

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 989 883**

51 Int. Cl.:

G10L 19/008 (2013.01)

H04S 3/02 (2006.01)

H03G 5/16 (2006.01)

G10L 19/02 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **04.03.2020 PCT/EP2020/055669**

87 Fecha y número de publicación internacional: **10.09.2020 WO20178321**

96 Fecha de presentación y número de la solicitud europea: **04.03.2020 E 20706774 (5)**

97 Fecha y número de publicación de la concesión europea: **04.09.2024 EP 3935630**

54 Título: **Mezclado descendente de audio**

30 Prioridad:

06.03.2019 EP 19161076

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
28.11.2024

73 Titular/es:

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

72 Inventor/es:

**REUTELHUBER, FRANZ;
EDLER, BERND;
FOTOPOULOU, ELENI;
MULTRUS, MARKUS;
MABEN, PALLAVI y
DISCH, SASCHA**

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 989 883 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

Mezclado descendente de audio

5 La presente invención está dirigida al procesamiento de señales de audio y, en particular, al mezclado descendente de señales multicanal o a la conversión de la resolución espectral de las señales de audio.

10 Aunque un flujo de bits codificado en estéreo usualmente será decodificado para ser reproducido en un sistema estéreo, no todos los dispositivos que son capaces de recibir un flujo de bits estéreo siempre serán capaces de producir una señal en estéreo. Un posible escenario sería la reproducción de la señal estéreo en un teléfono móvil con únicamente un altavoz mono. Con la llegada de los escenarios de comunicación móvil multicanal, respaldados por la nueva norma 3GPP IVAS, se requiere una mezcla estéreo-mono que no sufra retrasos adicionales y sea lo más eficiente posible desde el punto de vista de la complejidad, a la vez que proporcione la mejor calidad perceptiva posible más allá de lo que se puede conseguir con una simple mezcla pasiva.

15 Existen múltiples formas de convertir una señal estéreo en una señal mono. La forma más directa de hacerlo es mediante un mezclado descendente pasivo [1] en el dominio del tiempo, que genera una señal media sumando los canales izquierdo y derecho y escalando el resultado:

$$20 \quad \textit{Media} = \frac{(\textit{Izquierdo} + \textit{Derecho})}{2}$$

Además, entre los métodos de mezclado descendente más sofisticados (es decir, activos) basados en el dominio de tiempo se incluyen el escalamiento de energía en un esfuerzo por preservar la energía total de la señal [2] [3], la alineación de fase para evitar los efectos de cancelación [4] y la prevención de los efectos de filtro en peine mediante la supresión de coherencia [5].

30 Otro método es efectuar la corrección de energía en una manera dependiente de la frecuencia calculando factores de ponderación separados para múltiples bandas espectrales. Por ejemplo, esto se hace como parte del convertidor de formato MPEG-H [6], donde el mezclado descendente se efectúa en una representación de subbanda QMF híbrida o con un banco de filtros STFT de las señales con alineación de fase previa adicional de los canales. En el contexto IVAS, ya se usa un mezclado descendente (incluyendo ambas de la alineación de fase y de tiempo) para el Estéreo DFT en modo de velocidad de bits baja paramétrico, donde la ponderación y mezclado se aplican en el dominio de DFT [7].

35 La solución simple de un mezclado descendente de estéreo a mono pasivo en el dominio de tiempo después de decodificar las señales estéreo no es ideal, porque es sabido que en estado para reducir el número de canales audio puramente pasivo viene con ciertas desventajas, por ejemplo, efectos de cancelación de fase o pérdida general de energía, las cuales dependiendo del elemento pueden degradar severamente la calidad.

40 Otros métodos activos de mezclado descendente, basados exclusivamente en el dominio del tiempo, mitigan algunos de los problemas del mezclado descendente pasivo, pero siguen siendo subóptimos debido a la falta de ponderación en función de la frecuencia.

45 Con las restricciones implícitas para códecs de comunicaciones móviles como el IVAS en términos de retardo y complejidad, que tienen una etapa de procesamiento posterior dedicada como el convertidor de formato MPEG-H para aplicar un mezclado descendente por banda tampoco es una opción puesto que las transformaciones necesarias al dominio de la frecuencia y de regreso inevitablemente causarían un incremento en la complejidad y retraso.

50 Para un modo de códec estéreo que usa la codificación de transformación TCX con conmutación de bloques como en [8], pueden existir diferentes modos que pueden ser usados: por ejemplo, un bloque por trama con un tamaño de bloque de 20 ms (TCX20) y dos subbloques por trama con un tamaño de bloque de 10 ms (TCX10). Cada subbloque es un bloque completo TCX10 de 10 ms o nuevamente se subdivide en dos bloques de 5 ms (TCX5). La decisión de cuál de los modos usar se efectúa por cada canal independientemente del otro. Esto significa que es posible tomar decisiones diferentes entre los canales. Esto hace que sea imposible usar exactamente el mismo método de mezcla descendente que se usa en el codificador estéreo basado en la DFT descrito en [7] (ponderación de los canales en función de la banda y, a continuación, mezcla descendente mono de ambos en el dominio de la DFT) debido a las diferentes resoluciones de tiempo-frecuencia de las respectivas representaciones del dominio espectral.

60 El documento US 2016/0037279A1 divulga un método y un sistema que ajusta dinámicamente el audio de una señal de audio y vídeo para mejorar su calidad de sonido global y la inteligibilidad de los diálogos. Algunas realizaciones utilizan ganancia, equalización, compresión de la señal de audio y mejora espacial (reverberación) en canales individuales de una señal de audio multicanal. Para cada canal individual, se realiza un control de ganancia, una modulación de frecuencia y una nivelación/compresión. El canal central recibe un ligero aumento de frecuencia en la

gama típica del habla humana y una ligera reducción en el resto del espectro de audiofrecuencia. Por ejemplo, el aumento puede ser de aproximadamente 1 dB a 3 dB a 2850 Hz con un Q (ancho de banda) de hasta 250 ciclos. El usuario también puede seleccionar este realce a partir de preajustes preprogramados.

5 Anónimo: "Mezcla de audio (música grabada) - Wikipedia", 2 de marzo de 2019 (2019-03-02), XP055679711, URL:[https://web.archive.org/web/20190302214525/https://en.wikipedia.org/wiki/Audio_mixing_\(recorded_music\)](https://web.archive.org/web/20190302214525/https://en.wikipedia.org/wiki/Audio_mixing_(recorded_music)), recuperado el 2020-03-25 divulga varios aspectos del procesamiento de audio, como los procesos que afectan a una respuesta en frecuencia. Estos procesos limpian la señal de audio, mejoran la intercambiabilidad entre otras señales, ajustan el efecto de sonoridad y, en general, crean un sonido mucho más agradable o deliberadamente peor. Existen
10 dos procesos principales de respuesta en frecuencia: la ecualización y el filtrado. La descripción más sencilla de la ecualización es el proceso de alterar la respuesta en frecuencia de forma similar a lo que hacen los controles de tono en un equipo estéreo. Los ecualizadores profesionales diseccionan el espectro de audio en tres o cuatro partes que pueden denominarse controles de graves, medios- agudos y agudos. Los filtros se utilizan para eliminar determinadas frecuencias de la salida.

15 Es un objeto de la presente invención proporcionar un concepto mejorado para el procesamiento de señales de audio.

Este objeto se consigue mediante un mezclador descendente de la reivindicación 1, un método de mezcla descendente de la reivindicación 36, o un programa informático de la reivindicación 37.

20 El mezclador descendente comprende un estimador del valor de ponderación, un ponderador espectral, un convertidor y un mezclador conectado posteriormente. La conversión del dominio espectral en el dominio del tiempo se efectúa después de la ponderación espectral de la representación del dominio espectral de un primer canal y una ponderación de la representación del dominio espectral de un segundo canal y, según sea el caso, la ponderación espectral de las
25 representaciones del dominio espectral de canales adicionales. Las representaciones de dominio espectral ponderadas son convertidas de la representación del dominio espectral en una representación del tiempo del canal correspondiente. En el dominio del tiempo, la mezcla se realiza para obtener una señal de mezcla descendente como salida del mezclador descendente. Este procedimiento permite efectuar una ponderación de la calidad de audio útil y eficiente, pero no obstante alta en el dominio espectral, pero aún permite el procesamiento visual de los canales individuales en el dominio espectral en comparación con una situación, donde la ponderación del dominio espectral y el mezclado para reducir el número de canales de audio se efectúa como una sola operación. En esa situación, ya no es posible efectuar el procesamiento de canal individual, puesto que, después de la ponderación espectral y mezclado para reducir el número de canales de audio, solo existe la señal de mezclado para reducir el número de canales de
30 audio. De este modo, de acuerdo con este aspecto de la presente invención, se vuelve posible, no obstante, efectuar un procesamiento de canal individual en el dominio espectral, pero este procesamiento individual en el dominio espectral se efectúa después de la ponderación espectral.

En situaciones donde al menos dos canales tienen diferentes resoluciones de tiempo y frecuencia, el cálculo de los valores de ponderación por banda para al menos dos canales requiere convertir ya sea una o ambas representaciones del dominio espectral de al menos dos canales para las bandas individuales en representaciones correspondientes que tengan la misma resolución de tiempo o frecuencia. Pueden calcularse los valores de ponderación por banda. Sin embargo, en este aspecto, los valores de ponderación por banda no son aplicables a la representación del dominio espectral convertido o las dos o más representaciones espectrales combinadas. En su lugar, la ponderación espectral se aplica a la representación del dominio espectral original, de la cual se ha derivado la representación del dominio espectral combinada. De este modo, se asegura que las representaciones del dominio espectral ponderadas dependan únicamente de las representaciones del dominio espectral originales y únicamente los valores de ponderación que de todas maneras se basan en ciertas estimaciones de energía, preferiblemente usando una energía objetivo para una banda en los canales antes del mezclado y una energía objetivo para una banda de la señal de mezclado para reducir el número de canales de audio, se derivan de una o más representaciones del dominio espectral combinadas que son en algunos aspectos diferentes de las representaciones del dominio espectral originales.
50

Preferiblemente, el convertidor para convertir las representaciones del dominio espectral ponderadas en representaciones de tiempo tiene varios componentes. Un componente es el convertidor de frecuencia-tiempo real y un componente adicional es un procesamiento posterior por canal en el dominio del tiempo usando parámetros que se han transmitido vía, por ejemplo, información lateral con la señal de múltiples canales, de la cual provienen las representaciones del dominio espectral. De manera alternativa, el procesador posterior se aplica antes de la conversión de frecuencia-tiempo real. Los parámetros de control dirigen un procesamiento de dominio espectral de los canales individuales. Sin embargo, se prefiere tener primero el convertidor de frecuencia-tiempo y tener el procesador posterior para procesar posteriormente representaciones del dominio del tiempo de al menos dos canales usando parámetros de control por canal que se deriven de información lateral de la señal de múltiples canales o que sean generados o alimentados realmente en el mezclador mediante la entrada de usuario o cualquier otra generación de parámetros. Después de este procesamiento posterior del dominio del tiempo, el mezclador que realmente genera la señal de mezclado.
60

Este procedimiento proporciona un procesamiento de señales de audio de alta calidad debido a la aplicación de los valores de ponderación por banda a las representaciones del dominio espectral originales y, debido al hecho de que los valores de ponderación por banda de todos modos se basan en algún tipo de estimación de potencia o imagen se derivan de una o más representaciones del dominio espectral combinadas (creadas artificialmente). Por otro lado, se alcanza una alta flexibilidad de procesamiento debido al hecho de que cualquier procesamiento del dominio del tiempo o un dominio de la frecuencia probablemente requerido de canales individuales, pueda efectuarse aún, puesto que la etapa mezclado real es la última etapa en la cadena de procesamiento que ocurre, cuando todos los procesamientos de canal individual requeridos han sido aplicados. Además, este procedimiento es altamente eficiente, puesto que este procedimiento no requiere ningún mezclado de parámetros de control de lo que se requeriría, cuando la operación de mezclado sería la primera operación de procesamiento en la cadena de procesamiento.

De acuerdo con la presente invención, la operación de mezcla descendente se realiza usando ponderación espectral, donde los valores de ponderación por banda se calculan basándose en un valor de energía objetivo por banda, de modo que una energía en la banda de una señal de mezcla descendente esté en una relación predeterminada tal como igual o igual dentro de un rango de tolerancia de +/- 30% del valor más alto de las dos energías en las mismas bandas de los al menos dos canales. Los valores de ponderación por banda derivados por energía se aplican a representaciones del dominio espectral de al menos dos canales y la señal de mezcla se calcula usando las representaciones del dominio espectral ponderada de al menos dos canales ya sea en el dominio del tiempo como en el primer aspecto de la invención o en el dominio espectral según se requiera.

En un caso, donde las representaciones del dominio espectral son puramente reales, como en una transformación MDCT o cuando las representaciones del dominio espectral son puramente imaginarias, como cuando se aplica MDST (Transformación Sinusoidal Discreta Modificada), el estimador del valor de ponderación se configura para estimar, a partir de la representación del dominio espectral existente que es puramente real o puramente imaginaria, la otra representación del dominio espectral. Así, cuando existe una representación de dominio espectral de valor real, se estima la representación de dominio espectral imaginario, y cuando existe una representación de dominio espectral imaginario, se estima la representación de dominio espectral de valor real. Estos valores estimados se utilizan para calcular una energía del primer canal en la banda, para calcular una energía del segundo canal en la banda y para calcular términos mixtos entre los canales en función de un producto o una combinación lineal de valores espectrales de los al menos dos canales en la banda.

Este procedimiento de cálculo de los valores de ponderación por banda para una ponderación espectral en el contexto de una mezcla descendente puede aplicarse en el primer aspecto, en el que, entre la ponderación espectral y la mezcla descendente, se produce la transformada frecuencia-tiempo y algún procesamiento posterior en el dominio del tiempo. Con respecto al segundo aspecto de la invención, la representación del dominio espectral de uno o ambos canales que sean usados para calcular los valores de ponderación del dominio espectral de acuerdo con la característica de la energía objetivo se derivan ya sea de las representaciones del dominio espectral originales o se derivan de una o dos representaciones del dominio espectral combinadas que ya han sido generadas por la conversión de resolución espectral ilustrada con respecto al segundo aspecto de la invención o ilustrada con respecto al primer aspecto.

La mezcla descendente mediante ponderación espectral usando valores de ponderación por banda que se derivan basándose en un valor de energía objetivo por banda es, por un lado, muy eficaz debido a que la ponderación espectral puede realizarse fácilmente aplicando un mismo valor de ponderación a cada valor espectral de una banda, en particular, cuando se aplican anchos de banda motivados psicoacústicamente que aumentan desde anchos de banda pequeños a bajas frecuencias hasta anchos de banda elevados a altas frecuencias. Cuando, por ejemplo, se considera una banda alta que tiene, por ejemplo, 100 o más valores espectrales, únicamente se calcula un solo valor de ponderación para esta banda y este único valor de ponderación se aplica a cada valor espectral individual. Para este procedimiento, sólo se necesitan recursos computacionales moderados, ya que la ponderación mediante, por ejemplo, la multiplicación es un procedimiento de bajos recursos y bajo retardo y, al mismo tiempo, este procedimiento de aplicar el mismo valor de ponderación a cada valor espectral en una banda tiene un alto potencial para ser paralelizado mediante ciertos procesadores de hardware paralelos. Por otra parte, se obtiene una alta calidad de audio de la señal de mezcla descendente que está libre de cancelaciones de señal u otros artefactos que se producen, cuando los dos canales que se van a mezclar de forma descendente están en una relación de fase entre sí que es problemática en la mezcla descendente, es decir, cuando ambos canales están altamente correlacionados entre sí y tienen una cierta relación de fase.

