

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 709 248**

51 Int. Cl.:

**H04S 3/00**

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **18.12.2014 PCT/US2014/071100**

87 Fecha y número de publicación internacional: **09.07.2015 WO15102920**

96 Fecha de presentación y número de la solicitud europea: **18.12.2014 E 14824318 (1)**

97 Fecha y número de publicación de la concesión europea: **05.12.2018 EP 3090573**

54 Título: **Generación de audio binaural en respuesta a un audio multicanal que usa al menos una red de retardo de retroalimentación**

30 Prioridad:

**03.01.2014 US 201461923579 P**  
**29.04.2014 CN 201410178258**  
**05.05.2014 US 201461988617 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**15.04.2019**

73 Titular/es:

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

72 Inventor/es:

**YEN, KUAN-CHIEH;**  
**BREEBAART, DIRK J.;**  
**DAVIDSON, GRANT A.;**  
**WILSON, RHONDA;**  
**COOPER, DAVID M. y**  
**SHUANG, ZHIWEI**

74 Agente/Representante:

**ELZABURU, S.L.P**

**ES 2 709 248 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 audio binaural en respuesta a un audio multicanal que usa al menos una red de retardo de retroalimentación

**Referencia cruzada a la solicitud relacionada**

- 5 La presente invención reivindica la prioridad de la Solicitud de Patente China N° 201410178258,0 archivada el 29 de Abril de 2014; las Solicitudes de Patente Provisional de los EE.UU. N° 61/923.579 archivadas el 3 de Enero de 2014; y la Solicitud de Patente Provisional de los EE.UU N° 61/988.617 archivada el 5 de Mayo de 2014.

**Antecedentes de la invención**

## 1. Campo de la invención

- 10 La invención se relaciona con los métodos (a veces referidos como métodos de virtualización de los auriculares) y los sistemas para generar una señal binaural en respuesta a una señal de entrada de audio multi canal, mediante la aplicación de una respuesta al impulso binaural de una sala (BRIR) a cada canal de un conjunto de canales (por ejemplo, a todos los canales) de la señal de entrada. En algunas realizaciones, al menos una red de retardo de la retroalimentación (FDN) aplica una parte de reverberación tardía de una BRIR de mezcla descendente a una mezcla descendente de los canales.

## 2. Antecedentes de la invención

La virtualización de los auriculares (o la reproducción binaural) es una tecnología que pretende ofrecer una experiencia de sonido envolvente o un campo de sonido inmersivo usando auriculares estéreo estándar.

- 20 Los primeros virtualizadores de auriculares aplicaron una función de transferencia relacionada con la cabeza (HRTF) para transportar la información espacial en una reproducción binaural. Una HRTF es un conjunto de parejas de filtros dependientes de la dirección y la distancia que caracterizan cómo se transmite el sonido desde un punto específico en el espacio (ubicación de la fuente de sonido) a tanto los oídos de un oyente en un entorno anecoico. Señales espaciales esenciales tales como la diferencia de tiempo interaural (ITD), la diferencia de nivel interaural (ILD), el efecto de sombra de la cabeza, los picos y las muescas espectrales debido a las reflexiones del hombro y del pabellón auricular, se pueden percibir en el contenido binaural filtrado por la HRTF reproducido. Debido a las restricciones de tamaño de la cabeza humana, las HTRF no proporcionan señales suficientes o robustas con respecto a la distancia a la fuente más allá de aproximadamente un metro. Como resultado, los virtualizadores basados únicamente en una HRTF normalmente no consiguen una buena externalización o distancia percibida.

- 30 La mayoría de los eventos acústicos en nuestra vida diaria ocurren en entornos reverberantes en los que, además de la ruta directa (desde la fuente al oído) modelada por la HRTF, las señales de audio también llegan a los oídos de los oyentes a través de diversas rutas de reflexión. Las reflexiones introducen un impacto profundo a la percepción de un auditorio, tal como la distancia, el tamaño de la sala, y otros atributos del espacio. Para transportar esta información en la reproducción binaural, se necesita aplicar un virtualizador a la reverberación de la sala además de las señales en la HTRF de la ruta directa. Una respuesta al impulso binaural de una sala (BRIR) caracteriza la transformación de las señales de audio desde un punto específico en el espacio a los oídos del oyente en un entorno acústico específico. En teoría, las BRIR incluyen todas las señales acústicas respecto a la percepción espacial.

- 40 La Fig. 1 es un diagrama de bloques de un tipo de virtualizador de auriculares convencional que se configura para aplicar una respuesta al impulso binaural de una sala (BRIR) a cada canal de rango completo de frecuencias ( $X_1, \dots, X_N$ ) de una señal de entrada de audio multi canal. Cada uno de los canales  $X_1, \dots, X_N$  es un canal de altavoz correspondiente a una dirección de fuente diferente en relación a un supuesto oyente (esto es, la dirección de la ruta directa desde una supuesta posición de un altavoz correspondiente a la supuesta posición de un oyente), y cada uno de dichos canales es convolucionado por la BRIR para la correspondiente dirección de la fuente. La ruta acústica desde cada canal necesita ser simulada para cada oído. Por lo tanto, en el resto de este documento, el término BRIR se referirá a un par de respuestas al impulso asociadas con los oídos izquierdo y derecho. Por tanto, el subsistema 2 se configura para convolucionar el canal  $X_1$  con la  $BRIR_1$  (la BRIR para la dirección de la fuente correspondiente), el subsistema 4 se configura para convolucionar el canal  $X_N$  con la  $BRIR_N$  (la BRIR para la dirección de la fuente correspondiente), y así sucesivamente. La salida de cada subsistema BRIR (cada uno de los subsistemas 2, ..., 4) es una señal en el dominio del tiempo que incluye un canal izquierdo y un canal derecho. Las salidas del canal izquierdo de los subsistemas BRIR se mezclan en el elemento 6 de adición, y las salidas del canal derecho de los subsistemas BRIR se mezclan en el elemento 8 de adición. La salida del elemento 6 es el canal izquierdo, L, de la salida de la señal de audio binaural del virtualizador, y la salida del elemento 8 es el canal derecho, R, de la salida de la señal de audio binaural del virtualizador.

- 55 La señal de entrada del audio multi canal puede incluir también efectos de baja frecuencia (LFE) o un canal de graves, identificado en la Fig. 1 como el canal "LFE". De manera convencional, el canal LFE no se convolucionan con

- una BRIR, pero en su lugar se atenúa en la etapa 5 de ganancia de la Fig. 1 (por ejemplo, en -3dB o más) y la salida de la etapa 5 se mezcla de manera igualitaria (mediante los elementos 6 y 8) en cada uno de los canales de la señal de salida binaural del virtualizador. Puede ser necesaria una etapa adicional de retardo en la ruta LFE para alinear en el tiempo la salida de la etapa 5 con las salidas de los subsistemas BRIR (2,..., 4). De manera alternativa, el canal LFE se puede simplemente ignorar (esto es, no ser impuesto o procesado por el virtualizador). Por ejemplo, la realización de la Fig. 2 de la invención (que se describe más adelante) simplemente ignora cualquier canal LFE de la señal de entrada de audio multicanal procesado de este modo. Muchos auriculares de consumo no son capaces de reproducir de manera precisa un canal LFE.
- En algunos virtualizadores convencionales, la señal de entrada experimenta la transformación del dominio del tiempo al dominio de la frecuencia dentro del dominio QMF (filtro espejo en cuadratura), para generar los canales de los componentes de frecuencia en el dominio QMF. Estos componentes de frecuencia se someten al filtrado (por ejemplo, en las implementaciones en el dominio QMF de los subsistemas 2,..., 4 de la Fig. 1) en el dominio QMF y los componentes de frecuencia resultantes normalmente se transforman después de vuelta al dominio del tiempo (por ejemplo, en una etapa final de cada uno de los subsistemas 2,...,4 de la Fig. 1) para que la salida de audio del virtualizador sea una señal en el dominio del tiempo (por ejemplo, una señal binaural en el dominio del tiempo).
- En general, cada canal de rango completo de frecuencias de una entrada de señal de audio multicanal a un virtualizador de auriculares se supone que es indicativa del contenido de audio emitido desde una fuente de sonido en una ubicación conocida en relación con los oídos del oyente. El virtualizador de auriculares se configura para aplicar una respuesta al impulso binaural de una sala (BRIR) a cada uno de dichos canales de la señal de entrada. Cada BRIR se puede descomponer en dos partes: la respuesta directa y las reflexiones. La respuesta directa es la HRTF que corresponde a la dirección de llegada (DOA) de la fuente de sonido, ajustada con la ganancia y el retardo apropiados debido a la distancia (entre la fuente de sonido y el oyente), y de manera opcional aumentada con efectos de paralaje para las pequeñas distancias.
- La parte restante de la BRIR modela las reflexiones. Las reflexiones tempranas son normalmente las reflexiones primarias o secundarias y tienen una distribución temporal relativamente dispersa. La micro estructura (por ejemplo, la ITD y la ILD) de cada reflexión primaria o secundaria es importante. Para las reflexiones posteriores (el sonido reflejado desde más de dos superficies antes de incidir en el oyente), la densidad de eco aumenta con el aumento del número de reflexiones, y los micro atributos de las reflexiones individuales resultan difíciles de observar. Para las reflexiones cada vez más tardías, la macro estructura (por ejemplo, la tasa de decaimiento de la reverberación, la coherencia interaural, y la distribución espectral de la reverberación general) resulta más importante. Debido a esto, las reflexiones de manera adicional se pueden segmentar en dos partes: las reflexiones tempranas y las reverberaciones tardías.
- El retardo de la respuesta directa es la distancia a la fuente desde el oyente dividida entre la velocidad del sonido, y su nivel es (en ausencia de paredes o grandes superficies cercanas a la ubicación de la fuente) inversamente proporcional a la distancia a la fuente. Por otro lado, el retardo y el nivel de las reverberaciones tardías es generalmente insensible a la ubicación de la fuente. Debido a las consideraciones prácticas, los virtualizadores pueden elegir alinear en el tiempo las respuestas directas de las fuentes con las diferentes distancias, y/o comprimir su rango dinámico. Sin embargo, se debería mantener la relación temporal y de nivel entre la respuesta directa, las reflexiones tempranas, y la reverberación tardía dentro de una BRIR.
- La longitud efectiva de una BRIR típica se extiende a centenares de milisegundos o más en entornos más acústicos. La aplicación directa de las BRIR requiere la convolución con un filtro de miles de pulsaciones, que es caro desde el punto de vista computacional. Además, sin la parametrización, requeriría un gran espacio de memoria para almacenar las BRIR de las diferentes posiciones de la fuente para conseguir una resolución espacial suficiente. Por último pero no menos importante, las ubicaciones de la fuente de sonido pueden cambiar en el tiempo, y/o la posición y orientación del oyente puede variar en el tiempo. La simulación aproximada de dichos movimientos requiere respuestas al impulso de la BRIR variantes en el tiempo. La interpolación y aplicación apropiada de dichos filtros variantes en el tiempo puede ser un reto si las respuestas al impulso de estos filtros tienen muchas pulsaciones.
- Se puede usar un filtro que tiene la bien conocida estructura de filtro conocida como red de retardo de retroalimentación (FDN) para implementar un reverberador espacial que se configura para aplicar una reverberación simulada a uno o más canales de una señal de entrada de audio multi canal. La estructura de una FDN es simple. Comprende varios tanques de reverberación (por ejemplo el tanque de reverberación que comprende el elemento  $g_1$  de ganancia y la línea  $z^{-n_1}$  de retardo, en la FDN de la Fig. 4), teniendo cada tanque de reverberación un retardo y una ganancia. En una implementación típica de una FDN, las salidas desde todos los tanques de reverberación son mezcladas por una matriz de retroalimentación unitaria y las salidas de la matriz son retroalimentadas y sumadas con las entradas a los tanques de reverberación. Los ajustes de ganancia se pueden hacer a las salidas de los tanques de reverberación, y las salidas de los tanques de reverberación (o las versiones ajustadas en ganancia de estos) se pueden volver a mezclar de manera adecuada para una reproducción multicanal o binaural. La reverberación de sonido natural puede ser generada y aplicada por una FDN con huellas computacionales y de

memoria compactas. Las FDN se han usado por tanto en virtualizadores para aumentar la respuesta directa producida por la HRTF.

