

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 641 580**

51 Int. Cl.:

H04S 7/00

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **26.09.2014 PCT/US2014/057671**

87 Fecha y número de publicación internacional: **09.04.2015 WO15050785**

96 Fecha de presentación y número de la solicitud europea: **26.09.2014 E 14781030 (3)**

97 Fecha y número de publicación de la concesión europea: **30.08.2017 EP 3053359**

54 Título: **Generación de señal difusa adaptativa en un mezclador ascendente**

30 Prioridad:

03.10.2013 US 201361886554 P
22.11.2013 US 201361907890 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
10.11.2017

73 Titular/es:

DOLBY LABORATORIES LICENSING CORPORATION (100.0%)
1275 Market Street
San Francisco, CA 94103, US

72 Inventor/es:

SEEFELDT, ALAN J.;
VINTON, MARK S. y
BROWN, C. PHILLIP

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 641 580 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

Generación de señal difusa adaptativa en un mezclador ascendente.

Campo técnico

5 La presente descripción se refiere al procesamiento de datos de audio. En particular, la presente descripción se refiere al procesamiento de datos de audio que incluye tanto señales de audio difusas como señales de audio direccionales durante un proceso de mezcla ascendente.

Antecedentes

10 Un proceso conocido como mezcla ascendente implica derivar cierta cantidad M de canales de señal de audio de una cantidad más pequeña N de canales de señal de audio. Algunos dispositivos de procesamiento de audio con capacidad de mezcla ascendente (a los cuales se puede hacer referencia en la presente memoria como "mezcladores ascendentes") pueden, por ejemplo, producir 3, 5, 7, 9 o más canales de audio según 2 canales de audio de entrada. Algunos mezcladores ascendentes pueden analizar la fase y amplitud de dos canales de señal de entrada para determinar cómo el campo de sonido que representan pretende transmitir impresiones direccionales a un oyente. Un ejemplo de dicho dispositivo de mezcla ascendente es el descodificador Dolby® Pro Logic® II descrito en el documento de Gundry, "A New Active Matrix Decoder for Surround Sound" (19ª Conferencia AES, mayo de 2001).

15 Las señales de audio de entrada pueden incluir datos de audio difusos y/o direccionales. Con respecto a los datos de audio direccionales, un mezclador ascendente debe poder generar señales de salida para múltiples canales con el fin de proveer al oyente la sensación de uno o más componentes auditivos con ubicaciones y/o direcciones aparentes. Algunas señales de audio como, por ejemplo, las correspondientes a los disparos, pueden ser muy direccionales. Las señales de audio difusas como, por ejemplo, las correspondientes al viento, lluvia, ruido ambiente, etc., pueden tener poca o ninguna direccionalidad aparente. Cuando se procesan datos de audio que también incluyen señales de audio difusas, se debe proveer al oyente la percepción de un campo de sonido difuso envolvente correspondiente a las señales de audio difusas.

25 Compendio

Se proveen métodos mejorados para procesar señales de audio difusas. Algunas implementaciones implican un método para derivar M señales de audio difusas de N señales de audio para la presentación de un campo de sonido difuso, en donde M es mayor que N y es mayor que 2. Cada una de las N señales de audio puede corresponder a una ubicación espacial.

30 El método puede implicar recibir las N señales de audio, derivar porciones difusas de las N señales de audio y detectar instancias de condiciones de señal de audio transitoria. El método puede implicar procesar las porciones difusas de las N señales de audio para derivar las M señales de audio difusas. Durante instancias de condiciones de señal de audio transitoria, el procesamiento puede implicar distribuir las porciones difusas de las N señales de audio en mayor proporción a una o más de las M señales de audio difusas correspondientes a ubicaciones espaciales relativamente más cercanas a las ubicaciones espaciales de las N señales de audio y en menor proporción a una o más de las M señales de audio difusas correspondientes a ubicaciones espaciales relativamente más lejanas respecto a las ubicaciones espaciales de las N señales de audio.

35 El método puede implicar detectar instancias de condiciones de señal de audio no transitoria. Durante instancias de condiciones de señal de audio no transitoria, el procesamiento puede implicar distribuir las porciones difusas de las N señales de audio a las M señales de audio difusas de manera sustancialmente uniforme.

40 El procesamiento puede implicar aplicar una matriz de mezcla a las porciones difusas de las N señales de audio para derivar las M señales de audio difusas. La matriz de mezcla puede ser una matriz de distribución variable. La matriz de distribución variable se puede derivar de una matriz no transitoria más apropiada para usar durante condiciones de señal de audio no transitoria y de una matriz transitoria más apropiada para usar durante condiciones de señal de audio transitoria. En algunas implementaciones, la matriz transitoria se puede derivar de la matriz no transitoria. Cada elemento de la matriz transitoria puede representar una escala de un elemento de matriz no transitoria correspondiente. En algunas instancias, la escala puede ser una función de una relación entre una ubicación de canal de entrada y una ubicación de canal de salida.

45 El método puede implicar determinar un valor de señal de control transitoria. En algunas implementaciones, la matriz de distribución variable se puede derivar interpolando entre la matriz transitoria y la matriz no transitoria según, al menos en parte, el valor de señal de control transitoria. El valor de señal de control transitoria puede ser variable con el tiempo. En algunas implementaciones, el valor de señal de control transitoria puede variar de manera continua de un valor mínimo a un valor máximo. De manera alternativa, el valor de señal de control transitoria puede variar en un rango de valores discretos de un valor mínimo a un valor máximo.

En algunas implementaciones, determinar la matriz de distribución variable puede implicar computar la matriz de distribución variable según el valor de señal de control transitoria. Sin embargo, determinar la matriz de distribución variable puede implicar recuperar una matriz de distribución variable almacenada de un dispositivo de memoria.

5 El método puede implicar derivar el valor de señal de control transitoria en respuesta a las N señales de audio. El método puede implicar transformar cada una de las N señales de audio en B bandas de frecuencia y llevar a cabo la derivación, detección y procesamiento, de manera separada, para cada una de las B bandas de frecuencia. El método puede implicar tomar una panorámica de las porciones no difusas de las N señales de audio para formar M señales de audio no difusas y combinar las M señales de audio difusas con las M señales de audio no difusas para formar M señales de audio de salida.

10 En algunas implementaciones, el método puede implicar derivar K señales intermedias de las porciones difusas de las N señales de audio, en donde K es mayor que o igual a uno y es menor que o igual a M-N. Cada señal de audio intermedia se puede correlacionar, de forma psicoacústica con las porciones difusas de las N señales de audio. Si K es mayor que uno, cada señal de audio intermedia se puede decorrelacionar de forma psicoacústica con todas las otras señales de audio intermedias. En algunas implementaciones, derivar las K señales intermedias puede implicar un proceso de decorrelación que puede incluir uno o más de retardos, filtros pasa todo, filtros pseudoaleatorios o algoritmos de reverberación. Las M señales de audio difusas se pueden derivar en respuesta a las K señales intermedias así como a las N señales difusas.

20 Algunos aspectos de la presente descripción se pueden implementar en un aparato que incluye un sistema de interfaz y un sistema lógico. El sistema lógico puede incluir uno o más procesadores como, por ejemplo, procesadores de propósito general de un solo chip o multichip, procesadores digitales de señales (DSP, por sus siglas en inglés), circuitos integrados de aplicación específica (ASIC, por sus siglas en inglés), matrices de puertas programables en campo (FPGA, por sus siglas en inglés) u otros dispositivos de lógica programable, puerta discreta o lógica de transistor, componentes de hardware discretos y/o combinaciones de ellos. El sistema de interfaz puede incluir al menos una de una interfaz de usuario o una interfaz de red. El aparato puede incluir un sistema de memoria. El sistema de interfaz puede incluir al menos una interfaz entre el sistema lógico y el sistema de memoria.

30 El sistema lógico puede recibir, mediante el sistema de interfaz, N señales de audio de entrada. Cada una de las N señales de audio puede corresponder a una ubicación espacial. El sistema lógico puede derivar porciones difusas de las N señales de audio y detectar instancias de condiciones de señal de audio transitoria. El sistema lógico puede procesar las porciones difusas de las N señales de audio para derivar M señales de audio difusas, en donde M es mayor que N y es mayor que 2. Durante instancias de condiciones de señal de audio transitoria, el procesamiento puede implicar distribuir las porciones difusas de las N señales de audio en mayor proporción a una o más de las M señales de audio difusas correspondientes a ubicaciones espaciales relativamente más cercanas a las ubicaciones espaciales de las N señales de audio y en menor proporción a una o más de las M señales de audio difusas correspondientes a ubicaciones espaciales relativamente más lejanas respecto a las ubicaciones espaciales de las N señales de audio.

El sistema lógico puede detectar instancias de condiciones de señal de audio no transitoria. Durante instancias de condiciones de señal de audio no transitoria, el procesamiento puede implicar distribuir las porciones difusas de las N señales de audio a las M señales de audio difusas de manera sustancialmente uniforme.

40 El procesamiento puede implicar aplicar una matriz de mezcla a las porciones difusas de las N señales de audio para derivar las M señales de audio difusas. La matriz de mezcla puede ser una matriz de distribución variable. La matriz de distribución variable se puede derivar de una matriz no transitoria más apropiada para usar durante condiciones de señal de audio no transitoria y de una matriz transitoria más apropiada para usar durante condiciones de señal de audio transitoria. En algunas implementaciones, la matriz transitoria se puede derivar de la matriz no transitoria. Cada elemento de la matriz transitoria puede representar una escala de un elemento de matriz no transitoria correspondiente. En algunos ejemplos, la escala puede ser una función de una relación entre una ubicación de canal de entrada y una ubicación de canal de salida.

El sistema lógico puede determinar un valor de señal de control transitoria. En algunos ejemplos, la matriz de distribución variable se puede derivar interpolando entre la matriz transitoria y la matriz no transitoria según, al menos en parte, el valor de señal de control transitoria.

50 En algunas implementaciones, el sistema lógico puede transformar cada una de las N señales de audio en B bandas de frecuencia. El sistema lógico puede llevar a cabo la derivación, detección y procesamiento, de forma separada, para cada una de las B bandas de frecuencia.

55 El sistema lógico puede tomar una panorámica de las porciones no difusas de las N señales de audio de entrada para formar M señales de audio no difusas. El sistema lógico puede combinar las M señales de audio difusas con las M señales de audio no difusas para formar M señales de audio de salida.

5 El documento US2011/0081024 describe una técnica que reduce el suavizado de ganancia a lo largo de cortes espaciales de un contenido de audio si se detecta una señal de audio transitoria (como, p.ej., el comienzo de una batería). Dichos cortes espaciales contienen contenido de audio de ubicaciones perceptuales (como ubicaciones en escena de los intérpretes), es preciso ver D1, §36, 37, 72. El documento US7970144 detecta episodios de audio transitorios (p.ej. sonido de instrumentos tipo percusión) y controla, por consiguiente, la dirección de la panorámica (es preciso ver col.8,1.48-51) o una ganancia (es preciso ver col. 10,1.2-4).

10 Los métodos descritos en la presente memoria se pueden implementar mediante hardware, firmware, software almacenado en uno o más medios no transitorios y/o combinaciones de ellos. Los detalles de una o más implementaciones del objeto descrito en la presente memoria descriptiva se establecen en los dibujos anexos y en la descripción de más abajo. Otras características, aspectos y ventajas serán aparentes a partir de la descripción, dibujos y reivindicaciones. Es preciso notar que las dimensiones relativas de las siguientes figuras pueden no dibujarse a escala.

La invención se define en las reivindicaciones independientes. Las realizaciones preferidas están definidas en las reivindicaciones dependientes.

15 Breve descripción de los dibujos

La Figura 1 muestra un ejemplo de mezcla ascendente.

La Figura 2 muestra un ejemplo de un sistema de procesamiento de audio.

La Figura 3 es un diagrama de flujo que describe bloques de un método de procesamiento de audio que se puede llevar a cabo por un sistema de procesamiento de audio.

20 La Figura 4A es un diagrama de bloques que provee otro ejemplo de un sistema de procesamiento de audio.

La Figura 4B es un diagrama de bloques que provee otro ejemplo de un sistema de procesamiento de audio.

La Figura 5 muestra ejemplos de factores de escala para una implementación que implica una señal de entrada estéreo y una señal de salida de cinco canales.

25 La Figura 6 es un diagrama de bloques que muestra detalles adicionales de un procesador de señal difusa según un ejemplo.

La Figura 7 es un diagrama de bloques de un aparato que puede generar un conjunto de M señales de salida intermedias a partir de N señales de entrada intermedias.

La Figura 8 es un diagrama de bloques que muestra un ejemplo de decorrelación de señales intermedias seleccionadas.

30 La Figura 9 es un diagrama de bloques que muestra un ejemplo de componentes del decorrelador.

La Figura 10 es un diagrama de bloques que muestra un ejemplo alternativo de componentes del decorrelador.

La Figura 11 es un diagrama de bloques que provee ejemplos de componentes de un aparato de procesamiento de audio.

Los números de referencia y las designaciones iguales en los diversos dibujos indican elementos iguales.

35 Descripción de realizaciones a modo de ejemplo

La siguiente descripción se dirige a ciertas implementaciones con el fin de describir algunos aspectos innovadores de la presente descripción, así como ejemplos de contextos en los cuales dichos aspectos innovadores se pueden implementar. Sin embargo, las enseñanzas en la presente memoria descriptiva se pueden aplicar de varias maneras diferentes. Por ejemplo, mientras varias implementaciones de describen en términos de ambientes de reproducción particulares, las enseñanzas en la presente memoria son ampliamente aplicables a otros ambientes de reproducción conocidos, así como ambientes de reproducción que se pueden introducir en el futuro. Además, las implementaciones descritas se pueden implementar, al menos en parte, en varios dispositivos y sistemas como hardware, software, firmware, sistemas basados en la nube, etc. Por consiguiente, las enseñanzas de la presente descripción no pretenden limitarse a las implementaciones que se muestran en las figuras y/o que se describen en la presente memoria, sino que, en su lugar, tienen una amplia aplicabilidad.

45 La Figura 1 muestra un ejemplo de mezcla ascendente. En varios ejemplos descritos en la presente memoria, el sistema de procesamiento de audio 10 puede proveer una funcionalidad de mezclador ascendente y, en la presente memoria, se puede hacer referencia a este como un mezclador ascendente. En el presente ejemplo, el sistema de

procesamiento de audio 10 puede obtener señales de audio para cinco canales de salida designados como izquierdo (I), derecho (D), central (C), izquierdo-ambiente (IA) y derecho-ambiente (DA) mediante la mezcla ascendente de señales de audio para dos canales de entrada, los cuales son los canales izquierdo-entrada (I_e) y derecho-entrada (D_e) en el presente ejemplo. Algunos mezcladores ascendentes pueden producir diferentes cantidades de canales, p.ej., 3, 7, 9 o más canales de salida, a partir de 2 o de un número diferente de canales de entrada, p.ej., 3, 5 o más canales de entrada.

