

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 841 302**

51 Int. Cl.:

G11B 27/28 (2006.01)

G10L 19/02 (2013.01)

G11B 27/00 (2006.01)

G10L 21/04 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **17.05.2017 PCT/EP2017/061895**

87 Fecha y número de publicación internacional: **23.11.2017 WO17198737**

96 Fecha de presentación y número de la solicitud europea: **17.05.2017 E 17726861 (2)**

97 Fecha y número de publicación de la concesión europea: **04.11.2020 EP 3459078**

54 Título: **Aparato y método para procesar una señal de audio multicanal**

30 Prioridad:

20.05.2016 EP 16170723

14.07.2016 EP 16179531

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
08.07.2021

73 Titular/es:

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

72 Inventor/es:

**UHLE, CHRISTIAN;
KRATZ, MICHAEL;
KLOSE, PAUL;
LEONARD, TIMOTHY;
LUVIZOTTO, ANDRÉ y
SCHARRER, SEBASTIAN**

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 841 302 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 procesar una señal de audio multicanal

5 La invención se refiere a un aparato para procesar una señal de audio multicanal. La señal de audio multicanal comprende una pluralidad de es decir, al menos dos señales de canal. El aparato realiza una modulación en escala de tiempo de la señal de audio multicanal. La invención también se refiere a un método correspondiente y a un programa informático.

10 La modificación en escala de tiempo (TSM) se refiere al procesamiento para desacelerar o acelerar la reproducción de una señal de audio sin afectar a su tono. La TSM en combinación con la conversión de la tasa de muestra también permite cambiar el tono sin modificar el ritmo. El reto de la TSM es mantener todas las otras características de la señal de audio (excepto ya sea el ritmo o el tono) y, en particular, la calidad de sonido. El procesamiento no debe producir distorsiones audibles.

15 Para señales de entrada de canal individual, la característica más importante es el timbre. Para señales que tienen más de un canal, también se necesitan mantener las características espaciales. Las características espaciales comprenden la posición y ancho de la fuente de sonido directo y la difusividad del sonido ambiente. Pueden cuantificarse por diferencias a nivel entre canales (ICLD), diferencias de tiempo entre canales (ICTD), diferencias de fase entre canales (ICPD) y coherencia entre canales (ICC).

Existen dos enfoques fundamentalmente diferentes a la modificación en escala de tiempo. Uno se aplica en el dominio de tiempo y el otro se aplica en el dominio de frecuencia.

25 El procesamiento en el dominio de tiempo usa un esquema de superposición-adición sincronizadas (SOLA). La señal se corta en tramas de superposición y estas tramas se cambian y combinan para estirar o encoger la señal. La posición de cambio se calcula normalmente al maximizar una medida de similitud, por ejemplo, correlación entre la trama de señal y su copia cambiada.

30 Este método en el dominio de tiempo es de baja complejidad computacional. Produce buenos resultados para señales monofónicas (en contraste a las polifónicas), por ejemplo, voz o tonos de flauta, debido a que el desplazamiento de cambio puede determinarse como un múltiplo entero del período de la frecuencia fundamental a fin de evitar discontinuidades e interferencias destructivas en la señal de salida. En otras palabras, las tramas de señal cambiadas se añaden de una manera coherente en fase. Para entradas polifónicas con múltiples tonos que tienen diferentes frecuencias fundamentales, el cambio no puede determinarse de manera que se cumpla similaridad de forma de onda para todas las tonalidades (tonos).

35 Para muchas señales musicales, los mejores resultados en términos de la calidad de sonido se obtienen al aplicar el procesamiento en el dominio de frecuencia. Este método usa, por ejemplo, el esquema de codificador vocal de fase [1] tal como se muestra en el diagrama de bloques mostrado en la figura 1 y explicado brevemente a continuación.

La señal de audio de entrada $x(n)$ se transforma al dominio de frecuencia usando una transformada de Fourier de tiempo corto (STFT). De manera equivalente, puede usarse otro tipo de banco de filtros o transformada para lo cual puede aplicarse el procesamiento inverso con error de reconstrucción suficientemente pequeño.

45 En la realización mencionada, la señal de entrada $x(n)$ se corta en tramas de superposición y se calcula una transformada de Fourier discreta (DFT) para cada trama según la siguiente ecuación (1), produciendo una representación de la señal de transformada de Fourier de tiempo corto (STFT), también denominada coeficientes STFT (o coeficientes espectrales), según

$$X(m, k) = \sum_{n=0}^{N-1} x(mR_a + n)w_a(n)e^{-j\Omega_k n} \tag{1}$$

50 Un índice de trama de tiempo se denota por m , k es un índice de frecuencia discreta con $0 \leq k \leq N - 1$, y w_a es una función de ventana. La frecuencia angular normalizada Ω_k se da por $\Omega_k = 2 \pi k / N$. La DFT tiene tamaño N y R_a es un tamaño de salto de análisis.

Los índices para el tiempo y frecuencia se omiten en la descripción cuando sea posible por brevedad.

55 La señal de dominio de tiempo de salida $y(n)$ se calcula a partir de los coeficientes espectrales de salida $Y(m, k)$ en la fase de síntesis por medio de la inversa de la STFT que se realiza en dos etapas:

Primero, se calcula una transformada de Fourier discreta inversa para cada una de las M tramas según

$$y_m(n) = \frac{1}{N} \sum_{k=0}^{N-1} Y(m, k) e^{j\Omega_k n} \quad (2)$$

5 Segundo, se aplica un procedimiento de superposición-adición según

$$y(n) = \sum_{m=0}^M w_s(n - mR_s) y_m(n - mR_s) \quad (3)$$

10 con ventana de síntesis opcional $w_s(n)$ y el tamaño de salto de síntesis R_s .

Se logra una modificación en escala de tiempo al ajustar el tamaño de salto de síntesis R_s y el tamaño de salto de análisis R_a a diferentes valores: la señal se estira en tiempo si $R_a < R_s$ y se encoje si $R_a > R_s$.

15 La ventana de análisis w_a y la ventana de síntesis w_s se eligen de manera que si $R_a = R_s$ e $Y(m, k) = X(m, k)$, entonces la señal de entrada y de salida son idénticas.

Los coeficientes de transformada de Fourier de tiempo corto $X(m, k)$ para señales de entrada de valor real (que es el caso para señales de audio consideradas en el presente documento) son números de valor complejo que pueden expresarse en coordenadas polares por su magnitud $|X|$ y fase Φ_x como

$$20 \quad X = |X| \exp(j\Phi_x), \quad (4)$$

$$\Phi_x = \arg X, \quad (5)$$

25 donde $j = \sqrt{-1}$.

30 Si los dos tamaños de salto R_a y R_s difieren, es decir, si $R_a \neq R_s$, entonces las fases de $Y(m, k)$ necesitan modificarse de manera que se logra la "coherencia de fase horizontal". Esto significa que para una senoide de frecuencia constante, se superponen tramas sucesivas coherentemente sin discontinuidades ni cancelación de fase (interferencia destructiva).

35 El enfoque de codificador vocal de fase es apropiado para entradas polifónicas, por ejemplo, grabaciones musicales. Su desventaja es que la modificación de la fase puede producir una distorsión conocida como "borrosidad transitoria", es decir, la envoltura temporal de la señal se modifica de manera que se perciben inicios de notas como que tienen menos ataque y sonido menos percusivo. Puede aplicarse procesamiento adicional a la fase de salida para mitigar la borrosidad transitoria, por ejemplo, al aplicar un método llamado "bloqueo de fase" [2] o al reajustar la fase durante períodos de silencio [3].

40 Después de un procedimiento adecuado para modificar la fase, se obtiene una salida. Los coeficientes espectrales de la salida pueden escribirse con coordenadas polares como $Y = |Y| \exp(j\Phi_y)$ donde Φ_y denota la fase modificada. El procedimiento de cálculo de la fase Φ_y se denomina a continuación adaptación de fase (PA).

Se conocen varias maneras de procesar señales de entrada de audio de dos canales:

45 Una opción es realizar una mezcla descendente de la señal multicanal a una señal de un solo canal, es decir, añadir versiones a escala de todos los canales y procesar la señal de un solo canal (mono). El procesamiento de una mezcla descendente mono de la señal de entrada tiene la desventaja de que la información estereofónica se pierde y de este modo se reduce la calidad de sonido.