5 Por ejemplo, el comercialmente disponible virtualizador de auriculares Móvil Dolby incluye un reverberador que tienen una estructura basada en la FDN que es operable para aplicar reverberación a cada canal de una señal de audio de cinco canales (teniendo los canales frontal izquierdo, frontal derecho, central, envolvente izquierdo, y envolvente derecho) y para filtrar cada canal reverberado usando un par de filtros diferente de un conjunto de cinco pares de filtros de la función de transferencia relacionada con la cabeza ("HRTF"). El virtualizador de auriculares Móvil Dolby es operable también en respuesta a una señal de audio de dos canales, para generar una salida de audio binaural "reverberada" de dos canales (una salida de sonido envolvente virtual de dos canales a la que se ha aplicada reverberación). Cuando la salida binaural reverberada es procesada y reproducida por un par de auriculares, es percibida en el tímpano del oyente como un sonido filtrado por la HRTF, reverberado de los cinco altavoces en las posiciones frontal izquierda, frontal derecha, central, trasera izquierda (envolvente), y trasera derecha (envolvente). El virtualizador mezcla de manera ascendente una entrada de audio de dos canales mezclados de manera descendente (sin usar ningún parámetro de señal espacial recibido con la entrada de audio) para generar cinco canales de audio mezclados de manera ascendente, aplica la reverberación a los canales mezclados de manera ascendente, y mezcla de manera descendente las cinco señales de los canales reverberados para generar la salida reverberada de dos canales del virtualizador. La reverberación para cada canal mezclado de manera ascendente se filtra en un par diferente de filtros HRTF.

20 En un virtualizador, una FDN se puede configurar para conseguir un cierto tiempo de decaimiento de la reverberación y densidad de eco. Sin embargo, la FDN carece de la flexibilidad para simular la micro estructura de las reflexiones tempranas. Además, en los virtualizadores convencionales la sintonización y configuración de las FDN ha sido principalmente heurística.

25 Los virtualizadores de auriculares que no simulan todas las rutas de reflexión (tempranas y tardías) no pueden conseguir una externalización efectiva. Los inventores han reconocido que los virtualizadores que emplean las FDN que intentan simular todas las rutas de reflexión (tempranas y tardías) por lo general no tienen más que un éxito limitado en la simulación de tanto las reflexiones tempranas como de la reverberación tardía y en la aplicación de ambas a una señal de audio. Los inventores han reconocido también que los virtualizadores que emplean la FDN pero que no tienen la capacidad para controlar de manera apropiada los atributos acústicos espaciales tales como el tiempo de decaimiento de la reverberación, la coherencia interaural, y la relación directa a tardía, pueden conseguir un grado de externalización con el precio de introducir exceso de distorsión de timbre y reverberación.

30 Se hace también referencia al documento WO 2012/093352 A1, que describe un sistema de audio que comprende un receptor para recibir una señal de audio, tal como un objeto de audio o una señal de un canal de una señal multi canal espacial. Un circuito binaural genera una señal de salida binaural mediante el procesamiento de la señal de audio. El procesamiento es representativo de una función de transferencia binaural que proporciona una posición virtual de la fuente de sonido para la señal de audio. Un circuito de medición que genera los datos de medición indicativos de una característica del entorno acústico y un circuito de determinación que determina un parámetro del entorno acústico en respuesta a los datos de medición. El parámetro del entorno acústico puede ser normalmente un parámetro de reverberación, tal como un tiempo de reverberación. Un circuito de adaptación adapta la función de transferencia binaural en respuesta al parámetro del entorno acústico. Por ejemplo, la adaptación puede modificar un parámetro de reverberación para asemejarse más a las características de reverberación del entorno acústico.

### Breve descripción de la invención

45 En vista de lo anterior, la presente invención proporciona un método para generar una señal binaural en respuesta a un conjunto de canales de una señal de entrada de audio multi canal, y un sistema configurado para generar una señal binaural en respuesta a un conjunto de canales de una señal de entrada de audio multi canal, que tienen las características de las respectivas reivindicaciones independientes.

50 En una primera clase de realizaciones, la invención es un método para generar una señal binaural en respuesta a un conjunto de canales (por ejemplo, cada uno de los canales, o cada uno de los canales de rango completo de frecuencias) de una señal de entrada de audio multicanal, incluyendo los pasos de: (a) aplicar un respuesta al impulso binaural de una sala (BRIR) a cada canal de un conjunto (por ejemplo, convolucionando cada canal del conjunto con una BRIR correspondiente a dicho canal), generando de este modo unas señales filtradas, incluyendo el uso de al menos una red de retardo de retroalimentación (FDN) para aplicar una reverberación tardía común a una mezcla descendente (por ejemplo, una mezcla descendente monofónica) de los canales del conjunto; y (b) combinar las señales filtradas para generar la señal binaural. Se usa un banco de FDN para aplicar la reverberación tardía común a la mezcla descendente (por ejemplo, con cada FDN aplicando una reverberación tardía común a una banda de frecuencias diferente). La reverberación tardía común se ha generado para emular los macro atributos colectivos de las partes de reverberación tardías de al menos algunas (por ejemplo, todas) las BRIR de canal único.

Normalmente, el paso (a) incluye un paso de aplicación a cada canal del conjunto de una parte de "respuesta directa y reflexión temprana" de una BRIR de canal único para el canal.

El método para generar una señal binaural en respuesta a una señal de entrada de audio multicanal (o en respuesta a un conjunto de canales de dicha señal) es referido a veces en la presente memoria como método de “virtualización de auriculares”, y el sistema configurado para realizar dicho método es referido a veces en la presente memoria como “virtualizador de auriculares” (o “sistema de virtualización de auriculares” o “virtualizador binaural”).

5 En las realizaciones típicas en la primera clase, cada una de las FDN se implementa en el dominio del banco de filtros (por ejemplo, en el dominio del filtro espejo en cuadratura complejo híbrido (HCQMF) o en el dominio del filtro espejo en cuadratura (QMF), u otra transformada o dominio de sub banda que pueda incluir el diezmado), y en algunas de dichas realizaciones, los atributos acústicos espaciales dependientes de la frecuencia de la señal binaural son controlados mediante el control de la configuración de cada FDN empleada para aplicar la  
10 reverberación tardía. Normalmente, una mezcla descendente monofónica de los canales se usa como entrada de las FDN para un procesamiento binaural eficiente del contenido de audio de la señal multi canal. Las realizaciones típicas en la primera clase incluyen un paso de coeficientes de ajuste de la FDN correspondientes a los atributos dependientes de la frecuencia (por ejemplo, el tiempo de decaimiento de la reverberación, la coherencia interaural, la densidad modal, y la relación directa a tardía), por ejemplo, imponiendo los valores de control a la red de retardo  
15 de retroalimentación a un conjunto de al menos uno de entre la ganancia de entrada, las ganancias del tanque de reverberación, los retardos del tanque de reverberación, o los parámetros de la matriz de salida para cada FDN. Esto permite una mejor coincidencia de los entornos acústicos y unas salidas de sonido más naturales.

En una segunda clase de ejemplos, que no está cubierta por las reivindicaciones, la invención es un método para generar una señal binaural en respuesta a una señal de entrada de audio multi canal que tiene canales, aplicando  
20 una respuesta al impulso binaural de una sala (BRIR) a cada canal de un conjunto de canales de la señal de entrada (por ejemplo, cada uno de los canales de la señal de entrada o cada canal de rango completo de frecuencias de la señal de entrada), incluyendo: el procesamiento de cada canal del conjunto en una primera ruta de procesamiento configurada para modelar, y aplicar a dicho canal, una respuesta directa y una parte de reflexión temprana de una BRIR de canal único para el canal; y el procesamiento de una mezcla descendente (por ejemplo, una mezcla descendente monofónica (mono)) de los canales del conjunto en una segunda ruta de procesamiento (en paralelo con la primera ruta de procesamiento) configurada para modelar, y aplicar una reverberación tardía común a la  
25 mezcla descendente. Normalmente, la reverberación tardía común se ha generado para emular los macro atributos colectivos de las partes de reverberación tardías de al menos alguna (por ejemplo, todas) de las BRIR de canal único. Normalmente, la segunda ruta de procesamiento incluye al menos una FDN (por ejemplo, una FDN para cada una de las múltiples bandas de frecuencias). Normalmente, se usa una mezcla descendente mono como la entrada a todos los tanques de reverberación de cada FDN implementada mediante la segunda ruta de procesamiento. Normalmente, se proporcionan los mecanismos para el control sistemático de los macro atributos de cada FDN para simular mejor los entornos acústicos y producir una virtualización binaural de sonido más natural. Ya que la mayoría de dichos macro atributos son dependientes de la frecuencia, cada FDN se implementa normalmente en el dominio  
30 del filtro espejo en cuadratura complejo híbrido (HCQMF), el dominio de la frecuencia, el dominio, u otro dominio de banco de filtros, y se usa una FDN diferente o independiente para cada banda de frecuencias. El beneficio principal de implementar las FDN en el dominio del banco de filtros es permitir la aplicación de reverberación con las propiedades de reverberación dependientes de la frecuencia. En diversas realizaciones, las FDN se implementan en cualquiera de una amplia variedad de dominios de banco de filtros, usando cualquiera de entre una variedad de bancos de filtros, incluyendo, pero no limitado a los filtros espejo en cuadratura de valor complejo (QMF), los filtros de respuesta al impulso finita (filtros FIR), los filtros de respuesta al impulso infinita (filtros IIR), las transformadas de Fourier discretas (DFT), y las transformadas de coseno o seno (modificadas), las transformadas Wavelet, o los filtros cruzados. En una realización preferida, el banco de filtros o la transformada empleadas incluye el diezmado (por  
35 ejemplo, una disminución de la tasa de muestreo de la representación de la señal en el dominio de la frecuencia) para reducir la complejidad computacional del proceso de la FDN.

Algunas realizaciones en la primera clase implementan una o más de las siguientes características:

1. una implementación FDN en el dominio del banco de filtros (por ejemplo, en el dominio del filtro espejo en cuadratura complejo híbrido), o una implementación FDN en el dominio del banco de filtros híbridos y una  
40 implementación del filtro de reverberación tardía en el dominio del tiempo, que normalmente permite el ajuste independiente de los parámetros y/o los ajustes de la FDN para cada banda de frecuencias (que permite un control simple y flexible de los atributos acústicos dependientes de la frecuencia), por ejemplo, proporcionando la capacidad de variar los retardos del tanque de reverberación en diferentes bandas para cambiar la densidad modal como una función de la frecuencia;
2. El proceso de mezcla descendente específico, empleado para generar (desde la señal de audio de entrada multi canal) la señal mezclada de manera descendente (por ejemplo, mezcla descendente monofónica) procesada en la segunda ruta de procesamiento, depende de la distancia a la fuente de cada canal y el  
45 manejo de la respuesta directa para mantener un nivel apropiado y la relación de temporización entre las respuestas directa y tardía;

3. Se aplica un filtro paso todo (APF) en la segunda ruta de procesamiento (por ejemplo, en la entrada o la salida de un banco de FDN) para introducir la diversidad de fase y aumentar la densidad de eco sin cambiar el espectro y/o el timbre de la reverberación resultante;
- 5 4. Se implementan retardos fraccionales en la ruta de retroalimentación para cada FDN con una estructura multi tasa, de valor complejo, para superar los problemas relacionados con los retardos cuantificados para la cuadrícula de factor de disminución de muestras;
- 10 5. En las FDN, las salidas del tanque de reverberación se mezclan de manera descendente de manera lineal directamente dentro de los canales binaurales, usando los coeficientes de mezcla descendente de salida que se ajustan en base a la coherencia interaural deseada en cada banda de frecuencias. De manera opcional, la correspondencia de los tanques de reverberación a los canales de salida binaurales alterna entre bandas de frecuencias para conseguir un retardo balanceado entre los canales binaurales. También de manera opcional, se aplican factores de normalización a las salidas del tanque de reverberación para igualar sus niveles mientras conservan el retardo fraccional y la energía general,
- 15 6. El tiempo de decaimiento de la reverberación dependiente de la frecuencia y/o la densidad modal está controlado por las combinaciones apropiadas de ajustes de los retardos del tanque de reverberación y las ganancias en cada banda de frecuencias para simular salas reales;
- 20 7. Se aplica un factor de escala por banda de frecuencias (por ejemplo, en bien la entrada o la salida de la ruta de procesamiento relevante), para:  
 controlar la relación directa a tardía (DLR) dependiente de la frecuencia que coincide con la de una sala real (se puede usar un modelo simple para calcular el factor de escala requerido en base a la DLR objetivo y el tiempo de decaimiento, por ejemplo, T60);  
 proporcionar una atenuación de baja frecuencia para mitigar el exceso de artefactos de peinado y/o ruido de baja frecuencia; y/o  
 aplicar un conformado espectral de campo difuso a las respuestas de la FDN;
- 25 8. Se implementan modelos paramétricos simples para controlar los atributos dependientes de la frecuencia esenciales de la reverberación tardía, tales como el tiempo de decaimiento de la reverberación, la coherencia interaural, y/o la relación directa a tardía.

