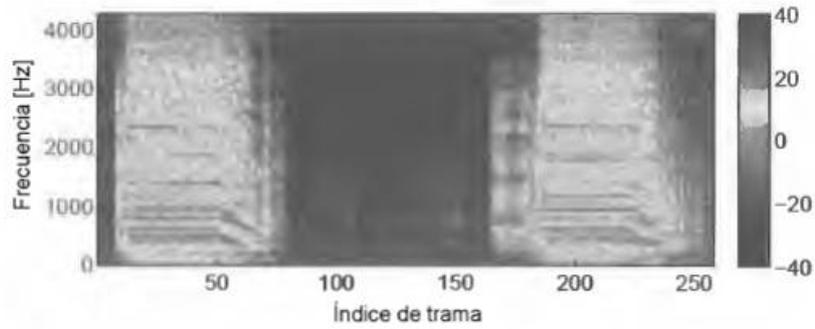


FIG 6C

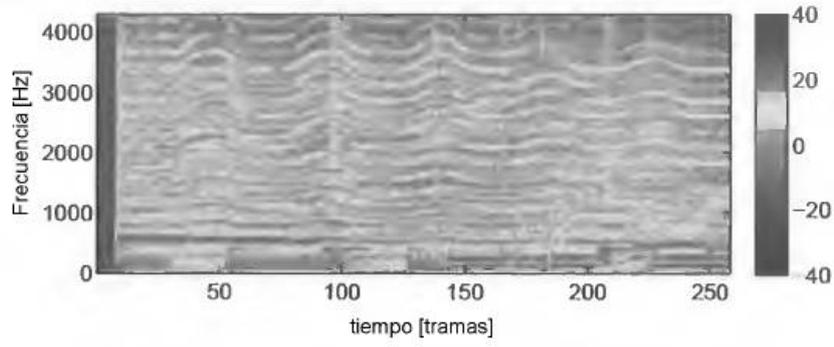


FIG 6D

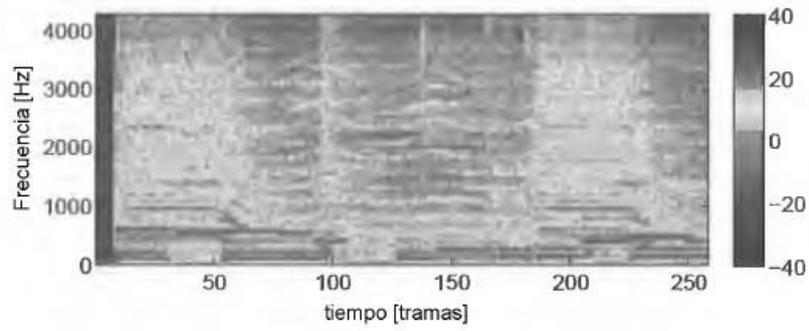


FIG 6E