50 Una opción diferente es procesar de manera independiente las señales de canal de entrada separadas. La principal desventaja del procesamiento de forma separada de cada señal de canal es que se introduce decorrelación arbitraria entre los canales que distorsiona la imagen estéreo. Debido a que la adaptación de fase de la modificación en escala de tiempo es un procesamiento dependiente de la señal, las relaciones entre las fases de los canales individuales no se conservan cuando son diferentes las correspondientes señales de canal. Las distorsiones de la información espacial pueden percibirse como difusión o ampliación de la imagen estéreo de fuentes de sonido directas (cantantes o solistas, por ejemplo).

La solicitud de patente WO 2008/046967 A1, NOKIA CORP [FI], OJALA PASI [FI], LAKANIEMI ARI [FI], VIROLAINEN JUSSI, 24 de abril de 2008, da a conocer el escalamiento de tiempo de señales de audio multicanal, que comprende el almacenamiento temporal de una señal de entrada de audio codificada que comprende al menos una señal de "suma" combinada de una pluralidad de canales de audio, y uno o más conjuntos correspondientes de parámetros de "información secundaria" que describen una imagen de sonido multicanal, es decir, indicaciones espaciales entre canales salientes codificadas de manera compacta como "información secundaria de BCC". La longitud de al menos una trama de audio se cambia añadiendo o eliminando un segmento de la señal combinada, y el uno o más conjuntos de parámetros de información secundaria se modifican en relación con el cambio de longitud de la al menos una trama de audio de la señal combinada.

Un objeto de la invención es mejorar los métodos existentes de modificación en escala de tiempo con respecto a la calidad de las características espaciales --por ejemplo, la imagen estéreo de la señal procesada.

El objeto se logra por un aparato así como por un método.

El objeto se logra por un aparato para procesar una señal de audio multicanal. La señal de audio multicanal comprende una pluralidad de, o al menos dos, señales de canal. El aparato se configura para realizar una modulación en escala de tiempo de la señal de audio multicanal, es decir, la señal de audio multicanal se desacelera o acelera sin afectar a su tono. El aparato comprende un adaptador de fase y un separador. El adaptador de fase se configura para proporcionar al menos una señal procesada al modificar una fase de una señal basada en una combinación de las señales de canal. El separador se configura para proporcionar señales separadas basadas en la al menos una señal procesada.

La invención mejora el procesamiento de señales y especialmente señales de audio o voz. Especialmente se afronta el problema de procesar señales estereofónicas de entrada que tienen dos o más canales.

Una ventaja de la invención es que las características espaciales de la señal de audio de entrada se conservan de manera que no se distorsiona la imagen estéreo percibida. En particular, las posiciones de las fuentes de sonido y la difusividad no se cambian debido al procesamiento TSM.

Por tanto, la invención afronta el problema de que las características espaciales de la señal de salida se distorsionan severamente en el estado de la técnica, que es más perceptible cuando se escuchan fuentes de sonido panoramizadas al centro de la imagen estéreo.

El aparato realiza la modificación en escala de tiempo de la señal de audio multicanal al realizar una adaptación de fase. En el estado de la técnica, se conocen diferentes procedimientos para modificar la fase (incluyendo el bloqueo de fase y otros medios).

Una realización para modificar la fase comprende modificar la fase de manera que la propagación de fase entre tramas adyacentes con el tamaño de salto de síntesis R_s es idéntico a la propagación de fase de la señal de entrada para el tamaño de salto de análisis R_a . Esto asegura que se mantenga la coherencia de fase horizontal (es decir, la evolución temporal de las fases en cada intervalo de frecuencia). Esto se logra al calcular la frecuencia instantánea en la trama de tiempo m dada la fase de entrada de la corriente y la trama de tiempo anterior y el tamaño de salto de análisis R_a y los parámetros STFT (transformada de Fourier de tiempo corto). La propagación de fase deseada se calcula usando la frecuencia instantánea y el tamaño de salto de síntesis R_s .

En una realización adicional, se mejora el método anterior por un método que comprende "bloqueo de fase". El bloqueo de fase tiene como finalidad mejorar la coherencia de fase vertical, es decir, mantener la relación de la fase entre intervalos adyacentes de frecuencia en cada trama. Esto mejora la calidad de sonido, puesto que se percibe, por ejemplo, cuando se procesan señales musicales que contienen transitorios o notas percusivas.

El adaptador de fase se configura para adaptar la fase de al menos una combinación de las señales de canal comprendidas por la señal de audio multicanal. Para esto, las señales que se van a procesar por el adaptador de fase se dan con coordenadas polares por una magnitud y una fase.

El separador después del adaptador de fase proporciona señales separadas basándose en la señal procesada, es decir, basándose en la señal con una fase modificada. El separador invierte la combinación de señales y extrae o genera señales separadas.

Por lo tanto, el aparato modifica la fase de al menos una combinación de señales de canal y proporciona señales modificadas individuales al separar la combinación adaptada en fase de señales de canal en señales separadas.

La adaptación de fase de la invención es más adecuada como una extensión a la modificación en escala de tiempo

que usa el codificador vocal de fase, es decir, procesamiento en el dominio de frecuencia. Sin embargo, también puede aplicarse como una extensión a TSM en el dominio de tiempo. Para este fin, la TSM en el dominio de tiempo se configura para procesar la al menos una señal que se obtiene al combinar las señales de canal comprendidas por la señal de audio multicanal. El separador después de la TSM de dominio de tiempo proporciona señales separadas basadas en la señal procesada.

En una realización, el adaptador de fase se configura para proporcionar N señales procesadas al modificar fases de N señales basándose en combinaciones de N señales de canal. El separador se configura para proporcionar N señales separadas basándose en las señales procesadas. N es un número de señales de canal comprendido por la señal de audio multicanal y es un número entero mayor que o igual a dos. En una realización, N es el número de todas las señales de canal comprendidas por la señal de audio multicanal. Por lo tanto, se convierten N señales de canal en N señales combinadas que, después de la adaptación de fase, se separan en N señales.

Según una realización, las N combinaciones de las señales de canal son combinaciones lineales de las señales de canal. En una realización, el adaptador de fase maneja diversas combinaciones de señales de canal, mientras que las combinaciones son combinaciones lineales, por ejemplo, la suma o la diferencia de señales de canal.

En una realización, N es igual a dos ($N = 2$) y las dos combinaciones de las dos señales de canal son una señal de diferencia y suma. En una realización diferente, N es mayor que dos.

En una realización adicional, el aparato comprende un transformador. El transformador se configura para proporcionar señales transformadas al transformar señales del dominio de tiempo al dominio de frecuencia. El transformador permite transformar señales del dominio de tiempo al dominio de frecuencia y, posteriormente, procesar estas señales en el dominio de frecuencia. Las señales que se van a transformar son, en una realización, las señales de canal comprendidas por la señal de audio multicanal.

Según una realización, el transformador se configura para proporcionar señales transformadas al aplicar una transformación de Fourier de tiempo corto. La transformada de Fourier de tiempo corto o, de manera alternativa, la transformada de Fourier de término reducido (STFT) es una transformada relacionada con Fourier de secciones locales de una señal conforme cambia con la etapa del tiempo. En la práctica, una señal más larga se divide en segmentos más cortos de igual longitud y la transformada de Fourier se calcula de manera separada en cada segmento.

En una realización adicional, el aparato comprende un combinador, en el que el combinador se configura para proporcionar señales combinadas basándose en las señales de canal. El transformador se configura para proporcionar señales transformadas basándose en las señales combinadas al aplicar una transformación. En el presente documento, las señales de canal se combinan y las combinaciones se transforman en el dominio de frecuencia.

En una realización diferente, la secuencia de combinador y transformador se invierte. En esta realización, el combinador comprendido por el aparato se configura para proporcionar señales combinadas basándose en señales transformadas proporcionadas por el transformador. En esta realización, el combinador recibe señales transformadas del transformador y las combina a fin de proporcionar señales combinadas. De esta manera, las señales de canal se transforman de manera individual en el dominio de frecuencia y se combinan en este dominio por el combinador.

En una realización adicional, el combinador se configura para proporcionar una señal de suma como una señal combinada al calcular una suma de dos señales transformadas. En esta realización, se usa una combinación lineal de las señales de canal mediante sus señales transformadas en la forma de una suma para proporcionar al menos una señal combinada. Si la señal de audio multicanal comprende, por ejemplo, dos señales de canal como señales estéreo, entonces ambas señales de canal se añaden entre sí para obtener una señal combinada que es una señal de suma. En una realización diferente, el combinador calcula una suma de señales de canal de al menos dos señales de canal.