Las señales de audio de entrada incluirán, en general, tanto datos de audio difusos como direccionales. Con respecto a los datos de audio direccionales, el sistema de procesamiento de audio 10 debe poder generar señales de salida direccionales que proveen al oyente 105 la sensación de uno o más componentes auditivos con ubicaciones y/o direcciones aparentes. Por ejemplo, el sistema de procesamiento de audio 10 puede aplicar un algoritmo de panorámica para crear una imagen fantasma o dirección aparente de sonido entre dos altavoces 110 mediante la reproducción de la misma señal de audio a través de cada uno de los altavoces 110.

Con respecto a los datos de audio difusos, sistema de procesamiento de audio 10 debe poder generar señales de audio difusas que proveen al oyente 105 la percepción de un campo de sonido difuso envolvente en el cual el sonido parece emanar de muchas (si no todas) direcciones alrededor del oyente 105. Un campo de sonido difuso de alta calidad normalmente no se puede crear mediante la simple reproducción de la misma señal de audio a través de múltiples altavoces 110 ubicados alrededor de un oyente. El campo de sonido resultante tendrá, en general, amplitudes que varían sustancialmente en diferentes ubicaciones auditivas, que con frecuencia cambian en grandes cantidades para cambios muy pequeños en la ubicación del oyente 105. Algunas posiciones dentro del área auditiva pueden parecer desprovistas de sonido para un oído pero no para el otro. El campo de sonido resultante puede parecer artificial. Por lo tanto, algunos mezcladores ascendentes pueden decorrelacionar las porciones difusas de las señales de salida, con el fin de crear la impresión de que las porciones difusas de las señales de audio se distribuyen, de manera uniforme, alrededor del oyente 105. Sin embargo, se ha observado que durante momentos "transitorios" o "de percusión" de la señal de audio de entrada, el resultado de la propagación de las señales difusas de manera uniforme a lo largo de todos los canales de salida puede ser un "esparcimiento" o "falta de golpe" percibido en el transitorio original. Ello puede ser especialmente problemático cuando varios de los canales de salida se encuentran espacialmente distantes de los canales de entrada originales. Tal es el caso, por ejemplo, con las señales ambiente derivadas de la entrada estéreo estándar.

Con el fin de abordar las cuestiones anteriores, algunas implementaciones descritas en la presente memoria proveen un mezclador ascendente que puede separar porciones difusas y no difusas o "directas" de las N señales de audio de entrada. El mezclador ascendente puede detectar instancias de condiciones de señal de audio transitoria. Durante las instancias de condiciones de señal de audio transitoria, el mezclador ascendente puede adicionar un control adaptativo a la señal a un proceso de expansión de señal difusa en el cual se producen M señales de audio. La presente descripción supone que el número N es mayor que o igual a uno, el número M es mayor que o igual a tres y el número M es mayor que el número N.

Según algunas de dichas implementaciones, el mezclador ascendente puede variar el proceso de expansión de señal difusa con el tiempo de modo que durante instancias de condiciones de señal de audio transitoria las porciones difusas de las señales de audio se pueden distribuir sustancialmente solo a canales de salida espacialmente cercanos a los canales de entrada. Durante instancias de condiciones de señal de audio no transitoria, las porciones difusas de señales de audio se pueden distribuir de manera sustancialmente uniforme. Con dicho enfoque, las porciones difusas de las señales de audio permanecen en la cercanía espacial de las señales de audio originales durante instancias de condiciones de señal de audio transitoria, con el fin de mantener el impacto de los transitorios. Durante instancias de condiciones de señal de audio no transitoria, las porciones difusas de las señales de audio se pueden extender de manera sustancialmente uniforme, con el fin de maximizar el envolvimiento.

La Figura 2 muestra un ejemplo de un sistema de procesamiento de audio. En la presente implementación, el sistema de procesamiento de audio 10 incluye un sistema de interfaz 205, un sistema lógico 210 y un sistema de memoria 215. El sistema de interfaz 205 puede, por ejemplo, incluir una o más interfaces de red, interfaces de usuario, etc. El sistema de interfaz 205 puede incluir una o más interfaces de bus serie universal (USB, por sus siglas en inglés) o interfaces similares. El sistema de interfaz 205 puede incluir interfaces inalámbricas o cableadas.

El sistema lógico 210 puede incluir uno o más procesadores como, por ejemplo, uno o más procesadores de propósito general de un solo chip o multichip, procesadores digitales de señales (DSP), circuitos integrados de aplicación específica (ASIC), matrices de puertas programables en campo (FPGA) u otros dispositivos de lógica programable, puerta discreta o lógica de transistor, componentes de hardware discretos, o combinaciones de ellos.

El sistema de memoria 215 puede incluir uno o más medios no transitorios como, por ejemplo, memoria de acceso aleatorio (RAM, por sus siglas en inglés) y/o memoria de solo lectura (ROM, por sus siglas en inglés). El sistema de memoria 215 puede incluir uno o más de otros tipos apropiados de medios de almacenamiento no transitorios como, por ejemplo, memoria flash, uno o más discos duros, etc. En algunas implementaciones, el sistema de interfaz 205 puede incluir al menos una interfaz entre el sistema lógico 210 y el sistema de memoria 215.

El sistema de procesamiento de audio 10 puede llevar a cabo uno o más de los varios métodos descritos en la presente memoria. La Figura 3 es un diagrama de flujo que describe bloques de un método de procesamiento de audio que se puede llevar a cabo por un sistema de procesamiento de audio. Por consiguiente, el método 300 que se describe en la Figura 3 también se describirá con referencia al sistema de procesamiento de audio 10 de la Figura 2. Al igual que con otros métodos descritos en la presente memoria, las funciones del método 300 no se llevan a cabo necesariamente en el orden que se muestra en la Figura 3. Además, el método 300 (y otros métodos provistos en la presente memoria) puede incluir más o menos bloques de los que se muestran o describen.

En el presente ejemplo, el bloque 305 de la Figura 3 implica recibir N señales de audio de entrada. Cada una de las N señales de audio puede corresponder a una ubicación espacial. Por ejemplo, para algunas implementaciones en las cuales $N=2$, las ubicaciones espaciales pueden corresponder a las supuestas ubicaciones de los canales de audio de entrada izquierdo y derecho. En algunas implementaciones, el sistema lógico 210 puede recibir, mediante el sistema de interfaz 205, las N señales de audio de entrada.

En algunas implementaciones, los bloques del método 300 se pueden llevar a cabo para cada una de las múltiples bandas de frecuencia. Por consiguiente, en algunas implementaciones, el bloque 305 puede implicar recibir datos de audio, correspondientes a las N señales de audio de entrada, que se han descompuesto en múltiples bandas de frecuencia. En implementaciones alternativas, el bloque 305 puede incluir un proceso de descomposición de los datos de audio de entrada en múltiples bandas de frecuencia. Por ejemplo, dicho proceso puede implicar algún tipo de banco de filtros como, por ejemplo, una transformada de Fourier de Tiempo Corto (STFT, por sus siglas en inglés) o Banco de Filtros Espejo en Cuadratura (QMF, por sus siglas en inglés).

En la presente implementación, el bloque 310 de la Figura 3 implica derivar porciones difusas de las N señales de audio de entrada. Por ejemplo, el sistema lógico 210 puede separar las porciones difusas de las porciones no difusas de las N señales de audio de entrada. Algunos ejemplos del presente proceso se proveen más abajo. En cualquier instante dado en el tiempo, el número de señales de audio correspondiente a las porciones difusas de las N señales de audio de entrada puede ser N, menos que N o más que N.

El sistema lógico 210 puede decorrelacionar señales de audio, al menos en parte. La correlación numérica de dos señales se puede calcular usando una variedad de algoritmos numéricos conocidos. Dichos algoritmos producen una medida de correlación numérica llamada un coeficiente de correlación que varía entre uno negativo y uno positivo. Un coeficiente de correlación con una magnitud igual o cercana a uno indica que dos señales se relacionan de forma cercana. Un coeficiente de correlación con una magnitud igual o cercana a cero indica que dos señales son, en general, independientes entre sí.

La correlación psicoacústica se refiere a las propiedades de correlación de las señales de audio que existen a lo largo de las subbandas de frecuencia que tienen el así llamado ancho de banda crítico. El poder de resolución de frecuencia del sistema auditivo humano varía con la frecuencia a lo largo del espectro de audio. El oído humano puede discernir componentes espectrales más cercanos entre sí en la frecuencia a frecuencias más bajas por debajo de alrededor de 500 Hz pero no tan cercanos entre sí dado que la frecuencia avanza hacia arriba hacia los límites de la audición. Se hace referencia al ancho de la presente resolución de frecuencia como un ancho de banda crítico, el cual varía con la frecuencia.

Se dice que dos señales de audio se decorrelacionan de forma psicoacústica una con respecto a la otra si el coeficiente de correlación numérica promedio a lo largo de los anchos de banda críticos psicoacústicos es igual o cercano a cero. La decorrelación psicoacústica se logra si el coeficiente de correlación numérica entre dos señales es igual o cercano a cero en todas las frecuencias. La decorrelación psicoacústica también puede lograrse incluso si el coeficiente de correlación numérica entre dos señales no es igual o cercano a cero en todas las frecuencias si la correlación numérica varía de modo que su promedio a lo largo de cada banda crítica psicoacústica es menos que la mitad del coeficiente de correlación máximo para cualquier frecuencia dentro de dicha banda crítica. Por consiguiente, la decorrelación psicoacústica es menos estricta que la decorrelación numérica en que dos señales se pueden considerar psicoacústicamente decorrelacionadas incluso si tienen algún grado de correlación numérica entre sí.

El sistema lógico 210 puede derivar K señales intermedias de las porciones difusas de las N señales de audio de modo que cada una de las K señales de audio intermedias se decorrelaciona psicoacústicamente con las porciones difusas de las N señales de audio. Si K es mayor que uno, cada una de las K señales de audio intermedias puede decorrelacionarse de forma psicoacústica con todas las otras señales de audio intermedias. Algunos ejemplos se describen más abajo.

En algunas implementaciones, el sistema lógico 210 también puede llevar a cabo las funciones descritas en los bloques 315 y 320 de la Figura 3. En el presente ejemplo, el bloque 315 implica detectar instancias de condiciones de señal de audio transitoria. Por ejemplo, el bloque 315 puede implicar detectar el comienzo de un cambio abrupto en la energía, p.ej., determinando si un cambio en la energía con el tiempo ha superado un umbral predeterminado. Por consiguiente, en la presente memoria se puede hacer referencia a la detección transitoria como detección de

comienzo. Más abajo se proveen ejemplos con referencia al módulo de detección de comienzo 415 de las Figuras 4B y 6. Algunos de dichos ejemplos implican la detección de comienzo en múltiples bandas de frecuencia. Por lo tanto, en algunas instancias, el bloque 315 puede implicar detectar una instancia de una señal de audio transitoria en algunas, pero no todas, las bandas de frecuencia.

5 Aquí, el bloque 320 implica procesar las porciones difusas de las N señales de audio para derivar las M señales de audio difusas. Durante instancias de condiciones de señal de audio transitoria, el procesamiento del bloque 320 puede implicar distribuir las porciones difusas de las N señales de audio en mayor proporción a una o más de las M señales de audio difusas correspondientes a ubicaciones espaciales relativamente más cercanas a las ubicaciones espaciales de las N señales de audio. El procesamiento del bloque 320 puede implicar distribuir las porciones difusas de las N señales de audio en menor proporción a una o más de las M señales de audio difusas correspondientes a ubicaciones espaciales relativamente más lejanas respecto a las ubicaciones espaciales de las N señales de audio. Un ejemplo se muestra en la Figura 5 y se describe más abajo. En algunas de dichas implementaciones, el procesamiento del bloque 320 puede implicar mezclar las porciones difusas de las N señales de audio y las K señales de audio intermedias para derivar las M señales de audio difusas. Durante instancias de condiciones de señal de audio transitoria, el proceso de mezcla puede implicar distribuir las porciones difusas de las señales de audio principalmente a señales de audio de salida que corresponden a canales de salida espacialmente cercanos a los canales de entrada. Algunas implementaciones implican también detectar instancias de condiciones de señal de audio no transitoria. Durante instancias de condiciones de señal de audio no transitoria, la mezcla puede implicar distribuir las señales difusas a canales de salida para las M señales de audio de salida de manera sustancialmente uniforme.

En algunas implementaciones, el procesamiento del bloque 320 puede implicar aplicar una matriz de mezcla a las porciones difusas de las N señales de audio y las K señales de audio intermedias para derivar las M señales de audio difusas. Por ejemplo, la matriz de mezcla puede ser una matriz de distribución variable que se deriva de una matriz no transitoria más apropiada para usar durante condiciones de señal de audio no transitoria y una matriz transitoria más apropiada para usar durante condiciones de señal de audio transitoria. En algunas implementaciones, la matriz transitoria se puede derivar de la matriz no transitoria. Según algunas de dichas implementaciones, cada elemento de la matriz transitoria puede representar una escala de un elemento de matriz no transitoria correspondiente. La escala puede, por ejemplo, ser una función de una relación entre una ubicación de canal de entrada y una ubicación de canal de salida.

30 Ejemplos más detallados del método 300 se proveen más abajo, incluidos, pero sin limitación, ejemplos de la matriz transitoria y la matriz no transitoria. Por ejemplo, varios ejemplos de bloques 315 y 320 se describen más abajo con referencia a las Figuras 4B-5.

La Figura 4A es un diagrama de bloques que provee otro ejemplo de un sistema de procesamiento de audio. Los bloques de la Figura 4A pueden, por ejemplo, implementarse por el sistema lógico 210 de la Figura 2. En algunas implementaciones, los bloques de la Figura 4A se pueden implementar, al menos en parte, por software almacenado en un medio no transitorio. En la presente implementación, el sistema de procesamiento de audio 10 puede recibir señales de audio para uno o más canales de entrada del trayecto de señal 19 y generar señales de audio a lo largo del trayecto de señal 59 para múltiples canales de salida. La pequeña línea que cruza el trayecto de la señal 19, así como las pequeñas líneas que cruzan los otros trayectos de señal, indican que dichos trayectos de señal pueden llevar señales para uno o más canales. Los símbolos N y M inmediatamente debajo de las pequeñas líneas de cruce indican que los diferentes trayectos de señal pueden llevar señales para los N y M canales, respectivamente. Los símbolos "x" e "y" inmediatamente debajo de algunas de las pequeñas líneas de cruce indican que los respectivos trayectos de señal pueden llevar un número no especificado de señales.

En el sistema de procesamiento de audio 10, el analizador de señal de entrada 20 puede recibir señales de audio para uno o más canales de entrada del trayecto de señal 19 y determinar qué porciones de las señales de audio de entrada representan un campo de sonido difuso y qué porciones de las señales de audio de entrada representan un campo de sonido que no es difuso. El analizador de señal de entrada 20 puede pasar las porciones de las señales de audio de entrada que se considera que representan un campo de sonido no difuso a lo largo del trayecto de señal 28 al procesador de señal no difusa 30. Aquí, el procesador de señal no difusa 30 puede generar un conjunto de M señales de audio que pretenden reproducir el campo de sonido no difuso a través de múltiples transductores acústicos como, por ejemplo, altavoces, y transmitir dichas señales de audio a lo largo del trayecto de señal 39. Un ejemplo de un dispositivo de mezcla ascendente que puede llevar a cabo este tipo de procesamiento es un decodificador Dolby Pro Logic II™.