Se analizan posteriormente realizaciones preferidas de la presente invención con respecto a las figuras anexas, en las cuales:

La Fig. 1 ilustra un mezclador descendente;

La Fig. 2 ilustra un mezclador descendente adicional;

La Fig. 3a ilustra una implementación del estimador del valor de ponderación;

La Fig. 3b ilustra otra implementación del estimador del valor de ponderación;

La Fig. 4a ilustra diferentes resoluciones de tiempo/frecuencia en distintos canales;

5 La Fig. 4b ilustra una representación espectral que muestra una resolución espectral alta, una resolución espectral media y una resolución espectral baja;

La Fig. 5a ilustra la estimación del valor de ponderación que da como resultado una baja resolución de frecuencia y una baja resolución de tiempo;

10 La Fig. 5b ilustra un procedimiento realizado por el estimador del valor de ponderación que da como resultado una alta resolución de frecuencia y una baja resolución de tiempo;

La Fig. 5c ilustra una implementación de la estimación del valor de ponderación que da como resultado una baja resolución de frecuencia y una alta resolución de tiempo;

La Fig. 5d ilustra un procedimiento adicional del estimador del valor de ponderación que da como resultado una alta resolución de frecuencia y una alta resolución de tiempo;

20 La Fig. 6 ilustra una realización de un aparato para convertir una resolución espectral;

La Fig. 7 ilustra una realización adicional del aparato para convertir una resolución espectral;

25 La Fig. 8 ilustra una realización del mezclador descendente de acuerdo con la invención;

La Fig. 9 ilustra una realización adicional del mezclador descendente.

La Fig. 1 ilustra un mezclador descendente.

30 La mezcladora comprende un estimador del valor de ponderación 100, un ponderador espectral 200 conectado con el estimador del valor ponderado 100 y una entrada para un primer canal o canal izquierdo y un segundo canal o canal derecho. El ponderador espectral 200 está conectado a un convertidor 300 para convertir las representaciones ponderadas del dominio espectral de los al menos dos canales en representaciones de tiempo de los al menos dos canales. Estas representaciones de tiempo se envían a un mezclador para mezclar las representaciones de tiempo de al menos dos canales y obtener una señal de mezcla descendente en el dominio del tiempo. Preferiblemente, el convertidor 300 comprende un convertidor frecuencia-tiempo 310 y un procesador posterior 320 conectado posteriormente. El convertidor frecuencia-tiempo 310 realiza en realidad la conversión de las representaciones del dominio espectral ponderado en el dominio del tiempo y el procesador posterior 320 que es una característica opcional realiza un procesamiento independiente del canal del primer canal y del segundo canal ya presentes en el dominio del tiempo usando parámetros de control para el canal izquierdo y el canal derecho, respectivamente. El convertidor 300 está configurado para generar, mediante el convertidor frecuencia-tiempo 310, las representaciones de tiempo en bruto usando un algoritmo de conversión espectro-tiempo y, adicionalmente, el convertidor 300 está configurado para procesar posteriormente, mediante el procesador posterior 320, las representaciones de tiempo en bruto individualmente y, en particular, en dirección de procesamiento de señal antes de la mezcla por el mezclador usando información de control separada para los canales para obtener las representaciones de tiempo de los al menos dos canales.

35 Preferiblemente, el procesador posterior 320 está configurado para efectuar, como la operación de procesamiento posterior, una filtración posterior de bajos, un procesamiento TCX-LTP (transformación de predicción a largo plazo de excitación codificada) o una síntesis LPC (codificación de predicción lineal). La ventaja del procesador posterior que opera sobre los canales ponderados espectralmente, pero que opera antes del mezclado real en la señal de mezcla es que los parámetros que están disponibles como parámetros separados para el canal izquierdo y derecho o, generalmente, para un canal individual de dos o más canales de la señal de múltiples canales puede usarse aún sin ningún parámetro de mezclado. De lo contrario, este procedimiento sería necesario cuando la mezcla descendente se realizaría junto con la ponderación espectral, de modo que a la salida del convertidor frecuencia-tiempo 310 ya existiera una señal de mezcla descendente en el dominio del tiempo.

50

55

60 Generalmente, la señal de múltiples canales puede comprender dos canales, es decir, el canal izquierdo y el canal derecho o la señal de múltiples canales comprende más de dos canales como tres canales o más. En esa situación, el estimador del valor de ponderación 100 se configura para calcular una pluralidad de primeros valores de ponderación por banda para una pluralidad de bandas de un primer canal de al menos dos canales y para calcular una segunda pluralidad de valores de ponderación por banda para la pluralidad de bandas de un segundo canal de al menos dos canales. Además, el estimador de valores de ponderación 100 está configurado para calcular la pluralidad de primeros

valores de ponderación en función de la banda para una pluralidad de bandas de un primer canal de la señal multicanal que tiene más de dos canales y para calcular una segunda pluralidad de valores de ponderación en función de la banda para la pluralidad de bandas de un segundo canal de los más de dos canales y para calcular la pluralidad adicional de valores de ponderación en función de la banda para la pluralidad de bandas de un tercer canal o incluso otro canal de los más de dos canales.

En particular, las representaciones en el dominio espectral de los al menos dos canales comprenden cada uno un conjunto de bandejas de frecuencias, donde los valores espectrales están asociadas a las bandejas de frecuencias. En particular, el estimador de valores de ponderación 100 está configurado para calcular los valores de ponderación por bandas, donde cada banda comprende uno, dos o más valores espectrales y, preferentemente, el número de bandejas de frecuencias por banda aumenta con las bandas que tienen una frecuencia central más alta, de modo que se obtiene la subdivisión psicoacústicamente motivada de las representaciones del dominio espectral en bandas con anchos de banda no uniformes.

En la Fig. 2 se ilustra una implementación preferida del mezclador descendente. La señal multicanal está disponible como un flujo de bits estéreo y se introduce en un decodificador estéreo 500 que se implementa preferentemente como un decodificador estéreo MDCT. Además, el estimador del valor de ponderación comprende un calculador del valor izquierdo 110, un calculador del valor derecho 112 y, adicionalmente, un estimador de la parte imaginaria 120 para el canal izquierdo y un estimador de la parte imaginaria 122 para el canal derecho. En la realización de la Fig. 2, el decodificador estéreo 500 es un decodificador estéreo MDCT, lo que significa que las representaciones espectrales decodificadas de los canales izquierdo y derecho tienen valores espectrales puramente reales, es decir, valores MDCT. Los estimadores imaginarios 120, 122 generarán valores espectrales puramente imaginarios, es decir, valores MDST (transformada sinusoidal discreta modificada). A partir de estos elementos de información, es decir, las representaciones espectrales y los valores espectrales estimados, se calculan los factores de ponderación y se envían al ponderador espectral 200 para que realice una ponderación por bandas, como se indica en la Fig. 2. Las representaciones ponderadas del dominio espectral se envían a los correspondientes convertidores frecuencia-tiempo 310 que se implementan como un convertidor IMDCT para cada canal. Además, se ilustra también un procesador posterior opcional 320 para cada canal y, los datos transformados y opcionalmente procesados posteriormente se introducen en el mezclador descendente DMX 400 para generar la señal de mezcla descendente en el dominio de tiempo que es, en la realización de la Fig. 2, una señal de salida mono, pero también puede ser una señal multicanal siempre que el número de uno o más canales de la señal de mezcla descendente sea inferior al número de canales de la señal multicanal antes de la mezcla descendente.

Alternativamente, cuando el decodificador multicanal o el decodificador estéreo 500 se implementa como un decodificador de valor imaginario tal como un decodificador MDST, los bloques 120, 122 estimarían datos puramente reales tales como valores MDCT. Así, en general, el estimador del valor de ponderación 100 está configurado para estimar una representación espectral imaginaria cuando la representación del dominio espectral es puramente real o para estimar la representación espectral real cuando la representación del dominio espectral original es puramente imaginaria. Además, el estimador del valor de ponderación 110 está configurado para estimar los valores de ponderación usando la representación espectral imaginaria estimada o la representación espectral real estimada, según sea el caso. Esto es particularmente útil para un cálculo de valores de ponderación espectral por banda que se basa en un valor de energía objetivo por banda, de modo que, de acuerdo con la invención, una energía en una banda de la señal de mezcla descendente está en una relación predeterminada con las energías en las mismas bandas de los al menos dos canales. Preferiblemente, la relación predeterminada es que la energía en una banda de la señal de mezcla descendente es la suma de las energías de las mismas bandas en los al menos dos canales. Sin embargo, también son útiles otras relaciones predeterminadas. A modo de ejemplo, la relación predeterminada puede alcanzar 75% a 125% de la suma de los dos canales como la energía de la banda correspondiente de la señal de mezcla. Sin embargo, en una realización más preferida, la relación predeterminada es la igualdad o la igualdad dentro de un rango de tolerancia de +/- 10%.

La Fig. 3a ilustra una implementación preferida del estimador de valor de ponderación 100. Particularmente, esta implementación es útil para calcular los valores de ponderación cuando las representaciones del dominio espectral de al menos dos canales tienen resoluciones de tiempo y frecuencia diferentes. Como se muestra en el bloque o etapa 130, el estimador de valor de ponderación 100 está configurado para comprobar, si las resoluciones de tiempo/frecuencia de las representaciones de dominio espectral del primer y segundo canal son diferentes entre sí. En caso de resoluciones de tiempo o frecuencia iguales, el estimador de valor de ponderación 100 está configurado para calcular los factores de ponderación de banda o los valores de ponderación de banda como se indica por w_L para el primer canal o canal izquierdo y como se indica por w_R para el segundo canal o canal derecho.

Alternativamente, cuando se determina por el estimador de valor de ponderación 100 en el bloque 130 que las resoluciones de tiempo o frecuencia no son iguales entre los canales izquierdo y derecho o primero y segundo para un cierto periodo de tiempo como se ilustra más adelante con respecto a la Fig. 4a, el estimador de valor de ponderación 100 está configurado para calcular 132 una o dos representaciones combinadas de dominio espectral. En particular, una primera representación en el dominio espectral de un primer canal de los al menos dos canales

tiene una primera resolución de tiempo o una primera resolución de frecuencia, y una segunda representación en el dominio espectral de un segundo canal de los al menos dos canales tiene una segunda resolución de tiempo o una segunda resolución de frecuencia, en las que la segunda resolución de tiempo es diferente de la primera resolución de tiempo o en las que la segunda resolución de frecuencia es diferente de la primera resolución de frecuencia. El estimador del valor de ponderación 100 se configura para convertir o calcular 132 la primera representación del dominio espectral en una representación del dominio espectral combinada que tiene una segunda resolución de tiempo o la segunda resolución de frecuencia y para calcular los valores de ponderación por banda usando la representación del dominio espectral combinada y la segunda representación del dominio espectral. Alternativamente, la segunda representación del dominio espectral se convierte en una representación combinada del dominio espectral que tiene la primera resolución de tiempo o la primera resolución de frecuencia y los valores de ponderación por banda se calculan usando la representación combinada del dominio espectral y la primera representación del dominio espectral. Alternativamente, cuando la primera representación en el dominio espectral de un primer canal tiene una primera resolución de tiempo o una primera resolución de frecuencia y la segunda representación en el dominio espectral de un segundo canal de los al menos dos canales tiene una segunda resolución de tiempo o una segunda resolución de frecuencia, en donde la segunda resolución de tiempo es diferente de la primera resolución de tiempo o en donde la segunda resolución de frecuencia es diferente de la primera resolución de tiempo, el estimador de valor de ponderación 100 está configurado para convertir o calcular 132 la primera representación del dominio espectral en una primera representación combinada del dominio espectral que tiene una tercera resolución de tiempo o una tercera resolución de frecuencia, en donde la tercera resolución de tiempo es diferente de la primera resolución de tiempo o de la segunda resolución de tiempo y en donde la tercera resolución de frecuencia es diferente de la primera resolución de frecuencia y/o de la segunda resolución de frecuencia. Además, la segunda representación espectral también se convierte en una segunda representación espectral combinada que tiene la tercera resolución de tiempo o la tercera resolución de frecuencia y los valores de ponderación por banda se calculan usando la primera representación espectral combinada y la segunda representación espectral. Dependiendo de la situación real, como se describe más adelante con respecto a las Fig. 5a a 5d, también puede existir la situación en la que los valores o factores de ponderación por banda calculados por el bloque 134 no se usen para ponderar espectralmente en realidad, sino que los factores de ponderación por banda derivados se calculen como se ilustra en 136 en la Fig. 3a.

En general, y suponiendo que el primer canal tiene una primera resolución de tiempo baja y una primera resolución de frecuencia alta, y suponiendo también que el segundo canal tiene una segunda resolución de tiempo alta y una segunda resolución de frecuencia baja, la funcionalidad del estimador de valores de ponderación 100 puede seleccionar una de las cuatro maneras diferentes de hacer el emparejamiento entre las resoluciones entre el primer y el segundo canal en el dominio espectral para calcular valores de ponderación en el dominio espectral para estos canales.

La Fig. 5a ilustra una primera realización, en la que los valores de ponderación por banda se calculan a partir de dos representaciones combinadas del dominio espectral en las que las dos representaciones combinadas del dominio espectral tienen una baja resolución de frecuencia y una baja resolución de tiempo.

En la segunda realización ilustrada en la Fig. 5b, sólo se calcula una única representación combinada del dominio espectral a partir de la representación de baja resolución de frecuencia, de modo que los valores de ponderación por banda se calculan a partir de un par de representaciones del dominio espectral que tienen tanto la resolución de alta frecuencia como la de baja resolución de tiempo.

La Fig. 5c ilustra una tercera realización adicional, en la que se calcula y usa una única representación combinada para el cálculo de los valores de ponderación por banda en el dominio espectral usando dos representaciones en el dominio espectral que tienen ambas la resolución de baja frecuencia y la alta resolución de tiempo.

En una cuarta realización ilustrada en la Fig. 5d, el estimador del valor de ponderación está configurado para calcular los valores de ponderación por banda usando dos representaciones combinadas que están ambas en un formato que muestra la alta resolución de frecuencia y la alta resolución de tiempo.

La Fig. 4a ilustra una situación en la que hay dos resoluciones diferentes (en tiempo y/o frecuencia) en el primer canal y en el segundo canal. La primera porción de la Fig. 4a muestra una trama que tiene un bloque largo en el primer canal y dos bloques cortos subsiguientes en el segundo canal. El bloque largo puede, por ejemplo, ser un bloque de TCX20. Los bloques cortos pueden ser dos bloques de TCX10 subsiguientes. Además, la Fig. 4a ilustra otra trama que se subdivide en dos subtramas A, B, donde, en el primer canal, la subtrama A tiene un bloque corto y, en el segundo canal, la subtrama también tiene un bloque corto. Sin embargo, en la subtrama B de la segunda trama de la Fig. 4a, el primer canal tiene un bloque corto y el segundo canal tiene dos bloques muy cortos, es decir, un bloque muy corto por cada subsubtrama. Los bloques muy cortos pueden, por ejemplo, ser bloques de TCX5. Generalmente, los bloques largos son más largos que los bloques cortos y los bloques cortos son más largos que los bloques muy cortos y, por supuesto, los bloques muy cortos son más cortos que los bloques largos. Naturalmente, no es necesario que un bloque largo tenga la misma longitud que dos bloques cortos. De manera alternativa, pueden existir tres bloques cortos que tengan una longitud combinada que sea igual a la longitud de un bloque largo o pueden existir cuatro bloques cortos

como un bloque muy corto por cada subtrama. También pueden existir otras subdivisiones, es decir, que dos bloques largos en el primer canal tengan una longitud combinada que sea igual a la longitud de tres bloques cortos en el segundo canal. Las longitudes de los bloques largos, cortos y muy cortos no necesariamente tienen que estar en una relación entera entre sí. Además, incluso puede haber más de tres longitudes de bloque diferentes, como más de tres longitudes de bloque o solo dos longitudes de bloque diferentes.