Según una realización, el combinador se configura para proporcionar una señal de diferencia como señal combinada al calcular una diferencia entre dos señales transformadas. En esta realización, la combinación lineal de señales es una diferencia entre dos señales. Las señales que se van a combinar son señales transformadas en el presente documento. En una realización diferente, el combinador calcula una diferencia entre dos señales de canal.

Según una realización, el combinador se configura para proporcionar señales combinadas al aplicar una matriz de mezclado, llamada \mathbf{g} , que tiene dimensiones N por N a una señal multicanal que comprende las N señales transformadas, según $\tilde{\mathbf{X}} = \mathbf{g}\mathbf{X}$

N es un número de señales de canal comprendidas por la señal de audio multicanal y $\tilde{\mathbf{X}}$ es la señal multicanal en matriz basada en las señales transformadas y representa, de esta manera, las señales combinadas en una forma general.

En otras palabras: el combinador se configura para proporcionar las señales combinadas al aplicar una matriz de mezclado que tiene dimensiones N por N a las N señales transformadas, preferiblemente en un formato en matriz, en el que las señales transformadas se basan en las N señales de canal que pertenecen a la señal de audio multicanal.

- 5 Puesto que la matriz de mezclado tiene el mismo número de columnas y filas, el número de señales que se van a combinar es igual al número de señales combinadas, tal como se muestra en la siguiente ecuación:

$$\begin{bmatrix} \tilde{X}_1 \\ \tilde{X}_2 \\ \vdots \\ \tilde{X}_N \end{bmatrix} = \begin{bmatrix} g_{1,1} & g_{1,2} & \dots & g_{1,N} \\ g_{2,1} & g_{2,2} & \dots & g_{2,N} \\ \vdots & \vdots & \vdots & \vdots \\ g_{N,1} & g_{N,2} & \dots & g_{N,N} \end{bmatrix} \cdot \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix}$$

- 10 La aplicación de la matriz de mezclado es equivalente a calcular la k-ésima señal de canal de la señal combinada al sumar todas las señales de canal de la señal de audio multicanal multiplicada por los elementos correspondientes de la k-ésima fila de la matriz, tal como se muestra en la siguiente ecuación:

$$\tilde{X}_k = g_{k,1}X_1 + g_{k,2}X_2 + \dots + g_{k,N}X_N$$

- 15 En una realización diferente, la matriz de mezclado definida se aplica a una señal basándose en las N señales de canal. La matriz de mezclado permite combinar cualquier número dado de señales, ya sean señales transformadas o de canal.

- 20 En una realización adicional, el adaptador de fase se configura para modificar las fases al aplicar un método de codificador vocal de fase. Un codificador vocal de fase es un codificador vocal que escala tanto los dominios de frecuencia como de tiempo de las señales de audio al usar información de fase.

- 25 Según una realización, el adaptador de fase se configura para proporcionar señales procesadas por sus coordenadas polares que tienen amplitudes y fases modificadas. Por tanto, en una realización con N = 2 señales de canal Y_s, Y_d , las señales procesadas se dan por: $Y_{s/d} = |Y_{s/d}| \exp(j\Phi_{s/d})$ con $\Phi_{s/d} = \arg Y_{s/d}$ y $j = \sqrt{-1}$.

- 30 El separador proporciona señales individuales basándose en las señales procesadas que se basan en las señales combinadas. Las señales proporcionadas por el separador se van a usar como señales de canal modificadas en escala de tiempo. Por lo tanto, según una realización, si las señales de audio multicanal comprenden N señales de canal, entonces el separador proporciona N señales separadas.

- 35 Según una realización, el separador se configura para proporcionar una señal separada basándose en una diferencia entre dos de las señales procesadas.

- En una realización adicional, el separador se configura para proporcionar una señal separada basándose en una suma de dos de las señales procesadas.

- 40 En una realización, el separador aplica un factor (por ejemplo, 0.5) a la amplitud de la combinación respectiva de las señales procesadas.

- 45 Según una realización, el separador se configura para proporcionar N señales separadas al aplicar una matriz de mezclado inversa que tiene dimensiones de N por N a una señal en matriz basándose en las N señales procesadas. N es un número de señales de canal comprendido por la señal de audio multicanal. Esta realización permite procesar cualquier número dado de señales procesadas y proporcionar el número correspondiente de señales separadas. La matriz de mezclado inversa es la inversa de la matriz de mezclado usada para obtener las señales combinadas.

- 50 Según una realización, el aparato comprende un corrector, en el que el corrector se configura para modificar las señales separadas al reemplazar amplitudes de las señales separadas con amplitudes basadas en amplitudes de las señales de canal comprendidas por la señal de audio multicanal.

Esta realización comprende procesar una combinación lineal de las señales de canal y aplicar un procedimiento de corrección de magnitud después de que las señales se han modificado en el dominio de frecuencia.

La novedad del método presentado está, por tanto, en esta realización dos veces:

5 1) aplicar el procesamiento de TSM a las combinaciones preferiblemente lineales, por ejemplo, a la señal de suma y a la señal de diferencia y

2) aplicar un procesamiento para restaurar la magnitud de la señal de salida en el dominio de tiempo-frecuencia a fin de restaurar las diferencias de nivel entre canales (ICLD).

10 En una realización adicional, el corrector se configura para reemplazar las amplitudes de las señales separadas con amplitudes de las señales transformadas correspondientes proporcionadas por un transformador, es decir, por las amplitudes de las señales de canal en el dominio de frecuencia.

15 El corrector, de esta manera, reemplaza las amplitudes de las señales separadas que tienen la fase adaptada por las amplitudes de las correspondientes señales antes de la adaptación. Por lo tanto, se restauran las amplitudes originales.

20 Las siguientes realizaciones se ocupan de señales en el dominio de frecuencia y permiten procesarlas en el dominio de tiempo.

En una realización, el aparato comprende un transformador inverso. El transformador inverso se configura para proporcionar señales de canal modificadas basándose en las señales separadas al aplicar una transformación inversa. La transformación inversa transforma en una realización señales del dominio de frecuencia al dominio de tiempo.

25 Según otra realización, el aparato comprende un transformador inverso. El transformador inverso se configura para proporcionar señales de canal modificadas y corregidas basándose en señales corregidas proporcionadas por el corrector al aplicar una transformación inversa.

30 Según una realización diferente, el transformador inverso se configura para aplicar una transformación de Fourier de tiempo corto inversa.

El transformador inverso se configura de esta manera para invertir la clase de transformación realizada en una etapa anterior a la adaptación de fase.

35 Según una realización, el aparato comprende un extractor, en el que el extractor se configura para proporcionar señales de canal comprendidas por la señal de audio multicanal. En esta realización, el aparato, por ejemplo, recibe la señal de audio multicanal y el extractor proporciona las señales de canal individuales. En una realización diferente, las señales de canal se presentan de manera separada al aparato.

40 La siguiente realización permite facilitar las etapas computacionales y los requisitos para las unidades.

Según una realización, el aparato se configura para realizar etapas en combinaciones de señales de canal basándose en una diferencia con menos precisión que en diferentes combinaciones de las señales de canal.

45 El objeto también se logra por un método para procesar una señal de audio multicanal.

El método comprende al menos las siguientes etapas:

50 proporcionar al menos una señal combinada basándose en señales de canal comprendidas por la señal de audio multicanal,

proporcionar una señal procesada al realizar una modulación en escala de tiempo de la señal combinada, y

55 proporcionar señales de canal modificadas basándose en una separación de las señales procesadas.

La modulación en escala de tiempo se realiza en una realización por una adaptación de fase.

60 Según una realización, el método comprende además modificar amplitudes de señales de canal modificadas al reemplazar las amplitudes con amplitudes basadas en amplitudes de las correspondientes señales de canal.

En esta realización, se presentan las siguientes etapas: Las señales de canal se combinan en señales combinadas. Las señales combinadas o señales basadas en las señales combinadas experimentan una adaptación de fase a fin de realizar la modificación en escala de tiempo. Las señales adaptadas en fase se separan en señales separadas. Las señales comprenden fases y amplitudes. Las amplitudes/magnitudes de estas señales se reemplazan por

amplitudes basadas en las señales de canal.