En el presente ejemplo, el analizador de señal de entrada 20 puede transmitir las porciones de las señales de audio de entrada correspondientes a un campo de sonido difuso a lo largo del trayecto de señal 29 al procesador de señal difusa 40. Aquí, el procesador de señal difusa 40 puede generar, a lo largo del trayecto de señal 49, un conjunto de M señales de audio correspondientes a un campo de sonido difuso. La presente descripción provee varios ejemplos de procesamiento de audio que se puede llevar a cabo por el procesador de señal difusa 40.

En la presente realización, el componente de suma 50 puede combinar cada una de las M señales de audio del procesador de señal no difusa 30 con una señal respectiva de las M señales de audio del procesador de señal difusa 40 para generar una señal de audio para uno respectivo de los M canales de salida. La señal de audio para cada canal de salida puede pretender dirigir un transductor acústico como, por ejemplo, un altavoz.

5 Varias implementaciones descritas en la presente memoria se dirigen a desarrollar y usar un sistema de ecuaciones de mezcla para generar un conjunto de señales de audio que pueden representar un campo de sonido difuso. En algunas implementaciones, las ecuaciones de mezcla pueden ser ecuaciones de mezcla lineales. Las ecuaciones de mezcla se pueden usar en el procesador de señal difusa 40, por ejemplo.

10 Sin embargo, el sistema de procesamiento de audio 10 es meramente un ejemplo de cómo se puede implementar la presente descripción. La presente descripción se puede implementar en otros dispositivos que pueden diferir en función o estructura de aquellos que se muestran y describen en la presente memoria. Por ejemplo, las señales que representan tanto porciones difusas como no difusas de un campo de sonido se pueden procesar por un solo componente. Algunas implementaciones para un procesador de señal difusa 40 distinto se describen más abajo, las que mezclan señales según un sistema de ecuaciones lineales definidas por una matriz. Varias partes de los
15 procesos para el procesador de señal difusa 40 y el procesador de señal no difusa 30 se pueden implementar por un sistema de ecuaciones lineales definidas por una sola matriz. Además, aspectos de la presente invención se pueden incorporar a un dispositivo sin incorporar también el analizador de señal de entrada 20, el procesador de señal no difusa 30 o el componente de suma 50.

20 La Figura 4B es un diagrama de bloques que provee otro ejemplo de un sistema de procesamiento de audio. Los bloques de la Figura 4B incluyen ejemplos más detallados de los bloques de la Figura 4A, según algunas implementaciones. Por consiguiente, los bloques de la Figura 4B pueden, por ejemplo, implementarse por el sistema lógico 210 de la Figura 2. En algunas implementaciones, los bloques de la Figura 4B se pueden implementar, al menos en parte, por software almacenado en un medio no transitorio.

25 Aquí, el analizador de señal de entrada 20 incluye un módulo de análisis estadístico 405 y un módulo de separación de señal 410. En la presente implementación, el procesador de señal difusa 40 incluye un módulo de detección de comienzo 415 y un módulo de expansión de señal difusa adaptativa 420. Sin embargo, en implementaciones alternativas, la funcionalidad de los bloques que se muestran en la Figura 4B se puede distribuir entre diferentes módulos. Por ejemplo, en algunas implementaciones, el analizador de señal de entrada 20 puede llevar a cabo las funciones del módulo de detección de comienzo 415.

30 El módulo de análisis estadístico 405 puede llevar a cabo varios tipos de análisis en la N señal de audio de entrada de canal. Por ejemplo, si $N = 2$, el módulo de análisis estadístico 405 puede computar un cálculo de la suma de la energía en las señales izquierda y derecha, la diferencia de la energía en las señales izquierda y derecha, y la parte real de la correlación cruzada entre las señales de entrada izquierda y derecha. Cada cálculo estadístico se puede acumular en un bloque de tiempo y en una banda de frecuencia. El cálculo estadístico se puede suavizar con el
35 tiempo. Por ejemplo, el cálculo estadístico se puede suavizar usando un integrador con fuga dependiente de la frecuencia como, por ejemplo, un filtro con respuesta al impulso infinita de primer orden (IIR, por sus siglas en inglés). El módulo de análisis estadístico 405 puede proveer datos de análisis estadístico a otros módulos, p.ej., al módulo de separación de señal 410 y/o al módulo de panorámica 425.

40 En la presente implementación, el módulo de separación de señal 410 puede separar las porciones difusas de las N señales de audio de entrada de las porciones no difusas o "directas" de las N señales de audio de entrada. El módulo de separación de señal 410 puede, por ejemplo, determinar que porciones altamente correlacionadas de las N señales de audio de entrada corresponden a señales de audio no difusas. Por ejemplo, si $N = 2$, el módulo de separación de señal 410 puede determinar, según datos de análisis estadístico del módulo de análisis estadístico 405, que la señal de audio no difusa es una porción altamente correlacionada de la señal de audio que está
45 contenida tanto en las entradas izquierda como derecha.

Según los mismos (o similares) datos de análisis estadístico, el módulo de panorámica 425 puede determinar que dicha porción de la señal de audio debe dirigirse a una ubicación apropiada, p.ej., como representando una fuente de audio localizada como, por ejemplo, una fuente puntual. El módulo de panorámica 425, u otro módulo del procesador de señal no difusa 30, puede producir M señales de audio no difusas correspondientes a las porciones
50 no difusas de las N señales de audio de entrada. El procesador de señal no difusa 30 puede proveer las M señales de audio no difusas al componente de suma 50.

El módulo de separación de señal 410 puede, en algunos ejemplos, determinar que las porciones difusas de las señales de audio de entrada son aquellas porciones de la señal que permanecen después de haber aislado las porciones no difusas. Por ejemplo, el módulo de separación de señal 410 puede determinar las porciones difusas de
55 la señal de audio computando la diferencia entre la señal de audio de entrada y la porción no difusa de la señal de audio. El módulo de separación de señal 410 puede proveer las porciones difusas de la señal de audio al módulo de expansión de señal difusa adaptativa 420.

Aquí, el módulo de detección de comienzo 415 puede detectar instancias de condiciones de señal de audio transitoria. En el presente ejemplo, el módulo de detección de comienzo 415 puede determinar un valor de señal de control transitoria y proveer el valor de señal de control transitoria al módulo de expansión de señal difusa adaptativa 420. En algunas instancias, el módulo de detección de comienzo 415 puede determinar si una señal de audio en cada una de las múltiples bandas de frecuencia incluye una señal de audio transitoria. Por consiguiente, en algunas instancias, el valor de señal de control transitoria determinado por el módulo de detección de comienzo 415 y provisto al módulo de expansión de señal difusa adaptativa 420 puede ser específico para una o más bandas de frecuencia particulares, pero no para todas las bandas de frecuencia.

En la presente implementación, el módulo de expansión de señal difusa adaptativa 420 puede derivar K señales intermedias de las porciones difusas de las N señales de audio de entrada. En algunas implementaciones, cada señal de audio intermedia puede decorrelacionarse de forma psicoacústica con las porciones difusas de las N señales de audio de entrada. Si K es mayor que uno, cada señal de audio intermedia se puede decorrelacionar de forma psicoacústica con todas las otras señales de audio intermedias.

En la presente implementación, el módulo de expansión de señal difusa adaptativa 420 puede mezclar porciones difusas de las N señales de audio y las K señales de audio intermedias para derivar M señales de audio difusas, en donde M es mayor que N y es mayor que 2. En el presente ejemplo, K es mayor que o igual a uno y es menor que o igual a M-N. Durante instancias de condiciones de señal de audio transitoria (determinadas, al menos en parte, según el valor de señal de control transitoria recibido del módulo de detección de comienzo 415), el proceso de mezcla puede implicar distribuir las porciones difusas de las N señales de audio en mayor proporción a una o más de las M señales de audio difusas correspondientes a ubicaciones espaciales relativamente más cercanas a ubicaciones espaciales de las N señales de audio, p.ej., más cercanas a las supuestas ubicaciones espaciales de los N canales de entrada. Durante instancias de condiciones de señal de audio transitoria, el proceso de mezcla puede implicar distribuir las porciones difusas de las N señales de audio en menor proporción a una o más de las M señales de audio difusas correspondientes a ubicaciones espaciales relativamente más lejanas respecto a las ubicaciones espaciales de las N señales de audio. Sin embargo, durante instancias de condiciones de señal de audio no transitoria, el proceso de mezcla puede implicar distribuir las porciones difusas de las N señales de audio a las M señales de audio difusas de manera sustancialmente uniforme.

En algunas implementaciones, el módulo de expansión de señal difusa adaptativa 420 puede aplicar una matriz de mezcla a las porciones difusas de las N señales de audio y las K señales de audio intermedias para derivar las M señales de audio difusas. El módulo de expansión de señal difusa adaptativa 420 puede proveer las M señales de audio difusas al componente de suma 50, el cual puede combinar las M señales de audio difusas con las M señales de audio no difusas, para formar M señales de audio de salida.

Según algunas de dichas implementaciones, la matriz de mezcla aplicada por el módulo de expansión de señal difusa adaptativa 420 puede ser una matriz de distribución variable que se deriva de una matriz no transitoria más apropiada para usar durante condiciones de señal de audio no transitoria y una matriz transitoria más apropiada para usar durante condiciones de señal de audio transitoria. Varios ejemplos de determinación de matrices transitorias y matrices no transitorias se proveen más abajo.

Según algunas de dichas implementaciones, la matriz transitoria se puede derivar de la matriz no transitoria. Por ejemplo, cada elemento de la matriz transitoria puede representar una escala de un elemento de matriz no transitoria correspondiente. La escala puede, por ejemplo, ser una función de una relación entre una ubicación de canal de entrada y una ubicación de canal de salida. En algunas implementaciones, el módulo de expansión de señal difusa adaptativa 420 puede interpolar entre la matriz transitoria y la matriz no transitoria según, al menos en parte, un valor de señal de control transitoria recibido del módulo de detección de comienzo 415.

En algunas implementaciones, el módulo de expansión de señal difusa adaptativa 420 puede computar la matriz de distribución variable según el valor de señal de control transitoria. Algunos ejemplos se proveen más abajo. Sin embargo, en implementaciones alternativas, el módulo de expansión de señal difusa adaptativa 420 puede determinar la matriz de distribución variable recuperando una matriz de distribución variable almacenada de un dispositivo de memoria. Por ejemplo, el módulo de expansión de señal difusa adaptativa 420 puede determinar qué matriz de distribución variable de múltiples matrices de distribución variable almacenadas recuperar del dispositivo de memoria, según al menos en parte el valor de señal de control transitoria.

El valor de señal de control transitoria será, en general, variable con el tiempo. En algunas implementaciones, el valor de señal de control transitoria puede variar de manera continua de un valor mínimo a un valor máximo. Sin embargo, en implementaciones alternativas, el valor de señal de control transitoria puede variar en un rango de valores discretos de un valor mínimo a un valor máximo.

Dejemos que $c(t)$ represente una señal de control transitorio variable con el tiempo que tiene valores de señal de control transitorios que varían continuamente entre los valores cero y uno. En el presente ejemplo, un valor de señal de control transitoria de uno indica que la señal de audio correspondiente es de naturaleza transitoria y un valor de

señal de control transitoria de cero indica que la señal de audio correspondiente es no transitoria. Dejemos que T represente una "matriz transitoria" más apropiada para usar durante instancias de condiciones de señal de audio transitoria, y dejemos que C represente una "matriz no transitoria" más apropiada para usar durante instancias de condiciones de señal de audio no transitoria. Varios ejemplos de la matriz no transitoria se describen más abajo. Una versión no normalizada de la matriz de distribución variable D(t) puede computarse como una interpolación preservadora de energía entre las matrices transitoria y no transitoria:

$$\mathbf{D}(t) = c(t)\mathbf{T} + \sqrt{1 - c^2(t)}\mathbf{C} \quad (\text{Ecuación 1})$$

Con el fin de mantener la energía relativa de la señal de salida difusa de M canal, dicha matriz no normalizada puede entonces normalizarse de modo que la suma de los cuadrados de todos los elementos de la matriz es igual a uno:

$$\bar{\mathbf{D}}(t) = \alpha(t)\mathbf{D}(t) \quad (\text{Ecuación 2a})$$

$$\alpha(t) = \sqrt{\frac{1}{\sum_{i=1}^M \sum_{j=1}^{N+K} D_{ij}^2(t)}} \quad (\text{Ecuación 2b})$$

En la ecuación 2b, $D_{ij}(t)$ representa el elemento en la *i*ésima fila y *j*ésima columna de la matriz de distribución no normalizada D(t). El elemento en la *i*ésima fila y *j*ésima columna de la matriz de distribución especifica la cantidad en la que el *j*ésimo canal difuso de entrada contribuye al *i*ésimo canal difuso de salida. El módulo de expansión de señal difusa adaptativa 420 puede entonces aplicar la matriz de distribución normalizada $\bar{\mathbf{D}}(t)$ a la señal de entrada difusa de N+K canal para generar la señal de salida difusa de M-canal.

Sin embargo, en implementaciones alternativas, el módulo de expansión de señal difusa adaptativa 420 puede recuperar la matriz de distribución normalizada $\bar{\mathbf{D}}(t)$ de una pluralidad almacenada de matrices de distribución normalizada $\bar{\mathbf{D}}(t)$ (p.ej., de una tabla de consulta) en lugar de recomputar la matriz de distribución normalizada $\bar{\mathbf{D}}(t)$ para cada nueva instancia temporal. Por ejemplo, cada una de las matrices de distribución normalizada $\bar{\mathbf{D}}(t)$ puede haberse computado previamente para un valor correspondiente (o rango de valores) de la señal de control $c(t)$.

Como se describe más arriba, la matriz transitoria T se puede computar como una función de C junto con las supuestas ubicaciones espaciales de los canales de entrada y salida. De manera específica, cada elemento de la matriz transitoria puede computarse como una escala del elemento de matriz no transitoria correspondiente. La escala puede, por ejemplo, ser una función de la relación de la ubicación del canal de salida correspondiente respecto a la de los canales de entrada. Mediante el reconocimiento de que el elemento en la *i*ésima fila y *j*ésima columna de la matriz de distribución especifica la cantidad en la que el *j*ésimo canal difuso de entrada contribuye al *i*ésimo canal difuso de salida, cada elemento de la matriz transitoria T se puede computar como

$$T_{ij} = \beta_i C_{ij} \quad (\text{Ecuación 3})$$

En la Ecuación 3, el factor de escala β_i se computa según la ubicación del *i*ésimo canal de la señal de salida de M canal con respecto a las ubicaciones de los N canales de la señal de entrada. En general, para los canales de salida cercanos a los canales de entrada, puede ser deseable que β_i sea cercano a uno. Dado que un canal de salida se vuelve espacialmente más distante de los canales de entrada, puede ser deseable que β_i se vuelva más pequeño.