La Fig. 4b ilustra una representación de un espectro con una alta resolución espectral en la primera línea. Los valores espectrales se indican en números enteros a lo largo de la línea de frecuencia y la Fig. 4b ilustra tres bandas subsiguientes b_1 , b_2 , b_3 , donde cada banda que representa frecuencias más altas es más ancha que cada banda que representa frecuencias más bajas. En la situación de alta resolución espectral, como en un espectro TCX20, la banda más baja b_1 tiene cuatro líneas espectrales o valores espectrales o bandejas espectrales. La segunda banda b_2 tiene, en la realización, ocho valores espectrales y la tercera banda espectral b_3 tiene doce bandejas espectrales. La transferencia o conversión de la alta resolución espectral a una representación de resolución espectral media tiene como resultado que, a partir de la representación espectral de alta resolución, los valores espectrales se combinan (o deciman) de modo que la resolución espectral media, como una resolución TCX10, tiene dos bandejas espectrales para la primera banda, cuatro bandejas espectrales para la segunda banda b_2 y seis bandejas espectrales para la tercera banda b_3 . Comparando de nuevo esta resolución espectral media con una representación de baja resolución espectral como la que se produce en un bloque TCX5, la primera banda sólo tendría una única bandeja de frecuencia, la segunda b_2 tendría dos bandejas de frecuencia y la tercera banda espectral b_3 tendría tres bandejas espectrales. La resolución espectral media podría convertirse en la resolución espectral baja combinando dos o más líneas espectrales adyacentes o por una operación de decimación.

Por otra parte, una representación de baja resolución espectral podría convertirse en una representación de mayor resolución mediante interpolación o copia o copia y filtrado, de modo que, por ejemplo, a partir de las dos bandejas espectrales de la primera banda b_1 para la resolución espectral media, podrían calcularse cuatro bandejas espectrales de alta resolución 1, 2, 3, 4, como se ilustra en la Fig. 4b.

El objetivo de este nuevo enfoque es proporcionar un método de mezcla descendente activo, sin retardo y en función de la banda, para la conversión de estéreo a mono, en el que sólo se realiza la ponderación en función de la banda de las bandas espectrales de los dos canales en el dominio de la frecuencia, mientras que la mezcla descendente real a una señal mono se realiza después de la transformación en el dominio del tiempo mediante la suma y el escalado de las dos señales ponderadas espectralmente.

En caso de que las representaciones en el dominio espectral de ambas señales tengan diferentes resoluciones de tiempo-frecuencia (es decir, un tamaño de bloque más corto para una señal), el cálculo de la ponderación se adapta combinando bandejas espectrales vecinas tanto de tiempo como espectralmente para que el cálculo de espectros cruzados pueda realizarse en las mismas regiones de tiempo-frecuencia.

Con este método, no es necesario que la resolución de tiempo-frecuencia de los dos canales estéreo sea uniforme, ya que se puede realizar una ponderación de banda de los canales si éstos difieren en este aspecto, mientras que la conversión crítica de estéreo a mono se realiza más tarde, cuando ambos canales ponderados espectralmente ya se han transformado de nuevo al dominio del tiempo.

Las realizaciones proporcionan una mezcla estéreo-mono optimizada y sin retardo en el lado del decodificador.

Los aspectos preferidos se refieren a una mezcla descendente activa con ponderación por bandas con etapas de ponderación (en el dominio de la frecuencia) y mezcla (en el dominio del tiempo) separadas.

Otros aspectos preferidos se refieren a una combinación de tiempo/espectral de bandejas de frecuencias para la correlación de espectros cruzados en el caso de canales con diferentes representaciones espectrales, donde estos aspectos pueden usarse por separado de los aspectos de mezcla descendente o junto con los aspectos de mezcla descendente.

A diferencia de los códecs estereofónicos paramétricos como [7], en los que sólo se transmite una señal central ya mezclada junto con varios parámetros laterales que representan la imagen estereofónica, en una aplicación estereofónica discreta basada en MDCT no hay mezcla disponible en el decodificador, en el que ambos canales siempre se codifican directamente con un codificador TCX. Por tanto, la mezcla descendente debe generarse íntegramente en el decodificador.

La Fig. 3b ilustra una implementación preferida del estimador de valor de ponderación 100 ilustrado en la Fig. 1. En la etapa 140, el estimador del valor de ponderación estima los correspondientes valores espectrales de valor real o imaginario por bandejas de frecuencia a partir del primer canal y el segundo canal o, alternativamente, a partir del primer canal y una representación de dominio espectral combinada o a partir del segundo canal y la representación de dominio espectral combinada o a partir de una primera representación de dominio espectral combinada y una

segunda representación de dominio espectral combinada. Generalmente, el estimador del valor de ponderación está configurado para calcular el primer valor de ponderación y el segundo valor de ponderación usando una energía del primer canal en la banda, y energía del segundo canal en la banda y un término mixto que depende de un producto o una combinación lineal de valores espectrales de los al menos dos canales en la banda. En la Fig. 3b, la energía del primer canal y la energía del segundo canal se calculan a modo de ejemplo en el bloque 140. Además, en el bloque 148 se calcula un término mixto que depende de un producto y en el bloque 146 se calcula otro término mixto que depende de una combinación lineal. Además, en el bloque 144 se calcula una "amplitud" por banda que corresponde a una raíz cuadrada de la potencia de las bandejas espectrales por banda.

Así, como se ilustra en la Fig. 3b, el primer valor de ponderación w_L se calcula a partir de las amplitudes por banda para ambos canales y en función de un término mixto y, preferentemente, el término mixto en función de la combinación lineal ilustrada en el bloque 146. Además, se prefiere que el vector de ponderación w_L por banda se calcule usando el valor de ponderación w_R por banda, es decir, para el otro canal. El valor para el otro canal, es decir, w_R por banda se calcula preferiblemente basándose en el término mixto que depende del producto ilustrado en 148 y las "amplitudes" por banda derivadas por el bloque 144 a partir de las potencias por banda en los canales correspondientes, tal como se determinan en el bloque 142.

Así, preferentemente, se usa como "amplitudes" la raíz cuadrada de una energía de valores espectrales sumados entre sí en la banda a partir de las representaciones en el dominio espectral de los al menos dos canales, pero también pueden usarse otras "amplitudes", como las "amplitudes" derivadas de las potencias por un exponente inferior a 1 y diferente de 1/2. Los valores espectrales de una banda se combinan linealmente, es decir, que se suman entre sí y se toma una raíz cuadrada de cualquier otro exponente con un exponente menor que 1 del valor resultante, donde, preferiblemente, las potencias para ambos canales en la banda se usan adicionalmente.

Como término mixto que representa el producto, puede determinarse también, por ejemplo, en el cálculo del bloque 148, un valor absoluto de un producto escalar complejo entre los valores espectrales en la banda del primer canal y los valores espectrales en la banda del segundo canal. Preferiblemente, se aplica la misma ponderación determinada por el ponderador espectral 200 a cada valor espectral de la banda de uno de los al menos dos canales y otra ponderación a cada valor espectral de la banda de otro canal de los al menos dos canales.

A continuación, se ilustra una realización preferida del cálculo de los factores de ponderación por banda que puede usar el estimador de valor de ponderación 100.

Dado que el uso de una mezcla descendente pasiva tiene sus desventajas, como se ha indicado anteriormente, el uso de un esquema de mezcla descendente activa conduce a mejoras significativas para muchos elementos. Añadir otra etapa de decodificación que incluya una transformada DFT para ambos canales después de la decodificación estéreo no es factible tanto por razones de complejidad como de retardo, por lo que el proceso de mezcla descendente se realiza como una combinación de procesamiento en el dominio MDCT y en el dominio de tiempo.

En primer lugar, se calculan las ponderaciones por banda y se aplican a las representaciones MDCT de ambos canales. Esto ocurre después del procesamiento estéreo (por ejemplo, MS inversa, etc.) y directamente antes de la retrotransformación IMDCT. Las ponderaciones se calculan con el mismo esquema que ya se usa en el codificador estéreo basado en DFT descrito en [7], apuntando a la energía del canal medio de fase rotada:

$$E_{objetivo} = \left| \frac{L + Re^{-j\varphi}}{2} \right|^2 = \frac{\langle L, L \rangle + \langle R, R \rangle + 2|\langle L, R \rangle|}{4} = \frac{|L|^2 + |R|^2 + 2|\langle L, R \rangle|}{4},$$

donde L y R representan las magnitudes espectrales de canal izquierdo y derecho. A partir de esta energía objetivo, la ponderación de los canales puede calcularse para cada banda espectral del siguiente modo:

$$w_R = \frac{1}{2\sqrt{2}} \frac{\sqrt{|L|^2 + |R|^2 + 2|\langle L, R \rangle|}}{|L| + |R|}$$

y

$$w_L = w_R + 1 - \frac{|L + R|}{|L| + |R|}$$

Estas ponderaciones o valores de ponderación por banda w_R y w_L se calculan por banda espectral, abarcando cada banda varias bandejas MDCT, empezando con un número bajo de bandejas para las bandas más bajas, por ejemplo 4, y aumentando hacia frecuencias más altas hasta varios o muchas bandejas para la banda más alta, por ejemplo 160.

Como los coeficientes MDCT transmitidos son sólo de valor real, los valores MDST complementarios que se requieren para la ponderación de preservación de energía se obtienen para cada canal mediante la estimación [9]

$$MDST_i = MDCT_{i+1} - MDCT_{i-1},$$

5

donde i especifica el número de bandeja espectral.

Usando esta estimación, $|L|/|R|$ se calculan para cada banda b como

$$10 \quad |L| = \sqrt{\sum_{i \in b} (MDCT_{i,l}^2 + MDST_{i,l}^2)}, \quad |R| = \sqrt{\sum_{i \in b} (MDCT_{i,r}^2 + MDST_{i,r}^2)}$$

$|L + R|$ se calcula como

$$15 \quad |L + R| = \sqrt{|L|^2 + |R|^2 + 2 \left(\sum_{i \in b} (MDCT_{i,l} MDCT_{i,r} + MDST_{i,l} MDST_{i,r}) \right)^2}$$

y $\langle L, R \rangle$ se calcula como la magnitud o valor absoluto del producto escalar complejo

$$\begin{aligned} & \langle L, R \rangle \\ &= \sqrt{\left(\sum_{i \in b} (MDCT_{i,l} MDCT_{i,r} + MDST_{i,l} MDST_{i,r}) \right)^2 + \left(\sum_{i \in b} (MDST_{i,l} MDCT_{i,r} - MDCT_{i,l} MDST_{i,r}) \right)^2} \end{aligned}$$

20

donde i especifica el número de bandeja dentro de la banda espectral b .

a pesar de que la transformada es diferente y las energías son sólo estimadas, las ponderaciones resultantes siguen conduciendo a una mezcla descendente similar al de [7].

25

En una segunda etapa, los dos canales ponderados se mezclan en el dominio de tiempo mediante una simple suma y escalado de los dos canales ponderados espectralmente.

Se hace referencia a la Fig. 2.

30

La razón de este enfoque combinado es doble: por un lado, al transformar ambos canales de nuevo al dominio de tiempo, el filtrado posterior, por ejemplo TCX-LTP -que también funciona en el dominio de tiempo- puede ejecutarse en ambos canales usando los parámetros (por ejemplo, el tono) extraídos de la codificación central de los canales individuales, evitando así la necesidad de intentar encontrar parámetros promediados que se ajusten a la mezcla descendente. En segundo lugar, y lo que es más importante, estéreo MDCT está configurado para permitir diferentes codificadores centrales y/o decisiones de solapamiento para los dos canales. Concretamente, esto significa que un canal puede codificarse con, por ejemplo, un bloque largo TCX20 (trama de 20 ms, mayor resolución de frecuencia, menor resolución de tiempo), mientras que el otro se codifica con, por ejemplo, dos bloques cortos TCX10 (2 subtramas de 10 ms, menor resolución de frecuencia, mayor resolución de tiempo), en los que uno o ambos bloques cortos pueden dividirse a su vez en dos subtramas TCX5 (2 x 5 ms). Esto hace prácticamente imposible una mezcla completa en el dominio de la frecuencia. Sin embargo, la ponderación por bandas puede realizarse directamente en el dominio MDCT.

35

Una realización ilustrada en la Fig. 5a funciona como sigue: Para el caso especial de núcleos diferentes en los dos canales, hay que adaptar ligeramente el cálculo de la correlación de espectros cruzados como parte del cálculo de la ponderación. Debido a las diferentes resoluciones de frecuencia y tiempo del TCX20 y el TCX10, no es posible calcular directamente el producto escalar entre la izquierda y la derecha. En su lugar, las bandejas de la TCMD deben combinarse de modo que cubran las mismas regiones de tiempo-frecuencia. Para TCX20 esto significa combinar siempre dos bandejas vecinas, mientras que para TCX10 cada bandeja de la primera subtrama debe combinarse con la misma bandeja de la subtrama siguiente, por ejemplo

50

$$MDCT_{i,l,combinado} = MDCT_{2i,l} + MDCT_{2i+1,l}$$

y

$$MDCT_{i,r,combinado} = MDCT_{i,k0,r} + MDCT_{i,k1,r}$$

si $MDCT_i$ es un espectro TCX20 MDCT y $MDCT_r$ un espectro TCX10 MDCT con 2 subtramas donde i especifica el número de bandeja espectral y $k0$ y $k1$ las subtramas TCX10. La misma combinación se realiza también con los espectros MDST estimados.

A continuación, se calcula la correlación de espectros cruzados $\langle L, R \rangle$ y/o el valor de $|L + R|$ con las bandejas combinadas resultantes. Esto conduce a una estimación de la correlación algo más gruesa, pero se ha comprobado que es totalmente suficiente.

Otra realización ilustrada en la Fig. 5b funciona como sigue: Para el caso especial de núcleos diferentes en los dos canales, hay que adaptar ligeramente el cálculo de la correlación de espectros cruzados como parte del cálculo de la ponderación. Debido a las diferentes resoluciones de frecuencia y tiempo del TCX20 y el TCX10, no es posible calcular directamente el producto escalar entre la izquierda y la derecha. Para hacerlo posible, el espectro de la (sub)trama con menor resolución espectral se convierte en una aproximación de un espectro con el doble de resolución espectral mediante cálculo:

$$MDCT_{2i,k0} = \frac{1}{2} (MDCT_{i,k1} + (-1)^i MDCT_{i,k0})$$

y

$$MDCT_{2i+1,k0} = \frac{1}{2} (MDCT_{i,k1} - (-1)^i MDCT_{i,k0})$$

donde i especifica el número de bandeja espectral y $k0$ y $k1$ las subtramas con resolución más baja. Esas resoluciones y sustracciones pueden verse como operaciones de filtración de paso alto y bajo que dividen una bandeja de resolución más baja en dos bandejas de resolución más altas donde la filtración depende si el número de bandeja i es par o impar (comenzando con $i = 0$ para la bandeja más baja).

Esto significa que si un canal es TCX20, el otro canal se convierte a la misma resolución espectral. Si una o ambas subtramas del otro canal se subdividen de nuevo en dos "sub-subtramas" TCX5, éstas se convierten primero a resolución TCX10 mediante el mismo filtrado antes de dividirse de nuevo para llegar a la representación final TCX20.

Aunque ninguno de los canales sea TCX20, puede ser necesaria la conversión a la resolución más alta para una o ambas subtramas en caso de que haya TCX10 en un canal y TCX5 en el otro. Como un ejemplo, si el canal izquierdo es TCX10 en la subtrama A y $2 \times$ TCX5 en la subtrama B, mientras el canal derecho es $2 \times$ TCX5 en la subtrama A y TCX10 en la subtrama B, ambos canales se convertirán para tener la resolución de TCX10 en ambas subtramas (convertir la subtrama B para el canal izquierdo, A para el canal derecho). Si en el mismo ejemplo el canal derecho también es TCX 10 para la subtrama A y $2 \times$ TCX5 para B, entonces no se efectúa la conversión; es decir que la subtrama A será mezclada de forma descendente para reducir el número de canales de audio con la resolución de TCX10, B con TCX5.

Las estimaciones de MDST y ponderaciones de canal finales se calculan entonces usando esos espectros convertidos. Las propias ponderaciones se aplican a los espectros de entrada originales, lo que significa que, en caso de conversión, cada ponderación calculada se aplica a todas las bandejas que cubren el mismo rango de frecuencias en la resolución original más baja para cada subtrama.

Al separar la etapa de ponderación de una mezcla descendente activa por bandas de la etapa de mezcla propiamente dicha, el nuevo método es capaz de producir una señal mono con las ventajas de la mezcla descendente activa, pero sin retardo ni complejidad adicionales e independientemente de la resolución de tiempo-frecuencia elegida para los canales individuales.