30 Los aspectos de la invención incluyen los métodos y sistemas que realizan (o se configuran para realizar, o soportan la realización de) la virtualización binaural de señales de audio (por ejemplo señales de audio cuyo contenido de audio consiste de canales de altavoces, y/o señales de audio basadas en objetos).

35 En otra clase de realizaciones, la invención es un sistema para generar una señal binaural en respuesta a un conjunto de canales de una señal de entrada de audio multi canal, incluyendo la aplicación de una respuesta al impulso binaural de una sala (BRIR) a cada canal del conjunto, generando de este modo unas señales filtradas, incluyendo mediante el uso de una red de retardo de retroalimentación única (FDN) para aplicar una reverberación tardía común a una mezcla descendente de los canales del conjunto; y la combinación de las señales filtradas para generar la señal binaural. La FDN se implementa en el dominio del tiempo. En algunas de dichas realizaciones, la FDN en el dominio del tiempo incluye:

un filtro de entrada que tiene una entrada acoplada para recibir la mezcla descendente, en donde el filtro de entrada se configura para generar una primera mezcla descendente filtrada en respuesta a la mezcla descendente;

40 un filtro paso todo, acoplado y configurado a una segunda mezcla descendente filtrada en respuesta a la primera mezcla descendente filtrada;

45 un subsistema de aplicación de la reverberación, que tiene una primera salida y una segunda salida, en donde el subsistema de aplicación de la reverberación comprende un conjunto de tanques de reverberación, teniendo cada uno de los tanques de reverberación un retardo diferente, y en donde el subsistema de aplicación de la reverberación se acopla y configura para genera un primer canal binaural no mezclado y un segundo canal binaural no mezclado en respuesta a la segunda mezcla descendente filtrada, para imponer el primer canal binaural no mezclado en la primera salida, y para imponer el segundo canal binaural no mezclado en la segunda salida; y

50 una etapa de filtrado y mezclado de coeficiente de correlación cruzada interaural (IACC) acoplada al subsistema de aplicación de la reverberación y configurada para generar un primer canal binaural mezclado y un segundo canal binaural mezclado en respuesta al primer canal binaural no mezclado y al segundo canal binaural no mezclado.

El filtro de entrada se puede implementar para generar (preferiblemente como una cascada de dos filtros configurados para generar) la primera mezcla descendente filtrada tal que cada BRIR tenga una relación directa a tardía (DLR) que coincida, al menos sustancialmente, con una DLR objetivo.

Cada tanque de reverberación se puede configurar para generar una señal retardada, y puede incluir un filtro de reverberación (por ejemplo, implementado como un filtro atenuador o una cascada de filtros atenuadores) acoplado y configurado para aplicar una ganancia a una señal que se propaga en dicho cada uno de los tanques de reverberación, para provocar que la señal retardada tenga una ganancia que coincida, al menos sustancialmente, con una ganancia de decaimiento objetivo para dicha señal retardada, en un esfuerzo para conseguir una característica de tiempo de decaimiento de reverberación objetivo (por ejemplo una característica  $T_{60}$ ) de cada BRIR.

En algunas realizaciones, el primer canal binaural no mezclado guía al segundo canal binaural no mezclado, los tanques de reverberación incluyen un primer tanque de reverberación configurado para generar una primera señal retardada que tenga el retardo más corto y un segundo tanque de reverberación configurado para generar una segunda señal retardada que tenga el segundo retardo más corto, en donde el primer tanque de reverberación se configura para aplicar una primera ganancia a la primera señal retardada, el segundo tanque de reverberación se configura para aplicar una segunda ganancia a la segunda señal retardada, siendo la segunda ganancia diferente de la primera ganancia, y la aplicación de la primera ganancia y de la segunda ganancia resulta en una atenuación del primer canal binaural no mezclado en relación con el segundo canal binaural no mezclado. Normalmente, el primer canal binaural no mezclado y el segundo canal binaural no mezclado son indicativos de una imagen estéreo que se vuelve a centrar. En algunas realizaciones, la etapa de filtrado y de mezclado IACC se configura para generar el primer canal binaural mezclado y el segundo canal binaural mezclado de manera tal que dicho primer canal binaural mezclado y dicho segundo canal binaural mezclado tengan una característica IACC que coincida al menos de manera sustancial con la característica IACC objetivo.

Las realizaciones típicas de la invención proporcionan una estructura simple y unificada para soportar tanto el audio de entrada consistente de canales de altavoces, y el audio de entrada basado en objetos. En las realizaciones en las que se aplica la BRIR a los canales de la señal de entrada que son canales de objetos, el procesamiento de la "respuesta directa y la reflexión temprana" realizado en cada canal de objeto supone una dirección de la fuente indicada mediante metadatos proporcionados con el contenido de audio del canal de objeto. En las realizaciones en las que se aplican las BRIR a los canales de la señal de entrada que son los canales de altavoz, el procesamiento de la "respuesta directa y la reflexión temprana" realizado en cada canal de altavoz supone una dirección a la fuente que corresponde con el canal de altavoz (esto es, la dirección de una ruta directa desde la posición supuesta del altavoz correspondiente hasta la posición supuesta del oyente). Independientemente de si los canales de entrada son canales de objeto o de altavoz, el procesamiento de la "reverberación tardía" es realizado en la mezcla descendente (por ejemplo, una mezcla descendente monofónica) de los canales de entrada y no supone ninguna dirección específica a la fuente para el contenido de audio de la mezcla descendente.

### Breve descripción de los dibujos

La FIG. 1 es un diagrama de bloques de un sistema de virtualización de auriculares convencional.

La FIG. 2 es un diagrama de bloques de un sistema que incluye una realización del sistema de virtualización de auriculares inventivo.

La FIG. 3 es un diagrama de bloques de otra realización del sistema de virtualización de auriculares inventivo.

La FIG. 4 es un diagrama de bloques de una FDN de un tipo incluido en una implementación típica del sistema de la Fig. 3.

La FIG. 5 es un gráfico del tiempo de decaimiento de la reverberación ( $T_{60}$ ) en milisegundos como una función de la frecuencia en Hz, que puede ser alcanzado mediante una realización del virtualizador inventivo para la que el valor de  $T_{60}$  en cada una de las dos frecuencias específicas ( $f_A$  y  $f_B$ ) se establece como sigue:  $T_{60,A} = 320$  ms con  $f_A = 10$  Hz, y  $T_{60,B} = 150$  ms con  $f_B = 2,4$  kHz.

La FIG. 6 es un gráfico de la coherencia Interaural (Coh) como una función de la frecuencia en Hz, que se puede alcanzar mediante una realización del virtualizador inventivo para el cual los parámetros de control  $Coh_{max}$ ,  $Coh_{min}$ , y  $f_C$  se establecen para tener los siguientes valores:  $Coh_{max} = 0,95$ ,  $Coh_{min} = 0,05$ , y  $f_C = 700$  Hz.

La FIG. 7 es un gráfico de la relación directa a tardía (DLR) con una distancia a la fuente de un metro, en dB, como una función de la frecuencia en Hz, que se puede alcanzar mediante una realización del virtualizador inventivo para el cual los parámetros de control  $DLR_{1K}$ ,  $DLR_{pendiente}$ ,  $DLR_{min}$ ,  $HPF_{pendiente}$ , y  $f_T$  se establecen para tener los siguientes valores  $DLR_{1K} = 18$  dB,  $DLR_{pendiente} = 6$  dB/10x frecuencia,  $DLR_{min} = 18$  dB,  $HPF_{pendiente} = 6$  dB/10x frecuencia, y  $f_T = 200$  Hz.

La FIG. 8 es un diagrama de bloques de otra realización de un subsistema de procesamiento de la reverberación tardía del sistema de virtualización de auriculares inventivo.

La FIG. 9 es un diagrama de bloques de una implementación en el dominio del tiempo de una FDN, de un tipo incluido en algunas realizaciones del sistema inventivo.

La FIG. 9A es un diagrama de bloques de un ejemplo de una implementación del filtro 400 de la FIG. 9.

La FIG. 9B es un diagrama de bloques de un ejemplo de una implementación del filtro 406 de la FIG. 9.

La FIG. 10 es un diagrama de bloques de una realización del sistema de virtualización de auriculares inventivo, en el que se implementa el subsistema 221 de procesamiento de la reverberación tardía en el dominio del tiempo.

La FIG. 11 es un diagrama de bloques de una realización de los elementos 422, 423, y 424 de la FDN de la FIG. 9.

5 La FIG. 11A es un gráfico de la respuesta (R1) en frecuencia de una implementación típica del filtro 500 de la Fig. 11, la respuesta (R2) en frecuencia de un filtro 501 de implementación típica de la Fig. 11, y la respuesta de los filtros 500 y 501 conectados en paralelo.

La FIG. 12 es un gráfico de un ejemplo de una característica IACC (curva "I") que se puede alcanzar mediante la implementación de la FDN de la Fig. 9, y una característica IACC objetivo (curva "I<sub>T</sub>").

10 La FIG. 13 es un gráfico de una característica T60 que se puede alcanzar mediante la implementación de la FDN de la Fig. 9, mediante la implementación apropiada de cada uno de los filtros 406, 407, 408, y 409 como un filtro atenuador.

La FIG. 14 es un gráfico de una característica T60 que se puede alcanzar mediante la implementación de la FDN de la Fig. 9, mediante la implementación apropiada de cada uno de los filtros 406, 407, 408, y 409 como una cascada de dos filtros atenuadores IIR.

15

### Notación y nomenclatura

A lo largo de esta descripción, incluyendo en las reivindicaciones, la expresión realizar una operación "en" una señal o datos (por ejemplo, filtrar, escalar, transformar, o aplicar ganancia a, la señal o los datos) se usa en un sentido amplio para denotar la realización de la operación directamente en la señal o los datos, o en una versión procesada de la señal o los datos (por ejemplo, en una versión de la señal que experimente un filtrado preliminar o un pre procesamiento antes de realizar la operación en sí).

20

A lo largo de esta descripción incluyendo en las reivindicaciones, la expresión "sistema" se usa en un sentido amplio para denotar un servicio, sistema, o subsistema. Por ejemplo, un subsistema que implementa un virtualizador puede ser referido como un sistema virtualizador, y un sistema que incluye dicho subsistema (por ejemplo, un sistema que genera X señales de salida en respuesta a las múltiples entradas, en el que el subsistema genera M de las entradas y las otras X – M entradas son recibidas desde una fuente externa) puede ser también referido como un sistema virtualizador (o un virtualizador).

25

A lo largo de esta descripción incluyendo en las reivindicaciones, el término "procesador" se usa en un sentido amplio para denotar un sistema o un dispositivo programable o de otra manera configurable (por ejemplo, con software o firmware) para realizar operaciones sobre los datos (por ejemplo, audio, o video u otros datos de imagen). Ejemplos de procesadores incluyen una matriz de puertas programables en campo (u otros circuitos integrados configurables o chipset), un procesador de señal digital programado y/o de otra manera configurado para realizar un procesamiento en serie sobre audio u otros datos de sonido, un procesador de propósito general programable u ordenador, y un chip microprocesador programable o chipset.