La Figura 5 muestra ejemplos de factores de escala para una implementación que implica una señal de entrada estéreo y una señal de salida de cinco canales. En el presente ejemplo, los canales de entrada se designan I_e y D_e , y los canales de salida se designan I, D, C, IA y DA. Las supuestas ubicaciones de canal y los valores a modo de ejemplo del factor de escala β_i se ilustran en la Figura 5. Vemos que para los canales de salida I, D y C, los cuales están espacialmente cerca de los canales de entrada I_e y D_e , el factor de escala β_i se ha establecido en uno en el presente ejemplo. Para los canales de salida IA y DA, los cuales se supone que están espacialmente más distantes de los canales de entrada I_e y D_e , el factor de escala β_i se ha establecido en 0,25 en el presente ejemplo.

5 Suponiendo que los canales de entrada I_e y D_e se ubican a menos y más 30 grados del plano mediano 505, entonces según algunas de dichas implementaciones $\beta_i = 0,25$ si el valor absoluto del ángulo del canal de salida del plano mediano 505 es más grande que 45 grados. De lo contrario, $\beta_i = 1$. Este ejemplo provee una estrategia simple para generar los factores de escala. Sin embargo, muchas otras estrategias son posibles. Por ejemplo, en algunas implementaciones, el factor de escala β_i puede tener un valor mínimo diferente y/o puede tener un rango de valores entre los valores mínimo y máximo.

10 La Figura 6 es un diagrama de bloques que muestra detalles adicionales de un procesador de señal difusa según un ejemplo. En la presente implementación, el módulo de expansión de señal difusa adaptativa 420 del procesador de señal difusa 40 incluye un módulo de decorrelador 605 y un módulo de matriz de distribución variable 610. En el presente ejemplo, el módulo de decorrelador 605 puede decorrelacionar N canales de señales de audio difusas y producir K canales de salida sustancialmente ortogonales al módulo de matriz de distribución variable 610. Según su uso en la presente memoria descriptiva, se considera que dos vectores son "sustancialmente ortogonales" entre sí si su producto escalar es menos que el 35% de un producto de sus vectores magnitudes. Ello corresponde a un ángulo entre vectores de alrededor de setenta grados a alrededor de 110 grados.

15 El módulo de matriz de distribución variable 610 puede determinar y aplicar una matriz de distribución variable apropiada, según al menos en parte un valor de señal de control transitoria recibido del módulo de detección de comienzo 415. En algunas implementaciones, el módulo de matriz de distribución variable 610 puede calcular la matriz de distribución variable, según al menos en parte el valor de señal de control transitoria. En implementaciones alternativas, el módulo de matriz de distribución variable 610 puede seleccionar una matriz de distribución variable almacenada, según al menos en parte el valor de señal de control transitoria, y recuperar la matriz de distribución variable seleccionada del dispositivo de memoria.

20 Mientras algunas implementaciones pueden funcionar en forma de banda ancha, puede ser preferible que el módulo de expansión de señal difusa adaptativa 420 funcione en múltiples bandas de frecuencia. De esta manera, se puede permitir que las bandas de frecuencia no asociadas a un transitorio permanezcan distribuidas de forma uniforme a lo largo de todos los canales, y maximizar así la cantidad de involucramiento mientras se preserva el impacto de los transitorios en las bandas de frecuencia apropiadas. Para lograr esto, el sistema de procesamiento de audio 10 puede descomponer la señal de audio de entrada en múltiples bandas de frecuencia.

25 Por ejemplo, el sistema de procesamiento de audio 10 puede aplicar algún tipo de banco de filtros como, por ejemplo, una transformada de Fourier de Tiempo Corto (STFT) o Banco de Filtros Espejo en Cuadratura (QMF). Para cada banda del banco de filtros, una instancia de uno o más componentes del sistema de procesamiento de audio 10 (p.ej., como se muestra en la Figura 4B o Figura 6) puede ejecutarse en paralelo. Por ejemplo, una instancia del módulo de expansión de señal difusa adaptativa 420 puede ejecutarse para cada banda del banco de filtros.

30 Según algunas de dichas implementaciones, el módulo de detección de comienzo 415 puede producir una señal de control transitoria multibanda que indica la naturaleza transitoria de las señales de audio en cada banda de frecuencia. En algunas implementaciones, el módulo de detección de comienzo 415 puede detectar aumentos de energía a lo largo del tiempo en cada banda y generar una señal de control transitoria correspondiente a dichos aumentos de energía. Dicha señal de control se puede generar a partir de la energía variable con el tiempo en cada banda de frecuencia, mezclada de forma descendente a lo largo de todos los canales de entrada. Dejemos que $E(b,t)$ represente esta energía en el tiempo t en la banda de frecuencia b , una versión de tiempo suavizado de dicha energía puede primero computarse usando un suavizador de un polo en un ejemplo:

$$E_s(b, t) = \alpha_s E_s(b, t - 1) + (1 - \alpha_s) E(b, t) \quad (\text{Ecuación 4})$$

35 En un ejemplo, el coeficiente de suavizado α_s se puede elegir para producir un tiempo de media caída de aproximadamente 200ms. Sin embargo, otros valores de coeficiente de suavizado pueden proveer resultados satisfactorios. A continuación, una señal transitoria en bruto $o(b, t)$ se puede computar restando el valor dB de la energía suavizada en un instante de tiempo previo del valor dB de la energía no suavizada en el instante de tiempo actual:

$$o(b, t) = 10 \log_{10}(E(b, t)) - 10 \log_{10}(E_s(b, t - 1)) \quad (\text{Ecuación 5})$$

40 Dicha señal transitoria en bruto puede entonces normalizarse para situarse entre cero y uno mediante el uso de límites de normalización transitorios o_{bajo} y o_{alto} .

$$\bar{o}(b, t) = \begin{cases} 1, & o(b, t) \geq o_{\text{alto}} \\ \frac{o(b, t) - o_{\text{bajo}}}{o_{\text{alto}} - o_{\text{bajo}}}, & o_{\text{bajo}} < o(b, t) < o_{\text{alto}} \\ 0, & o(b, t) \leq o_{\text{bajo}} \end{cases} \quad (\text{Ecuación 6})$$

Se ha descubierto que los valores de $o_{\text{bajo}} = 3\text{dB}$ y $o_{\text{alto}} = 9\text{dB}$ funcionan bien. Sin embargo, otros valores pueden producir resultados aceptables. Finalmente, la señal de control transitoria $c(b, t)$ se puede computar. En un ejemplo, la señal de control transitoria $c(b, t)$ se puede computar mediante el suavizado de la señal transitoria normalizada con un ataque infinito, filtro de suavizado de un polo de liberación lenta:

$$c(b, t) = \begin{cases} \bar{o}(b, t), & \bar{o}(b, t) \geq c(b, t - 1) \\ \alpha_r c(b, t - 1), & \text{de lo contrario} \end{cases} \quad (\text{Ecuación 7})$$

Se ha descubierto que un coeficiente de liberación α_r que produce un tiempo de media caída de aproximadamente 200ms funciona bien. Sin embargo, otros valores de coeficiente de liberación pueden proveer resultados satisfactorios. En el presente ejemplo, la señal de control transitoria $c(b, t)$ resultante de cada banda de frecuencia instantáneamente se eleva a uno cuando la energía en dicha banda muestra un aumento significativo, y luego gradualmente se reduce a cero a medida que la energía de la señal se reduce. La variación proporcional subsiguiente de la matriz de distribución en cada banda produce una modulación que se percibe transparente del campo de sonido difuso, lo cual mantiene tanto el impacto de los transitorios como el envolvimiento total.

A continuación se muestran algunos ejemplos de formación y aplicación de la matriz no transitoria C, así como de métodos y procesos relacionados.

Primer método de derivación

Con referencia, nuevamente, a la Figura 4A, en el presente ejemplo el procesador de señal difusa 40 genera a lo largo del trayecto 49 un conjunto de M señales mezclando los N canales de señales de audio recibidos del trayecto 29 según un sistema de ecuaciones lineales. Para una mejor descripción en la siguiente descripción, se hace referencia a las porciones de los N canales de señales de audio recibidas del trayecto 29 como señales de entrada intermedias y se hace referencia a los M canales de señales intermedias generados a lo largo del trayecto 49 como señales de salida intermedias. Dicha función de mezcla incluye el uso de un sistema de ecuaciones lineales que se puede representar por una multiplicación de matriz, por ejemplo como se muestra más abajo:

$$\vec{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_M \end{bmatrix} = \begin{bmatrix} C_{1,1} & \cdots & C_{1,N+K} \\ \vdots & \ddots & \vdots \\ C_{M,1} & \cdots & C_{M,N+K} \end{bmatrix} \cdot \begin{bmatrix} X_1 \\ \vdots \\ X_{N+K} \end{bmatrix} = C \cdot \vec{X} \quad \text{para } 1 \leq K \leq (M-N) \quad (\text{Ecuación 8})$$

En la Ecuación 8, \vec{X} representa un vector de columna correspondiente a N+K señales obtenidas de las N señales de entrada intermedias; C representa una M x (N+K) matriz o matriz de coeficientes de mezcla; e \vec{Y} representa un vector de columna correspondiente a las M señales de salida intermedias. La función de mezcla se puede llevar a cabo en señales representadas en el dominio temporal o dominio de la frecuencia. La siguiente descripción hace una mención más particular de las implementaciones de dominio temporal.

Como se muestra en la expresión 1, K es mayor que o igual a uno y menor que o igual a la diferencia (M-N). Como resultado, el número de señales X_i y el número de columnas en la matriz C se encuentra entre N+1 y M. Los coeficientes de la matriz C se pueden obtener de un conjunto de N+K vectores de magnitud de unidad en un M-espacio dimensional que son sustancialmente ortogonales entre sí. Según se advierte más arriba, se considera que dos vectores son "sustancialmente ortogonales" entre sí si su producto escalar es menos que el 35% de un producto de sus magnitudes.

Cada columna en la matriz C puede tener M coeficientes que corresponden a los elementos de uno de los vectores en el conjunto. Por ejemplo, los coeficientes que se encuentran en la primera columna de la matriz C corresponden a uno de los vectores V en el conjunto cuyos elementos se denotan como (V_1, \dots, V_M) de modo que $C_{1,1} = p \cdot V_1, \dots, C_{M,1} = p \cdot V_M$, donde p representa un factor de escala usado para escalar los coeficientes de matriz, según se desee. De manera alternativa, los coeficientes en cada columna j de la matriz C se pueden escalar mediante diferentes

factores de escala p_j . En muchas aplicaciones, los coeficientes se escalan de modo que la norma de Frobenius de la matriz es igual a o se encuentra dentro del 10% de \sqrt{N} . Aspectos adicionales de la escala se describen más abajo.

5 El conjunto de N+K vectores se puede derivar de cualquier manera que se desee. Un método crea una M x M matriz G de coeficientes con valores pseudoaleatorios que tienen una distribución gaussiana, y calcula la descomposición del valor singular de dicha matriz para obtener tres M x M matrices denotadas aquí como U, S y V. Las matrices U y V pueden ser, ambas, matrices unitarias. La matriz C se puede obtener seleccionando N+K columnas de la matriz U o de la matriz V y escalando los coeficientes en dichas columnas para lograr una norma de Frobenius igual a o dentro del 10% de \sqrt{N} . Un método que relaja algunos de los requisitos para la ortogonalidad se describe más abajo.

10 La correlación numérica de dos señales se puede calcular usando una variedad de algoritmos numéricos conocidos. Dichos algoritmos producen una medida de correlación numérica llamada un coeficiente de correlación que varía entre uno negativo y uno positivo. Un coeficiente de correlación con una magnitud igual o cercana a uno indica que dos señales se relacionan de manera cercana. Un coeficiente de correlación con una magnitud igual o cercana a cero indica que dos señales son, en general, independientes entre sí.

15 Las N+K señales de entrada se pueden obtener decorrelacionando las N señales de entrada intermedias una con respecto a la otra. En algunas implementaciones, la decorrelación puede ser lo que en la presente memoria se llama una "decorrelación psicoacústica", la cual se describe brevemente más arriba. La decorrelación psicoacústica es menos estricta que la decorrelación numérica en que dos señales se pueden considerar psicoacústicamente decorrelacionadas incluso si tienen algún grado de correlación numérica entre sí.

20 La decorrelación psicoacústica se puede lograr usando retardos u otros tipos de filtros, algunos de los cuales se describen más abajo. En muchas implementaciones, N de las N+K señales X_i se pueden tomar directamente de las N señales de entrada intermedias sin usar retardos o filtros para lograr la decorrelación psicoacústica ya que dichas N señales representan un campo de sonido difuso y se decorrelacionarán, probablemente, de forma psicoacústica.

Segundo método de derivación

25 Si las señales generadas por el procesador de señal difusa 40 se combinan con otras señales que representan un campo de sonido no difuso según el primer método de derivación descrito más arriba, la combinación de señales resultante puede, algunas veces, generar artefactos no deseables. En algunas instancias, dichos artefactos pueden resultar porque el diseño de la matriz C no representa, de forma adecuada, posibles interacciones entre las porciones difusas y no difusas de un campo de sonido. Como se describe más arriba, la distinción entre difuso y no difuso no siempre se encuentra definida. Por ejemplo, con referencia a la Figura 4A, el analizador de señal de entrada 20 puede generar algunas señales a lo largo del trayecto 28 que representan, en cierto grado, un campo de sonido difuso y puede generar señales a lo largo del trayecto 29 que representan un campo de sonido no difuso en cierto grado. Si el generador de señal difusa 40 destruye o modifica el carácter no difuso del campo de sonido representado por las señales en el trayecto 29, pueden ocurrir artefactos no deseables o distorsiones audibles en el campo de sonido que se produce a partir de las señales de salida generadas a lo largo del trayecto 59. Por ejemplo, si la suma de las M señales procesadas difusas en el trayecto 49 con las M señales procesadas no difusas en el trayecto 39 provoca la cancelación de algunos componentes de señal no difusa, ello puede degradar la impresión subjetiva que, de lo contrario, se lograría.

40 Una mejora se puede lograr diseñando la matriz C para que represente la naturaleza no difusa del campo de sonido que se procesa por el procesador de señal no difusa 30. Ello se puede llevar a cabo identificando primero una matriz E que representa, o se supone que representa, el procesamiento de codificación que procesa M canales de señales de audio para crear los N canales de señales de audio de entrada recibidos del trayecto 19, y luego derivando una inversa de dicha matriz, p.ej., como se describe más abajo.