También permite el uso del procesamiento posterior adicional en el dominio del tiempo (por ejemplo, filtración posterior TCX-LTP que usa información de tono) en ambos canales sin necesidad de una mezcla descendente de parámetros específica.

La Fig. 5a ilustra la primera alternativa, en la que se generan dos representaciones combinadas del dominio espectral. La primera representación de dominio espectral combinada se calcula sumando dos bandejas vecinas de la

representación de dominio espectral de alta resolución ilustrada a la izquierda de la Fig. 5a para obtener la primera representación de dominio espectral combinada.

5 Además, las dos representaciones de resolución espectral baja ilustradas en TCX10 en el centro de la Fig. 5a se combinan entre sí para obtener la segunda representación combinada del dominio espectral. El estimador del valor de ponderación 100 está configurado para calcular los factores de ponderación izquierdo y derecho w_L y w_R a partir de estas dos representaciones combinadas del dominio espectral.

10 Con respecto a la ponderación espectral realmente realizada por el ponderador espectral 200, el factor de ponderación para el canal izquierdo se aplica a la representación original del canal izquierdo, es decir, la representación TCX20 ilustrada a la izquierda de la Fig. 5a. Además, los valores de ponderación por banda para el canal derecho representado por dos bloques TCX10 subsiguientes en el tiempo se aplican a ambos bloques TCX10. El mismo valor de ponderación por banda se aplica a las bandas correspondientes de los dos bloques TCX10 subsiguientes en el tiempo que se ilustran en el centro de la Fig. 5a.

15 en la segunda alternativa ilustrada en la Fig. 5b, sólo se calcula una única representación combinada en el dominio espectral, como se ilustra para varios casos diferentes. Cuando, por ejemplo, una subtrama del primer canal tiene dos tramas muy cortas, como TCX5, y la subtrama siguiente tiene una única trama TCX10, y cuando el segundo canal tiene, por ejemplo, dos tramas TCX10, la representación combinada en el dominio espectral se calcula para la primera subtrama mientras que, para la segunda subtrama, el primer y el segundo canal ya están en la representación TCX10.

20 En este ejemplo, el ponderador espectral 200 está configurado para aplicar los factores de ponderación de alta resolución espectral a las bandas correspondientes en las subtramas que representan cada una, cinco milisegundos, por ejemplo. Además, los factores de ponderación de alta resolución se aplican a las correspondientes representaciones originales del dominio espectral del otro canal que tiene una trama TCX10 corta en la primera subtrama A, por ejemplo.

25 Alternativamente, la situación es tal que el primer canal tiene una representación ilustrada a la izquierda de la Fig. 5b y el segundo canal tiene una representación ilustrada a la derecha de la Fig. 5b, la representación del primer canal se convierte en una única representación combinada del dominio espectral a través de las dos etapas de la izquierda en la Fig. 5b hacia el centro y del centro en la Fig. 5b hacia la derecha. La resolución de frecuencia se usa para calcular los factores de ponderación y los factores de ponderación correspondientes se aplican a la representación de alta resolución de frecuencia y baja resolución de tiempo del segundo canal que tendría una resolución ilustrada a la derecha en la Fig. 5b, y los mismos valores para una banda se aplicarían a todas las subtramas individuales A, B y la subtrama siguiente ilustrada por D y C en la Fig. 5b.

30 La Fig. 5c ilustra otra alternativa en la que los valores de ponderación del dominio real se calculan a partir de una representación de baja resolución de frecuencia y de alta resolución de tiempo. El primer canal es, por ejemplo, una representación TCX20 y el segundo canal es, por ejemplo, una secuencia de dos representaciones TCX10. En contraste con la alternativa ilustrada en la Fig. 5b, la representación combinada es ahora una representación de alta resolución de tiempo y baja resolución de frecuencia ilustrada en la esquina superior derecha de la Fig. 5c. Los factores de ponderación del dominio espectral se calculan a partir de la representación combinada, por un lado, y de la representación original del dominio espectral del segundo canal, ilustrada en la esquina inferior izquierda de la Fig. 5c.

35 Se obtienen dos conjuntos de valores de ponderación por banda, es decir, uno para cada subtrama. Estos valores se aplican a las subtramas correspondientes del segundo canal. Sin embargo, debido al hecho de que el primer canal sólo tiene una única representación en el dominio espectral para toda la trama, los valores de ponderación derivados del dominio espectral se calculan como se ilustra en el bloque 136 de la Fig. 3a. Un procedimiento para calcular un valor de ponderación derivado del dominio espectral consiste en realizar una suma ponderada de los correspondientes valores de ponderación de una misma banda para las dos (o más) subtramas, en la que cada valor de ponderación se pondera, por ejemplo, mediante 0,5 en la suma ponderada que da lugar a una operación de promediado. Otra alternativa sería calcular una media aritmética o geométrica de los valores de ponderación para las dos subtramas o cualquier otro procedimiento para obtener un único valor de ponderación a partir de dos valores de ponderación para una banda en una trama. Una opción podría ser simplemente seleccionar uno de los dos valores y despreciar el otro, etc.

40 Además, para calcular la representación combinada en el dominio espectral a partir del primer canal, se puede usar un procedimiento como el comentado anteriormente con respecto a la Fig. 5a, es decir, se pueden sumar dos valores espectrales vecinos para reducir la resolución espectral. Esto también se ilustra en la Fig. 4b, donde la alta resolución espectral que tiene un cierto número de valores espectrales en una banda puede reducirse a la resolución espectral media que tiene un número menor de valores espectrales en la misma banda. Además, para duplicar los valores espectrales de las dos subtramas ilustradas en la esquina superior derecha de la Fig. 5c se pueden usar, por ejemplo, los mismos valores espectrales (de baja resolución espectral) para una banda en ambas subtramas, o se puede realizar algún tipo de decimación ponderado usando valores anteriores o posteriores, según el caso.

La Fig. 5d ilustra una realización adicional en la que el primer canal tiene una representación de alta frecuencia y baja resolución de tiempo, como una representación TCX20, y el segundo canal tiene una representación de baja frecuencia y alta resolución de tiempo, como una secuencia de dos tramas cortas, como dos tramas TCX10. La primera representación combinada del dominio espectral es una representación de alta resolución de frecuencia y alta resolución de tiempo y la segunda representación combinada del dominio espectral es, además, una representación de alta resolución de frecuencia y alta resolución de tiempo. El procedimiento ilustrado en la Fig. 5d puede, por ejemplo, realizarse de forma que, a partir del primer canal, se calcule la primera representación combinada en el dominio espectral tomando los mismos valores espectrales, pero ahora para dos marcos de tiempo posteriores ilustrados por TCX10. Alternativamente, también se puede realizar algún tipo de procesamiento de interpolación, etc., para duplicar el número de tramas de modo que, a partir de una trama TCX20, se calculen dos tramas TCX10 posteriores. Además, el segundo canal ya tiene la resolución de tiempo correcta, pero hay que duplicar la resolución de frecuencia. Para ello, se puede realizar un procedimiento de una línea inferior a una línea superior en la Fig. 4b, es decir, se puede procesar el valor espectral en una bandeja de frecuencia de la representación TCX10 para tener el mismo valor espectral para un par de bandejas de frecuencia. Para tener una energía correcta, se puede realizar algún tipo de ponderación. Alternativa o adicionalmente, se puede realizar algún tipo de interpolación avanzada para que las bandejas de frecuencia adyacentes entre sí en la segunda representación combinada del dominio espectral no tengan necesariamente exactamente el mismo valor espectral, sino valores diferentes. Los valores de ponderación del dominio espectral son calculados por el estimador de valores de ponderación 100 a partir de la primera representación combinada del dominio espectral y la segunda representación combinada del dominio espectral que se derivan de datos de alta resolución de frecuencia y alta resolución de tiempo.

El ponderador espectral 200 está configurado para aplicar los correspondientes valores de ponderación del dominio espectral al segundo canal, donde para cada subtrama existe un conjunto de valores de ponderación por banda. Con el fin de ponderar los datos TCX20 del primer canal, el estimador de valor de ponderación 100 está configurado para calcular de nuevo factores de ponderación de banda derivada 136, ya que sólo se requiere un conjunto de factores de ponderación de dominio espectral para ponderar la representación de dominio espectral de resolución de alta frecuencia y resolución de tiempo (TCX20) del primer canal. Un procedimiento de combinación para calcular los valores de ponderación por banda derivados podría ser, por ejemplo, un promedio.

La Fig. 6 ilustra un aparato para convertir una resolución espectral de una representación en el dominio espectral de un canal que comprende al menos dos subtramas, en la que cada subtrama comprende una pluralidad de valores espectrales que representan un tamaño de una bandeja de tiempo y un tamaño de bandeja de frecuencia. El calculador de valor espectral 160 incluido en el aparato para convertir, de acuerdo con el segundo aspecto, comprende un primer combinador de modo 170 y un segundo combinador de modo 180. Preferiblemente, el primer combinador de modo funciona como un procesador de paso bajo y el segundo combinador de modo funciona como un procesador de paso alto. El calculador del valor espectral combina, por medio del combinador del primer modo, los valores espectrales pertenecientes a la misma bandeja de frecuencia por cada subtrama de la representación del dominio espectral para obtener un primer grupo de valores espectrales combinados y el combinador del segundo modo 180 combina valores espectrales pertenecientes a la misma bandeja de frecuencia de cada subtrama de la representación del dominio espectral en un segundo modo para obtener un segundo grupo de valores espectrales combinados, donde el segundo modo es diferente del primer modo y donde el primer grupo de valores espectrales combinados y el segundo grupo de valores espectrales combinados representan unas representaciones del dominio espectral combinada que tiene un tamaño de bandeja de tiempo diferente y un tamaño de bandeja de frecuencia diferente. Una implementación preferida de este cálculo se describe e ilustra con respecto a la Fig. 5b donde, en una ilustración, la secuencia de A_2 , A_1 y B_2 , B_1 se convierte en una representación de alta resolución espectral pero que ahora tiene una baja resolución de tiempo como se ilustra por F_2 , E_2 por un lado y F_1 , E_1 por el otro.

Alternativamente, la Fig. 5b también ilustra la situación en la que las al menos dos subtramas se ilustran en el gráfico central de la Fig. 5b como dos subtramas de 10 ms subsiguientes en el tiempo y en la que la representación de alta resolución espectral y baja resolución de tiempo se ilustra a la derecha de la Fig. 5b. Preferiblemente, además se efectúa en el primer modo y una sustracción se efectúa en el segundo modo. Además, se prefiere que ambos procedimientos también comprendan una función promedio. Además, el calculador del valor espectral 160 en la Fig. 6 se configura aplicando cualquiera del primer modo o el segundo modo que comprende una ponderación usando un signo de ponderación, donde el calculador del valor espectral se configura para establecer el signo de ponderación de acuerdo con un número de bandeja de frecuencia de la misma bandeja de frecuencia. Además, el calculador de valor espectral está, como se ilustra en la Fig. 5b, configurado para transformar una bandeja de resolución baja en dos bandejas de resolución alta, donde la primera manera se usa para un número de bandeja par y la segunda para un número de bandeja impar.

La Fig. 7 ilustra una implementación adicional del aparato para convertir una resolución espectral. Además del combinador de resolución espectral 160, el aparato para convertir una resolución espectral puede comprender elementos adicionales. Los elementos adicionales son, por ejemplo, un procesador espectral 500 y/o un calculador de datos de procesamiento 190 y/o un procesador espectral adicional 220. En la implementación con el procesador

espectral 500, la representación de dominio espectral convertida que ha sido convertida sin ninguna operación de transformación inversa y hacia adelante y, por lo tanto, ha sido generada con bajos recursos computacionales y bajo retardo puede ser procesada posteriormente sola o, por ejemplo, junto con otra representación espectral que tenga la misma segunda resolución espectral. Esto puede realizarse, por ejemplo, para algún tipo de mezcla descendente. La
5 representación de alta resolución de frecuencia y baja resolución de tiempo ilustrada a la derecha de la Fig. 5b no sólo puede usarse para calcular los datos de procesamiento, sino que en realidad se procesa posteriormente para usos adicionales u otros usos alternativos como, por ejemplo, la mezcla descendente o cualquier tipo de representación de audio en una etapa de procesamiento posterior.

10 Por otra parte, el procedimiento discutido anteriormente con respecto a la Fig. 1 y la Fig. 5b, es que la representación del dominio espectral con la segunda resolución espectral, es decir, la "representación combinada del dominio espectral" sólo se usa para calcular algún tipo de datos de procesamiento, tales como valores de ponderación para un canal izquierdo y un canal derecho o, dicho de manera general, para un primer y un segundo canal de una señal multicanal. Los datos de procesamiento generados usando la representación del dominio espectral que se ha
15 convertido a una alta resolución espectral sólo se usan para calcular los datos de procesamiento, pero esta representación del dominio espectral no se procesa posteriormente por sí misma. En su lugar, usando los datos de procesamiento, como los valores de ponderación, se procesa espectralmente la representación espectral de dominio de entrada original con la primera resolución espectral, como se ilustra en el bloque 220. Para ello, se prefiere usar, por ejemplo, otra representación en el dominio espectral con la primera resolución, como para una operación de mezcla
20 descendente, por ejemplo, que se produzca en el dominio espectral.

La Fig. 8 ilustra una realización de la presente invención que funciona como mezclador descendente para mezclar de forma descendente una señal multicanal que tiene al menos dos canales. El mezclador descendente comprende un
25 estimador de valor de ponderación 100 para estimar los valores de ponderación por banda para los al menos dos canales, donde el estimador de valor de ponderación está configurado para calcular los valores de ponderación por banda basándose en el valor de energía objetivo por banda, de modo que una energía en la banda de una señal de mezcla descendente esté en una relación predeterminada con las energías en las mismas bandas de los al menos dos canales. Preferiblemente, el estimador de valor de ponderación 100 se implementa como se ilustra en la Fig. 3b y como se discute en el contexto de la Fig. 3b. El mezclador descendente comprende además un ponderador espectral
30 200 y un mezclador 400 conectado posteriormente para calcular la señal de mezcla descendente usando las representaciones ponderadas en el dominio espectral de los dos canales como mínimo.

La Fig. 9 ilustra una implementación adicional del mezclador descendente de la Fig. 8. El ponderador espectral 200
35 está configurado preferentemente para recibir datos de control para el primer y/o el segundo canal. Además, el ponderador espectral está configurado para aplicar los datos de control para uno de cuatro pares diferentes de datos de entrada. El primer par de datos de entrada puede ser la representación del dominio espectral del primer canal y la representación del dominio espectral del segundo canal, como se ilustra a la izquierda en la Fig. 9. La segunda alternativa puede ser la representación en el dominio espectral del primer canal y la representación en el dominio espectral combinada derivada como, por ejemplo, se discute con respecto a las Fig. 5b, 5c. Además, la otra alternativa
40 puede ser un par de datos que representen la representación del dominio espectral del segundo canal y una única representación combinada del dominio espectral, como también se ha comentado anteriormente con respecto a las Fig. 5b, 5c. Otra alternativa puede ser que el ponderador espectral 200 aplique las ponderaciones espectrales a una primera representación de dominio espectral combinada y a una segunda representación espectral combinada, como se ilustra, con respecto a la Fig. 5a o 5d. Los datos de control para el primer y/o segundo canal pueden ser, por
45 ejemplo, los valores de ponderación w_L por un lado y w_R por otro, pero también pueden ser cualquier otro dato de control usado para realizar cualquier tipo de ponderación espectral.

Otro elemento del mezclador descendente es, en una realización, un sumador 480 que calcula una representación de dominio espectral añadida, es decir, una representación de dominio espectral descendente en el dominio espectral.
50 Se puede usar un procesador de señal mono 490 que es, por ejemplo, controlado por cualquier dato o que es, por ejemplo, implementado como un convertidor de frecuencia-tiempo como se ha discutido anteriormente con respecto al bloque 310 de la Fig. 1 o Fig. 2.