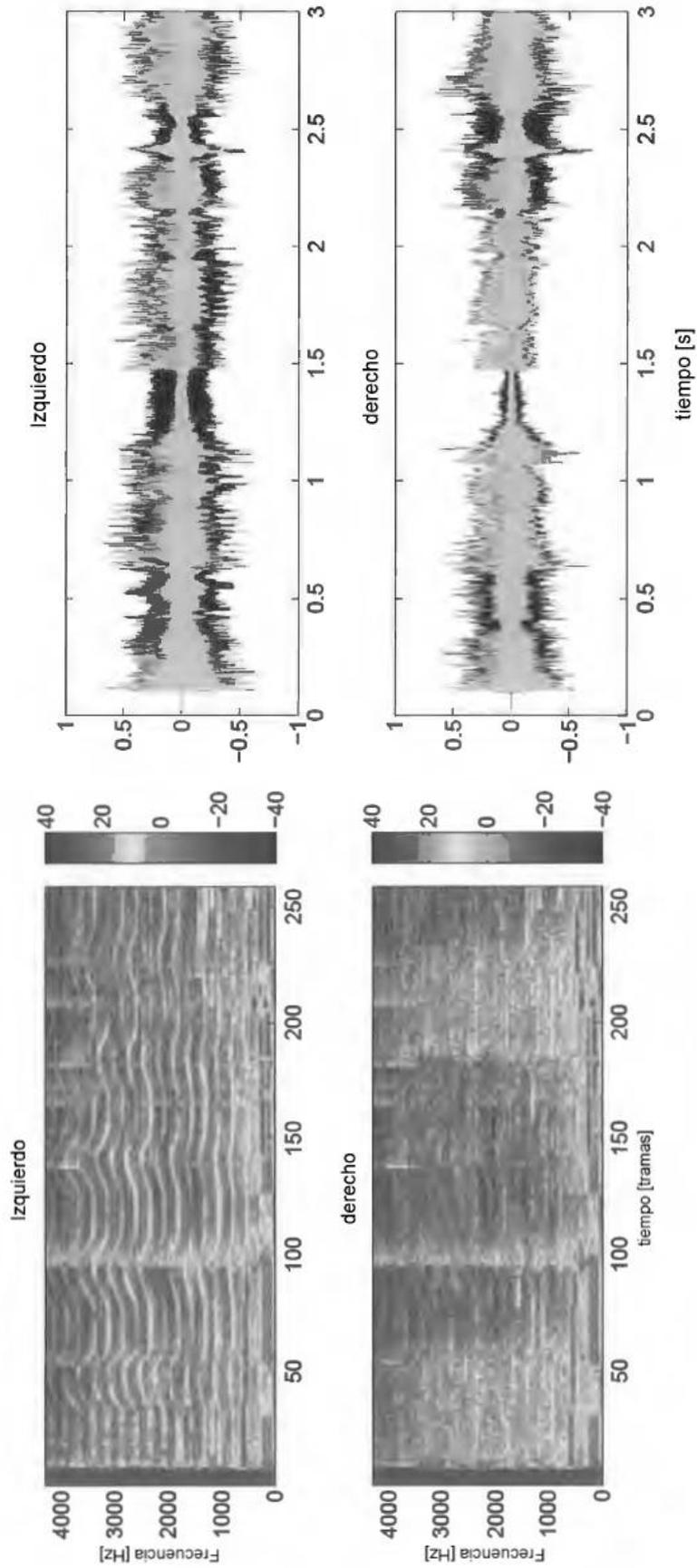


FIG 8

FIG 7

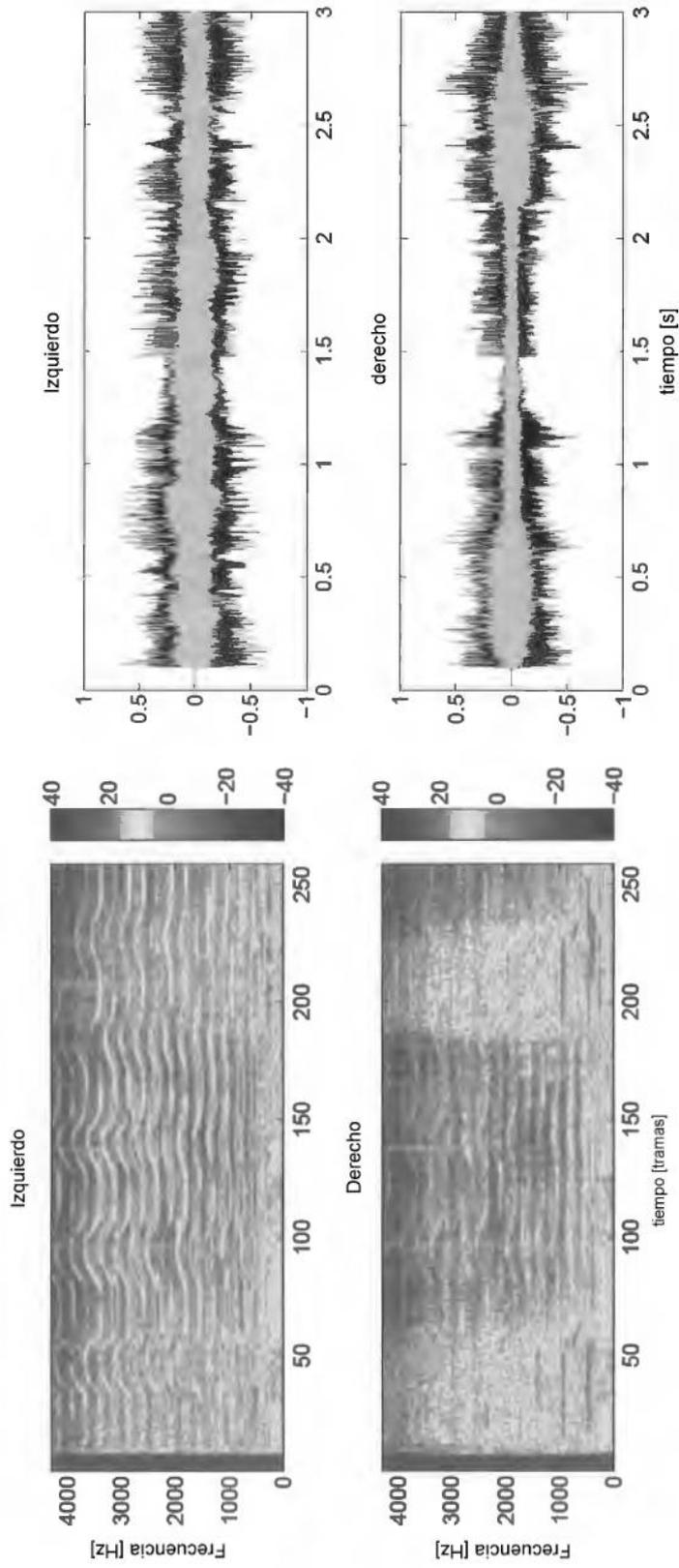


FIG 10

FIG 9

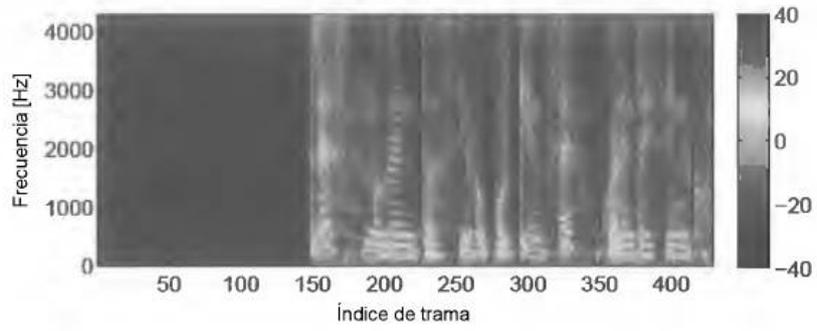


FIG 11A