45 Un ejemplo de una matriz E es una 5 x 2 matriz que se usa para submezclar cinco canales, I, C, D, IA, DA, en dos canales denotados como izquierdo-total (I_T) y derecha total (D_T). Las señales para los canales I_T y D_T son un ejemplo de las señales de audio de entrada para dos (N=2) canales que se reciben del trayecto 19. En el presente ejemplo, el dispositivo 10 se puede usar para sintetizar cinco (M=5) canales de señales de audio de salida que pueden crear un campo de sonido que se percibe similar al (si no sustancialmente idéntico al) campo de sonido que se podría haber creado a partir de las cinco señales de audio originales.

50 Un ejemplo de una 5 x 2 matriz E que se puede usar para codificar las señales de canal I_T y D_T de las señales de canal I, C, D, IA y DA se muestra en la siguiente expresión:

$$E = \begin{bmatrix} 1 & \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{3}}{2} & \frac{-1}{2} \\ 0 & \frac{\sqrt{2}}{2} & 1 & \frac{-1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \quad (\text{Ecuación 9})$$

Una M x N matriz pseudoinversa B se puede derivar de la N x M matriz E mediante el uso de técnicas numéricas conocidas como, por ejemplo, aquellas implementadas en software numérico como, por ejemplo, la función "pinv" en Matlab®, disponible en The MathWorks™, Natick, Massachusetts, o la función "Pseudoinverse" en Mathematica®, disponible en Wolfram Research, Champaign, Illinois. La matriz B puede no ser óptima si sus coeficientes crean una diafonía no deseada entre cualquiera de los canales, o si los coeficientes son números imaginarios o complejos. La matriz B se puede modificar para eliminar dichas características no deseables. La matriz B también se puede modificar para lograr una variedad de efectos artísticos deseados cambiando los coeficientes para enfatizar las señales para los altavoces seleccionados. Por ejemplo, los coeficientes se pueden cambiar para aumentar la energía en las señales destinadas a la reproducción a través de altavoces para los canales izquierdo y derecho y para reducir la energía en señales destinadas a la reproducción a través de los altavoces para el canal central. Los coeficientes en la matriz B se pueden escalar de modo que cada columna de la matriz representa un vector de magnitud de unidad en un espacio M dimensional. Los vectores representados por las columnas de la matriz B no necesitan ser sustancialmente ortogonales entre sí.

5

10

15 Un ejemplo de una 5 x 2 matriz B se muestra en la siguiente expresión:

$$B = \begin{bmatrix} 0,65 & 0 \\ 0,40 & 0,40 \\ 0 & 0,65 \\ 0,60 & -0,24 \\ -0,24 & 0,60 \end{bmatrix} \quad (\text{Ecuación 10})$$

Una matriz como, por ejemplo, la de la Ecuación 10 se puede usar para generar un conjunto de M señales de salida intermedias a partir de las N señales de entrada intermedias por la siguiente función:

$$\vec{Y} = B \cdot \vec{X} \quad (\text{Ecuación 11})$$

20 La Figura 7 es un diagrama de bloques de un aparato que puede generar un conjunto de M señales de salida intermedias a partir de N señales de entrada intermedias. El mezclador ascendente 41 puede, por ejemplo, ser un componente del procesador de señal difusa 40, p.ej. como se muestra en la Figura 4A. En el presente ejemplo, el mezclador ascendente 41 recibe las N señales de entrada intermedias de los trayectos de señal 29-1 y 29-2 y mezcla dichas señales según un sistema de ecuaciones lineales para generar un conjunto de M señales de salida intermedias a lo largo de los trayectos de señal 49-1 y 49-5. Las cajas dentro del mezclador ascendente 41 representan la multiplicación o amplificación de señales por coeficientes de la matriz B según el sistema de ecuaciones lineales.

25

Aunque la matriz B se puede usar sola, el rendimiento se puede mejorar usando una M x K matriz de aumento adicional A, donde $1 \leq K \leq (M-N)$. Cada columna en la matriz A puede representar un vector de magnitud de unidad en un espacio M dimensional que es sustancialmente ortogonal a los vectores representados por las N columnas de la matriz B. Si K es mayor que uno, cada columna puede representar un vector que también es sustancialmente ortogonal a los vectores representados por todas las otras columnas en la matriz A.

30

Los vectores para las columnas de la matriz A se pueden derivar en una variedad de formas. Por ejemplo, se pueden usar las técnicas mencionadas más arriba. Otros métodos implican escalar los coeficientes de la matriz de

aumento A y la matriz B, p.ej., como se explica más abajo, y concatenar los coeficientes para producir la matriz C. En un ejemplo, la escala y concatenación se pueden expresar de forma algebraica como:

$$C = \left[\beta \cdot B \mid \alpha \cdot A \right] \quad (\text{Ecuación 12})$$

5 En la Ecuación 12, "|" representa una concatenación horizontal de las columnas de la matriz B y matriz A, α representa un factor de escala para los coeficientes de la matriz A, y β representa un factor de escala para los coeficientes de la matriz B.

En algunas implementaciones, los factores de escala α y β se pueden elegir de modo que la norma de Frobenius de la matriz compuesta C es igual a o se encuentra dentro del 10% de la norma de Frobenius de la matriz B. La norma de Frobenius de la matriz C se puede expresar como:

$$\|C\|_F = \sqrt{\sum_i \sum_j |c_{ij}|^2} \quad (\text{Ecuación 13})$$

10

En la Ecuación 13, c_{ij} representa el coeficiente de matriz en la fila i y columna j .

Si cada una de las N columnas en la matriz B y cada una de las K columnas en la matriz A representa un vector de magnitud de unidad, la norma de Frobenius de la matriz B es igual a \sqrt{N} y la norma de Frobenius de la matriz A es igual a \sqrt{K} . Para el presente caso, se puede mostrar que si la norma de Frobenius de la matriz C se establece igual a \sqrt{N} , entonces los valores para los factores de escala α y β se relacionan entre sí como se muestra en la siguiente expresión:

15

$$\alpha = \sqrt{\frac{N \cdot (1 - \beta^2)}{K}} \quad (\text{Ecuación 14})$$

20

Después de establecer el valor del factor de escala β , el valor para el factor de escala α se puede calcular a partir de la Ecuación 14. En algunas implementaciones, el factor de escala β se puede seleccionar de modo que las señales mezcladas por los coeficientes en las columnas de la matriz B se dan al menos con 5 dB de mayor peso que las señales mezcladas por coeficientes en las columnas de la matriz de aumento A. Una diferencia en el peso de al menos 6 dB se puede lograr limitando los factores de escala de modo que $\alpha < \frac{1}{2} \beta$. Diferencias mayores o menores en el peso de escala para las columnas de la matriz B y la matriz A se pueden usar para lograr un balance acústico deseado entre los canales de audio.

25

De manera alternativa, los coeficientes en cada columna de la matriz de aumento A se pueden escalar de forma individual como se muestra en la siguiente expresión:

$$C = \left[\beta \cdot B \mid \alpha_1 \cdot A_1 \quad \alpha_2 \cdot A_2 \quad \cdots \quad \alpha_K \cdot A_K \right] \quad (\text{Ecuación 15})$$

30

En la Ecuación 15, A_j representa la columna j de la matriz de aumento A y α_j representa el factor de escala respectivo para la columna j . Para dicha alternativa, podemos elegir valores arbitrarios para cada factor de escala α_j , siempre que cada factor de escala satisfaga el límite $\alpha_j < \frac{1}{2} \beta$. En algunas implementaciones, los valores de los coeficientes α_j y β se eligen para asegurar que la norma de Frobenius de C sea aproximadamente igual a la norma de Frobenius de la matriz B.

35

Cada una de las señales que se mezclan según la matriz de aumento A se puede procesar de modo que se decorrelacionan de forma psicoacústica de las N señales de entrada intermedias y de todas las otras señales que se mezclan según la matriz de aumento A. La Figura 8 es un diagrama de bloques que muestra un ejemplo de decorrelación de señales intermedias seleccionadas. En el presente ejemplo, dos (N=2) señales de entrada intermedias, cinco (M=5) señales de salida intermedias y tres (K=3) señales decorrelacionadas se mezclan según la

matriz de aumento A. En el ejemplo que se muestra en la Figura 8, las dos señales de entrada intermedias se mezclan según la matriz inversa básica B, representada por el bloque 41. Las dos señales de entrada intermedias se decorrelacionan por el decorrelador 43 para proveer tres señales decorrelacionadas que se mezclan según la matriz de aumento A, lo cual se representa por el bloque 42.

5 El decorrelador 43 se puede implementar en una variedad de formas. La Figura 9 es un diagrama de bloques que muestra un ejemplo de componentes del decorrelador. La implementación que se muestra en la Figura 9 puede lograr la decorrelación psicoacústica mediante un retardo de las señales de entrada en cantidades variables. Los retardos en el rango de uno a veinte milisegundos son apropiados para muchas aplicaciones.

10 La Figura 10 es un diagrama de bloques que muestra un ejemplo alternativo de componentes del decorrelador. En el presente ejemplo, se procesa una de las señales de entrada intermedias. Una señal de entrada intermedia se distribuye en dos trayectos de procesamiento de señal diferentes que aplican filtros a sus respectivas señales en dos subbandas de frecuencia superpuestas. El trayecto de frecuencia más baja incluye un filtro de cambio de fase 61 que filtra su señal de entrada en una primera subbanda de frecuencia según una primera respuesta al impulso y un filtro paso bajo 62 que define la primera subbanda de frecuencia. El trayecto de frecuencia más alta incluye un retardo dependiente de la frecuencia 63 implementado por un filtro que filtra su señal de entrada en una segunda subbanda de frecuencia según una segunda respuesta al impulso que no es igual a la primera respuesta al impulso, un filtro paso alto 64 que define la segunda subbanda de frecuencia y un componente de retardo 65. Las salidas del retardo 65 y del filtro paso bajo 62 se combinan en el nodo de suma 66. La salida del nodo de suma 66 es una señal que se decorrelaciona de forma psicoacústica con respecto a la señal de entrada intermedia.

20 La respuesta de fase del filtro de cambio de fase 61 puede ser dependiente de la frecuencia y puede tener una distribución bimodal en la frecuencia con picos sustancialmente iguales a noventa grados positivos y negativos. Una implementación ideal del filtro de cambio de fase 61 tiene una respuesta de magnitud de unidad y una respuesta de fase que alterna o cambia entre noventa grados positivos y noventa grados negativos en los bordes de dos o más bandas de frecuencia dentro de la banda de paso del filtro. Un cambio de fase se puede implementar por una transformada dispersa de Hilbert que tiene una respuesta al impulso que se muestra en la siguiente expresión:

$$H_s(k) = \begin{cases} 2/k'\pi & \{ \text{impar } k' = k/S \} \\ 0 & \{ \text{de lo contrario} \} \end{cases} \quad (\text{Ecuación 16})$$

30 La respuesta al impulso de la transformada dispersa de Hilbert se trunca, preferiblemente, para una longitud seleccionada para optimizar el rendimiento del decorrelador mediante un equilibrio de una compensación entre el rendimiento transitorio y la suavidad de la respuesta de la frecuencia. La cantidad de cambios de fase se puede controlar por el valor del parámetro S. Dicho parámetro se debe elegir para equilibrar una compensación entre el grado de decorrelación y la longitud de respuesta al impulso. Una respuesta al impulso más larga se puede requerir a medida que el valor del parámetro S aumenta. Si el valor del parámetro S es demasiado pequeño, el filtro puede proveer una decorrelación insuficiente. Si el parámetro S es demasiado grande, el filtro puede marcar los sonidos transitorios durante un intervalo de tiempo lo suficientemente largo para crear artefactos objetables en la señal decorrelacionada.

La capacidad de equilibrar dichas características se puede mejorar mediante la implementación del filtro de cambio de fase 21 para tener un espaciado no uniforme en la frecuencia entre los cambios de fase adyacentes, con un espaciado más estrecho en frecuencias más bajas y un espaciado más amplio en frecuencias más altas. En algunas implementaciones, el espaciado entre los cambios de fase adyacentes es una función logarítmica de frecuencia.

40 El retardo dependiente de la frecuencia 63 se puede implementar por un filtro que tiene una respuesta al impulso igual a una secuencia sinusoidal de longitud finita $h[n]$ cuya frecuencia instantánea se reduce de manera monótona de π a cero en la duración de la secuencia. Dicha secuencia se puede expresar como:

$$h[n] = G \sqrt{|\omega'(n)|} \cos(\phi(n)), \text{ para } 0 \leq n < L \quad (\text{Ecuación 17})$$

45 En la Ecuación 17, $\omega(n)$ representa la frecuencia instantánea, $\omega'(n)$ representa la primera derivada de la frecuencia instantánea, G representa un factor de normalización, $\phi(n) = \int_0^n \omega(t) dt$ representa una fase instantánea, y L representa la longitud del filtro de retardo. En algunos ejemplos, el factor de normalización G se puede establecer en un valor de modo que:

$$\sum_{n=0}^{L-1} h^2 [n] = 1 \quad (\text{Ecuación 18})$$

Un filtro con dicha respuesta al impulso puede, a veces, generar artefactos "con chirrido" cuando se aplica a señales de audio con transitorios. Dicho efecto se puede reducir agregando un término tipo ruido al término de fase instantánea como se muestra en la siguiente expresión:

$$h[n] = G \sqrt{|\omega'(n)|} \cos(\phi(n) + N(n)), \text{ para } 0 \leq n < L \quad (\text{Ecuación 19})$$

Si el término tipo ruido es una secuencia de ruido blanco gaussiano con una varianza que es una fracción pequeña de π , los artefactos que se generan filtrando los transitorios que sonarán más como ruido que como chirridos y la relación deseada entre el retardo y la frecuencia se puede lograr aún.

Las frecuencias de corte del filtro paso bajo 62 y del filtro paso alto 64 se pueden elegir para que sean de aproximadamente 2,5 kHz, de modo que no exista ningún intervalo entre las bandas de paso de los dos filtros y de modo que la energía espectral de sus salidas combinadas en la región cercana a la frecuencia de corte donde las bandas de paso se superponen sea sustancialmente igual a la energía espectral de la señal de entrada intermedia en dicha región. La cantidad de retardo impuesta por el retardo 65 se puede establecer de modo que el retardo de propagación de los trayectos de procesamiento de señal de frecuencia más alta y frecuencia más baja sea aproximadamente igual a la frecuencia de corte.

El decorrelador se puede implementar de diferentes maneras. Por ejemplo, uno o ambos del filtro paso bajo 62 y filtro paso alto 64 pueden preceder al filtro de cambio de fase 61 y al retardo dependiente de la frecuencia 63, respectivamente. El retardo 65 se puede implementar por uno o más componentes de retardo ubicados en los trayectos de procesamiento de señal, según se desee.

La Figura 11 es un diagrama de bloques que provee ejemplos de componentes de un sistema de procesamiento de audio. En el presente ejemplo, el sistema de procesamiento de audio 1100 incluye un sistema de interfaz 1105. El sistema de interfaz 1105 puede incluir una interfaz de red como, por ejemplo, una interfaz de red inalámbrica. De manera alternativa, o adicional, el sistema de interfaz 1105 puede incluir una interfaz de bus serial universal (USB) u otra interfaz.