Las realizaciones del aparato también pueden realizarse por etapas del método y correspondientes realizaciones del método. Por lo tanto, las explicaciones dadas para las realizaciones del aparato también se mantienen para el método.

5 El objeto también se logra por un programa informático para realizar, cuando se ejecuta en un ordenador o un procesador, el método según cualquiera de las realizaciones anteriores.

10 La invención se explicará a continuación con respecto a las figuras adjuntas y las realizaciones representadas en las figuras adjuntas, en las cuales:

la figura 1 muestra un diagrama de bloques de un método de modificación en escala de tiempo en el dominio de frecuencia según el estado de la técnica,

15 la figura 2 ilustra un diagrama de bloques del método de modificación en escala de tiempo de la invención para señales de audio de entrada con dos canales,

la figura 3 proporciona esquemáticamente una realización del aparato,

20 la figura 4 proporciona una realización diferente del aparato y

la figura 5 ilustra una versión general de la realización de la figura 4.

25 Un diagrama de bloques de una realización del método de la invención se representa en la figura 2 para el ejemplo de señales de entrada que tienen dos canales (por ejemplo, el canal izquierdo y derecho de sonido estéreo).

La señal de audio de entrada es $\mathbf{x} = [x_1 \ x_2]^T$, donde x_1 denota la primera señal de canal y x_2 denota la segunda señal de canal. Una representación de transformada de Fourier de tiempo corto (STFT) se calcula para x_1 y x_2 , produciendo X_1 y X_2 , respectivamente, en la etapa STFT.

30 Según la invención, una señal de suma X_s y una señal de diferencia X_d se calculan en la etapa llamada MS a partir de las señales de canal (en el presente documento: las señales de canal transformadas X_1 y X_2) de la señal de audio multicanal según

35
$$X_s = (X_1 + X_2), \tag{6}$$

$$X_d = (X_1 - X_2). \tag{7}$$

40 En una realización diferente, se cambian el orden de STFT y la combinación de las señales en la etapa MS. Esto tiene en cuenta que el cálculo de las señales de suma y de diferencia también puede realizarse en el dominio de tiempo y que la STFT puede calcularse a partir de las señales de suma y de diferencia del dominio de tiempo. Sin embargo, es ventajoso aplicar el orden tal como se muestra en la figura 2 para reducir la carga computacional.

45 La señal de suma X_s y la señal de diferencia X_d entonces se procesan por medio de un método de adaptación de fase (PA), por ejemplo, al usar el método de codificador vocal de fase descrito en [2], o cualquier otro método de modificación en escala de tiempo. La señal de suma y la señal de diferencia procesadas se denotan por Y_s e Y_d , respectivamente.

50 A esto le sigue una transformación inversa en la etapa MS inversa, por ejemplo, al transformar las señales Y_s e Y_d por una transformada de Fourier de tiempo corto inversa.

Las señales procesadas para el primero y el segundo canal se obtienen en la realización mostrada al aplicar las siguientes ecuaciones (8) y (9), es decir, el procesamiento inverso de las ecuaciones (6) y (7).

55 Por tanto, las dos señales separadas Y_1 e Y_2 se dan en la realización mostrada por:

$$Y_1 = 0.5 (Y_s + Y_d), \tag{8}$$

$$Y_2 = 0.5 (Y_s - Y_d). \tag{9}$$

60 En una forma general, el cálculo de la señal de suma y la señal de diferencia (como ejemplos de las señales combinadas) se expresan en la notación de matriz como

$$\tilde{\mathbf{X}} = \mathbf{g}\mathbf{X} \tag{10}$$

con coeficientes STFT de la señal de entrada $\mathbf{X} = [X_1 \dots X_N]^T$, un número de canales N, coeficientes STFT de la señal en matriz $\tilde{\mathbf{X}} = [\tilde{X}_1 \dots \tilde{X}_N]^T$, y una matriz de mezclado g de tamaño N por N.

5 Por ejemplo, la señal de suma X_s y la señal de diferencia X_d , tal como se da por las ecuaciones (6) y (7), se obtienen para N = 2 al ajustar

$$\mathbf{g} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \text{ y } \tilde{\mathbf{X}} = [X_s X_d]^T.$$

10 El procesamiento inverso de la operación de matriz, es decir, el cálculo de las señales de canal $\mathbf{Y} = [Y_1 \dots Y_N]^T$ a partir de la señal en matriz $\tilde{\mathbf{Y}} = [\tilde{Y}_1 \dots \tilde{Y}_N]^T$ se obtiene a partir de

$$\mathbf{Y} = \mathbf{g}^{-1} \tilde{\mathbf{Y}} \tag{11}$$

15 Donde $\tilde{\mathbf{Y}}$ es el resultado del procesamiento de PA aplicado a $\tilde{\mathbf{X}}$ y \mathbf{g}^{-1} es la inversa de la matriz g. Con esta generalización, el método propuesto también puede aplicarse a señales que tienen más de dos canales.

Una etapa muy ventajosa de la realización mostrada es corregir la magnitud de los coeficientes espectrales de manera que los coeficientes espectrales de valor complejo resultantes tengan la fase del resultado de la etapa de separación usando las ecuaciones (8) y (9) y la magnitud de X_1 y X_2 de las señales de canal transformadas.

20 Las señales separadas se dan en coordenadas polares como:

$$Y_{1/2} = |Y_{1/2}| \exp(j\Phi_{Y,1/2})$$

con $\Phi_{y,1/2} = \arg Y_{1/2}$ y $j = \sqrt{-1}$.

25 Por tanto, para las señales de entrada de dos canales mostradas, las señales corregidas Z_1 y Z_2 después de la etapa MC se dan por:

$$Z_1 = |X_1| \exp(j\Phi_{Y,1}), \tag{12}$$

$$Z_2 = |X_2| \exp(j\Phi_{Y,2}). \tag{13}$$

Esta etapa asegura que las diferencias de nivel entre canales (ICLD) de la señal de audio se conserven.

35 Por lo tanto, en esta etapa, las magnitudes de las señales procesadas se sustituyen por las amplitudes originales de las señales de canal transformadas.

Para el caso general de señales de entrada que tienen más de dos canales, cada señal de salida después de la operación de matriz inversa, es decir, después de la separación en señales separadas después de la adaptación de fase individual, se modifica en una realización de manera que su magnitud se reemplaza por la magnitud de la señal de canal correspondiente, preferiblemente transformada (antes de la formación en matriz, es decir, antes del cálculo de las combinaciones de diferentes señales de canal).

45 El procesamiento de PA tiene una complejidad computacional considerable. Comprende diversas etapas de procesamiento que pueden implementarse con precisión reducida a fin de reducir la carga computacional.

Por ejemplo, el cálculo de las coordenadas polares (magnitud y fase) de un número complejo dado por sus coordenadas cartesianas (componente real e imaginario) puede implementarse con mayor o menor precisión. Frecuentemente, los cálculos con menor precisión tienen menos costes computacionales pero introducen un error.

50 La carga computacional puede reducirse al aprovecharse del hecho de que para señales de audio típicas (por ejemplo, grabaciones musicales o grabaciones de radiodifusión) la señal de suma tiene mayor energía que la señal de diferencia. Los errores que resultan de las aproximaciones en el cálculo tienen un efecto menor cuando se introducen en la señal de diferencia y mayor efecto cuando se producen en la señal de suma. Esto puede aprovecharse al aplicar métodos con menor precisión para calcular la señal de diferencia Y_d y métodos con mayor precisión cuando se calcula la señal de suma Y_s .

Otro medio para reducir la carga computacional es saltarse ciertas etapas de procesamiento. Por ejemplo, el bloqueo

de fase puede saltarse cuando se calcula la señal de diferencia Y_d . El bloqueo de fase se refiere a una etapa de procesamiento adicional para mejorar la calidad de sonido. Para reducir la carga computacional, las fases para la señal de diferencia Y_d se calculan de manera que la propagación de fase se mantiene sin aplicar el procesamiento de bloqueo de fase.

5 Otro medio para reducir la carga computacional es aplicar el procesamiento con alta calidad (gran precisión incluyendo todas las etapas de procesamiento tal como el bloqueo de fase) solo hasta un valor máximo de frecuencia. Para una señal digital muestreada a 48 kHz, por ejemplo, el procesamiento de alta calidad solo se aplica a las bandas de frecuencia hasta un valor máximo de 10 kHz. La frecuencia máxima hasta la cual se aplica el procesamiento con la mejor calidad posible puede reducirse adicionalmente para el cálculo de la señal de diferencia.