Cabe destacar que los tres aspectos pueden utilizarse por separado, pero también pueden combinarse de manera
55 ventajosa entre sí. En particular, la implementación del estimador de valor de ponderación de acuerdo con la Fig. 8 puede aplicarse en el estimador de valor de ponderación 100 del primer aspecto ilustrado en la Fig. 1. Además, el convertidor de resolución espectral ilustrado en la Fig. 6 se implementa preferentemente mediante el estimador de valor de ponderación 100 de la Fig. 1 en la alternativa ilustrada en la Fig. 5b generando una representación de dominio espectral de alta resolución/baja resolución a partir de dos subtramas de alta resolución de tiempo y baja resolución espectral. Además, la funcionalidad del primer aspecto ilustrado en la Fig. 1, en particular con respecto al cálculo de los datos de procesamiento, puede implementarse mediante el calculador de datos de procesamiento 190 y el procesador espectral adicional 220 ilustrados en la Fig. 7, y el mezclador 400 puede implementarse, como alternativa
60 a la Fig. 9, de tal manera que el mezclador 400 para calcular la señal de mezcla descendente aplique la funcionalidad del convertidor 300 ilustrado en la Fig. 1 antes de realizar la suma real muestra a muestra en el dominio del tiempo.

Así, todas las realizaciones específicas definidas en una reivindicación dependiente para uno de los tres aspectos pueden aplicarse también a cualquier otro aspecto de los tres aspectos en la definición de la reivindicación dependiente correspondiente.

- 5 Así, queda claro que, dependiendo de la aplicación, los tres aspectos pueden aplicarse por separado o combinarse entre sí, ya sea combinando dos de los tres aspectos cualesquiera o combinando los tres aspectos.

Se mencionará aquí que todas las alternativas o aspectos como se analizó anteriormente y todos los aspectos como se definen por reivindicaciones independientes en las siguientes reivindicaciones se pueden usar individualmente, es decir, sin ninguna otra alternativa u objeto que la alternativa, objeto o reivindicación independiente contemplada. Sin embargo, en otras realizaciones, se pueden combinar dos o más de las alternativas o los aspectos o las reivindicaciones independientes entre sí y, en otras realizaciones, se pueden combinar todos los aspectos o alternativas y todas las reivindicaciones independientes entre sí.

- 10
15 Una señal de audio codificada de forma inventiva puede almacenarse en un medio de almacenamiento digital o en un medio de almacenamiento no transitorio, o puede transmitirse en un medio de transmisión, como un medio de transmisión inalámbrico o un medio de transmisión por cable, como Internet.

20 Si bien se han descrito algunos aspectos en el contexto de un aparato, es obvio que estos aspectos también representan una descripción del método correspondiente, en el cual un bloque o dispositivo corresponde a una etapa del método o a una característica de una etapa del método. De manera análoga, los aspectos descritos en el contexto de una etapa del método también representan una descripción de un bloque o elemento correspondiente o de una característica de un aparato correspondiente.

25 Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención se pueden implementar en hardware o en software. La implementación se puede realizar usando un medio de almacenamiento digital, por ejemplo, un disquete, un DVD, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tenga señales de control electrónicamente legibles almacenadas en el mismo, que cooperan (o sean capaces de cooperar) con un sistema de computadora programable de manera que se realiza el método respectivo.

30 Algunas realizaciones de acuerdo con la invención comprenden un soporte de datos que tiene señales de control electrónicamente legibles, que son capaces de cooperar con un sistema informático programable, de tal manera que se realiza uno de los métodos inventivos descritos en el presente documento.

35 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 inventivos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede ser almacenado, por ejemplo, en un portador legible por una máquina.

40 Otras realizaciones comprenden el programa de computadora para realizar uno de los métodos descritos en la presente, almacenado en un portador legible por máquina o un medio de almacenamiento no transitorio.

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

45 Una realización adicional de los métodos inventivos es, por lo tanto, un soporte de datos (o 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 de acuerdo con la invención descrita en el presente documento.

50 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 de acuerdo con la invención aquí descrita. El flujo de datos o la secuencia de señales pueden estar configurados, por ejemplo, para ser transferida a través de una conexión de comunicación de datos, por ejemplo a través de Internet.

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

60 Una realización adicional comprende un ordenador que tiene instalado en él el programa informático para llevar a cabo uno de los métodos descritos en la presente.

En algunas realizaciones, se puede usar un dispositivo lógico programable (por ejemplo, una matriz de puertas programable en campo) para llevar a cabo algunas o todas las funcionalidades de los métodos descritos en la presente. En algunas realizaciones, una matriz de puertas programables en campo puede cooperar con un microprocesador

para llevar a cabo uno de los métodos descritos en la presente. Por lo general, los métodos son ejecutados preferentemente por cualquier aparato de hardware.

5 Las realizaciones anteriormente descritas son meramente ilustrativas de los principios de la presente invención. Se entiende que modificaciones y variaciones de las disposiciones y los detalles descritos en el presente documento resultarán evidentes para los expertos en la técnica. Es la intención, por consiguiente, limitarse sólo por el alcance de las reivindicaciones de patente que siguen y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones en el presente documento.

10 **Referencias**

[1] UIT-R BS.775-2, *Multichannel Stereophonic Sound System With And Without Accompanying Picture*, 07/2006.

15 [2] F. Baumgarte, C. Faller y P. Kroon, "Audio Coder Enhancement using Scalable Binaural Cue Coding with Equalized Mixing", en la 116ª Convención de la AES, Berlín, 2004.

[3] G. Stoll, J. Groh, M. Link, J. Deigmöller, B. Runow, M. Keil, R. Stoll, M. Stoll y C. Stoll, "Method for Generating a Downward-Compatible Sound Format". Patente estadounidense US 2012/0 014 526, 2012.

20 [4] M. Kim, E. Oh y H. Shim, "Stereo audio coding improved by phase parameters", en la 129ª Convención de la AES, San Francisco, 2010.

[5] A. Adami, E. Habets y J. Herre, "Down-mixing using coherence suppression", en IEEE International Conference on Acoustics, Speech and Signal Processing, Florencia, 2014.

25 [6] ISO/IEC 23008-3:, Information technology - High efficiency coding and media delivery in heterogeneous environments - Parte 3: 3D audio, 2019.

30 [7] S. Bayer, C. Borß, J. Bütthe, S. Disch, B. Edler, G. Fuchs, F. Ghido y M. Multrus, "DOWNMIXER AND METHOD FOR DOWNMIXING AT LEAST TWO CHANNELS AND MULTICHANNEL ENCODER AND MULTICHANNEL DECODER". WO2018086946.

[8] 3GPP TS 26.445, Codecfor Enhanced Voice Services (EVS); Detailed algorithmic description.

35 [9] S. Chen, H. Ruimin y S. Zhang, "Estimating spatial cues for audio coding in MDCT domain", en IEEE International Conference on Multimedia and Expo, Nueva York, 2009.

REIVINDICACIONES

1. Mezclador descendente para la mezcla descendente de una señal de audio multicanal que tiene al menos dos canales, que comprende:
 - 5 un estimador de valor de ponderación (100) para estimar valores de ponderación por banda para los al menos dos canales, en el que el estimador de valor de ponderación (100) está configurado para calcular los valores de ponderación por banda basándose en un valor de energía objetivo por banda, de modo que una energía en la banda de una señal de audio de mezcla descendente esté en una relación predeterminada con las energías en las mismas bandas de los al menos dos canales;
 - 10 un ponderador espectral (200) para ponderar las representaciones espectrales de al menos dos canales usando los valores de ponderación por banda;
 - 15 un convertidor (300) para convertir las representaciones ponderadas del dominio espectral de los al menos dos canales en representaciones de tiempo de los al menos dos canales; y
 - 20 un mezclador (400) para mezclar las representaciones de tiempo de los al menos dos canales para obtener la señal de audio de mezcla descendente.
2. Mezclador descendente de la reivindicación 1, en el que el estimador del valor de ponderación (100) está configurado para calcular una pluralidad de primeros valores de ponderación en función de la banda para una pluralidad de bandas de un primer canal de los al menos dos canales y para calcular una pluralidad de segundos valores de ponderación en función de la banda para la pluralidad de bandas de un segundo canal de los al menos dos canales, o
 - 25 en el que la señal de audio multicanal tiene más de dos canales y en el que el estimador del valor de ponderación (100) está configurado para calcular una pluralidad de primeros valores de ponderación en función de la banda para una pluralidad de bandas de un primer canal de los más de dos canales, para calcular una pluralidad de segundos valores de ponderación en función de la banda para la pluralidad de bandas de un segundo canal de los más de dos canales y para calcular otra pluralidad de valores de ponderación en función de la banda para la pluralidad de bandas de otro canal de los más de dos canales.
3. Mezclador descendente de la reivindicación 1 o 2,
 - 35 en el que las representaciones en el dominio espectral de los al menos dos canales comprenden cada uno un conjunto de bandejas de frecuencias, en el que los valores espectrales están asociados a las bandejas de frecuencias,
 - 40 en el que el estimador del valor de ponderación (100) está configurado para calcular los valores de ponderación por bandas, en el que cada banda comprende uno, dos o más bandejas de frecuencias, o
 - 45 en el que el número de bandejas de frecuencia por banda aumenta con las bandas que tienen una frecuencia central más alta.
4. Mezclador descendente de una de las reivindicaciones anteriores, que comprende además:
 - 50 un decodificador de núcleo (500) para decodificar una señal codificada, teniendo la señal codificada representaciones del dominio espectral codificadas de al menos dos canales originales, en el que el decodificador central está configurado para generar las representaciones del dominio espectral a partir de las representaciones del dominio espectral codificadas.
5. Mezclador descendente de acuerdo con una de las reivindicaciones anteriores,
 - 55 en el que las representaciones del dominio espectral son puramente reales,
 - 60 en el que el estimador del valor de ponderación (100) está configurado para estimar (120, 122) una representación del dominio espectral imaginaria cuando la representación del dominio espectral es puramente real, y
 - en el que el estimador del valor de ponderación (100) está configurado para estimar los valores de ponderación por banda usando el dominio espectral imaginario estimado.
6. Mezclador descendente de una de las reivindicaciones 1 a 4,

- en el que las representaciones del dominio espectral son puramente imaginarias,
- 5 en el que el estimador del valor de ponderación (100) está configurado para estimar una representación del dominio espectral real cuando la representación del dominio espectral es puramente imaginaria, y
- en el que el estimador del valor de ponderación (100) está configurado para estimar los valores de ponderación por bandas usando la representación espectral real estimada.
- 10 7. Mezclador descendente de la reivindicación 1, en el que el estimador del valor de ponderación (100) está configurado para calcular un primer valor de ponderación de los valores de ponderación por bandas para una banda de un primer canal de los al menos dos canales,
- 15 en el que el estimador del valor de ponderación (100) está configurado para calcular un segundo valor de ponderación de los valores de ponderación por banda para la banda de un segundo canal de los al menos dos canales, y
- 20 en el que el estimador del valor de ponderación (100) está configurado para calcular el primer valor de ponderación y el segundo valor de ponderación usando una energía del primer canal en la banda, una energía del segundo canal en la banda, y un término mixto que depende de un producto o una combinación lineal de valores espectrales de los al menos dos canales en la banda.
8. Mezclador descendente de la reivindicación 7,
- 25 en el que el estimador del valor de ponderación (100) está configurado para calcular, como término mixto dependiente de la combinación lineal, una raíz cuadrada de una energía de valores espectrales sumados entre sí en la banda a partir de las representaciones del dominio espectral de los al menos dos canales, donde la banda comprende una pluralidad de valores espectrales, o para calcular, como término mixto dependiente del producto, un valor absoluto de un producto punto complejo entre los valores espectrales en la banda del primer canal y los valores espectrales en la banda del segundo canal de los al menos dos canales.
- 30 9. Mezclador descendente de la reivindicación 1,
- 35 en el que cada banda de un primer y un segundo canal de la pluralidad de al menos dos canales tiene una pluralidad de valores espectrales, en el que el ponderador espectral (200) está configurado para aplicar la misma ponderación a cada valor espectral en la banda de uno de los al menos dos canales, y para aplicar otra ponderación a cada valor espectral en la banda de otro canal de los al menos dos canales.
- 40 10. Mezclador descendente de una de las reivindicaciones anteriores,
- en el que las representaciones del dominio espectral ponderado son espectros MDCT (transformada discreta del coseno modificada), y
- 45 en el que el convertidor (300) está configurado para realizar, para cada canal de la pluralidad de canales, una transformada MDCT inversa usando una operación de ventana de síntesis y una operación de superposición-añadido.
- 50 11. Mezclador descendente de una de las reivindicaciones anteriores,
- en el que el mezclador (400) está configurado para aplicar una suma muestra a muestra de las representaciones de tiempo de los al menos dos canales, o
- 55 en el que el mezclador (400) está configurado para aplicar una suma muestra a muestra de las representaciones de tiempo de los al menos dos canales y una operación de escalado aplicada a un resultado de la suma muestra a muestra o aplicada a las entradas en la suma muestra a muestra.
12. Mezclador descendente de una de las reivindicaciones anteriores,
- 60 en el que el convertidor (300) está configurado para generar (310) representaciones de tiempo en bruto usando un algoritmo espectro-tiempo, y
- procesar posteriormente (320) las representaciones de tiempo en bruto individualmente, en dirección de procesamiento de señal antes de la mezcla por el mezclador (400), usando información de control separada

para los al menos dos canales para obtener las representaciones de tiempo.

13. Mezclador descendente de la reivindicación 12,
5 en el que el convertidor (300) está configurado para realizar, como procesado posterior (320), un filtrado posterior de graves, un procesado TCX-LTP (transformación de predicción a largo plazo de excitación codificada) o una síntesis LPC (codificación de predicción lineal) individualmente para cada representación de tiempo.
- 10 14. Mezclador descendente de la reivindicación 1,
15 en la que una primera representación en el dominio espectral de un primer canal de los al menos dos canales tiene una primera resolución de tiempo o frecuencia,
20 en la que una segunda representación en el dominio espectral de un segundo canal de los al menos dos canales tiene una segunda resolución de tiempo o frecuencia, en la que la segunda resolución de tiempo o frecuencia es diferente de la primera resolución de tiempo o frecuencia, y
25 en el que el estimador del valor de ponderación (100) está configurado para calcular los valores de ponderación por banda de modo que una resolución de frecuencia de las bandas asociadas con los valores de ponderación por banda sea inferior a la primera resolución de frecuencia y a la segunda resolución de frecuencia o sea igual a la menor de la primera y la segunda resolución de frecuencia.
15. Mezclador descendente de la reivindicación 14,
30 en la que la primera representación del dominio espectral tiene una primera pluralidad de valores espectrales en una banda,
35 en la que la segunda representación del dominio espectral tiene una segunda pluralidad de valores espectrales en la banda, siendo la segunda pluralidad de valores espectrales superior a la primera pluralidad, y
40 en el que el estimador del valor de ponderación (100) está configurado para combinar dos o más valores espectrales de la segunda pluralidad de valores espectrales o seleccionar, a partir de la segunda pluralidad de valores espectrales, un subconjunto de valores espectrales,
45 para calcular un término mixto en función de productos o combinaciones lineales de valores espectrales de los al menos dos canales de la banda usando un resultado de la combinación de los dos o más valores espectrales o usando el subconjunto de valores espectrales, y
50 para calcular los valores de ponderación por banda usando el término mixto.
16. Mezclador descendente de acuerdo con una de las reivindicaciones 1 a 13,
55 en la que una primera representación en el dominio espectral comprende una pluralidad de primeros valores espectrales que representan un primer tamaño de bandeja de tiempo y un primer tamaño de bandeja de frecuencia,
60 en la que una segunda representación del dominio espectral comprende una pluralidad de valores espectrales que representan un segundo tamaño de bandeja de tiempo y un segundo tamaño de bandeja de frecuencia,
65 en el que el tamaño de la primera bandeja de tiempo es mayor que el de la segunda, o en el que el tamaño de la primera bandeja de frecuencia es menor que el de la segunda,
70 en la que el estimador del valor de ponderación (100) está configurado para combinar una pluralidad de valores espectrales de la primera representación del dominio espectral para obtener una primera representación combinada del dominio espectral en la que un tamaño combinado de la bandeja de frecuencia es igual al segundo tamaño de la bandeja de frecuencia, o para combinar una pluralidad de valores espectrales de la segunda representación del dominio espectral para obtener una primera representación combinada del dominio espectral en la que un tamaño combinado de la bandeja de tiempo es igual al primer tamaño de la bandeja de tiempo.
17. Mezclador descendente de la reivindicación 16,