El sistema de procesamiento de audio 1100 incluye un sistema lógico 1110. El sistema lógico 1110 puede incluir un procesador como, por ejemplo, un procesador de propósito general de un solo chip o multichip. El sistema lógico 1110 puede incluir un procesador digital de señales (DSP), un circuito integrado de aplicación específica (ASIC), una matriz de puertas programables en campo (FPGA) u otro dispositivo de lógica programable, puerta discreta o lógica de transistor, o componentes de hardware discretos, o combinaciones de ellos. El sistema lógico 1110 se puede configurar para controlar los otros componentes del sistema de procesamiento de audio 1100. Aunque en la Figura 11 no se muestran interfaces entre los componentes del sistema de procesamiento de audio 1100, el sistema lógico 1110 se puede configurar con interfaces para la comunicación con los otros componentes. Los otros componentes pueden o no configurarse para la comunicación entre sí, según corresponda.

El sistema lógico 1110 se puede configurar para llevar a cabo la funcionalidad de procesamiento de audio, incluidos, pero sin limitación, los tipos de funcionalidad descritos en la presente memoria descriptiva. En algunas de dichas implementaciones, el sistema lógico 1110 se puede configurar para funcionar (al menos en parte) según el software almacenado en uno o más medios no transitorios. Los medios no transitorios pueden incluir memoria asociada al sistema lógico 1110 como, por ejemplo, memoria de acceso aleatorio (RAM, por sus siglas en inglés) y/o memoria de solo lectura (ROM, por sus siglas en inglés). Los medios no transitorios pueden incluir memoria del sistema de memoria 1115. El sistema de memoria 1115 puede incluir uno o más tipos apropiados de medios de almacenamiento no transitorios como, por ejemplo, memoria flash, un disco duro, etc.

El sistema de visualización 1130 puede incluir uno o más tipos apropiados de visualización, según la manifestación del sistema de procesamiento de audio 1100. Por ejemplo, el sistema de visualización 1130 puede incluir una pantalla de cristal líquido, una pantalla plasma, una pantalla biestable, etc.

El sistema de entrada de usuario 1135 puede incluir uno o más dispositivos configurados para aceptar la entrada de un usuario. En algunas implementaciones, el sistema de entrada de usuario 1135 puede incluir una pantalla táctil que cubre una pantalla del sistema de visualización 1130. El sistema de entrada de usuario 1135 puede incluir un ratón, un control del cursor, un sistema de detección de gestos, un *joystick*, una o más GUI (interfaz gráfica de usuario) y/o menús presentados en el sistema de visualización 1130, botones, un teclado, interruptores, etc. En algunas implementaciones, el sistema de entrada de usuario 1135 puede incluir el micrófono 1125: un usuario puede

proveer comandos de voz al sistema de procesamiento de audio 1100 mediante el micrófono 1125. El sistema lógico se puede configurar para el reconocimiento de voz y para controlar al menos algunas funciones del sistema de procesamiento de audio 1100 según dichos comandos de voz. En algunas implementaciones, el sistema de entrada de usuario 1135 se puede considerar una interfaz de usuario y, por lo tanto, como parte del sistema de interfaz 1105.

- 5 El sistema de energía 1140 puede incluir uno o más dispositivos de almacenamiento de energía apropiados como, por ejemplo, una batería de níquel-cadmio o una batería de litio-ion. El sistema de energía 1140 se puede configurar para recibir energía de una salida eléctrica.

- 10 Varias modificaciones a las implementaciones descritas en la presente descripción pueden ser aparentes para aquellos con experiencia ordinaria en la técnica. Los principios generales definidos en la presente memoria se pueden aplicar a otras implementaciones sin apartarse del alcance de la presente memoria descriptiva. Por consiguiente, las reivindicaciones no pretenden limitar las implementaciones que se muestran en la presente memoria descriptiva, sino que se les otorgará el más amplio alcance congruente con la presente descripción, los principios y las características innovadoras descritas en la presente memoria descriptiva.

REIVINDICACIONES

1. Un método para derivar M señales de audio difusas de N señales de audio para la presentación de un campo de sonido difuso, en donde M es mayor que N y es mayor que 2, y en donde el método comprende:
- 5 recibir las N señales de audio, en donde cada una de las N señales de audio corresponde a una ubicación espacial (305);
- derivar porciones difusas de las N señales de audio (310);
- detectar instancias de condiciones de señal de audio transitoria (315) en las N señales de audio; y
- 10 procesar las porciones difusas de las N señales de audio para derivar las M señales de audio difusas, en donde durante instancias de condiciones de señal de audio transitoria, el procesamiento comprende distribuir las porciones difusas de las N señales de audio en mayor proporción a una o más de las M señales de audio difusas correspondientes a ubicaciones espaciales relativamente más cercanas a las ubicaciones espaciales de las N señales de audio y en menor proporción a una o más de las M señales de audio difusas correspondientes a ubicaciones espaciales relativamente más lejanas respecto a las ubicaciones espaciales de las N señales de audio (320).
- 15 2. El método de la reivindicación 1, que además comprende detectar instancias de condiciones de señal de audio no transitoria, en donde durante instancias de condiciones de señal de audio no transitoria el procesamiento implica distribuir las porciones difusas de las N señales de audio a las M señales de audio difusas de manera sustancialmente uniforme.
- 20 3. El método de la reivindicación 2, en donde el procesamiento implica aplicar una matriz de mezcla a las porciones difusas de las N señales de audio para derivar las M señales de audio difusas.
4. El método de la reivindicación 3, en donde la matriz de mezcla es una matriz de distribución variable (D(t)) que se deriva de una matriz no transitoria más apropiada para usar durante condiciones de señal de audio no transitoria y una matriz transitoria más apropiada para usar durante condiciones de señal de audio transitoria.
- 25 5. El método de la reivindicación 4, que además comprende determinar un valor de señal de control transitoria, en donde la matriz de distribución variable se deriva interpolando entre la matriz transitoria y la matriz no transitoria según, al menos en parte, el valor de señal de control transitoria.
6. El método de la reivindicación 5, en donde el valor de señal de control transitoria es variable con el tiempo, puede variar de manera continua de un valor mínimo a un valor máximo, o puede variar en un rango de valores discretos de un valor mínimo a un valor máximo.
- 30 7. El método de cualquiera de las reivindicaciones 5-6, en donde determinar la matriz de distribución variable implica computar la matriz de distribución variable según el valor de señal de control transitoria, o recuperar una matriz de distribución variable almacenada de un dispositivo de memoria.
8. El método de cualquiera de las reivindicaciones 1-7, en donde el método además comprende:
- 35 derivar K señales intermedias de las porciones difusas de las N señales de audio de modo que cada señal de audio intermedia se decorrelaciona de forma psicoacústica con las porciones difusas de las N señales de audio y, si K es mayor que uno, se decorrelaciona de forma psicoacústica con todas las otras señales de audio intermedias, en donde K es mayor que o igual a uno y es menor que o igual a M-N, en donde derivar las K señales intermedias implica, de forma opcional, un proceso de decorrelación que incluye uno o más de retardos, filtros de todos los pasos, filtros pseudoaleatorios o algoritmos de reverberación, y/o en donde las M señales de audio difusas se
- 40 derivan, de forma opcional, en respuesta a las K señales intermedias así como a las N señales difusas.
9. Un aparato, que comprende:
- un sistema de interfaz (1105); y
- un sistema lógico (1110) que puede:
- 45 recibir, mediante el sistema de interfaz, N señales de audio de entrada, en donde cada una de las N señales de audio corresponde a una ubicación espacial (305);
- derivar porciones difusas de las N señales de audio (310);
- detectar instancias de condiciones de señal de audio transitoria (315) en las N señales de audio de entrada; y

- 5 procesar las porciones difusas de las N señales de audio para derivar M señales de audio difusas, en donde M es mayor que N y es mayor que 2, y en donde durante instancias de condiciones de señal de audio transitoria, el procesamiento comprende distribuir las porciones difusas de las N señales de audio en mayor proporción a una o más de las M señales de audio difusas correspondientes a ubicaciones espaciales relativamente más cercanas a las ubicaciones espaciales de las N señales de audio y en menor proporción a una o más de las M señales de audio difusas correspondientes a ubicaciones espaciales relativamente más lejanas respecto a las ubicaciones espaciales de las N señales de audio (320).
- 10 10. El aparato de la reivindicación 9, en donde el sistema lógico puede detectar instancias de condiciones de señal de audio no transitoria y en donde durante instancias de condiciones de señal de audio no transitoria el procesamiento implica distribuir las porciones difusas de las N señales de audio a las M señales de audio difusas de manera sustancialmente uniforme.
11. El aparato de la reivindicación 10, en donde el procesamiento implica aplicar una matriz de mezcla a las porciones difusas de las N señales de audio para derivar las M señales de audio difusas.
- 15 12. El aparato de la reivindicación 11, en donde la matriz de mezcla es una matriz de distribución variable que se deriva de una matriz no transitoria más apropiada para usar durante condiciones de señal de audio no transitoria y una matriz transitoria más apropiada para usar durante condiciones de señal de audio transitoria.
13. El aparato de la reivindicación 12, en donde la matriz transitoria se deriva de la matriz no transitoria, en donde cada elemento de la matriz transitoria representa una escala de un elemento correspondiente de la matriz no transitoria, y en donde la escala es, de manera opcional, una función de una relación entre una ubicación de canal de entrada y una ubicación de canal de salida.
- 20 14. El aparato de cualquiera de las reivindicaciones 12-13, en donde el sistema lógico puede determinar un valor de señal de control transitoria, en donde la matriz de distribución variable se deriva interpolando entre la matriz transitoria y la matriz no transitoria según, al menos en parte, el valor de señal de control transitoria.
15. El aparato de cualquiera de las reivindicaciones 9-14, en donde el sistema lógico puede:
- 25 transformar cada una de las N señales de audio en B bandas de frecuencia; y
- llevar a cabo la derivación, detección y procesamiento, de forma separada, para cada una de las B bandas de frecuencia, en donde el sistema lógico puede, de manera opcional:
- tomar una panorámica de las porciones no difusas de las N señales de audio de entrada para formar M señales de audio no difusas; y
- 30 combinar las M señales de audio difusas con las M señales de audio no difusas para formar M señales de audio de salida.
16. Un medio no transitorio que tiene software allí almacenado, el software incluye instrucciones para controlar al menos un aparato para:
- 35 recibir N señales de audio de entrada, en donde cada una de las N señales de audio corresponde a una ubicación espacial (305);
- derivar porciones difusas de las N señales de audio (310);
- detectar instancias de condiciones de señal de audio transitoria (315) en las N señales de audio de entrada; y
- 40 procesar las porciones difusas de las N señales de audio para derivar M señales de audio difusas, en donde M es mayor que N y es mayor que 2, y en donde durante instancias de condiciones de señal de audio transitoria, el procesamiento comprende distribuir las porciones difusas de las N señales de audio en mayor proporción a una o más de las M señales de audio difusas correspondientes a ubicaciones espaciales relativamente más cercanas a las ubicaciones espaciales de las N señales de audio y en menor proporción a una o más de las M señales de audio difusas correspondientes a ubicaciones espaciales relativamente más lejanas respecto a las ubicaciones espaciales de las N señales de audio (320).