La figura 3 muestra una realización del aparato 1 que realiza una modificación en escala de tiempo de una señal de audio multicanal 100. La señal de audio multicanal 100 puede comprender más de dos señales de canal. La modificación en escala de tiempo se logra por una adaptación de fase.

15 Un extractor 2 recupera las señales de canal X_1 , X_2 comprendidas por la señal de audio multicanal 100. En la realización mostrada, solo hay dos señales de canal. No obstante, la invención no se limita a dos señales de canal.

20 En una realización diferente, y no mostrada, las señales de canal x_1 , x_2 se proporcionan de manera separada al aparato 1. Por lo tanto, para esta realización no se requiere extractor.

Las señales de canal x_1 , x_2 son señales en el dominio de tiempo y se presentan, en la realización mostrada, a un transformador 3. El transformador 3 transforma las señales de canal x_1 , x_2 en el dominio de frecuencia y de esta manera proporciona señales transformadas X_1 , X_2 .

25 Las señales transformadas X_1 , X_2 se presentan a un combinador 4. El combinador 4 combina las señales transformadas X_1 , X_2 , en el ejemplo dado, al aplicar combinaciones lineales de estas, por ejemplo, al calcular una suma X_s de las señales transformadas y una diferencia X_d entre estas. El número de señales combinadas en una realización es idéntico al número de señales de canal.

30 En una realización diferente, no mostrada, la secuencia del transformador 3 y el combinador 4 se invierten. Esto implica que el combinador 4 combina las señales de canal y el transformador 3 transforma las señales combinadas en esta realización diferente y no mostrada.

35 En la realización representada, las señales combinadas, que son señales de canal transformadas combinadas, X_s y X_d se presentan al adaptador de fase 5.

40 El adaptador de fase 5 modifica las fases de las señales combinadas X_s y X_d y proporciona señales procesadas Y_s e Y_d . Las señales procesadas Y_s e Y_d tienen fases adaptadas que reflejan la modificación en escala de tiempo apropiada de las señales combinadas X_s y X_d . Por lo tanto, las señales combinadas o bien se desaceleran o bien se aceleran.

45 A fin de obtener señales de canal modificadas en escala de tiempo, las señales procesadas Y_s e Y_d se separan por el separador 6, proporcionando señales separadas Y_1 e Y_2 que experimentan una transformación inversa por un transformador inverso 7. Las señales de canal modificadas resultantes y_1 e y_2 son señales de tiempo y tienen la escala de tiempo deseada.

La figura 4 muestra una realización diferente del aparato 1.

50 La estructura del aparato 1 mostrada en la figura 4 es similar a la realización mostrada en la figura 3. La diferencia entre ambas realizaciones se da por las unidades después del separador 6. Para la explicación de los otros elementos, véase la descripción de la figura 3.

55 El separador 6 en la realización mostrada en la figura 4 también proporciona las señales separadas Y_1 e Y_2 . Estas señales de dominio de frecuencia Y_1 e Y_2 se presentan en esta realización a un corrector 8 que está aguas arriba del transformador inverso 7.

60 El corrector 8 reemplaza las amplitudes de las señales separadas Y_1 e Y_2 por las amplitudes de las correspondientes señales transformadas X_1 y X_2 , es decir, con las amplitudes o magnitudes antes de la adaptación de fase y especialmente antes de la combinación de las señales de canal.

Las señales modificadas en amplitud o corregidas resultantes Z_1 y Z_2 (compárense las ecuaciones (12) y (13)) se presentan al transformador inverso 7 y se transforman en el dominio de tiempo como señales de canal modificadas y corregidas z_1 , z_2 .

A fin de permitir la corrección, el transformador 3 se conecta al corrector 8.

5 En una realización adicional, no mostrada, la secuencia del transformador 3 y el combinador 4 se cambia y el transformador 3, de esta manera, transforma las señales combinadas. Para la corrección de las señales separadas Y_1, Y_2 , el corrector 8 se refiere, por lo tanto, a transformaciones adicionales de las correspondientes señales de canal x_1, x_2 .

La realización de la figura 5 es una versión generalizada de la mostrada en la figura 4.

10 En el presente documento, la señal de audio multicanal 100 comprende N señales de canal x_1, x_2, \dots, x_N donde N es un número entero mayor de dos.

15 Las señales de canal x_1, x_2, \dots, x_N que son señales de tiempo recuperadas por el extractor 2 se presentan al transformador 3, que proporciona las señales transformadas en el dominio de frecuencia X_1, X_2, \dots, X_N , que se dan en el presente documento por un vector X (de manera alternativa dado por X).

El siguiente combinador 4 proporciona un vector de las combinaciones lineales $\tilde{X} = [\tilde{X}_1 \dots \tilde{X}_N]^T$ al aplicar la ecuación (10).

20 Sigue el adaptador de fase 5 que proporciona señales procesadas (en el presente documento dadas por un vector: \tilde{Y}) que están separadas por un separador 6. Las señales separadas Y_1, Y_2, \dots, Y_N (en el presente documento dado por un vector Y) se corrigen con respecto a sus amplitudes por el corrector 8. Las señales corregidas Z_1, Z_2, \dots, Z_N (dado en el presente documento por un vector Z) se presentan al transformador inverso 7 que produce N señales de canal modificadas y corregidas z_1, z_2, \dots, z_N . El cambio de letras mayúsculas a letras minúsculas indica la transformación del dominio de frecuencia al dominio de tiempo.

30 Aunque algunos aspectos se han descrito en el contexto de un aparato, está claro que estos aspectos también representan una descripción del método correspondiente, donde un bloque o dispositivo corresponde a una etapa de método o una característica de una etapa de método. De manera análoga, los aspectos descritos en el contexto de una etapa de método también representan una descripción de un bloque o elemento o característica correspondiente de un aparato correspondiente. Algunas o todas las etapas del método pueden ejecutarse por (o al usar) un aparato de hardware tal como, por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, algunos de unos o más de las etapas de método más importantes pueden ejecutarse por este aparato.

35 La señal transmitida o codificada de la invención puede almacenarse en un medio de almacenamiento digital o puede transmitirse en un medio de transmisión, tal como un medio de transmisión inalámbrica o un medio de transmisión alámbrica tal como Internet.

40 Dependiendo de ciertos requisitos de implementación, pueden implementarse realizaciones de la invención en hardware o en software. La implementación puede realizarse usando un medio de almacenamiento digital, por ejemplo, un disco flexible, un DVD, un Blu-Ray, un CD, una ROM, una PROM y EPROM, una EEPROM o una memoria FLASH, que tienen señales de control electrónicamente legibles almacenadas en los mismos, que actúan conjuntamente (o son capaces de actuar conjuntamente) con un sistema informático programable de manera que se realice el método respectivo. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.

50 Algunas realizaciones según la invención comprenden un portador de datos que tiene señales de control electrónicamente legibles, que son capaces de actuar conjuntamente con un sistema informático programable, de manera que se realice uno de los métodos descritos en el presente documento.

55 En general, las realizaciones de la presente invención pueden implementarse como un producto de programa informático con un código de programa, siendo el código de programa 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 legible por máquina.

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

60 En otras palabras, una realización del método de la invención 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.

Una realización adicional del método de la invención es, por lo tanto, un portador de datos (o un medio de almacenamiento no transitorio, tal como un medio de almacenamiento digital, o un medio legible por ordenador) que

comprende, grabado en el mismo, el programa informático para realizar uno de los métodos descritos en el presente documento. El portador de datos, el medio de almacenamiento digital o el medio grabado son normalmente tangibles y/o no transitorios.

5 Por lo tanto, una realización adicional del método de la invención es un flujo de datos o una secuencia de señales que representa el programa informático para realizar uno de los métodos descritos en el presente documento. El flujo de datos o la secuencia de señales puede, por ejemplo, configurarse para transferirse mediante una conexión de comunicación de datos, por ejemplo, mediante Internet.

10 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.

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

15 Una realización adicional según la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, de manera electrónica u óptica) un programa informático para realizar uno de los métodos descritos en el presente documento a un receptor. El receptor puede ser, por ejemplo, un ordenador, un dispositivo móvil, un dispositivo de memoria o similar. El aparato o sistema puede, por ejemplo, comprender un servidor de archivos para transferir el programa informático al receptor.