- 5 en el que el estimador del valor de ponderación (100) está configurado para usar la primera representación combinada del dominio espectral o la segunda representación combinada del dominio espectral para estimar los valores de ponderación por bandas para un primer canal de los al menos dos canales y un segundo canal de los al menos dos canales, comprendiendo la estimación un cálculo de un término mixto en bandas y un cálculo de energías en las bandas, y
- 10 en el que el ponderador espectral (200) está configurado para aplicar los valores de ponderación por bandas para el primer canal de los al menos dos canales a los valores espectrales de la primera representación del dominio espectral en las bandas correspondientes y para aplicar los valores de ponderación por bandas para el segundo canal de los al menos dos canales a los valores espectrales de la segunda representación del dominio espectral en las bandas correspondientes.
- 15 18. Mezclador descendente de acuerdo con una de las reivindicaciones 1 a 15,
- en la que una primera representación en el dominio espectral de un primer canal comprende una pluralidad de primeros valores espectrales que representan un primer tamaño de bandeja de tiempo y un primer tamaño de bandeja de frecuencia,
- 20 en la que una segunda representación en el dominio espectral de un segundo canal comprende al menos dos subtramas, en la que cada subtrama comprende una pluralidad de valores espectrales que representan un segundo tamaño de bandeja de tiempo y un segundo tamaño de bandeja de frecuencia,
- 25 en el que el tamaño de la primera bandeja de tiempo es mayor que el de la segunda, o en el que el tamaño de la primera bandeja de frecuencia es menor que el de la segunda,
- en el que el estimador del valor de ponderación (100) está configurado
- 30 para combinar valores espectrales pertenecientes a la misma bandeja de frecuencias de cada subtrama de la segunda representación en el dominio espectral de una primera manera para obtener un primer grupo de valores espectrales combinados, y
- 35 para combinar valores espectrales pertenecientes a la misma bandeja de frecuencias de cada subtrama de la segunda representación en el dominio espectral de una segunda manera para obtener un segundo grupo de valores espectrales combinados, siendo la segunda manera diferente de la primera,
- 40 en la que el primer grupo de valores espectrales combinados y el segundo grupo de valores espectrales combinados representan una representación combinada en el dominio espectral que tiene el primer tamaño de bandeja de tiempo y el primer tamaño de bandeja de frecuencia, y
- usar los valores espectrales de la representación combinada del dominio espectral y de la primera representación del dominio espectral para estimar los valores de ponderación por bandas.
- 45 19. Mezclador descendente de acuerdo con la reivindicación 18,
- en el que el estimador del valor de ponderación (100) está configurado para realizar una suma y sustracción de la primera manera y la otra de las sumas y sustracciones de la segunda manera.
- 50 20. Mezclador descendente de acuerdo con la reivindicación 18 o 19, en el que el estimador del valor de ponderación (100) está configurado para realizar una función de promedio de la primera manera y de la segunda manera.
- 55 21. Mezclador descendente de acuerdo con una de las reivindicaciones 18 a 20, en el que el estimador de valor de ponderación (100) está configurado para aplicar la primera manera o la segunda manera que comprende una ponderación usando un signo de ponderación, en donde el estimador de valor de ponderación (100) está configurado para establecer el signo de ponderación de acuerdo con un número de bandeja de frecuencia de la misma bandeja de frecuencia.
- 60 22. Mezclador descendente de acuerdo con una de las reivindicaciones 18 a 21, en el que el estimador de valor de ponderación (100) está configurado para aplicar, como primera manera, uno de un filtrado de paso alto y un filtrado de paso bajo, y como segunda manera, el otro de un filtrado de paso alto y un filtrado de paso bajo.
23. Mezclador descendente de acuerdo con una de las reivindicaciones 18 a 22, en el que el estimador de valor de ponderación (100) está configurado para transformar una bandeja de menor resolución en dos bandejas

de mayor resolución, en donde la primera manera se usa para un número de bandejas par de una primera bandeja de mayor resolución de las dos bandejas de mayor resolución y la segunda manera se usa para un número de bandeja impar de una segunda bandeja de mayor resolución de las dos bandejas de mayor resolución.

5

24. Mezclador descendente de acuerdo con una de las reivindicaciones 18 a 23,

10

en la que la primera representación en el dominio espectral del primer canal comprende una trama TCX20, en la que la segunda representación en el dominio espectral del segundo canal comprende dos subtramas TCX10, en la que el estimador del valor de ponderación (100) está configurado para calcular una representación combinada en el dominio espectral TCX20 a partir de las dos subtramas TCX10, o bien

15

en la que la primera representación del dominio espectral del primer canal comprende una trama TCX20, en la que la segunda representación del dominio espectral del segundo canal comprende una subtrama TCX10 y dos subtramas TCX5, en la que el estimador del valor de ponderación (100) está configurado para calcular una primera representación combinada del dominio espectral TCX10 a partir de las dos subtramas TCX5 y para calcular una segunda subtrama TCX20 combinada a partir de la primera representación combinada del dominio espectral TCX10 y de la subtrama TCX10, o bien

20

en la que la primera representación en el dominio espectral del primer canal comprende una subtrama TCX10, en la que la segunda representación en el dominio espectral del segundo canal comprende dos subtramas TCX5, y en la que el estimador del valor de ponderación (100) está configurado para calcular una representación combinada en el dominio espectral TCX10 a partir de las dos subtramas TCX5,

25

en la que la expresión TCX20 indica una primera porción con una primera longitud de tiempo, en la que la expresión TCX10 indica una segunda porción con una segunda longitud de tiempo, y en la que la expresión TCX5 indica una tercera porción con una tercera longitud de tiempo, en la que la primera longitud de tiempo es mayor que la segunda longitud de tiempo o que la tercera longitud de tiempo, o en la que la segunda longitud de tiempo es menor que la primera longitud de tiempo o mayor que la segunda longitud de tiempo, y en la que la tercera longitud de tiempo es menor que la primera longitud de tiempo o menor que la segunda longitud de tiempo.

30

25. Mezclador descendente de acuerdo con una de las reivindicaciones 18 a 24, en el que el estimador del valor de ponderación (100) está configurado para aplicar la primera manera basándose en la siguiente ecuación:

35

$$MDCT_{2i,k0} = \frac{1}{2} (MDCT_{i,k1} + (-1)^i MDCT_{i,k0})$$

, o

40

en el que el estimador del valor de ponderación (100) está configurado para aplicar la segunda manera basándose en la siguiente ecuación:

$$MDCT_{2i+1,k0} = \frac{1}{2} (MDCT_{i,k1} - (-1)^i MDCT_{i,k0}),$$

45

en el que i especifica un número de bandeja espectral y k_0 y k_1 especifican subtramas de la segunda representación en el dominio espectral del segundo canal, y

50

en el que $MDCT_{2i,k0}$ y $MDCT_{2i+1,k0}$ indican valores espectrales de la representación combinada en el dominio espectral y $MDCT_{i,k1}$ y $MDCT_{i,k0}$ indican valores espectrales de una segunda subtrama k_1 y una primera subtrama k_0 , respectivamente.

26. Mezclador descendente de acuerdo con la reivindicación 1,

55

en la que una primera representación en el dominio espectral de un primer canal de los al menos dos canales tiene una primera resolución de tiempo o una primera resolución de frecuencia, en la que una segunda representación en el dominio espectral de un segundo canal de los al menos dos canales tiene una segunda resolución de tiempo o una segunda resolución de frecuencia, en la que la segunda resolución de tiempo es diferente de la primera resolución de tiempo o en la que la segunda resolución de frecuencia es diferente de la primera resolución de frecuencia, y

60

en la que el estimador del valor de ponderación (100) está configurado para convertir (132) la primera representación del dominio espectral en una representación combinada del dominio espectral con la segunda

5 resolución de tiempo o la segunda resolución de frecuencia, y para calcular los valores de ponderación por bandas usando la representación combinada del dominio espectral y la segunda representación del dominio espectral, o para convertir la segunda representación del dominio espectral en una representación combinada del dominio espectral con la primera resolución de tiempo o la primera resolución de frecuencia, y para calcular los valores de ponderación por bandas usando la representación combinada del dominio espectral y la primera representación del dominio espectral, o

10 en la que una primera representación en el dominio espectral de un primer canal de los al menos dos canales tiene una primera resolución de tiempo o una primera resolución de frecuencia, en la que una segunda representación en el dominio espectral de un segundo canal de los al menos dos canales tiene una segunda resolución de tiempo o una segunda resolución de frecuencia, en la que la segunda resolución de tiempo es diferente de la primera resolución de tiempo o en la que la segunda resolución de frecuencia es diferente de la primera resolución de frecuencia, y

15 en el que el estimador del valor de ponderación (100) está configurado

20 para convertir (132) la primera representación en el dominio espectral en una primera representación combinada en el dominio espectral que tenga una tercera resolución de tiempo o una tercera resolución de frecuencia,

en el que la tercera resolución de tiempo es diferente de la primera resolución de tiempo o la segunda resolución de tiempo y en donde la tercera resolución de frecuencia es diferente de la primera resolución de frecuencia o la segunda resolución de frecuencia,

25 para convertir (132) la segunda representación en el dominio espectral en una segunda representación combinada en el dominio espectral que tenga la tercera resolución de tiempo o la tercera resolución de frecuencia, y

30 para calcular (134) los valores de ponderación por banda usando la primera representación combinada del dominio espectral y la segunda representación combinada del dominio espectral.

27. Mezclador descendente de acuerdo con la reivindicación 26,

35 en el que el segundo canal comprende, para una porción de tiempo específica (TCX20), la segunda representación del dominio espectral,

en el que el primer canal comprende, para la porción de tiempo específica (2xTCX10), dos o más de las primeras representaciones del dominio espectral,

40 en la que el estimador del valor de ponderación (100) está configurado para convertir las dos o más primeras representaciones del dominio espectral en la representación combinada del dominio espectral que tiene la misma resolución de tiempo y de frecuencia que la segunda representación del dominio espectral, y para calcular los valores de ponderación por bandas usando la representación combinada del dominio espectral y la segunda representación del dominio espectral, y

45 en el que el ponderador espectral (200) está configurado para ponderar la segunda representación de dominio espectral usando los valores de ponderación por bandas, y para ponderar cada primera representación de dominio espectral de las dos o más primeras representaciones de dominio espectral usando los mismos valores de ponderación por bandas.

50 28. Mezclador descendente de acuerdo con la reivindicación 27,

55 en el que el estimador del valor de ponderación (100) está configurado para sumar valores espectrales de una misma frecuencia de las dos o más primeras representaciones del dominio espectral a fin de obtener un primer valor espectral de la representación combinada del dominio espectral, y para restar valores espectrales de una misma frecuencia de las dos o más primeras representaciones del dominio espectral a fin de obtener un segundo valor espectral de la representación combinada del dominio espectral que sea superior y adyacente en frecuencia al primer valor espectral de la representación combinada del dominio espectral, y

60 en el que el ponderador espectral (200) está configurado para ponderar una banda que tiene las mismas frecuencias en cada primera representación de dominio espectral de las dos o más primeras representaciones de dominio espectral usando el mismo valor de ponderación por banda.

29. Mezclador descendente de acuerdo con la reivindicación 26,

en el que el segundo canal comprende, para una porción de tiempo específica (TCX20), la segunda representación del dominio espectral,

5 en el que el primer canal comprende, para la porción de tiempo específica (2xTCX10), dos o más primeras representaciones del dominio espectral,

en el que el estimador del valor de ponderación (100) está configurado

10 para convertir la segunda representación del dominio espectral en dos o más representaciones combinadas del dominio espectral que tengan la misma resolución de tiempo y frecuencia que las dos o más primeras representaciones del dominio espectral,

15 para calcular los primeros valores de ponderación por banda usando una primera representación de dominio espectral combinada de las dos o más representaciones de dominio espectral combinadas y una primera representación de dominio espectral de las dos o más primeras representaciones de dominio espectral,

20 para calcular los segundos valores de ponderación por banda usando una segunda representación combinada en el dominio espectral de las dos o más representaciones combinadas en el dominio espectral y una segunda primera representación en el dominio espectral de las dos o más primeras representaciones en el dominio espectral, y

en el que el ponderador espectral (200) está configurado

25 para ponderar la segunda representación espectral usando valores de ponderación derivados (136) de los valores de ponderación primero y segundo,

para ponderar la primera representación del dominio espectral de las dos o más primeras representaciones del dominio espectral usando los primeros valores de ponderación por bandas y

30 para ponderar la segunda representación del primer dominio espectral de las dos o más representaciones del primer dominio espectral usando los segundos valores de ponderación por bandas.

30. Mezclador descendente de acuerdo con la reivindicación 29,

35 en el que el estimador del valor de ponderación (100) está configurado para sumar valores espectrales de pares de frecuencias de la segunda representación del dominio espectral con el fin de obtener un valor espectral sumado y para copiar el valor espectral sumado con el fin de obtener un valor espectral combinado para cada una de las dos o más representaciones combinadas del dominio espectral, y

40 en el que el ponderador espectral (200) está configurado para combinar (136) un valor de ponderación para una banda específica de los primeros valores de ponderación por banda con un valor de ponderación para la banda específica de los segundos valores de ponderación por banda para obtener un valor de ponderación derivado para la banda específica de los valores de ponderación por banda derivados.

45 31. Mezclador descendente de acuerdo con la reivindicación 26,

en el que el segundo canal comprende, para una porción de tiempo específica (TCX20), la segunda representación del dominio espectral,

50 en el que el primer canal comprende, para la porción de tiempo específica (2xTCX10), dos o más primeras representaciones del dominio espectral,

en el que el estimador del valor de ponderación (100) está configurado

55 para convertir la segunda representación del dominio espectral en dos o más representaciones combinadas del dominio espectral que tengan la misma resolución de tiempo que las dos o más primeras representaciones del dominio espectral, y que tengan la misma resolución de frecuencia que la segunda representación del dominio espectral,

60 para calcular los primeros valores de ponderación por banda usando una primera representación de dominio espectral combinada de las dos o más representaciones de dominio espectral combinadas y una primera representación de dominio espectral de las dos o más primeras representaciones de dominio espectral,

para calcular los segundos valores de ponderación por banda usando una segunda representación combinada en el dominio espectral de las dos o más representaciones combinadas en el dominio espectral y una segunda primera representación en el dominio espectral de las dos o más primeras representaciones en el dominio espectral, y

5 en el que el ponderador espectral (200) está configurado

para ponderar la segunda representación del dominio espectral usando valores de ponderación por bandas derivados (136) de los valores de ponderación por bandas primero y segundo,

10 para ponderar la primera representación de dominio espectral de las dos o más primeras representaciones de dominio espectral usando los primeros valores de ponderación por bandas, y

15 para ponderar la segunda representación del primer dominio espectral de las dos o más representaciones del primer dominio espectral usando los segundos valores de ponderación por bandas.

32. Mezclador descendente de acuerdo con la reivindicación 31,

20 en el que el estimador del valor de ponderación (100) está configurado para realizar un sobremuestreo de uno o más valores espectrales a fin de obtener valores espectrales sobremuestreados para frecuencias adyacentes de la segunda representación del dominio espectral y para copiar a los valores espectrales sobremuestreados a fin de obtener valores espectrales combinados para cada una de las dos o más representaciones combinadas del dominio espectral, y

25 en el que el ponderador espectral (200) está configurado para combinar (136) un valor de ponderación para una banda específica de los primeros valores de ponderación por banda con un valor de ponderación para la banda específica de los segundos valores de ponderación por banda para obtener un valor de ponderación derivado para la banda específica de los valores de ponderación por banda derivados.