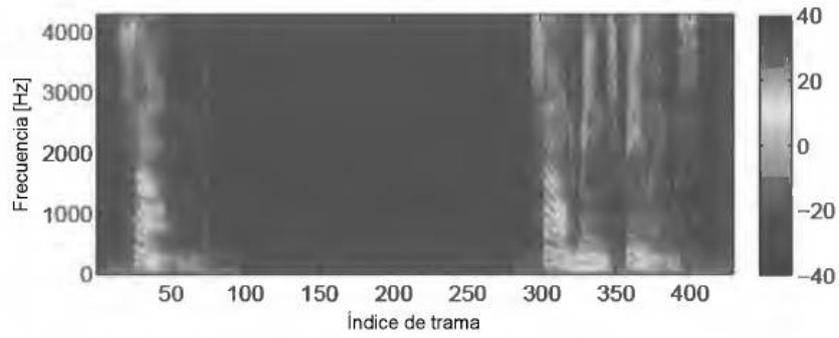


FIG 11B

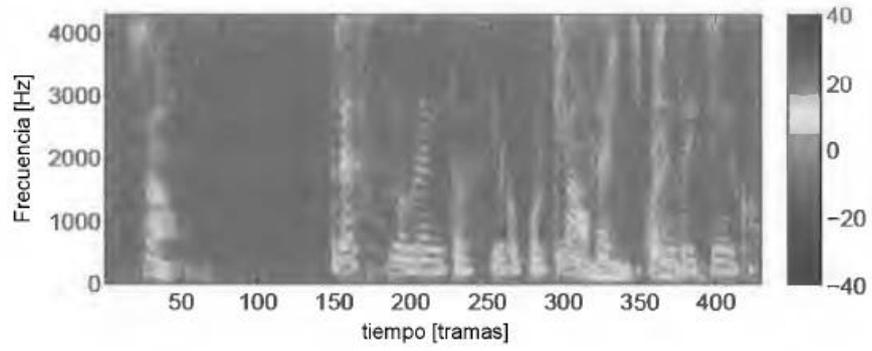


FIG 11C

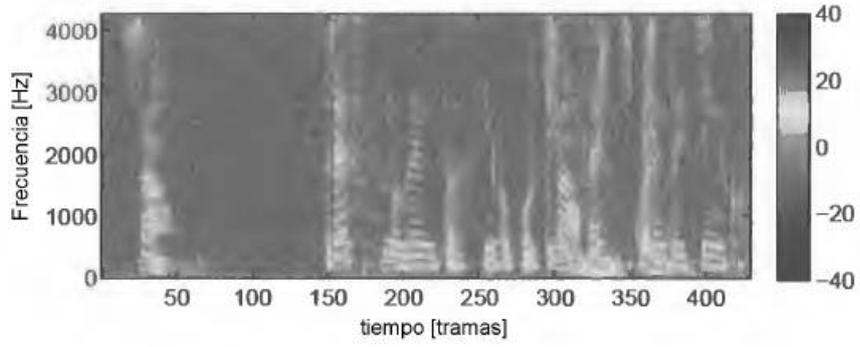


FIG 11D

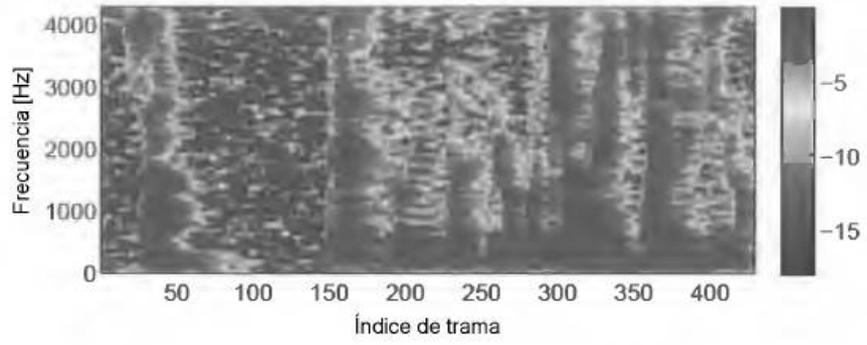