20 En algunas realizaciones, puede usarse un dispositivo lógico programable (por ejemplo, una matriz de puertas de campo programable) para realizar algunas o todas las funcionalidades de los métodos descritos en el presente documento. En algunas realizaciones, una matriz de puertas de campo programable puede actuar conjuntamente con un microprocesador con el fin de realizar uno de los métodos descritos en el presente documento. En general, los métodos se realizan preferiblemente por cualquier aparato de hardware.

25 Las realizaciones descritas anteriormente son meramente ilustrativas de los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles descritos en el presente documento serán evidentes para otros expertos en la técnica. Por lo tanto, se propone que se limiten solo por el alcance de las reivindicaciones de patente inminentes y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones en el presente documento.

35 **Bibliografía**

[1] M. Dolson, "The Phase Vocoder: A Tutorial", Computer Music Journal, vol. 10, páginas 14 - 27, 1986.

[2] J. Laroche y M. Dolson, "Improved Phase Vocoder Time-Scale Modification of Audio", IEEE Transaction on Speech and Audio Processing, vol. 7, n.º 3, páginas 323 - 332, 1999.

40 [3] T. Karrer, E. Lee, y J. Borchers, "PhaVoRIT: A Phase Vocoder for Real-Time Interactive Time-Stretching," en Proc. of ICMC, 2006.

REIVINDICACIONES

1. Aparato (1) para procesar una señal de audio multicanal (100) que comprende una pluralidad de señales de canal (x_1, x_2), en el que el aparato (1) se configura para realizar una modificación en escala de tiempo de la señal de audio multicanal (100), y en el que el aparato (1) comprende:
- 5 un combinador (4) para proporcionar señales combinadas (X_s, X_d) basándose en las señales de canal (x_1, x_2), en el que el combinador (4) se configura para proporcionar una señal de suma (X_s) al calcular una suma de dos señales de canal (x_1, x_2) y para proporcionar una señal de diferencia (X_d) al calcular una diferencia de dichas dos señales de canal (x_1, x_2);
- 10 un adaptador de fase (5), en el que el adaptador de fase (5) se configura para proporcionar al menos una señal procesada (Y_s) para realizar modificación en escala de tiempo al modificar una fase de la señal de suma (X_s), y opcionalmente una señal de diferencia procesada (Y_d) al modificar una fase de la señal de diferencia, en el que el adaptador de fase (5) se configura para modificar la fase al aplicar un método de codificador vocal de fase, en el que el adaptador de fase (5) se configura para modificar la fase al aplicar un bloqueo de fase, o en el que el adaptador de fase (5) se configura para modificar la fase de modo que se mantenga una coherencia de fase horizontal, y
- 15 un separador (6), en el que el separador (6) se configura para proporcionar señales separadas (Y_1, Y_2) basándose en la al menos una señal procesada (Y_s) y la señal de diferencia (X_d) o la señal de diferencia opcionalmente procesada (Y_d), en el que el separador (6) se configura para invertir la combinación de las señales de canal, aplicada por el combinador (4).
- 20 2. Aparato (1) según la reivindicación 1, en el que el adaptador de fase (5) se configura para proporcionar N señales procesadas (Y_s, Y_d) al modificar fases de N señales (X_s, X_d) basándose en combinaciones de N señales de canal (x_1, x_2),
- 25 en el que el separador (6) se configura para proporcionar N señales separadas (Y_1, Y_2) basándose en la señal procesada (Y_s, Y_d), y
- 30 en el que N es un número de señales de canal (x_1, x_2) comprendido por la señal de audio multicanal (100).
3. Aparato (1) según la reivindicación 2, en el que las N combinaciones de las señales de canal (x_1, x_2) son combinaciones lineales de las señales de canal (x_1, x_2).
- 35 4. Aparato (1) según una cualquiera de las reivindicaciones 1 a 3, en el que el aparato (1) comprende un transformador (3), y
- 40 en el que el transformador (3) se configura para proporcionar señales transformadas (X_1, X_2) al transformar señales del dominio de tiempo al dominio de frecuencia.
5. Aparato (1) según la reivindicación 4, en el que el transformador (3) se configura para aplicar una transformación de Fourier de tiempo corto.
- 45 6. Aparato (1) según una cualquiera de las reivindicaciones 4 o 5,
- 50 en el que el combinador (4) se configura para proporcionar señales combinadas (X_s, X_d) basándose en las señales de canal (x_1, x_2), y
- en el que el transformador (3) se configura para proporcionar señales transformadas basándose en las señales combinadas al aplicar una transformación.
- 55 7. Aparato (1) según la reivindicación 4 o 5,
- en el que el combinador (4) se configura para proporcionar señales combinadas (X_s, X_d) basándose en señales transformadas (X_1, X_2) proporcionadas por el transformador (3).
- 60 8. Aparato (1) según la reivindicación 7, en el que el combinador (4) se configura para proporcionar una señal de suma (X_s) al calcular una suma de dos señales transformadas (X_1, X_2).
9. Aparato (1) según la reivindicación 7 u 8, en el que el combinador (4) se configura para proporcionar una señal de diferencia (X_d) al calcular una diferencia entre dos señales transformadas (X_1, X_2).

10. Aparato (1) según la reivindicación 7, en el que el combinador (4) se configura para proporcionar las señales combinadas (X_s , X_d) al aplicar una matriz de mezclado (g) que tiene dimensiones de N por N a las N señales transformadas (X_1 , X_2) basándose en las N señales de canal (x_1 , x_2) que pertenecen a la señal de audio multicanal (100), y
- 5 en el que N es un número de señales de canal (x_1 , x_2) comprendido por la señal de audio multicanal (100).
11. Aparato (1) según una cualquiera de las reivindicaciones 4 o 5, en el que el transformador (3) tiene un tamaño de salto de análisis R_a ,
- 10 en el que el aparato comprende, además, un transformador inverso (7), en el que el transformador inverso (7) tiene un tamaño de salto de síntesis R_s , y
- 15 en el que el tamaño de salto de síntesis R_s y el tamaño de salto de análisis R_a se ajustan a diferentes valores, de modo que la modificación en escala de tiempo es un estiramiento de señal en tiempo, cuando $R_a < R_s$, o de modo que la modificación en escala de tiempo es un encogimiento de señal en tiempo, cuando $R_a > R_s$.
12. Aparato (1) según una cualquiera de las reivindicaciones 2 a 11, en el que el separador (6) se configura para proporcionar una señal separada (Y_1 , Y_2) basándose en una diferencia entre dos de las señales procesadas (Y_s , Y_d).
- 20 13. Aparato (1) según una cualquiera de las reivindicaciones 2 a 12, en el que el separador (6) se configura para proporcionar una señal separada (Y_1 , Y_2) basándose en una suma de dos de las señales procesadas (Y_s , Y_d).
- 25 14. Aparato (1) según una cualquiera de las reivindicaciones 2 a 13, en el que el separador (6) se configura para proporcionar N señales separadas (Y_1 , Y_2) al aplicar una matriz de mezclado inverso (g^{-1}) que tiene dimensiones de N por N a una señal en matriz basada en las N señales procesadas (Y_s , Y_d), y
- 30 en el que N es un número de señales de canal (x_1 , x_2) comprendido por la señal de audio multicanal (100).
15. Aparato (1) según una cualquiera de las reivindicaciones 1 a 14, en el que el aparato (1) comprende un corrector (8), y
- 35 en el que el corrector (8) se configura para modificar las señales separadas (Y_1 , Y_2) al reemplazar amplitudes de las señales separadas (Y_1 , Y_2) con amplitudes basadas en amplitudes de las señales de canal (x_1 , x_2).
16. Aparato (1) según la reivindicación 15, en el que el corrector (8) se configura para reemplazar las amplitudes de las señales separadas (Y_1 , Y_2) con amplitudes de señales transformadas (X_1 , X_2) proporcionadas por un transformador (3).
- 40 17. Aparato (1) según una cualquiera de las reivindicaciones 1 a 16, en el que el aparato (1) comprende un transformador inverso (7), y
- 45 en el que el transformador inverso (7) se configura para proporcionar señales de canal modificadas (y_1 , y_2) basándose en las señales separadas (Y_1 , Y_2) al aplicar una transformación inversa.
18. Aparato (1) según la reivindicación 15 o 16, en el que el aparato (1) comprende un transformador inverso (7), y
- 50 en el que el transformador inverso (7) se configura para proporcionar señales de canal modificadas y corregidas (z_1 , z_2) basándose en señales corregidas (Z_1 , Z_2) proporcionadas por el corrector (8) al aplicar una transformación inversa.
- 55 19. Aparato (1) según la reivindicación 17 o 18, en el que el transformador inverso (7) se configura para aplicar una transformación de Fourier de tiempo corto inversa.
20. Aparato (1) según una cualquiera de las reivindicaciones 1 a 19, en el que el aparato (1) comprende un extractor (2), y
- 60 en el que el extractor (2) se configura para proporcionar señales de canal (x_1 , x_2) comprendidas por la señal de audio multicanal (100).
21. Aparato (1) según una cualquiera de las reivindicaciones 1 a 20, en el que el aparato (1) se configura para

realizar etapas en combinaciones de señales de canal (x_1, x_2) basándose en una diferencia entre señales con menos precisión que en diferentes combinaciones de señales.