30

A lo largo de esta descripción incluyendo en las reivindicaciones, la expresión "banco de filtros de análisis" se usa en un sentido amplio para denotar un sistema (por ejemplo, un subsistema) configurado para aplicar una transformación (por ejemplo una transformación desde el dominio del tiempo al dominio de la frecuencia) a una señal en el dominio del tiempo para generar valores (por ejemplo, componentes de frecuencia) indicativos del contenido de la señal en el dominio del tiempo, en cada una de un conjunto de bandas de frecuencias. A lo largo de esta descripción incluyendo en las reivindicaciones, la expresión "dominio del banco de filtros" se usa en un sentido amplio para denotar el dominio de los componentes de frecuencia generados mediante una transformación o un banco de filtros de análisis (por ejemplo, el dominio en el que dichas componentes de frecuencia se procesan). Ejemplo de los dominios del banco de filtros incluyen (pero no se limitan a) el dominio de la frecuencia, el dominio del filtro espejo en cuadratura (QMF), y el dominio del filtro espejo en cuadratura complejo híbrido (HCQMF). Ejemplos de la transformada que se puede aplicar mediante un análisis del banco de filtros incluyen (pero no se limitan a) la transformada de coseno discreta (DCT), la transformada de coseno discreta modificada (MDCT), la transformada de Fourier discreta (DFT), y la transformada wavelet. Ejemplos de bancos de filtros de análisis incluyen (pero no se limitan a) los filtros espejo en cuadratura (QMF), los filtros de respuesta al impulso finita (filtros FIR), los filtros de respuesta al impulso infinita (filtros IIR), los filtros cruzados, y los filtros que tienen otras estructuras multi tasa adecuadas.

35

40

45

A lo largo de esta descripción incluyendo en las reivindicaciones, el término "metadatos" se refiere a datos separados y diferentes de los datos de audio correspondientes (el contenido de audio de un flujo de bits que también incluye los metadatos). Los metadatos se asocian con los datos de audio, e indican al menos una característica de los datos de audio (por ejemplo, que tipo o tipos de procesamiento se han realizado ya, o se deberían realizar, en los datos de audio, o la trayectoria de un objeto indicada por los datos de audio). La asociación de los metadatos con los datos de audio es sincrónica en el tiempo. Por tanto, los metadatos actuales (lo recibidos o actualizados más

50

55

recientemente) pueden indicar que los datos de audio correspondientes tienen a la vez una característica indicada y/o comprenden los resultados de un tipo indicado de procesamiento de los datos de audio.

5 A lo largo de esta descripción incluyendo en las reivindicaciones, el término “acopla” o “acoplado” se usa para indicar bien una conexión directa o indirecta. Por tanto, si un primer dispositivo se acopla con un segundo dispositivo, esa conexión puede ser a través de una conexión directa, o a través de una conexión indirecta a través de otros dispositivos y conexiones.

A lo largo de esta descripción incluyendo en las reivindicaciones, las siguientes expresiones tienen las siguientes definiciones:

10 altavoz o megafonía se usan de manera sinónima para denotar cualquier transductor emisor de sonido. Esta definición incluye los altavoces implementados como múltiples transductores (por ejemplo, el woofer y el tweeter);

alimentación de altavoz: una señal de audio a ser aplicada directamente a un altavoz, o una señal de audio que se ha de aplicar a un amplificador y un altavoz en serie;

15 canal (“o canal de audio”): una señal de audio monofónica. Dicha señal se puede reproducir normalmente de tal manera que sea equivalente a la aplicación de la señal directamente a un altavoz en una posición deseada o nominal. La posición deseada puede ser estática, como normalmente es el caso con altavoces físicos, o dinámica;

programa de audio: un conjunto de uno o más canales de audio (al menos un canal de altavoz y/o al menos un canal de objeto) y opcionalmente también los metadatos asociados (por ejemplo, los metadatos que describen una presentación de audio espacial deseada);

20 canal del altavoz (o “canal de alimentación del altavoz”): un canal de audio que se asocia con un altavoz designado (en una posición deseada o nominal), o con una zona de altavoz designada dentro de una configuración de altavoz definida. Se reproduce un canal de altavoz de tal manera como para ser equivalente a la aplicación de la señal de audio directamente al altavoz designado (en la posición deseada o nominal) o a un altavoz en la zona de altavoz designado;

25 canal de objeto: un canal de audio indicativo del sonido emitido por una fuente de audio (a veces referida como un “objeto” de audio). Normalmente, un canal de objeto determina una descripción de la fuente de audio paramétrica (por ejemplo, los metadatos indicativos de la descripción de la fuente de audio paramétrica se incluyen en o se proporcionan con el canal de objeto). La descripción de la fuente puede determinar el sonido emitido por la fuente (como una función del tiempo), la posición evidente (por ejemplo, las coordenadas espaciales 3D) de la fuente como función del tiempo, y de manera opcional al menos un parámetro adicional (por ejemplo, el tamaño o la anchura de la fuente evidente) que caracterizan la fuente;

30 programa de audio: basado en objeto: un programa de audio que comprende un conjunto de uno o más canales de objeto (y de manera opcional también comprende al menos un canal de altavoz) y de manera opcional también los metadatos asociados (por ejemplo, los metadatos indicativos de una trayectoria de un objeto de audio que emite el sonido indicado por un canal de objeto, o de otra forma los metadatos indicativos de una presentación de audio espacial deseada del sonido indicado por un canal de objeto, o los metadatos indicativos de una identificación de al menos un objeto de audio que es una fuente de sonido indicada mediante un canal de objeto); y

40 reproducir: el proceso de convertir un programa de audio en una o más alimentaciones de altavoz, o el proceso de convertir un programa de audio en una o más alimentaciones de altavoz y convertir la alimentación o alimentaciones de altavoz en sonido usando uno o más altavoces (en el último caso, el procesamiento es referido a veces en la presente memoria como procesamiento “por” el altavoz u altavoces). Un canal de audio se puede procesar de manera trivial (“en” una posición deseada) aplicando la señal directamente a un altavoz físico en la posición deseada, o uno o más canales de audio se puede procesar usando una de entre una variedad de técnicas de virtualización diseñadas para ser sustancialmente equivalente (para el oyente) a dicho procesamiento trivial. En el último caso, cada canal de audio se puede convertir en una o más alimentaciones de altavoz a ser aplicadas al

45 altavoz o altavoces en las ubicaciones conocidas, que en general son posiciones diferentes de la posición deseada, de manera tal que el sonido emitido por el altavoz o altavoces en respuesta a la alimentación o alimentaciones será percibida como que se emite desde la posición deseada. Ejemplos de dichas técnicas de virtualización incluyen el procesamiento binaural a través de los auriculares (por ejemplo, usando el procesamiento de Auriculares Dolby que simula hasta 7.1 canales de sonido envolvente para la portadora de auriculares) y la síntesis de campo de onda.

50 La notación de que una señal de audio multi canal es una señal de canal “x.y” o “x.y.z” denota en la presente memoria que la señal tiene “x” canales de altavoces de frecuencias completas (correspondiente a los altavoces posicionados nominalmente en el plano horizontal de los supuestos oídos del oyente), “y” canales LFE (o de subwoofer), y opcionalmente también “z” canales de altavoces superiores de frecuencias completas (correspondientes a los altavoces posicionados por encima de la supuesta cabeza del oyente, por ejemplo, en o

55 cerca del techo de la sala).

La expresión "IACC" denota en la presente memoria el coeficiente de correlación cruzada interaural en su sentido habitual, que es una medida de la diferencia entre los tiempos de llegada de la señal de audio a los oídos de un oyente, normalmente indicado mediante un número en un rango desde un primer valor que indica que las señales que llegan son iguales en magnitud y exactamente fuera de fase, a un valor intermedio que indica que las señales que llegan no tienen ninguna similitud, a un valor máximo que indica que las señales que llegan son idénticas teniendo la misma amplitud y fase.

### Descripción detallada de las realizaciones preferidas

Muchas realizaciones de la presente invención son tecnológicamente posibles. Será evidente para aquellos expertos en la técnica a partir de la presente descripción cómo implementarlas. Las realizaciones del sistema y el método inventivo serán descritas con referencia a las Fig. 2-14.

La FIG. 2 es un diagrama de bloques de un sistema (20) que incluye una realización del sistema de virtualización de auriculares inventivo. El sistema de virtualización de auriculares (a veces referido como virtualizador) se configura para aplicar una respuesta al impulso binaural de una sala (BRIR) a N canales de rango completo de frecuencias ( $X_1, \dots, X_N$ ) de una señal de entrada de audio multi canal. Cada uno de los canales  $X_1, \dots, X_N$  (que pueden ser canales de altavoz o canales de objeto) corresponden a una dirección y distancia a la fuente específica en relación a un supuesto oyente, y el sistema de la Fig. 2 se configura para convolucionar cada uno de dichos canales con una BRIR para la dirección y distancia a la fuente correspondiente.

El sistema 20 puede ser un decodificador que se acopla para recibir un programa de audio codificado, y que incluye un subsistema (no mostrado en la Fig. 2) acoplado y configurado para decodificar el programa mediante la recuperación de los N canales de rango completo de frecuencias ( $X_1, \dots, X_N$ ) de este y para proporcionarles los elementos 12, ..., 14 y 15 del sistema de virtualización (que comprende los elementos 12, ..., 14, 15, 16, y 18, acoplados como se muestra). El decodificador puede incluir subsistemas adicionales, algunos de los cuales realizan funciones no relacionadas con la función de virtualización realizada por el sistema de virtualización, y algunas de las cuales pueden realizar funciones relacionadas con la función de virtualización. Por ejemplo, las últimas funciones pueden incluir la extracción de los metadatos desde el programa codificado, y la provisión de los metadatos a un subsistema de control de la virtualización que emplea los metadatos para controlar los elementos del sistema virtualizador.

El subsistema 12 (con el subsistema 15) se configura para convolucionar el canal  $X_1$  con la  $BRIR_1$  (la BRIR para la dirección y distancia a la fuente correspondiente), el subsistema 14 (con el subsistema 15) se configura para convolucionar el canal  $X_N$  con la  $BRIR_N$  (la BRIR para la dirección a la fuente correspondiente), y así sucesivamente para cada uno de los otros N-2 subsistemas BRIR. La salida de cada uno de los subsistemas 12, ..., 14, y 15 es una señal en el dominio del tiempo que incluye un canal izquierdo y un canal derecho. Además los elementos 16 y 18 se acoplan a las salidas de los elementos 12, ..., 14, y 15. Además el elemento 16 se configura para combinar (mezclar) las salidas del canal izquierdo de los subsistemas BRIR, y además el elemento 18 se configura para combinar (mezclar) las salidas del canal derecho de los subsistemas BRIR. La salida del elemento 16 es el canal izquierdo, L, de la señal de audio binaural emitida desde el virtualizador de la Fig. 2, y la salida del elemento 18 es el canal derecho, R, de la señal de audio binaural emitida desde el virtualizador de la Fig. 2.

Las características importantes de las realizaciones típicas de la invención son evidentes a partir de la comparación de la realización de la Fig. 2 del virtualizador de auriculares inventivo con el virtualizador de auriculares convencional de la Fig. 1. Por propósitos de comparación, suponemos que los sistemas de la Fig. 1 y de la Fig. 2 se configuran para que, cuando se imponga la misma entrada de audio multi canal a cada uno de ellos, los sistemas apliquen una BRIR, que tenga la misma parte de respuesta directa y de reflexión temprana (esto es, la EBRIR relevante de la Fig. 2) para cada canal de rango completo de frecuencias,  $X_i$ , de la señal de entrada (aunque no necesariamente con el mismo grado de éxito). Cada  $BRIR_i$  aplicada por el sistema de la Fig. 1 o de la Fig. 2 se puede descomponer en dos partes: una parte de respuesta directa y de reflexión temprana (por ejemplo, una de las partes  $EBIR_1, \dots, EBIR_N$  aplicadas por los subsistemas 12-14 de la Fig. 2), y una parte de reverberación tardía. La realización de la Fig. 2 (y otras realizaciones típicas de la invención suponen que las partes de reverberación tardía de las BRIR de canal único,  $BRIR_i$ , se pueden compartir a través de las direcciones de la fuente y por tanto todos los canales, y así aplicar la misma reverberación tardía (esto es, una reverberación tardía común) a una mezcla descendente de todos los canales de rango completo de frecuencias de la señal de entrada. Esta mezcla descendente puede ser una mezcla descendente monofónica (mono) de todos los canales, pero de manera alternativa puede ser una mezcla descendente estéreo o multi canal obtenido a partir de los canales de entrada (por ejemplo, a partir de un subconjunto de los canales de entrada).