FIG 12A

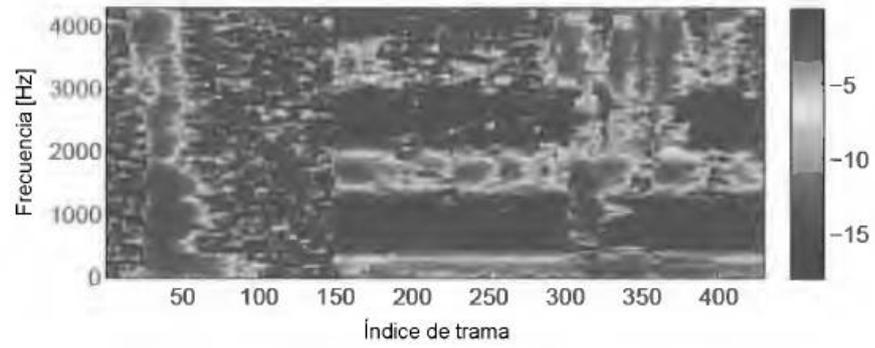


FIG 12B

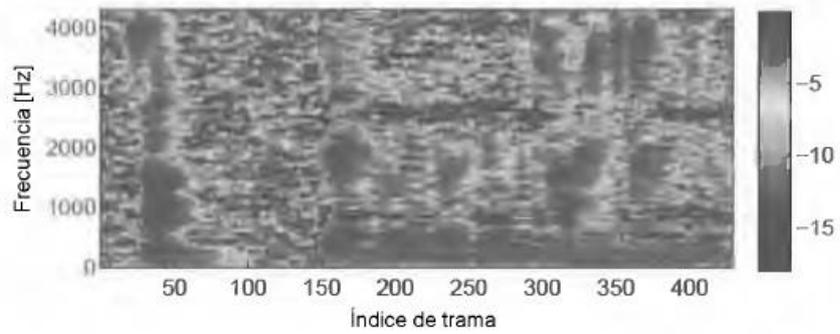


FIG 12C

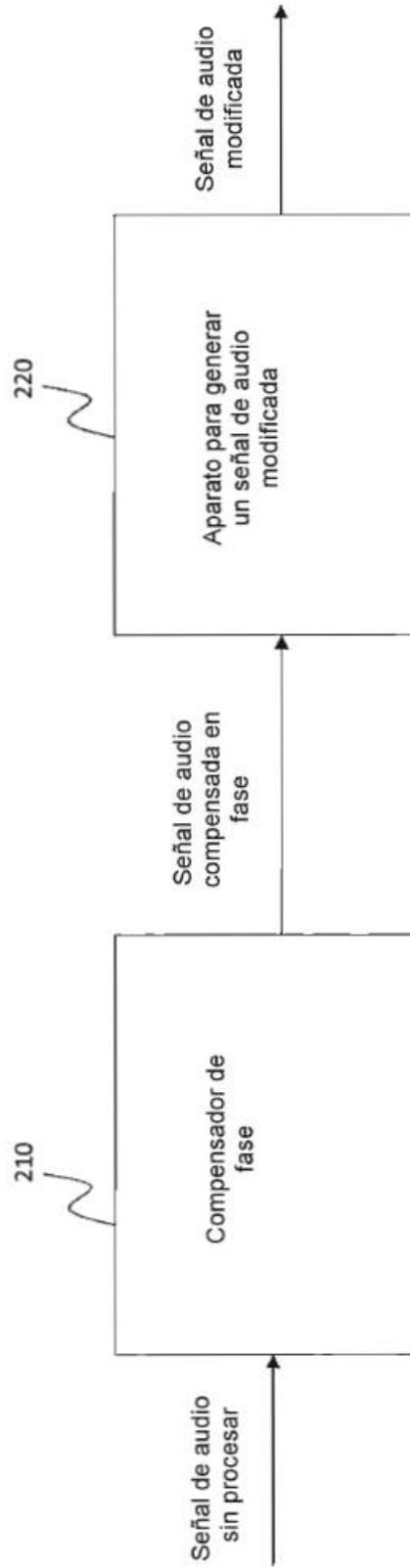


FIG 13