30 33. Mezclador descendente de acuerdo con la reivindicación 26,

en el que el segundo canal comprende, para una porción de tiempo específica (TCX20), la segunda representación del dominio espectral,

35 en el que el primer canal comprende, para la porción de tiempo específica (2xTCX10), dos o más de las primeras representaciones del dominio espectral,

en el que el estimador del valor de ponderación (100) está configurado

40 para convertir las dos o más primeras representaciones del dominio espectral en la primera representación combinada del dominio espectral que tenga la misma resolución de tiempo que la segunda representación del dominio espectral,

45 para convertir las segundas representaciones del dominio espectral en la segunda representación combinada del dominio espectral que tenga la misma resolución de frecuencia que las dos o más primeras representaciones del dominio espectral, y

50 para calcular los valores de ponderación por banda usando la primera representación combinada del dominio espectral y la segunda representación combinada del dominio espectral, y

55 en el que el ponderador espectral (200) está configurado para ponderar la segunda representación de dominio espectral usando los valores de ponderación por bandas, y para ponderar cada primera representación de dominio espectral de las dos o más primeras representaciones de dominio espectral usando los mismos valores de ponderación por bandas.

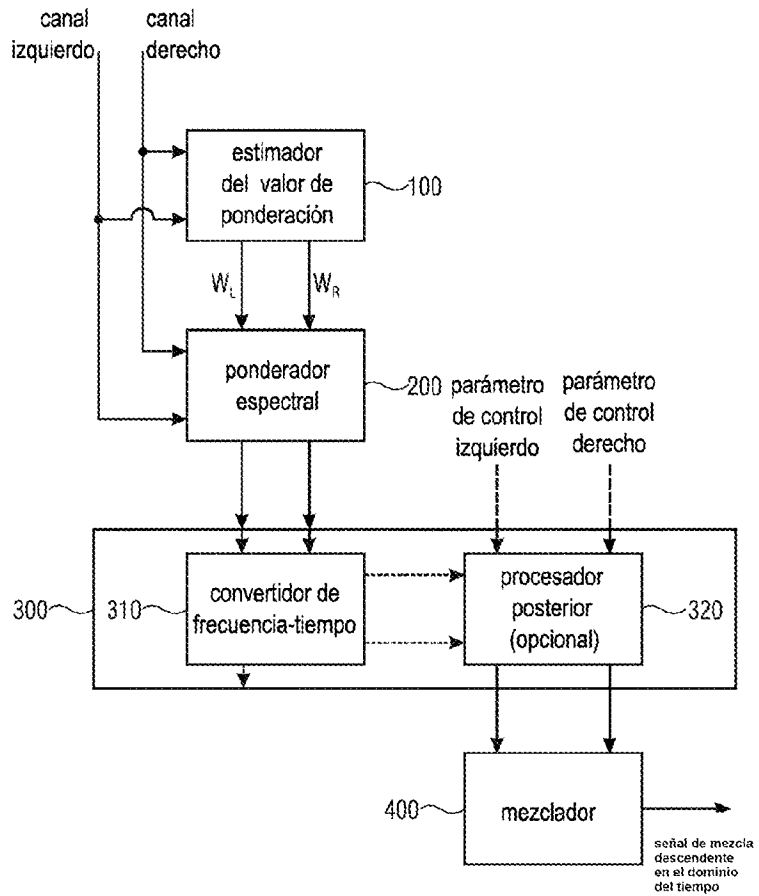
34. Mezclador descendente de acuerdo con la reivindicación 33,

60 en la que el estimador del valor de ponderación (100) está configurado para sumar valores espectrales de pares de frecuencias de la segunda representación del dominio espectral para obtener la segunda representación combinada del dominio espectral, y para sumar valores espectrales de la misma frecuencia de las dos o más de las primeras representaciones del dominio espectral para obtener la primera representación combinada del dominio espectral, y

en el que el ponderador espectral (200) está configurado para ponderar una banda que tiene las mismas

frecuencias en cada primera representación de dominio espectral de las dos o más primeras representaciones de dominio espectral usando el mismo valor de ponderación por banda.

- 5 35. Mezclador descendente de acuerdo con una de las reivindicaciones anteriores,
- 10 en el que el estimador del valor de ponderación (100) está configurado para calcular una pluralidad de primeros valores de ponderación en función de la banda para una pluralidad de bandas de un primer canal de los al menos dos canales usando una primera regla de cálculo que depende de al menos dos de los valores espectrales de una primera representación del dominio espectral del primer canal, los valores espectrales de una segunda representación del dominio espectral del segundo canal, valores espectrales de una única representación de dominio espectral combinado derivado de los valores espectrales de la primera representación de dominio espectral o de la segunda representación de dominio espectral, valores espectrales de una primera representación de dominio espectral combinado derivado de los valores espectrales de la primera representación de dominio espectral, y valores espectrales de una segunda representación de dominio espectral combinado derivado de los valores espectrales de la segunda representación de dominio espectral, y
- 15
- 20 en el que el estimador de valores de ponderación (100) está configurado para calcular una pluralidad de segundos valores de ponderación por bandas para una pluralidad de bandas del primer canal de los al menos dos canales usando una segunda regla de cálculo que depende de al menos dos de la pluralidad de primeros valores de ponderación por bandas, los valores espectrales de la primera representación de dominio espectral del primer canal, los valores espectrales de la segunda representación de dominio espectral del segundo canal, los valores espectrales de la única representación combinada del dominio espectral derivados de los valores espectrales de la primera representación del dominio espectral o de la segunda representación del dominio espectral, los valores espectrales de una primera representación combinada del dominio espectral derivados de los valores espectrales de la primera representación del dominio espectral, y los valores espectrales de una segunda representación combinada del dominio espectral derivados de los valores espectrales de la segunda representación del dominio espectral, en la que la segunda regla de cálculo es diferente de la primera regla de cálculo.
- 25
- 30 36. Método de mezcla descendente de una señal de audio multicanal que tiene al menos dos canales, que comprende:
- 35 estimación de los valores de ponderación por banda para los dos canales como mínimo, en la que la estimación comprende el cálculo de los valores de ponderación por banda basándose en un valor de energía objetivo por banda, de modo que una energía en la banda de una señal de audio de mezcla descendente esté en una relación predeterminada con las energías en las mismas bandas de los dos canales como mínimo;
- 40 ponderación de las representaciones espectrales de al menos dos canales usando los valores de ponderación por banda;
- 45 conversión de las representaciones espectrales ponderadas de al menos dos canales en representaciones de tiempo de al menos dos canales; y
- mezclar las representaciones de tiempo de los al menos dos canales para obtener la señal de audio de mezcla descendente.
- 50 37. Programa informático para realizar, cuando se ejecuta en un ordenador o un procesador, el método de la reivindicación 36.



(MEZCLADOR DESCENDENTE)

Figura 1

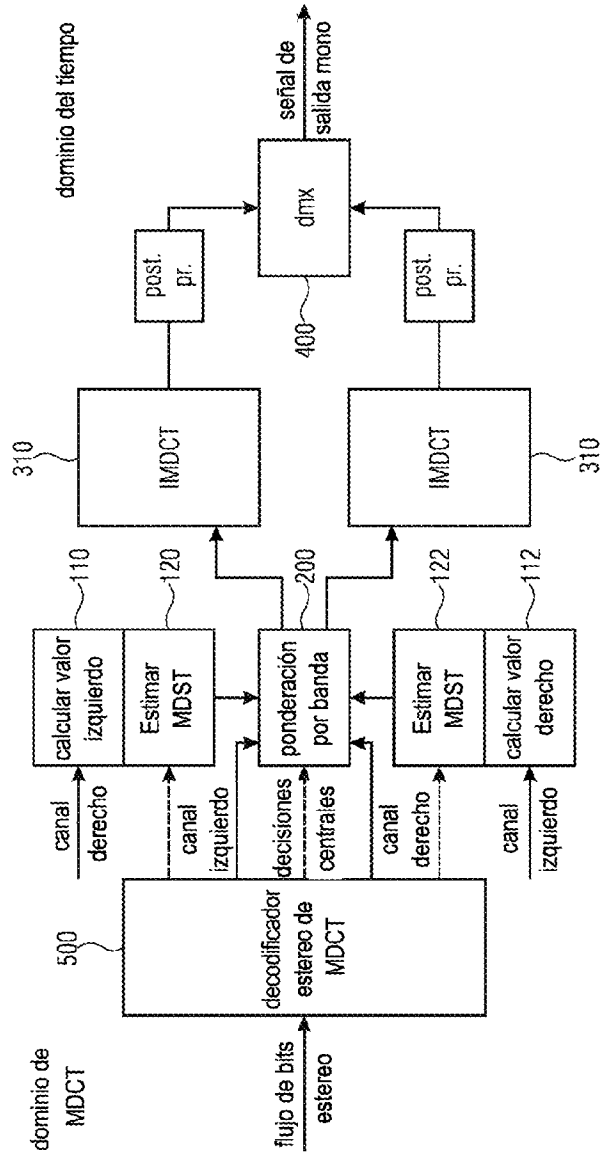


Figura 2

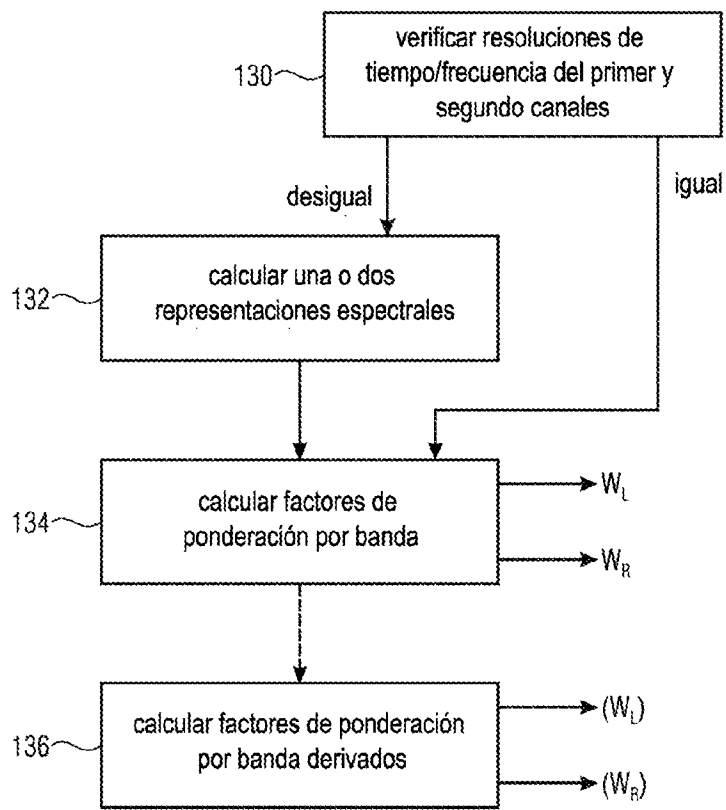


Figura 3a

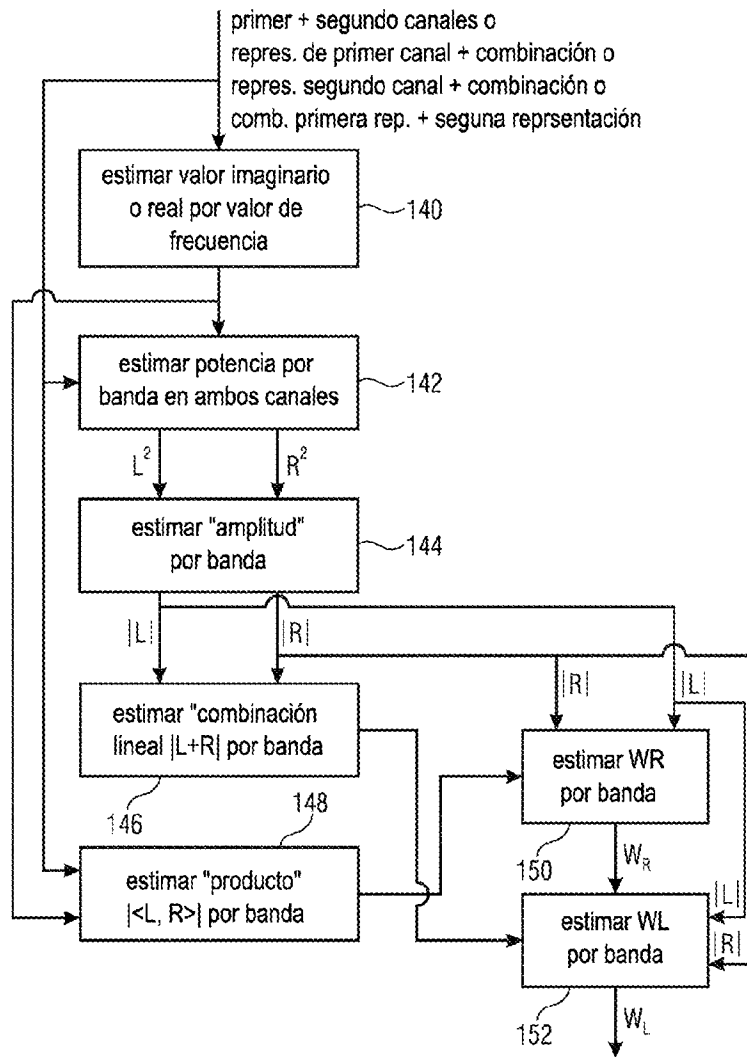
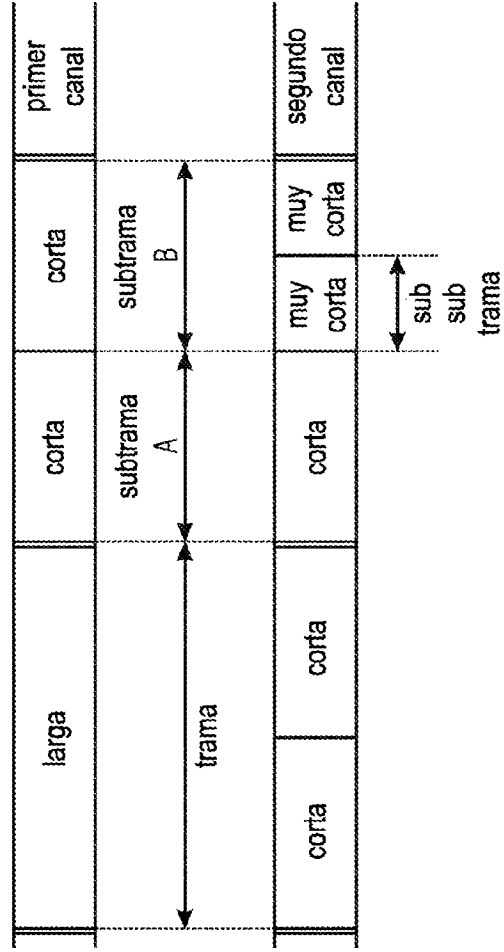