Más específicamente, el subsistema 12 de la Fig. 2 se configura para convolucionar el canal  $X_1$  de la señal de entrada con la  $EBRIR_1$  (la parte de la BRIR de respuesta directa y reflexión temprana para la dirección de fuente correspondiente), el subsistema 14 se configura para convolucionar el canal  $X_N$  con la  $EBRIR_N$  (la parte BRIR de respuesta directa y reflexión temprana para la dirección de fuente correspondiente), y así sucesivamente. El subsistema 15 de reverberación tardía de la Fig. 2 se configura para generar una mezcla descendente mono de todos los canales de rango completo de frecuencias de la señal de entrada, y para convolucionar la mezcla descendente con la LBRIR (una reverberación tardía común para todos los canales que se mezclan de manera

descendente). La salida de cada subsistema BRIR del virtualizador de la Fig. 2 (cada uno de los subsistemas 12,..., 14, y 15) incluye un canal izquierdo y un canal derecho (de una señal binaural generada a partir del correspondiente canal de altavoz o mezcla descendente). Las salidas del canal izquierdo de los subsistemas BRIR se combinan (se mezclan) en el elemento 16 de adición, y las salidas del canal derecho de los subsistemas BRIR se combinan (se mezclan) en el elemento 18 de adición

El elemento 16 de adición se puede implementar para simplificar la suma correspondiente a las muestras del canal binaural Izquierdo (las salidas del canal Izquierdo de los subsistemas 12,..., 14, y 15) para generar el canal Izquierdo de la señal de salida binaural, suponiendo que se implementan los ajustes de nivel y los alineamientos de tiempo apropiados en los subsistemas 12,..., 14, y 15. De manera similar, el elemento 18 de adición se puede implementar también para simplificar la suma correspondiente a las muestras del canal binaural Derecho (las salidas del canal Derecho de los subsistemas 12,..., 14, y 15) para generar el canal Derecho de la señal de salida binaural, suponiendo que se implementan los ajustes de nivel y los alineamientos de tiempo apropiados en los subsistemas 12,..., 14, y 15.

El subsistema 15 de la Fig. 2 se puede implementar de cualquiera de entre una variedad de maneras, pero normalmente incluye al menos una red de retardo de retroalimentación configurada para aplicar la reverberación tardía común a una mezcla descendente monofónica de los canales de la señal de entrada impuestos a éste. Normalmente, donde cada uno de los subsistemas 12,..., 14 aplica una parte de respuesta directa y reflexión temprana (EBRIR<sub>i</sub>) de una BRIR de canal único para el canal ( $X_i$ ) que éste procesa, la reverberación tardía común se ha generado para emular los macro atributos colectivos de las partes de reverberación tardía de al menos alguno (por ejemplo, todos) de las BRIR de canal único (cuya "partes de respuesta directa y reflexión temprana" son aplicadas por los subsistemas 12,...,14). Por ejemplo, una implementación del subsistema 15 tiene la misma estructura que el subsistema 200 de la Fig. 3, que incluye un banco de redes (203, 204,..., 205) de retardo de la retroalimentación configurado para aplicar una reverberación tardía común a una mezcla descendente monofónica de los canales de la señal de entrada impuesta a ésta.

Los subsistemas 12,..., 14 de la Fig. 2 se pueden implementar de cualquiera de entre una variedad de maneras (en ya sea el dominio del tiempo o el dominio del banco de filtros), con la implementación preferida para cualquier aplicación específica dependiendo de diversas consideraciones, tales como (por ejemplo) el rendimiento, la computación, y la memoria. En una implementación ejemplar, cada uno de los subsistemas 12,..., 14 se configura para convolucionar el canal impuesto a éste con un filtro FIR correspondiente a las respuestas directa y temprana asociadas con el canal, con la ganancia y el retardo establecidos de manera apropiada para que las salidas de los subsistemas 12,..., 14 se puedan simplificar y combinar de manera eficiente con aquellas del subsistema 15.

La FIG. 3 es un diagrama de bloques de otra realización del sistema de virtualización de auriculares inventivo. La realización de la Fig. 3 es similar a la de la Fig. 2, con dos señales en el dominio del tiempo (canal izquierdo y derecho) que se emiten desde un subsistema 100 de procesamiento de la respuesta directa y la reflexión temprana, y dos señales en el dominio del tiempo (canal izquierdo y derecho) que se emiten desde el subsistema 200 de procesamiento de la reverberación tardía. El elemento 210 de adición se acopla a las salidas de los subsistemas 100 y 200. El elemento 210 se configura para combinar (mezclar) las salidas del canal izquierdo de los subsistemas 100 y 200 para generar el canal izquierdo, L, de la salida de la señal de audio binaural del virtualizador de la Fig. 3, y para combinar (mezclar) las salidas del canal derecho de los subsistemas 100 y 200 para generar el canal derecho, R, de la salida de la señal de audio binaural del virtualizador de la Fig. 3. El elemento 210 se puede implementar para simplificar la suma correspondiente a las muestras del canal izquierdo emitidas desde los subsistemas 100 y 200 para generar el canal izquierdo de la señal de salida binaural, y para simplificar la suma correspondiente a las muestras del canal derecho emitidas desde los subsistemas 100 y 200 para generar el canal derecho de la señal de salida binaural, suponiendo que se implementan los ajustes de nivel y alineamientos de tiempo apropiados en los subsistemas 100 y 200.

En el sistema de la Fig. 3, los canales,  $X_i$ , de la señal de audio multi canal se dirigen a, y experimentan el procesamiento en, dos rutas de procesamiento paralelas: una a través del subsistema 100 de procesamiento de la respuesta directa y de la reflexión temprana; el otro a través del subsistema 200 de procesamiento de la reverberación tardía. El sistema de la Fig. 3 se configura para aplicar una BRIR<sub>i</sub> a cada canal,  $X_i$ . Cada BRIR<sub>i</sub> se puede descomponer en dos partes: una parte de respuesta directa y reflexión temprana (aplicada por el subsistema 100), y una parte de reverberación tardía (aplicada por el subsistema 200). En funcionamiento, el subsistema 100 de procesamiento de la respuesta directa y la reflexión temprana genera por tanto las partes de respuesta directa y reflexión temprana de la señal de audio binaural que se emite desde el virtualizador, y el subsistema 200 de procesamiento de la reverberación tardía ("generador de reverberación tardía") genera por tanto la parte de reverberación tardía de la señal de audio binaural que se emite desde el virtualizador. Las salidas de los subsistemas 100 y 200 se mezclan (mediante el subsistema 210 de adición) para generar la señal de audio binaural, que se impone normalmente desde el subsistema 210 a un sistema de procesamiento (no mostrado) en el que se experimenta el procesamiento binaural para la reproducción mediante los auriculares.

Normalmente, cuando se procesa y reproduce mediante un par de auriculares, una señal de audio binaural normal emitida desde el elemento 210 se percibe en el tímpano del oyente como un sonido desde los "N" altavoces (donde

$N \geq 2$  y  $N$  normalmente es igual a 2, 5 o 7) en cualquiera de una amplia variedad de posiciones, incluyendo las posiciones en frente de, detrás de, y por encima del oyente. La reproducción de las señales de salida generadas en el funcionamiento del sistema de la Fig. 3 pueden dar al oyente la experiencia de que el sonido viene de más de dos (por ejemplo, cinco o siete) fuentes “envolventes”. Al menos algunas de esas fuentes son virtuales.

5 El subsistema 100 de respuesta directa y reflexión temprana se puede implementar de cualquier variedad de maneras (en ya sea el dominio del tiempo o el dominio del banco de filtros), con la implementación preferida para cualquier aplicación específica dependiendo de diversas consideraciones, tales como (por ejemplo) el rendimiento, la computación, y la memoria. En una implementación ejemplar, el subsistema 100 se configura para convolucionar cada canal impuesto a éste con un filtro FIR correspondiente a las respuestas directa y temprana asociadas con el canal, con la ganancia y el retardo establecidos de manera apropiada para que la salida de los subsistemas 100 se pueda combinar de manera simple y eficiente (en el elemento 210) con estos subsistemas 200.

10 Como se muestra en la Fig. 3, el generador 200 de reverberación tardía incluye un subsistema 201 de mezcla descendente, un banco de filtros 202, un banco de FDN (FDN 203, 204,..., 205), y un banco de filtros 207 de síntesis, acoplado como se muestra. El subsistema 201 se configura para mezclar de manera descendente los canales de la señal de entrada multi canal en una mezcla descendente mono, y el banco de filtros 202 de análisis se configura para aplicar una transformación a la mezcla descendente mono para dividir la mezcla descendente mono en “K” bandas de frecuencias, donde K es un número entero. Los valores en el dominio del banco de filtros (salida del banco de filtros 202) en cada una de las diferentes bandas de frecuencias se imponen a una diferente de las FDN 203, 204,..., 205 (hay “K” de estas FDN, cada una acoplada y configurada para aplicar una parte de reverberación tardía de una BRIR a los valores en el dominio del banco de filtros impuestos a éste). Los valores en el dominio del banco de filtros preferiblemente son diezmados en el tiempo para reducir la complejidad computacional de las FDN.

15 En principio, cada canal de entrada (al subsistema 100 y al subsistema 201 de la Fig. 3) se puede procesar en su propia FDN (o banco de FDN) para simular la parte de reverberación tardía de su BRIR. A pesar del hecho de que la parte de reverberación tardía de las BRIR asociadas con las diferentes ubicaciones de la fuente de sonido son normalmente muy diferentes en términos de diferencias cuadráticas medias en las respuestas al impulso, sus atributos estadísticos tales como su espectro energía media, su estructura de decaimiento de energía, la densidad modal, la densidad de picos y similares a menudo son muy similares. Por lo tanto, la parte de reverberación tardía de un conjunto de BRIR normalmente es perceptualmente bastante similar a lo largo de los canales y por consiguiente, es posible usar una FDN o un banco de FDN común (por ejemplo, las FDN 203, 204,..., 205) para simular la parte de reverberación tardía de dos o más BRIR. En las realizaciones típicas, se emplea una FDN (o un banco de FDN) común tal, y la entrada a ésta está comprendida de una o más mezclas descendentes construidas a partir de los canales de entrada. En la implementación ejemplar de la Fig. 2, la mezcla descendente es una mezcla descendente monofónica (impuesta en la salida del subsistema 201) de todos los canales de entrada.

20 Con referencia a la realización de la Fig. 2, cada una de las FDN 203, 204,..., 205, se implementa en el dominio del banco de filtros, y se acopla y configura para procesar una banda de frecuencias diferente de los valores emitidos desde el banco de filtros 202 de análisis, para generar las señales reverberadas izquierda y derecha para cada banda. Para cada banda, la señal reverberada izquierda es una secuencia de valores en el dominio del banco de filtros, y la señal reverberada derecha es otra secuencia de valores en el dominio del banco de filtros. El banco de filtros 207 de síntesis se acopla y configura para aplicar una transformada desde el dominio de la frecuencia al dominio del tiempo a las  $2K$  secuencias de los valores en el dominio del banco de filtros (por ejemplo, los componentes de frecuencia en el dominio QMF) emitidos desde las FDN, y para hacer valer los valores transformados en una señal en el dominio del tiempo del canal izquierdo (indicativa del contenido de audio de la mezcla descendente mono a la que se aplica la reverberación tardía) y una señal en el dominio del tiempo del canal derecho (también indicativa del contenido de audio de la mezcla descendente mono a la que se aplica la reverberación tardía). Estas señales del canal izquierdo y el canal derecho se emiten al elemento 210.