- 5 22. Método para procesar una señal de audio multicanal (100) que comprende una pluralidad de señales de canal (x_1, x_2), en el que el método se configura para realizar una modificación en escala de tiempo de la señal de audio multicanal (100), comprendiendo el método:
- 10 proporcionar señales combinadas (X_s, X_d) basándose en las señales de canal (x_1, x_2), en el que las señales combinadas comprenden una señal de suma (X_s) calculada a partir de una suma de dos señales de canal (x_1, x_2) y una señal de diferencia (X_d) calculada a partir de una diferencia de dichas dos señales de canal (x_1, x_2),
- 15 proporcionar al menos una señal procesada (Y_s) para realizar modificación en escala de tiempo al modificar una fase de la señal de suma (X_s), y opcionalmente una señal de diferencia procesada (Y_d) al modificar una fase de la señal de diferencia, en el que proporcionar la señal procesada comprende modificar la fase al aplicar un método de codificador vocal de fase, o modificar la fase al aplicar un bloqueo de fase, o modificar la fase de modo que se mantenga una coherencia de fase horizontal, y
- 20 proporcionar señales separadas (Y_1, Y_2) basándose en la al menos una señal procesada (Y_s) y la señal de diferencia (X_d) o la señal de diferencia opcionalmente procesada (Y_d), en el que
- proporcionar las señales separadas (Y_1, Y_2) comprende invertir la combinación de las señales de canal, aplicada al proporcionar las señales combinadas (X_s, X_d).
- 25 23. Método según la reivindicación 22, que comprende además:
- modificar amplitudes de señales de canal modificadas al reemplazar las amplitudes de señales de canal modificadas con amplitudes basadas en amplitudes de señales de canal correspondientes.
- 30 24. Programa informático para realizar, cuando se ejecuta en un ordenador o un procesador, el método según la reivindicación 22 o 23.