Figura 3b



larga p. ej. TCX20
 corta p. ej. TCX10
 muy corta p. ej. TCX5

Figura 4a

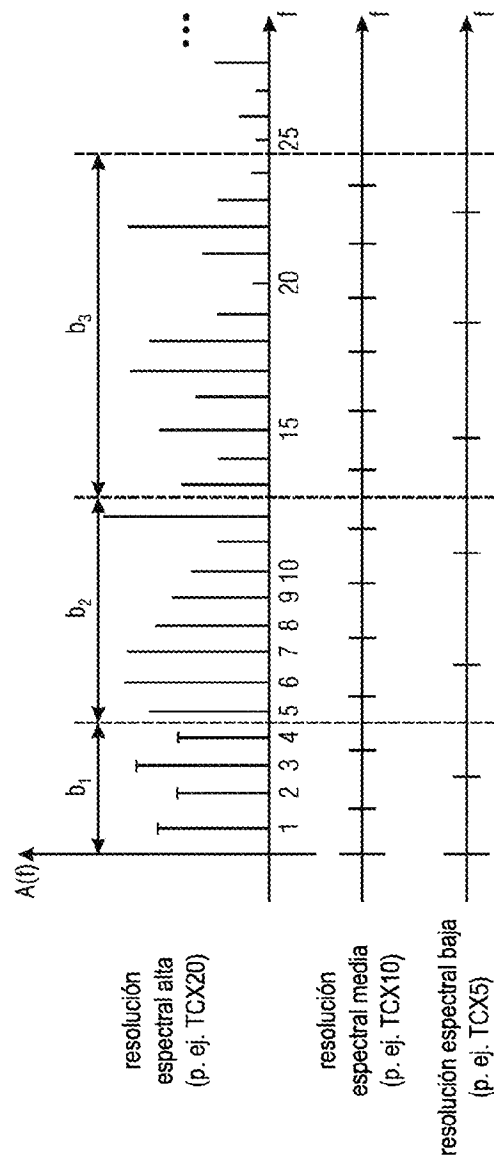
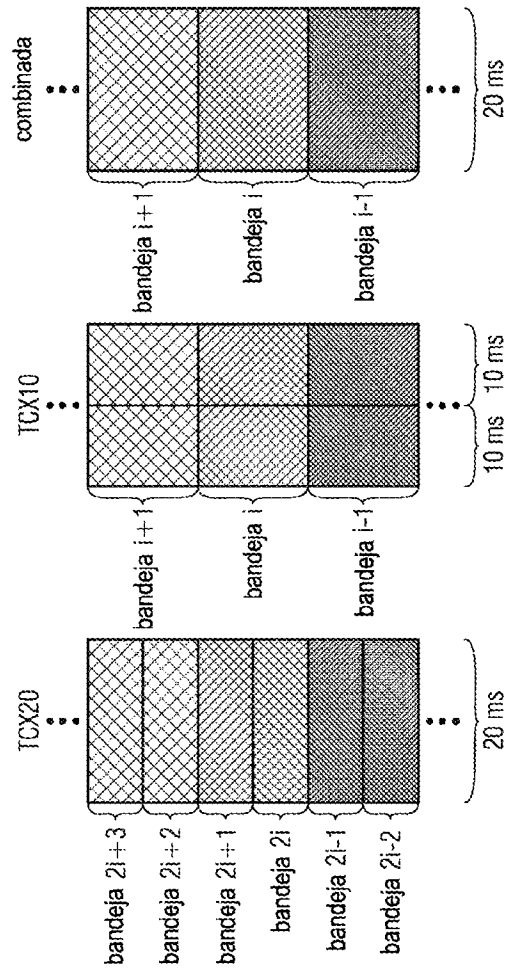
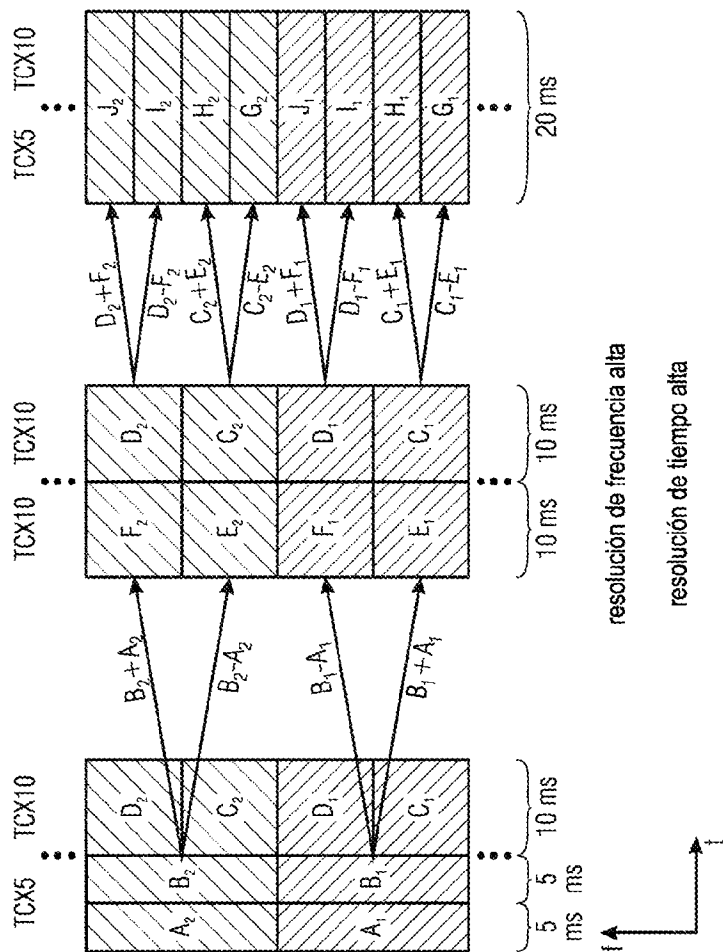


Figura 4b



resolución de frecuencia baja
 resolución de tiempo baja

Figura 5a



resolución de frecuencia alta
 resolución de tiempo alta

Figura 5b

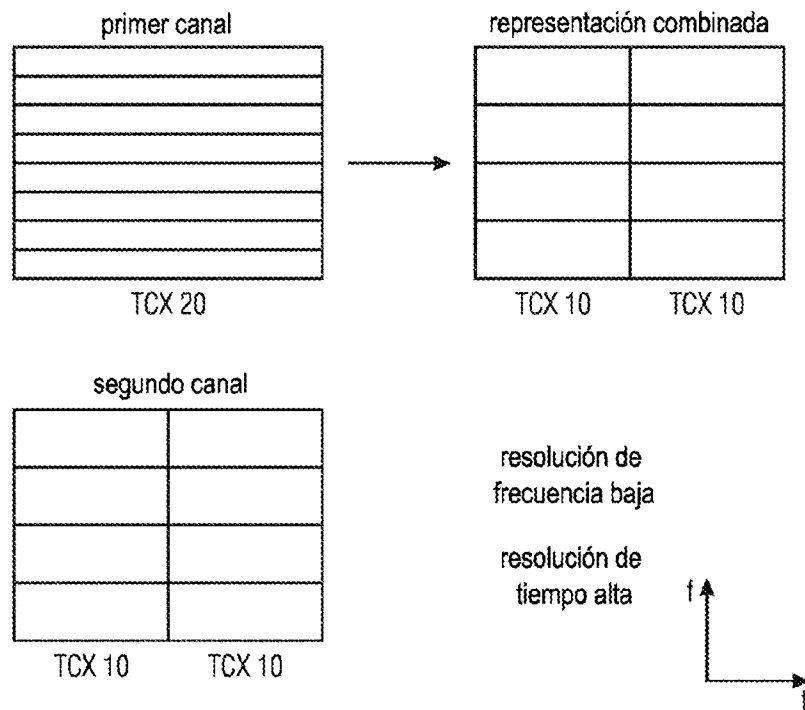


Figura 5c

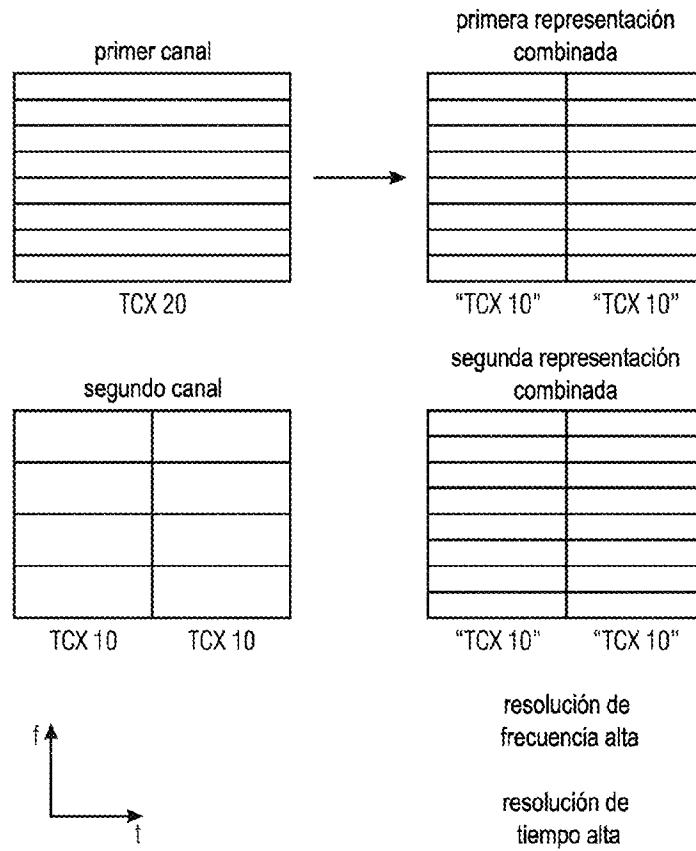
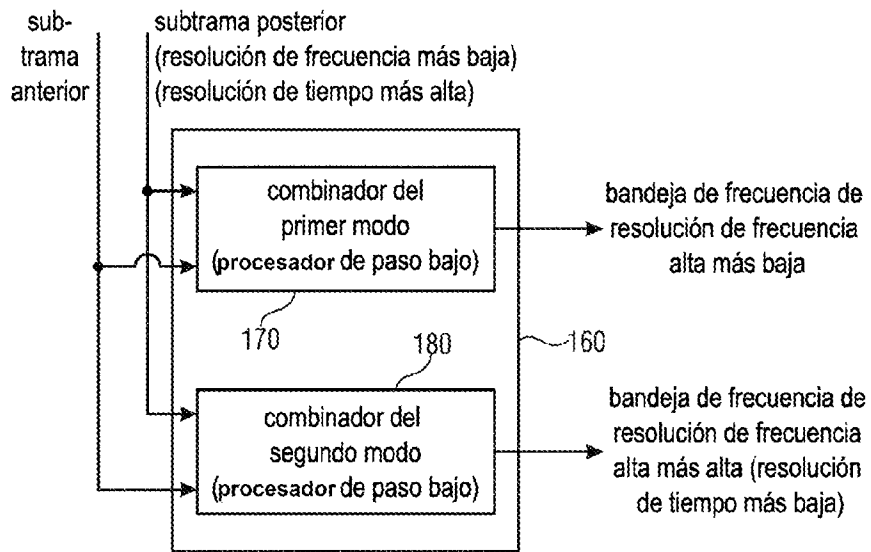


Figura 5d



(CONVERTIDOR DE RESOLUCIÓN ESPECTRAL)

Figura 6

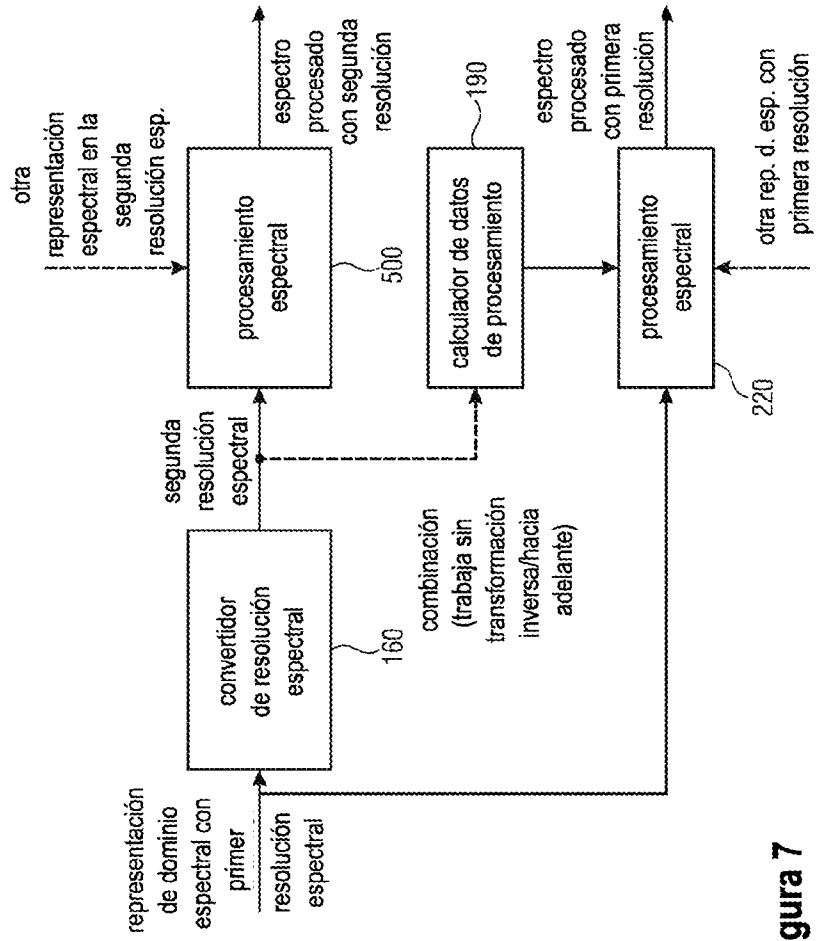


Figura 7

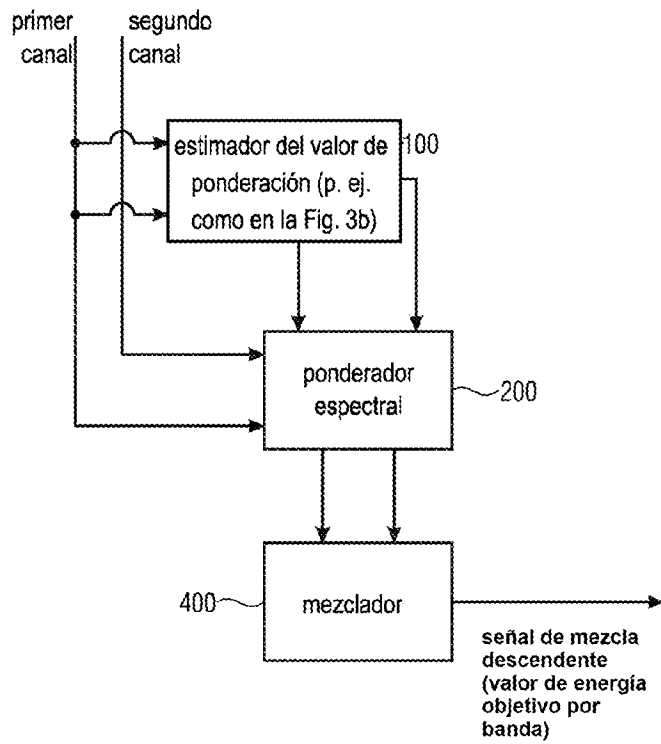


Figura 8

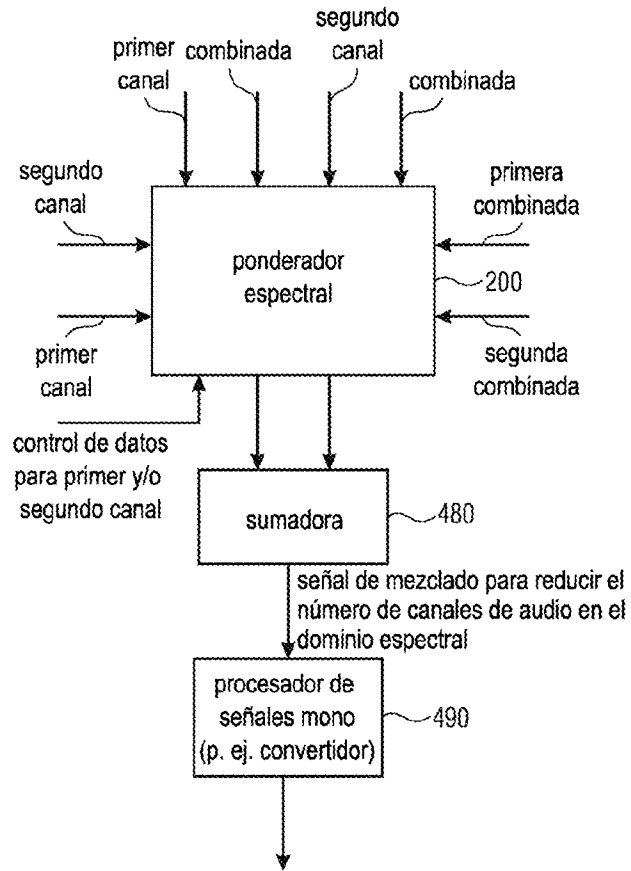


Figura 9