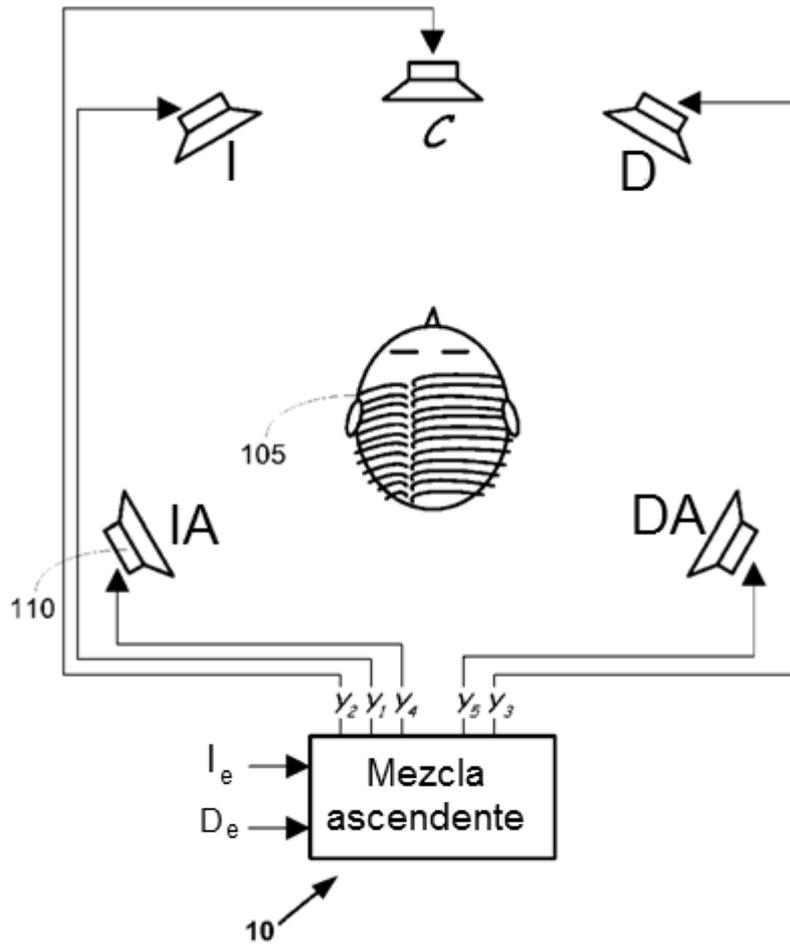


Figura 1

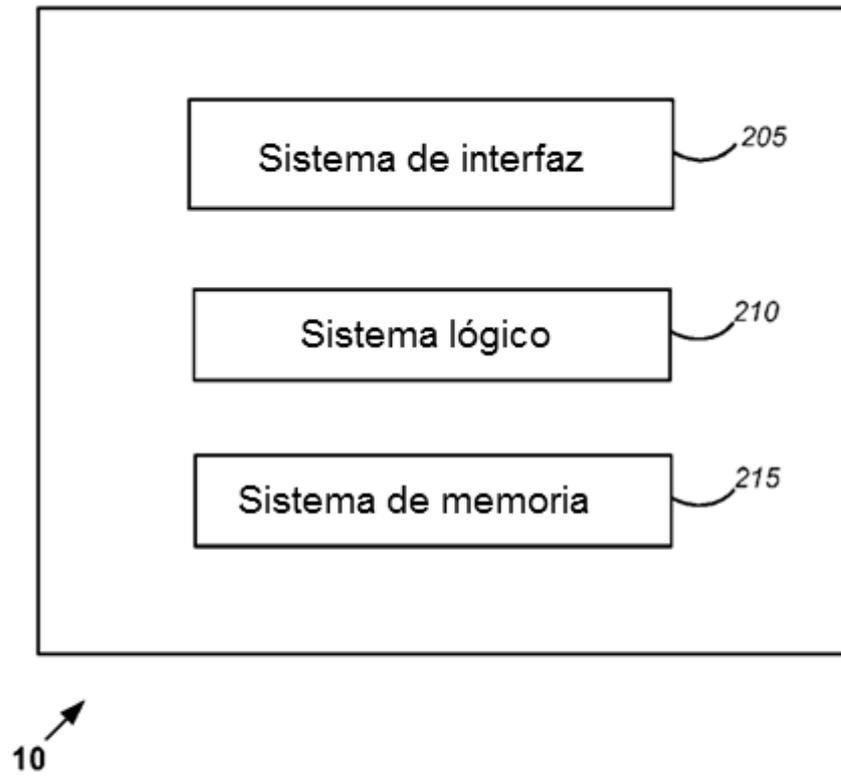
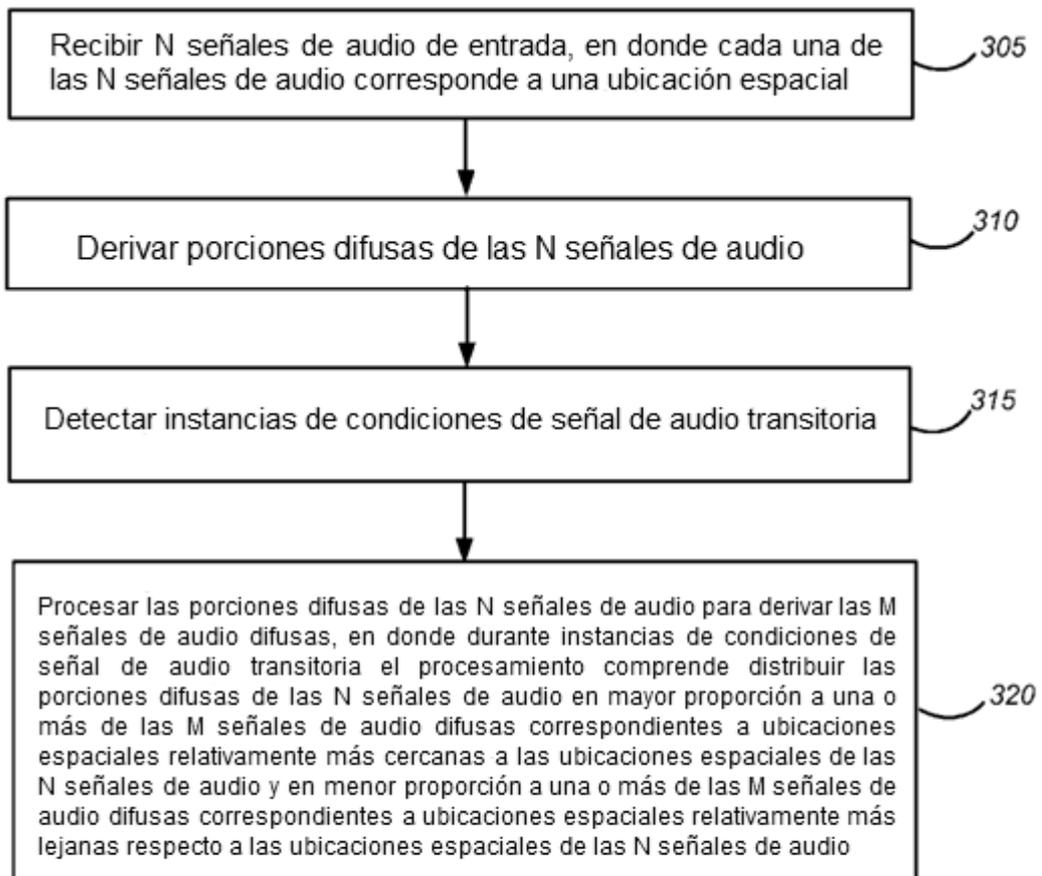