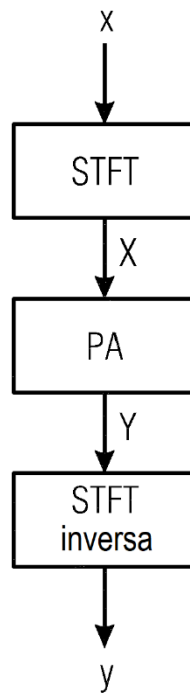


Fig. 1

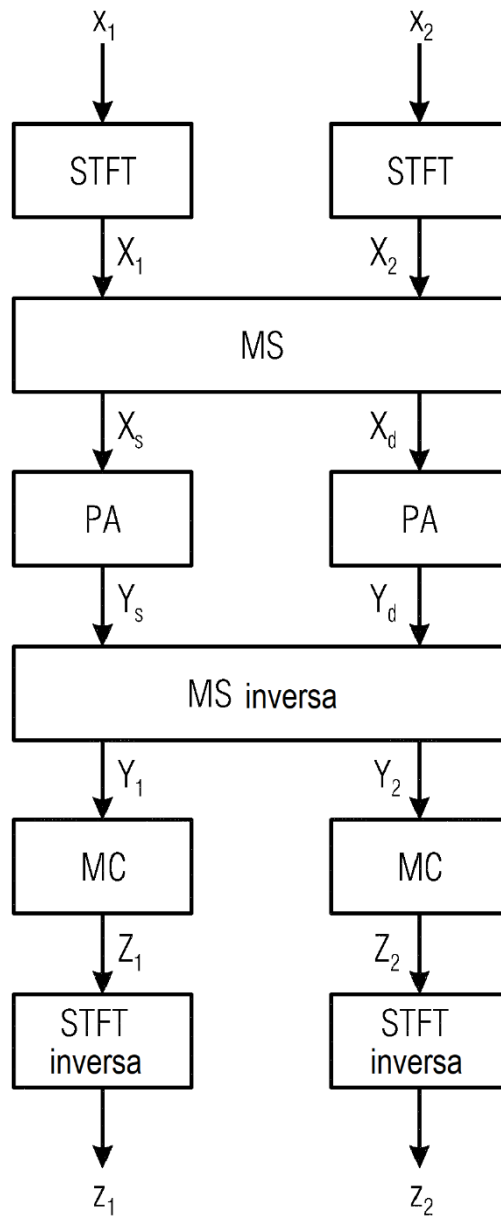


Fig. 2

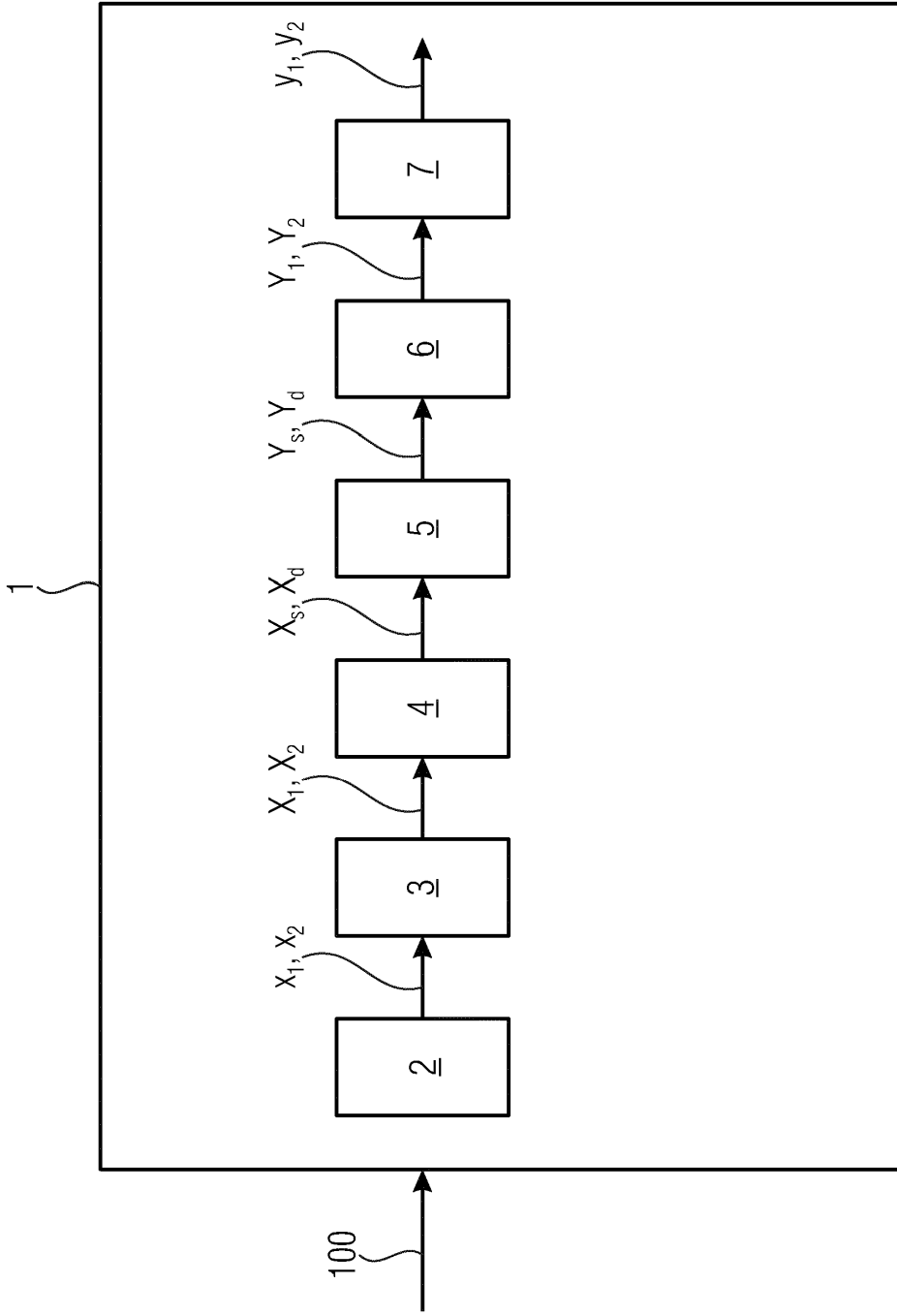


Fig. 3

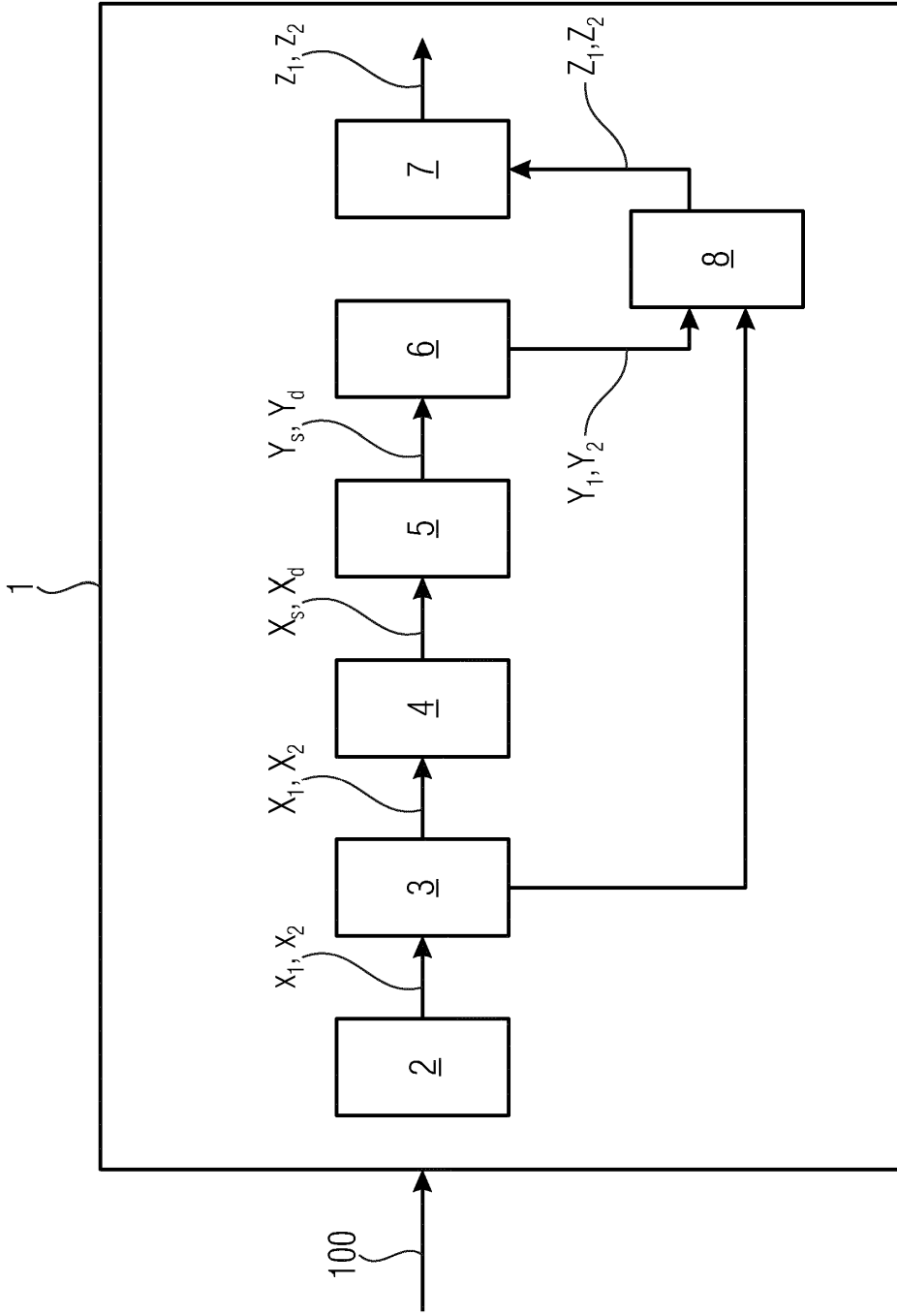


Fig. 4

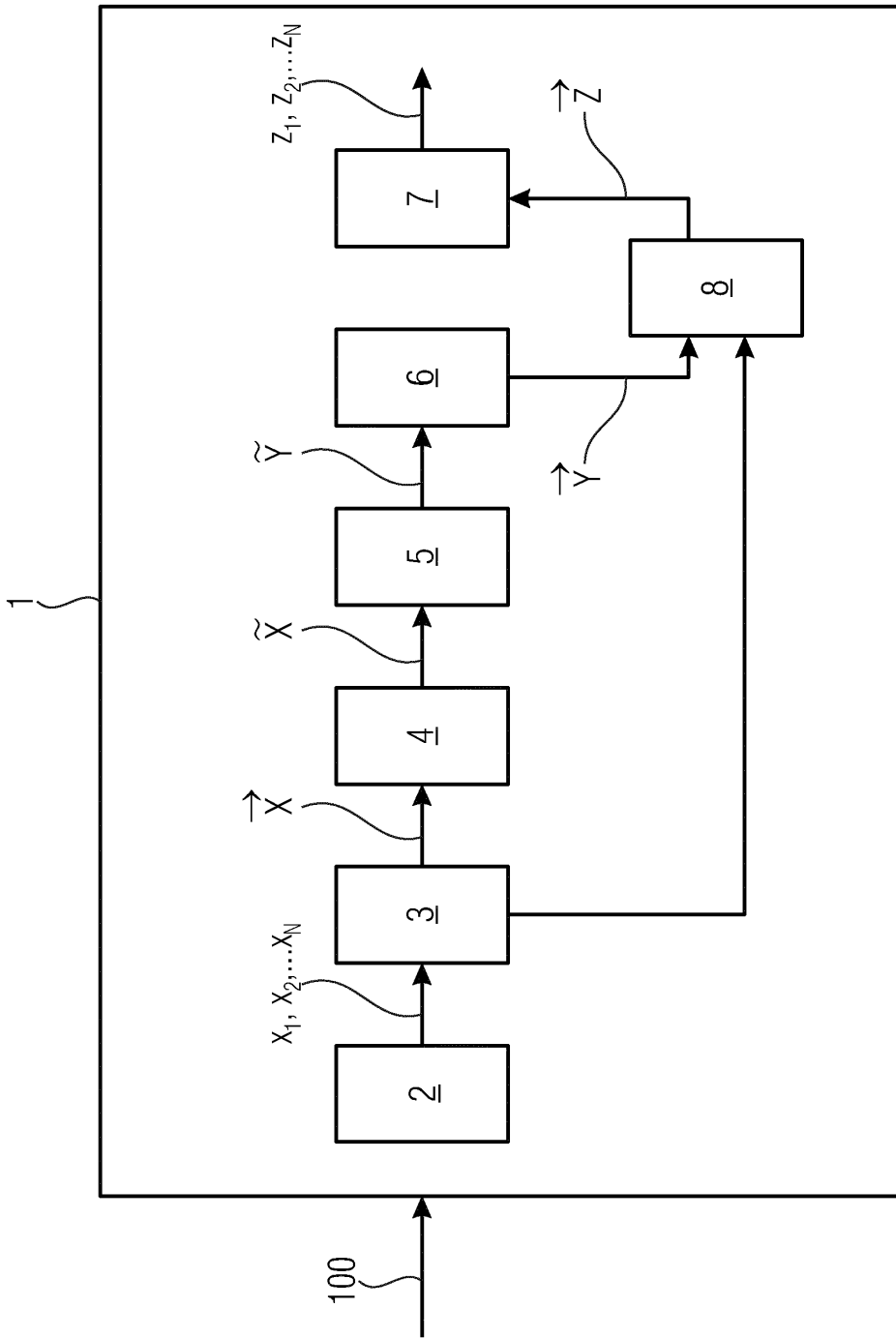


Fig. 5