25 En una implementación típica cada una de las FDN 203, 204,..., 205, se implementa en el dominio QMF, y el banco de filtros 202 transforma la mezcla descendente mono del subsistema 201 en el dominio QMF (por ejemplo, en el dominio del filtro espejo en cuadratura complejo híbrido (HCQMF)), para que la señal impuesta desde el banco de filtros 202 a una salida de cada FDN 203, 204,..., 205 sea una secuencia de componentes de frecuencia en el dominio QMF. En dicha implementación, la señal impuesta desde el banco de filtros 202 a la FDN 203 es una secuencia de componentes de frecuencia en el dominio QMF en una primera banda de frecuencias, la señal impuesta desde el banco de filtros 202 a la FDN 204 es una secuencia de componentes de frecuencia en el dominio QMF en una segunda banda de frecuencias, y la señal impuesta desde el banco de filtros 202 a la FDN 205 es una secuencia de componentes de frecuencia en el dominio QMF en una “K-ésima” banda de frecuencias. Cuando el banco de filtros 202 de análisis se implementa así, el banco de filtros 207 de síntesis se configura para aplicar una transformada desde el dominio QMF al dominio del tiempo a las  $2K$  secuencias de salida de los componentes de frecuencia en el dominio QMF de las FDN, para generar las señales en el dominio del tiempo con reverberación tardía del canal izquierdo y del canal derecho que se emiten al elemento 210.

Por ejemplo, si  $K=3$  en el sistema de la Fig. 3, entonces hay seis entradas al banco de filtros 207 de síntesis (canales izquierdo y derecho, que comprenden muestras en el dominio de la frecuencia o en el dominio QMF, emitidas desde cada FDN 203, 204, y 205) y dos salidas desde 207 (canales izquierdo y derecho, cada uno consistente de muestras en el dominio del tiempo). En este ejemplo, el banco de filtros 207 normalmente se

- 5 implementarían como dos bancos de filtros de síntesis: uno (al que se impondrían los tres canales izquierdos de las FDN 203, 204, y 205) configurado para generar la salida de la señal del canal izquierdo en el dominio del tiempo desde el banco de filtros 207; y un segundo (al que se impondrían los tres canales derechos de las FDN 203, 204, y 205) configurado para generar la salida de la señal del canal derecho en el dominio del tiempo desde el banco de filtros 207.
- 10 De manera opcional, el subsistema 209 de control se acopla a cada una de las FDN 203, 204, ..., 205, y se configura para imponer los parámetros de control a cada una de las FDN para determinar la parte de reverberación tardía (LBRIR) que es aplicada por el subsistema 200. Más adelante se describen ejemplos de dichos parámetros de control. Se contempla que en algunas implementaciones el subsistema 209 de control sea operable en tiempo real (por ejemplo, en respuesta a los comandos de usuario impuestos a este por un dispositivo de entrada) para
- 15 implementar la variación en tiempo real de la parte de reverberación tardía (LBRIR) aplicada por el subsistema 200 a la mezcla descendente monofónica de los canales de entrada.

Por ejemplo, si la señal de entrada al sistema de la Fig. 2 es una señal de 5.1 canales (cuyos canales de rango completo de frecuencias están en el siguiente orden de canales: L, R, C, Ls, Rs), todos los canales de rango completo de frecuencias tienen la misma distancia a la fuente, y el subsistema 201 de mezcla descendente se puede

- 20 implementar como la siguiente matriz de mezcla descendente, que simplemente suma los canales de rango completo de frecuencias para formar una mezcla descendente mono:

$$D = [1 \ 1 \ 1 \ 1 \ 1]$$

Después del filtrado paso todo (en el elemento 301 en cada una de las FDN 203, 204, ..., y 205), la mezcla descendente mono mezcla de manera ascendente los cuatros tanques de reverberación de manera que se conserve la energía:

- 25

$$U = \begin{bmatrix} 1/\sqrt{4} \\ 1/\sqrt{4} \\ 1/\sqrt{4} \\ 1/\sqrt{4} \end{bmatrix}$$

De manera alternativa (como un ejemplo), podemos elegir asignar los canales del lado izquierdo a los dos primeros tanques de reverberación, los canales del lado derecho a los dos últimos tanques de reverberación, y el canal central a todos los tanques de reverberación. En este caso, el subsistema 201 de mezcla descendente se

- 30 implementarían para formar dos señales de mezcla descendente:

$$D = \begin{bmatrix} 1 & 0 & 1/\sqrt{2} & 1 & 0 \\ 0 & 1 & 1/\sqrt{2} & 0 & 1 \end{bmatrix}$$

En este ejemplo, la mezcla ascendente de los tanques de reverberación (en cada una de las FDN 203, 204, ..., y 205) es:

$$U = \begin{bmatrix} 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} \end{bmatrix}$$

- 35 Ya que existen dos señales de mezcla descendente, el filtrado paso todo (en el elemento 301 en cada una de las FDN 203, 204, ..., y 205) necesita ser aplicado dos veces. Se introduciría diversidad para las respuestas tardías de (L, Ls), (R, Rs) y C a pesar de que todos ellos tienen los mismos macro atributos. Cuando los canales de la señal de entrada tienen diferentes distancias a la fuente, aún sería necesario que se aplicaran los retardos y las ganancias apropiadas en el proceso de mezcla descendente.

- 40 A continuación describiremos consideraciones para implementaciones específicas del subsistema 201 de mezcla descendente, y de los subsistemas 100 y 200 del virtualizador de la Fig. 3.

El proceso de mezcla descendente implementado por el subsistema 201 depende de la distancia a la fuente (entre la fuente de sonido y la supuesta posición del oyente) para cada canal a ser mezclado, y el manejo de la respuesta directa. El retardo de la respuesta directa  $t_d$  es:

$$t_d = d / v_s$$

donde  $d$  es la distancia entre la fuente de sonido y el oyente y  $v_s$  es la velocidad del sonido. Además, la ganancia de la respuesta directa es proporcional a  $1/d$ . Si estas reglas se conservan en el manejo de las respuestas directas de los canales con diferentes distancias a la fuente, el subsistema 201 puede implementar una mezcla descendente directa de todos los canales, ya que el retardo y el nivel de la reverberación tardía generalmente es insensible a la ubicación de la fuente.

Debido a consideraciones prácticas, los virtualizadores (por ejemplo, el subsistema 100 del virtualizador de la Fig. 3) se pueden implementar para alinear en el tiempo las respuestas directas para los canales de entrada que tienen diferentes distancias a la fuente. Para mantener el retardo relativo entre la respuesta directa y la reverberación tardía para cada canal, un canal con distancia a la fuente  $d$  se debería retrasar en  $(d_{\max} - d)/v_s$  antes de ser mezclado de manera descendente con otros canales. Aquí  $d_{\max}$  denota la máxima distancia posible a la fuente.

Los virtualizadores (por ejemplo el subsistema 100 del virtualizador de la Fig. 3) se pueden implementar para comprimir el rango dinámico de las respuestas directas. Por ejemplo, la respuesta directa para un canal con una distancia  $d$  a la fuente se puede escalar por un factor de  $d^{-\alpha}$ , donde  $0 \leq \alpha \leq 1$ , en lugar de  $d^{-1}$ . Para mantener el nivel de diferencia entre la respuesta directa y la reverberación tardía, se puede necesitar implementar el subsistema 201 de mezcla descendente para escalar un canal con una distancia a la fuente  $d$  por un factor de  $d^{1-\alpha}$  antes de mezclarlo de manera descendente con los otros canales escalados.

La red de retardo de retroalimentación de la Fig. 4 es una implementación ejemplar de la FDN 203 (o 204, o 205) de la Fig. 3. Aunque el sistema de la Fig. 4 tiene cuatro tanques de reverberación (cada uno incluyendo una etapa de ganancia,  $g_i$ , y una línea de retardo,  $z^{-n_i}$ , acoplada a la salida de la etapa de ganancia) las variaciones respecto al sistema (y otras FDN empleadas en las realizaciones del virtualizador inventivo) implementan más de o menos de cuatro tanques de reverberación.

La FDN de la Fig. 4 incluye un elemento 300 de ganancia de entrada, un filtro 301 paso todo (APF) acoplado a la salida del elemento 300, los elementos 302, 303, 304, y 305 de adición acoplados a la salida del APF 301, y cuatro tanques de reverberación (que comprenden cada uno un elemento de ganancia,  $g_k$  (uno de los elementos 306), una línea de retardo,  $z^{-M_k}$  (uno de los elementos 307) acoplado a éste, y un elemento de ganancia,  $1/g_k$  (uno de los elementos 309) acoplado a este, donde  $0 \leq k-1 \leq 3$ ) cada uno acoplado a la salida de uno diferente de los elementos 302, 303, 304, y 305. La matriz 308 unitaria se acopla a las salidas de las líneas 307 de retardo, y se configura para imponer una salida de retroalimentación a una segunda entrada de cada uno de los elementos 302, 303, 304, y 305. Las salidas de los dos elementos 309 de ganancia (del primer y segundo tanques de reverberación) se imponen a las entradas del elemento 310 de adición, y la salida del elemento 310 se impone a una entrada de la matriz 312 de mezcla de salida. Las salidas de los otros dos elementos 309 de ganancia (del tercer y cuarto tanques de reverberación) se imponen a las entradas del elemento 311 de adición, y la salida del elemento 311 se impone a la otra entrada de la matriz 312 de mezcla de salida.

El elemento 302 se configura para añadir la salida de la matriz 308 que corresponde a la línea de retardo  $z^{-n_1}$  (esto es, para aplicar la retroalimentación desde la salida de la línea de retardo  $z^{-n_1}$  a través de la matriz 308) a la entrada del primer tanque de reverberación. El elemento 303 se configura para añadir la salida de la matriz 308 que corresponde a línea de retardo  $z^{-n_2}$  (esto es, para aplicar la retroalimentación desde la salida de la línea de retardo  $z^{-n_2}$  a través de la matriz 308) a la entrada del segundo tanque de reverberación. El elemento 304 se configura para añadir la salida de la matriz 308 que corresponde a línea de retardo  $z^{-n_3}$  (esto es, para aplicar la retroalimentación desde la salida de la línea de retardo  $z^{-n_3}$  a través de la matriz 308) a la entrada del tercer tanque de reverberación. El elemento 305 se configura para añadir la salida de la matriz 308 que corresponde a línea de retardo  $z^{-n_4}$  (esto es, para aplicar la retroalimentación desde la salida de la línea de retardo  $z^{-n_4}$  a través de la matriz 308) a la entrada del cuarto tanque de reverberación.

El elemento 300 de ganancia de entrada de la FDN de la Fig. 4 se acopla para recibir una banda de frecuencias de la señal de mezcla descendente monofónica transformada (una señal en el dominio del banco de filtros) que es una salida del banco de filtros 202 de análisis de la Fig.3. El elemento 300 de ganancia de entrada aplica un factor de ganancia (escalado)  $G_{\text{entrada}}$ , a la señal en el dominio del banco de filtros impuesta a éste. De manera colectiva, los factores  $G_{\text{entrada}}$  de escala (implementados por todas las FDN 203, 204, ..., 205 de la Fig. 3) para todas las bandas de frecuencias controlan el conformado espectral y el nivel de la reverberación tardía. El establecimiento de las ganancias de entrada,  $G_{\text{entrada}}$ , en todas las FDN del virtualizador de la Fig. 3 a menudo tiene en cuenta los siguientes objetivos:

una relación directa a tardía (DLR), de la BRIR aplicada a cada canal, que coincida con salas reales;

una atenuación de baja frecuencia necesaria para mitigar el exceso de artefactos de combinación y/o ruido de baja frecuencia; y

hacer coincidir la envolvente de los espectros de campo difuso.

Si suponemos que la respuesta directa (aplicada por el subsistema 100 de la Fig.3) proporciona una ganancia unitaria en todas las bandas de frecuencias, se puede alcanzar una DLR (relación de energía) específica ajustando  $G_{entrada}$  para ser:

$$G_{entrada} = \text{raiz cuadrada} (\ln(10^6)/(T60 * DLR)),$$

5 donde T60 es el tiempo de decaimiento de la reverberación definido como el tiempo que la reverberación necesita para decaer 60 dB (está determinado mediante los retardos de reverberación y las ganancias de reverberación discutidas anteriormente), y "ln" denota la función logarítmica natural.