Figura 2



300 ↗

Figura 3

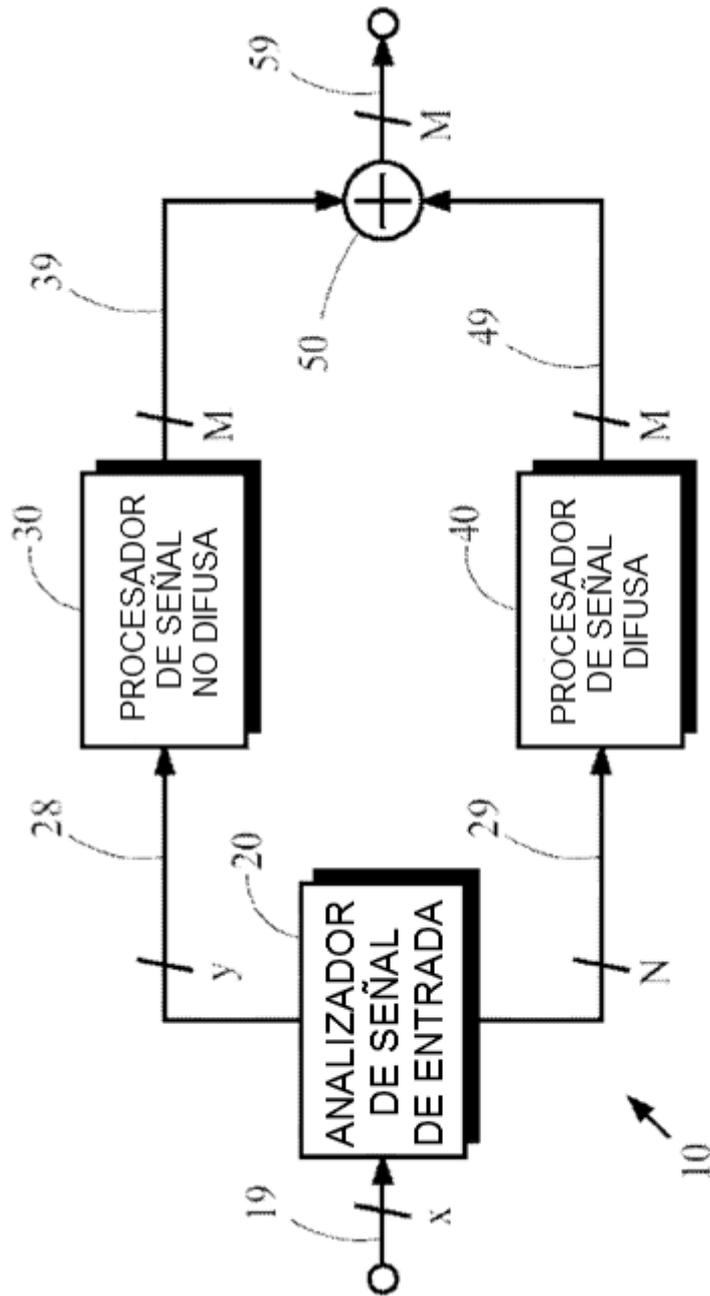


Figura 4A

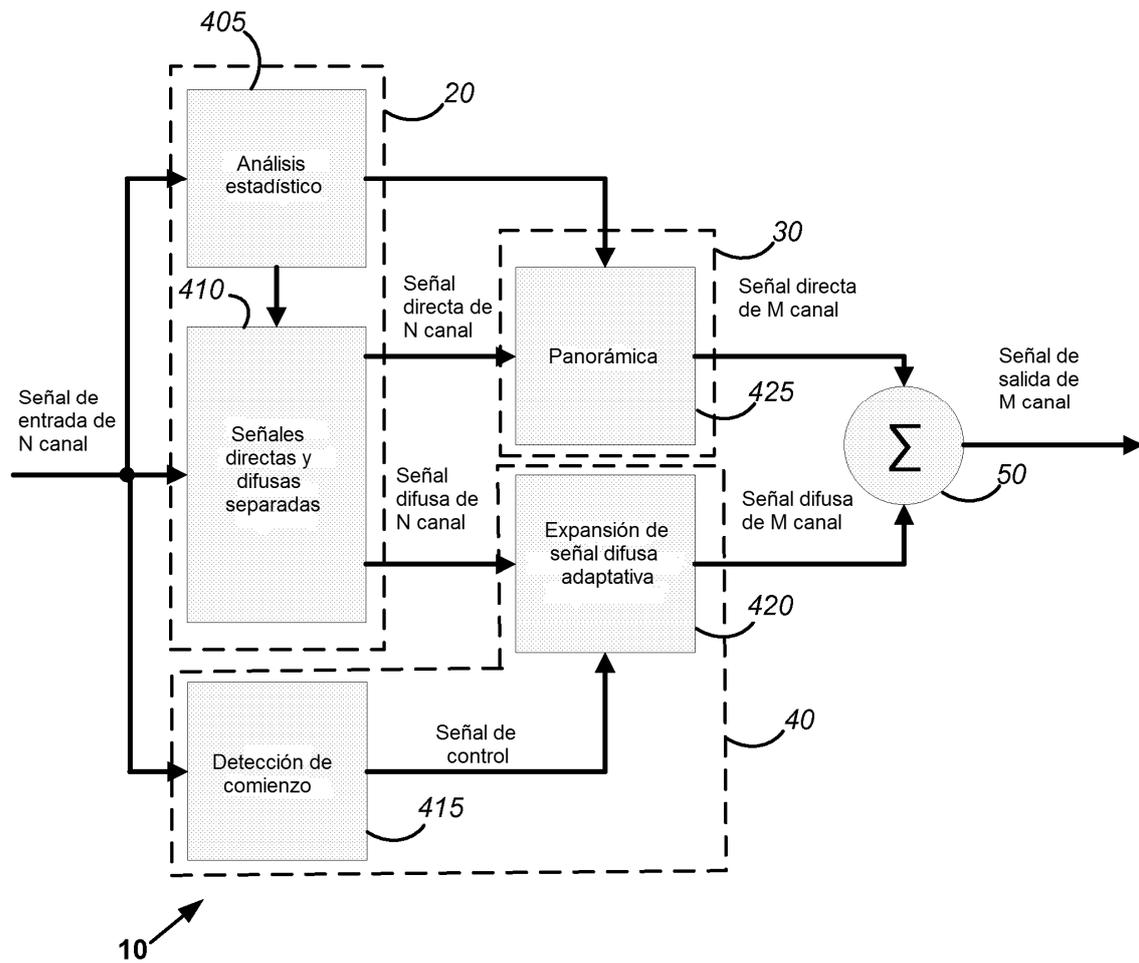


Figura 4B

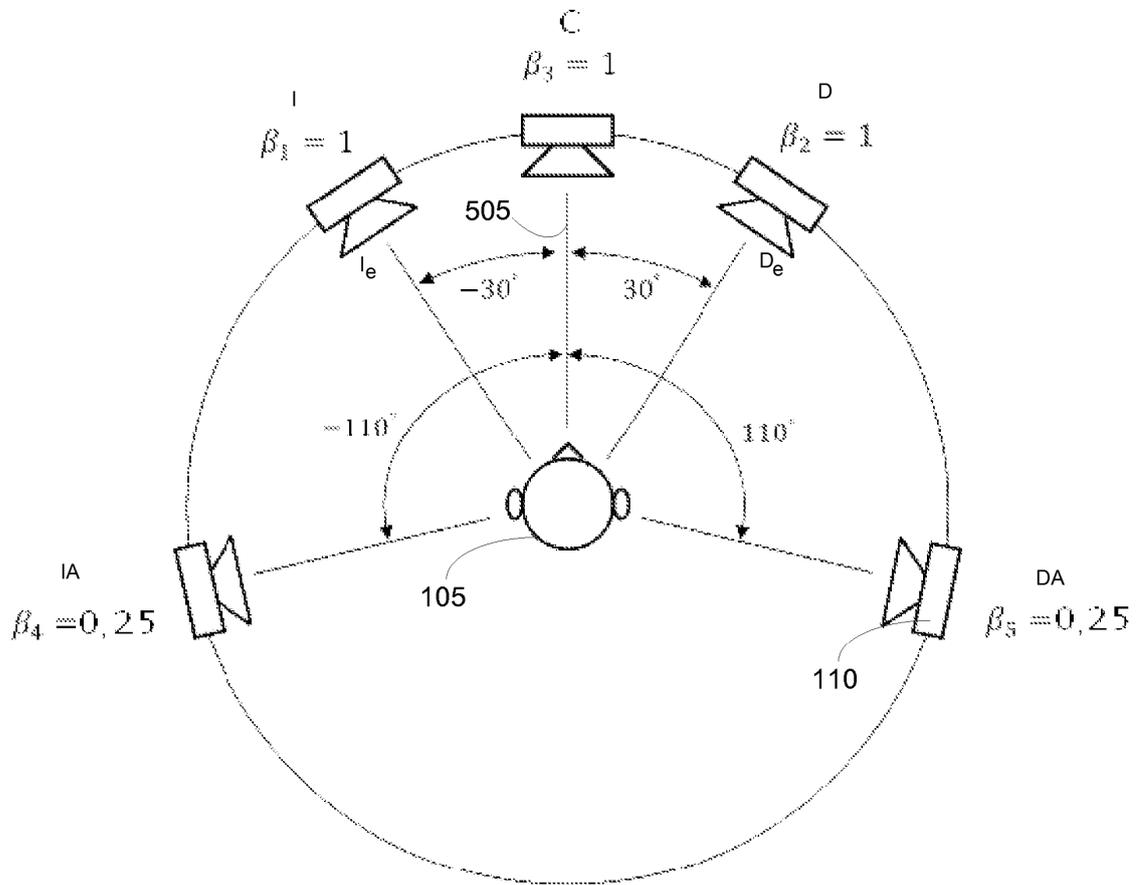


Figura 5

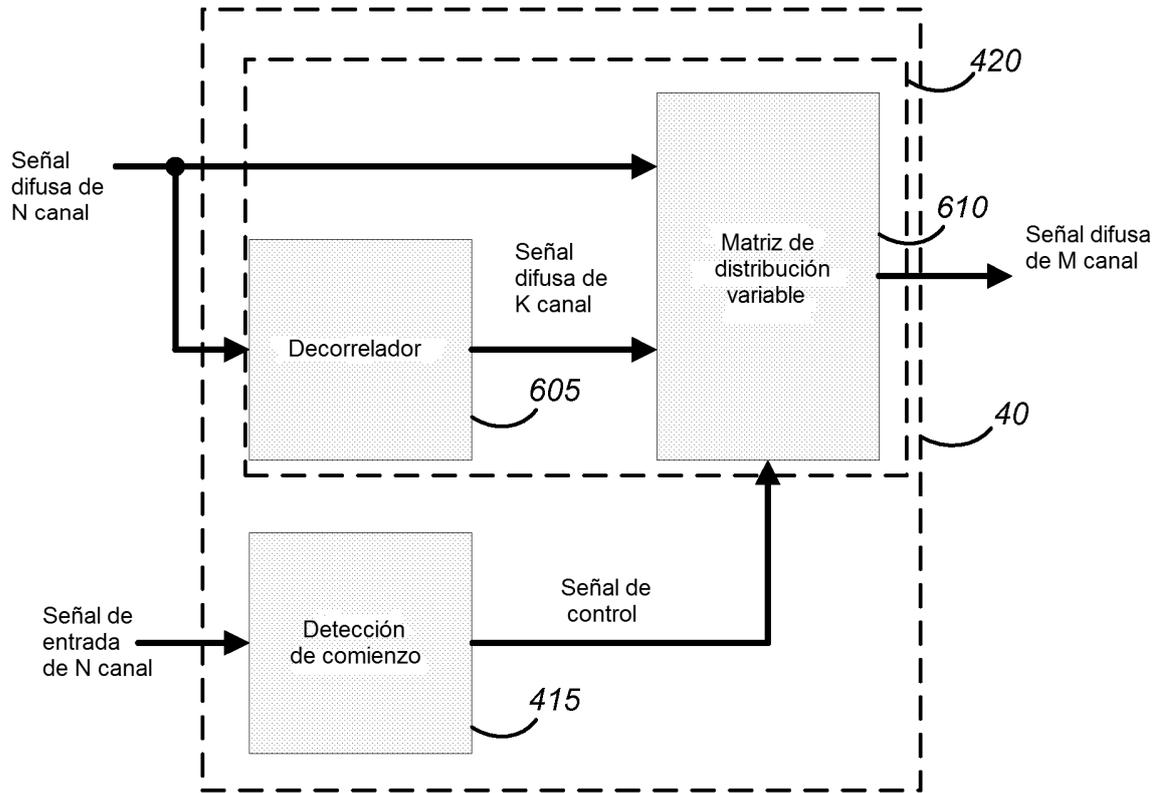


Figura 6

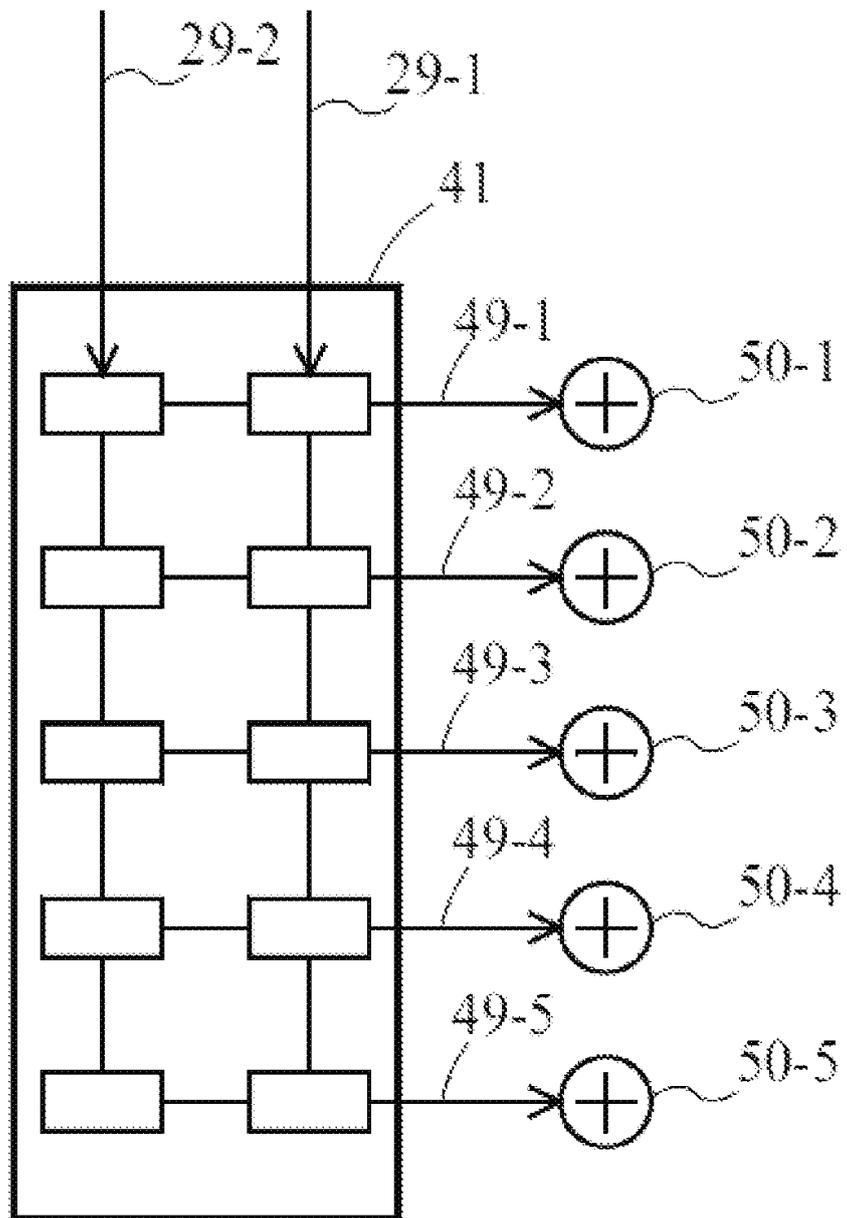


Figura 7

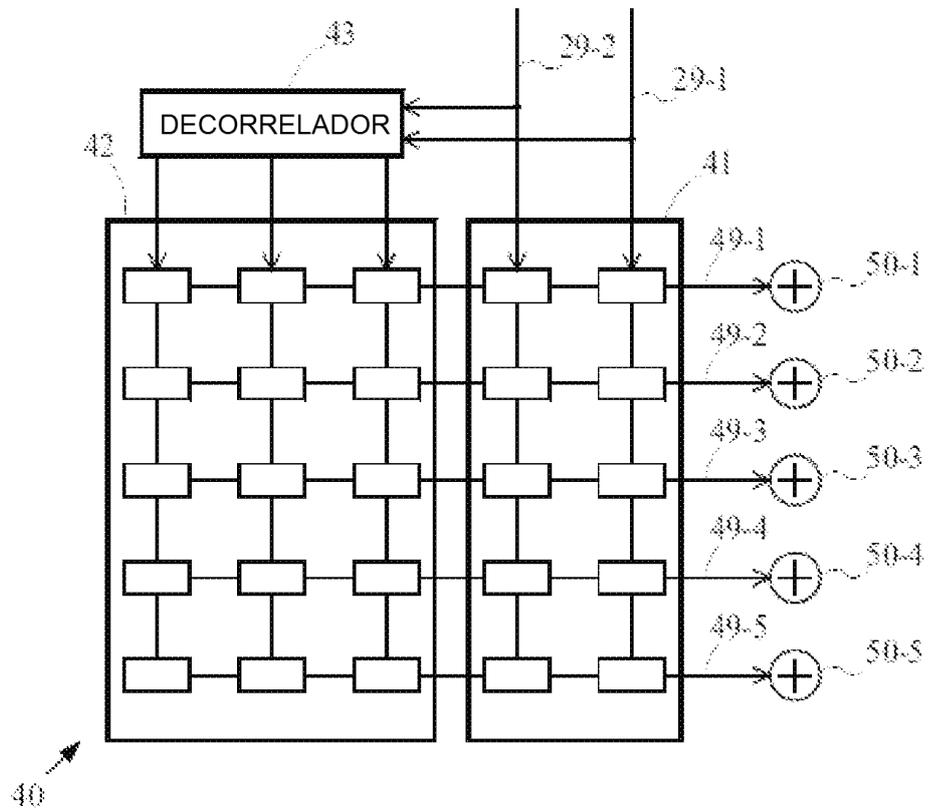


Figura 8

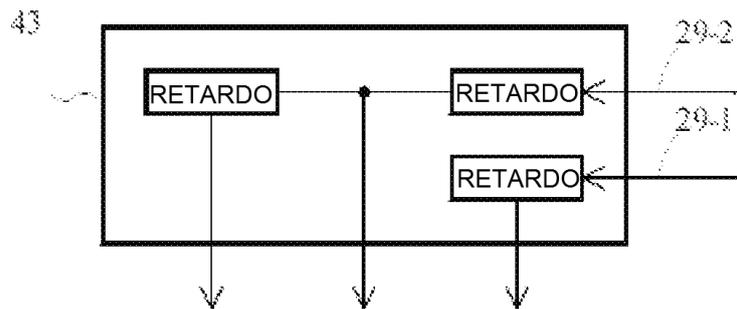


Figura 9

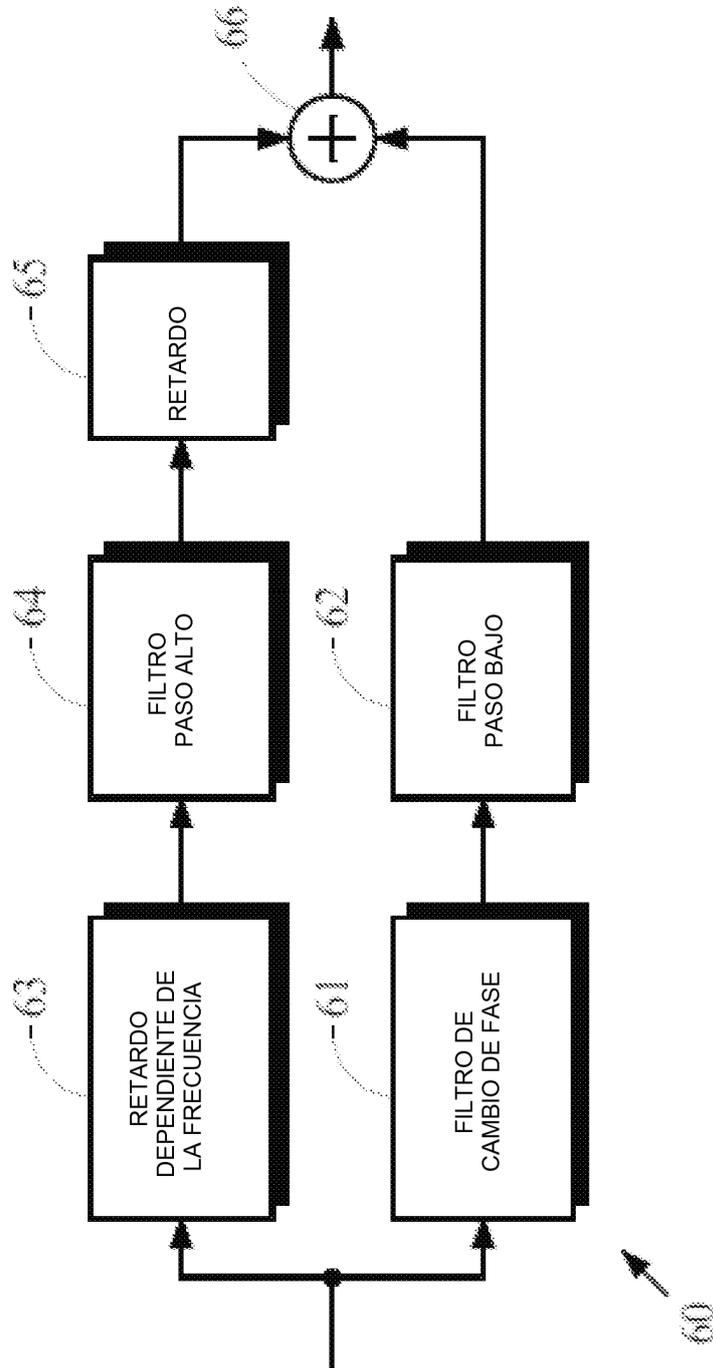


Figura 10

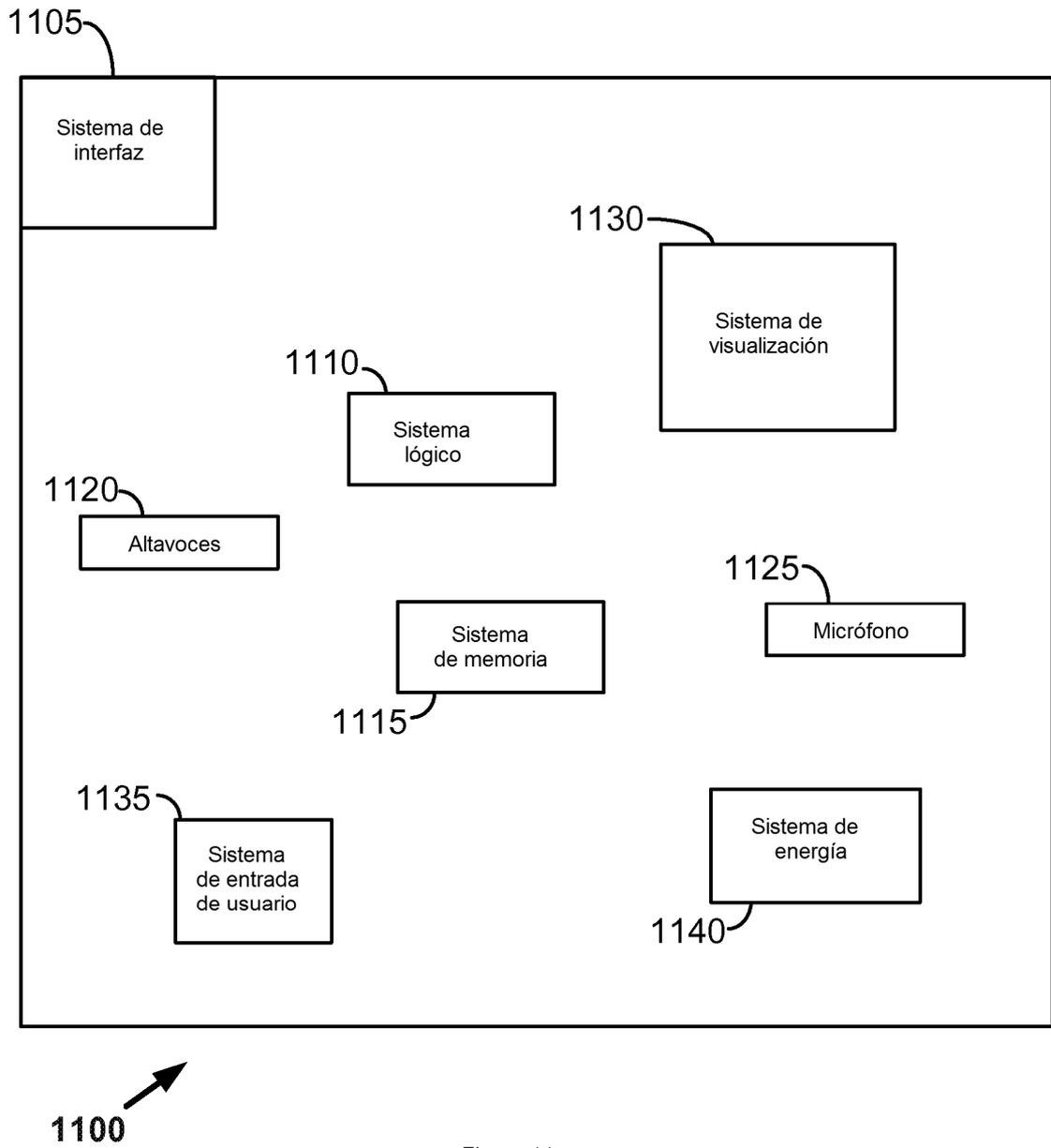


Figura 11