El factor de ganancia de entrada,  $G_{entrada}$ , puede ser dependiente del contenido que se esté procesando. Una aplicación de dicha dependencia del contenido es asegurar que la energía de la mezcla descendente en cada segmento de tiempo/frecuencia es igual a la suma de las energías de las señales de canal individual que se están mezclando de manera descendente, independientemente de cualquier correlación que pueda existir entre las señales del canal de entrada. En este caso, el factor de ganancia de entrada puede ser (o se puede multiplicar por) un término similar o igual a:

$$\sqrt{\frac{\sum_i \sum_j x_i^2(j)}{\sum_j y^2(j)}}$$

15 en el que i es un índice sobre todas las muestras de mezcla descendente de una franja o sub banda de tiempo/frecuencia, y(i) son las muestras de la mezcla descendente para la franja, y  $x_i(j)$  es la señal de entrada (para el canal  $X_i$ ) impuesta a la entrada del subsistema 201 de mezcla descendente.

En una implementación en el dominio QMF típica de la FDN de la Fig. 4, la señal impuesta desde la salida del filtro 301 paso todo (APF) a las entradas de los tanques de reverberación es una secuencia de componentes de frecuencia en el dominio QMF. Para generar una salida de la FDN de sonido más natural, se aplica el APF 301 a la salida del elemento 300 de ganancia para introducir una diversidad de fase y aumentar la densidad de eco. De manera alternativa, o adicional, se pueden aplicar uno o más filtros de retardo paso todo a: las entradas individuales al subsistema 201 de mezcla descendente (de la Fig. 3) antes que se mezclen en el subsistema 201 y sean procesadas por la FDN; o en las rutas de alimentación hacia adelante o retroalimentación del tanque de reverberación representadas en la Fig. 4 (por ejemplo, además o en lugar de las líneas  $z^{mk}$  de retardo en cada tanque de reverberación; o las salidas de la FDN (esto es, a las salidas de la matriz 312 de salida).

En la implementación de los retardos del tanque de reverberación,  $z^{n_i}$ , los retardos  $n_i$  de reverberación deberían ser mutuamente números primos para evitar que los modos de reverberación se alineen en la misma frecuencia. La suma de los retardos debería ser suficientemente grande para proporcionar una densidad modal suficiente para evitar una salida de sonido artificial. Pero los retardos más cortos deben ser lo suficientemente cortos para evitar el exceso de salto de tiempo entre la reverberación tardía y los otros componentes de la BRIR.

Normalmente, las salidas del tanque de reverberación se distribuyen a bien el canal binaural izquierdo o derecho. Normalmente, los conjuntos de salidas del tanque de reverberación que se distribuyen a los dos canales binaurales son iguales en número y mutuamente exclusivos. Se desea también equilibrar el tiempo de los dos canales binaurales. Por lo que si la salida del tanque de reverberación con el menor retardo va a un canal binaural, la que tenga el segundo retardo más corto iría al otro canal.

Los retardos del tanque de reverberación pueden ser diferentes a lo largo de las bandas de frecuencias para cambiar la densidad modal como una función de la frecuencia. De manera general, las bandas de frecuencias inferiores requieren una mayor densidad modal, por tanto los retardos más largos del tanque de reverberación.

40 Las amplitudes de las ganancias del tanque de reverberación,  $g_i$ , y los retardos del tanque de reverberación determinan conjuntamente el tiempo de decaimiento de la reverberación de la FDN de la Fig. 4:

$$T_{60} = -3n_i / \log_{10}(|g_i|) / F_{FRM}$$

45 donde  $F_{FRM}$  es la tasa de tramas del banco de filtros 202 (de la Fig. 3). Las fases de las ganancias del tanque de reverberación introducen retardos fraccionales para superar los problemas relacionados con los retardos del tanque de reverberación que se cuantifican para la cuadrícula de factor de disminución de muestras del banco de filtros.

La matriz 308 de retroalimentación unitaria proporciona una mezcla descendente uniforme entre los tanques de reverberación en la ruta de retroalimentación.

Para igualar los niveles de las salidas del tanque de reverberación, los elementos 309 de ganancia aplican una ganancia de normalización,  $1/|g_i|$  a la salida de cada tanque de reverberación, para eliminar el impacto del nivel de

las ganancias del tanque de reverberación a la vez que mantiene los retardos fraccionales introducidos por sus fases.

La matriz 312 de mezcla descendente de salida (también definida como  $M_{\text{salida}}$ ) es una matriz 2 x 2 configurada para mezclar de manera descendente los canales binaurales no mezclados (las salidas de los elementos 310 y 311, respectivamente) desde la distribución inicial para alcanzar los canales binaurales izquierdo y derecho de salida (las señales L y R impuestas en la salida de la matriz 312) que tienen la coherencia interaural deseada. Los canales binaurales no mezclados están cerca de ser no correlados después de la distribución inicial ya que no consisten de ninguna salida común del tanque de reverberación. Si la coherencia interaural deseada es  $Coh$ , donde  $|Coh| \leq 1$ , la matriz 312 de mezcla de salida se puede definir como:

$$M_{\text{salida}} = \begin{bmatrix} \cos \beta & \text{sen} \beta \\ \text{sen} \beta & \cos \beta \end{bmatrix}, \text{ donde } \beta = \arcsen(Coh)/2$$

Ya que los retardos del tanque de reverberaciones son diferentes, uno de los canales binaurales no mezclados guiaría al otro constantemente. Si la combinación de los retardos del tanque de reverberación y el patrón de distribución es idéntico a lo largo de las bandas de frecuencias, se produciría un sesgo de la imagen de sonido. Este sesgo se puede mitigar si el patrón de distribución se alterna a lo largo de las bandas de frecuencias de manera tal que los canales binaurales mezclados se guíen y arrastren el uno al otro en bandas de frecuencias alternas. Esto se puede alcanzar mediante la implementación de la matriz 312 de mezcla de salida para tener la forma que se expone en el párrafo anterior en las bandas de frecuencias impares (esto es, en la primera banda de frecuencias (procesada por la FDN 203 de la Fig. 3), la tercera banda de frecuencias, y así sucesivamente), y para tener la siguiente forma en las bandas de frecuencias pares (esto es, en la segunda banda de frecuencias (procesada por la FDN 204 de la Fig. 3), la cuarta banda de frecuencias, y así sucesivamente):

$$M_{\text{salida, alt}} = \begin{bmatrix} \text{sen} \beta & \cos \beta \\ \cos \beta & \text{sen} \beta \end{bmatrix}$$

donde la definición de  $\beta$  sigue siendo la misma. Se debería observar que la matriz 312 se puede implementar para ser idéntica en las FDN para todas las bandas de frecuencias, pero el orden de canal de sus entradas se puede conmutar para alternar una de las bandas de frecuencias (por ejemplo, la salida del elemento 310 se puede imponer a la primera entrada de la matriz 312 y la salida del elemento 311 se puede imponer a la segunda entrada de la matriz 312 en las bandas de frecuencias impares, y la salida del elemento 311 se puede imponer a la primera entrada de la matriz 312 y la salida del elemento 310 se puede imponer a la segunda entrada de la matriz 312 en las bandas de frecuencias pares).

En el caso de que las bandas de frecuencias se solapen (parcialmente), la anchura del rango de frecuencias sobre el cual se alterna la forma de la matriz 312 puede aumentar (por ejemplo, se podría alternar una vez por cada dos o tres bandas consecutivas), o el valor de  $\beta$  en las expresiones anteriores (para la forma de la matriz 312) se puede ajustar para asegurar que la coherencia media iguale el valor deseado para compensar el solapamiento espectral de bandas de frecuencias consecutivas.

Si los tributos acústicos objetivos anteriormente definidos  $T_{60}$ ,  $Coh$ , y  $DLR$  son conocidos para la FDN para cada banda de frecuencias específica en el virtualizador inventivo, cada una de las FDN (cada una de las cuales puede tener la estructura mostrada en la Fig. 4) se puede configurar para alcanzar los atributos objetivos. De manera específica, en algunas realizaciones la ganancia de entrada ( $G_{\text{entrada}}$ ) y las ganancias del tanque de reverberación y los retardos ( $g_i$  y  $n_i$ ) y los parámetros de la matriz  $M_{\text{salida}}$  de salida para cada FDN se pueden establecer (por ejemplo, mediante valores de control impuestos al respecto por el subsistema 209 de control de la Fig. 3) para alcanzar los atributos objetivo de acuerdo con las relaciones descritas en la presente memoria. En la práctica, el establecimiento de los atributos dependientes de la frecuencia mediante modelos con parámetros de control simples es a menudo suficiente para generar reverberación tardía de sonido natural que coincida con los entornos acústicos específicos.

A continuación describimos un ejemplo de cómo se puede determinar un tiempo ( $T_{60}$ ) de decaimiento de reverberación objetivo para la FDN para cada banda de frecuencias específica de una realización del virtualizador inventivo, determinando el tiempo ( $T_{60}$ ) de decaimiento de reverberación objetivo para cada una de entre un pequeño número de bandas de frecuencias. El nivel de respuesta de la FDN decae de manera exponencial en el tiempo.  $T_{60}$  es inversamente proporcional al factor de decaimiento,  $df$  (definido como dB de decaimiento por unidad de tiempo):

$$T_{60} = 60 / df.$$

El factor de decaimiento,  $df$ , depende de la frecuencia y generalmente aumenta linealmente en comparación con la escala de frecuencia logarítmica, por lo que el tiempo de decaimiento de la reverberación es también una función de

la frecuencia que generalmente disminuye según aumenta la frecuencia. Por lo tanto, si se determinan (por ejemplo, se establecen) los valores de  $T_{60}$  para los dos puntos de frecuencia, se determina la curva  $T_{60}$  para todas las frecuencias. Por ejemplo, si los tiempos de decaimiento de reverberación para los puntos de frecuencia  $f_A$  y  $f_B$  son  $T_{60,A}$  y  $T_{60,B}$ , respectivamente, la curva  $T_{60}$  se define como:

$$T_{60}(f) = \frac{T_{60,A} T_{60,B} \log(f_B / f_A)}{T_{60,A} \log(f / f_A) - T_{60,B} \log(f / f_B)}$$

La Figura 5 muestra un ejemplo de la curva  $T_{60}$  que se puede alcanzar mediante una realización del virtualizador inventivo para el que el valor  $T_{60}$  en cada una de las dos frecuencias específicas ( $f_A$  y  $f_B$ ) se establece:  $T_{60,A} = 320$  ms con  $f_A = 10$  Hz, y  $T_{60,B} = 150$  ms con  $f_B = 2,4$  kHz.

A continuación describimos un ejemplo de cómo se puede alcanzar la coherencia Interaural objetivo (Coh) para la FDN para cada banda específica de frecuencias de una realización del virtualizador inventivo mediante el establecimiento de un pequeño número de parámetros de control. La coherencia Interaural objetivo (Coh) de la reverberación tardía en gran parte sigue el patrón de un campo de sonido difuso. Se puede modelar mediante una función seno hasta una frecuencia  $f_C$  de corte, y una constante por encima de la frecuencia de corte. Un modelo simple para la curva Coh es:

$$Coh(f) = \begin{cases} Coh_{\min} + (Coh_{\max} - Coh_{\min}) \sin(f / f_C) & , f \leq f_C \\ Coh_{\min} & , f \geq f_C \end{cases}$$

donde los parámetros  $Coh_{\min}$  y  $Coh_{\max}$  satisfacen  $-1 \leq Coh_{\min} < Coh_{\max} \leq 1$ , y controlan el rango de la Coh. La frecuencia  $f_C$  de corte óptima depende del tamaño de la cabeza del oyente. Una  $f_C$  demasiado grande lleva a una imagen de la fuente de sonido internalizada, mientras que un valor demasiado pequeño lleva a una imagen de la fuente de sonido dispersa o dividida. La Figura 6 es un ejemplo de una curva Coh que se puede alcanzar mediante una realización del virtualizador inventivo para el que los parámetros de control  $Coh_{\max}$ ,  $Coh_{\min}$ , y  $f_C$  se establecen para tener los siguientes valores:  $Coh_{\max} = 0,95$ ,  $Coh_{\min} = 0,05$ , y  $f_C = 700$  Hz.

A continuación describiremos un ejemplo de cómo se puede alcanzar una relación directa a tardía (DLR) objetivo para la FDN para cada banda específica de frecuencias de una realización del virtualizador inventivo estableciendo un pequeño número de parámetros de control. La relación Directa a tardía (DLR), en dB, generalmente aumenta linealmente en comparación con la escala de frecuencias logarítmicas. Se puede controlar mediante el establecimiento de la  $DLR_{1K}$  (DLR en dB @ 1 kHz) y la  $DLR_{pend}$  (en dB por 10x frecuencia). Sin embargo, una DLR baja en el rango de frecuencias inferior a menudo resulta en excesivos artefactos de peinado. Para mitigar los artefactos, se añaden dos mecanismos de modificación al control de la DLR:

una DLR mínima de suelo,  $DLR_{\min}$  (en dB); y

un filtro paso alto definido mediante una frecuencia de transición,  $f_T$ , y la pendiente de la curva de atenuación por debajo de ésta,  $HPF_{pend}$  (en dB por 10x frecuencia).

La curva DLR resultante se define como:

$$DLR(f) = \max(DLR_{1K} + DLR_{pend} \log_{10}(f / 1000) , DLR_{\min}) + \min(HPF_{pend} \log_{10}(f / f_T) , 0)$$

se debería observar que la DLR cambia con la distancia a la fuente incluso en el mismo entorno acústico. Por tanto, tanto la  $DLR_{1K}$  como la  $DLR_{\min}$  de la presente memoria son los valores para una distancia a la fuente nominal, tal como 1 metro. La Figura 7 es un ejemplo de una curva DLR para una distancia a la fuente de 1 metro alcanzada mediante una realización del virtualizador inventivo con los parámetros de control  $DLR_{1K}$ ,  $DLR_{pend}$ ,  $DLR_{\min}$ ,  $HPF_{pend}$ , y  $f_T$  fijados para tener los siguientes valores:  $DLR_{1K} = 18$  dB,  $DLR_{pend} = 6$  dB/10x frecuencia,  $DLR_{\min} = 18$  dB,  $HPF_{pend} = 6$  dB/10x frecuencia, y  $f_T = 200$  Hz.

Las variaciones sobre las realizaciones descritas en la presente memoria tienen una o más de las siguientes características:

las FDN del virtualizador inventivo se implementan en el dominio del tiempo, o tienen una implementación híbrida con la captura de la respuesta al impulso basada en la FDN y el filtrado de la señal basada en la FIR.

el virtualizador inventivo se implementa para permitir la aplicación de la compensación de energía como una función de la frecuencia durante la ejecución del paso de mezcla descendente que genera la señal de entrada mezclada de manera descendente para el subsistema de procesamiento de la reverberación tardía; y

el virtualizador inventivo se implementa para permitir el control manual o automático de los atributos de la reverberación tardía aplicada en respuesta a factores externos (esto es, en respuesta al establecimiento de los parámetros de control).

5 Para las aplicaciones en las que la latencia del sistema es crítica y el retardo provocado por los bancos de filtros de análisis y síntesis es prohibitivo, la estructura de la FDN en el dominio del banco de filtros de las realizaciones típicas del virtualizador inventivo se puede trasladar al dominio del tiempo, y cada estructura FDN se puede implementar en el dominio del tiempo en una clase de realizaciones del virtualizador. En las implementaciones en el dominio del tiempo, los subsistemas que aplican el factor ( $G_{\text{entrada}}$ ) de ganancia de entrada, las ganancias ( $g_i$ ) del tanque de reverberación, y las ganancias ( $1/|g_i|$ ) se reemplazan por filtros con respuestas de amplitud similares para permitir  
10 controles dependientes de la frecuencia. La matriz ( $M_{\text{salida}}$ ) de mezcla de salida se reemplaza también por una matriz de filtros. Diferente a los otros filtros, la respuesta de fase de esta matriz de filtros es crítica ya que la conservación de la energía y la coherencia interaural se pueden ver afectados por la respuesta de fase. Los retardos del tanque de reverberación en una implementación en el dominio del tiempo pueden necesitar variarse ligeramente (de sus valores en una implementación en el dominio del banco de filtros) para evitar compartir el paso del banco de filtros como un factor común. Debido a las diversas limitaciones, el rendimiento de las implementaciones en el dominio de la frecuencia de las FDN del virtualizador inventivo no coincide exactamente con las implementaciones en el dominio del banco de filtros del mismo.

Con referencia a al Fig. 8, a continuación describiremos una implementación híbrida (dominio del banco de filtros y dominio del tiempo) del subsistema de procesamiento de la reverberación tardía inventivo del virtualizador inventivo.  
20 Esta implementación híbrida del subsistema de procesamiento de la reverberación tardía inventiva es una variación del subsistema 200 de procesamiento de la reverberación tardía de la Fig. 4, que implementa una captura de la respuesta al impulso basada en una FDN y un filtrado de señal basado en un FIR.

La realización de la Fig. 8 incluye los elementos 201, 202, 203, 204, 205, y 207 que son idénticos a los elementos numerados de manera idéntica del subsistema 200 de la Fig. 3. La descripción anterior de estos elementos no se repetirá con referencia a la Fig. 8. En la realización de la Fig. 8, el generador 211 de impulso unitario se acopla para imponer una señal de entrada (un pulso) al banco de filtros 202 de análisis. Un filtro 208 LBRIR (entrada mono, salida estéreo) implementado como un filtro FIR aplica la parte de reverberación tardía apropiada de la BRIR (la LBRIR) a la salida mezclada de manera descendente monofónica del subsistema 201. Por tanto, los elementos 211, 202, 203, 204, 205, y 207 son una cadena lateral de procesamiento al filtro 208 LBRIR.

30 Siempre que el ajuste de la LBRIR de la parte de la reverberación tardía se ha de modificar, el generador 211 de impulsos se opera para imponer un impulso unitario al elemento 202, y la salida resultante desde el banco de filtros 207 se captura y se impone al filtro 208 (para establecer el filtro 208 para aplicar la nueva LBRIR determinada por la salida del banco de filtros 207). Para acelerar el intervalo de tiempo desde el cambio del ajuste de la LBRIR al momento en el que la nueva LBRIR tiene efecto, las muestras de la nueva LBRIR pueden empezar a reemplazar la vieja LBRIR según están disponibles. Para acortar la latencia inherente de las FDN, se pueden descartar los ceros iniciales de la LBRIR. Estas opciones proporcionan flexibilidad y permiten que la implementación híbrida proporcione una mejora del rendimiento potencial (en relación al proporcionado por la implementación en el dominio del banco de filtros), con el coste de un cómputo añadido al filtrado de la FIR.

40 Para las aplicaciones en las que la latencia del sistema es crítica, pero la carga computacional es menos importante, el procesador de reverberación tardía en el dominio del banco de filtros de la cadena lateral (por ejemplo, el implementado por los elementos 211, 202, 203, 204,..., 205, y 207 de la Fig. 8) se puede usar para capturar la respuesta al impulso FIR efectiva a aplicar por el filtro 208. El filtro 208 FIR puede implementar esta respuesta FIR capturada y aplicarla directamente a la mezcla descendente mono de los canales de entrada (durante la virtualización de los canales de entrada).

45 Los diversos parámetros de la FDN y por tanto los atributos resultantes de la reverberación tardía se pueden sintonizar de manera manual y posteriormente cablear en una realización del subsistema de procesamiento de la reverberación tardía inventivo, por ejemplo por medio de uno o más pre ajustes que se pueden ajustar (por ejemplo, operando el subsistema 209 de control de la Fig. 3) por el usuario del sistema. Sin embargo, dada la descripción de alto nivel de la reverberación tardía, su relación con los parámetros de la FDN, y la capacidad para modificar su comportamiento, se conciben una amplia variedad de métodos para controlar las diversas realizaciones del procesador de reverberación tardía basado en la FDN, incluyendo (pero no limitado a ) los siguientes:

- 55 1. El usuario final puede controlar de manera manual los parámetros de la FDN, por ejemplo por medio de una interfaz de usuario en un elemento de presentación (por ejemplo, implementado por una realización del subsistema 209 de control de la Fig. 3) o conmutar los pre ajustes usando controles físicos (por ejemplo, implementados por una realización del subsistema 209 de control de la Fig. 3). De esta manera, el usuario final puede adaptar la simulación de la sala según el gusto, el entorno, o el contenido;
2. el autor del contenido de audio a ser virtualizado puede proporcionar los ajustes o los parámetros deseados que se transportan con el contenido en sí, por ejemplo mediante metadatos proporcionados con la señal de audio de entrada. Dichos metadatos se pueden analizar y emplear (por ejemplo, mediante una realización

del subsistema 209 de control de la Fig. 3) para controlar los parámetros relevantes de la FDN. Los metadatos pueden ser por tanto indicativos de propiedades tales como el tiempo de reverberación, el nivel de reverberación, la relación directa a reverberación, y así sucesivamente, y estas propiedades pueden ser variantes en el tiempo, señalizadas por metadatos variantes en el tiempo;

- 5 3. Un dispositivo de reproducción puede ser consciente de su ubicación o entorno, por medio de uno o más sensores. Por ejemplo, un dispositivo móvil puede usar redes GSM, el sistema de posicionamiento global (GPS), puntos de acceso WiFi conocidos, o cualquier otro servicio de ubicación para determinar dónde está el dispositivo. Posteriormente, los datos indicativos de la ubicación y/o el entorno se pueden emplear (por ejemplo, mediante una realización del subsistema 209 de control de la Fig. 3) para controlar los parámetros relevantes de la FDN. Por tanto los parámetros de la FDN se pueden modificar en respuesta a la ubicación del dispositivo, por ejemplo para imitar el entorno físico:
- 10 4. En relación a la ubicación del dispositivo de reproducción, se puede usar un servicio en la nube o un medio social para derivar los ajustes más comunes que están usando los consumidores en un cierto entorno. De manera adicional, los usuarios pueden subir sus ajustes actuales a un servicio en la nube o medio social, en asociación con la ubicación (conocida) para hacerlos disponibles a otros usuarios, o a ellos mismos;
- 15 5. Un dispositivo de reproducción puede contener otros sensores tales como una cámara, un sensor de luz, un micrófono, un acelerómetro, un giroscopio, para determinar la actividad del usuario y el entorno en el que el usuario está, para optimizar los parámetros de la FDN para esa actividad y/o entorno concreto;
- 20 6. Los parámetros de la FDN pueden ser controlados por el contenido de audio. Los algoritmos de clasificación de audio, o el contenido anotado de manera manual pudiendo indicar si los segmentos de audio comprenden palabras, música, efectos de sonido, silencio y similares. Los parámetros de la FDN se pueden ajustar según dichas etiquetas. Por ejemplo, la relación directa a reverberación se puede reducir para un diálogo para mejorar la inteligibilidad del diálogo. De manera adicional, se puede usar análisis de video para determinar la ubicación de un segmento de video actual, y los parámetros de la FDN se pueden ajustar por consiguiente para simular más estrechamente el entorno representado en el video; y/o
- 25 7. un sistema de reproducción de estado sólido puede usar unos ajustes de la FDN diferentes a un dispositivo móvil, por ejemplo los ajustes pueden ser dependientes del dispositivo. Un sistema de estado sólido presente en una sala de estar puede simular un escenario típico de sala de estar (bastante reverberante) con fuentes alejadas, mientras que un dispositivo móvil puede reproducir contenido más cerca del oyente.

30 Algunas implementaciones del virtualizador inventivo incluyen las FDN (por ejemplo, una implementación de la FDN de la Fig. 4) que se configuran para aplicar un retardo fraccional así como un retardo de muestra entera. Por ejemplo, en una implementación tal se conecta un elemento de retardo fraccional en cada tanque de reverberación en serie con una línea de retardo que aplica un retardo entero igual a un número entero de periodos de muestra (por ejemplo, cada elemento de retardo fraccional se posiciona después o de otra manera en serie con una de las líneas de retardo). El retardo fraccional se puede aproximar mediante un desplazamiento de fase (multiplicación compleja de unidad) en cada banda de frecuencias que corresponda a una fracción del periodo de muestra:  $f = \tau/T$ , donde  $f$  es la fracción de retardo,  $\tau$  es el retardo deseado para la banda y  $T$  es el periodo de muestra para la banda. Es bien sabido cómo aplicar el retardo fraccional en el contexto de aplicar reverberación en el dominio QMF.

40 En una primera clase de ejemplos, que no están cubiertos por las reivindicaciones, la invención es un método de virtualización de auriculares para generar una señal binaural en respuesta a un conjunto de canales (por ejemplo, cada uno de los canales, o cada uno de los canales de rango completo de frecuencias) de una señal de entrada de audio multi canal, que incluye los pasos de: (a) aplicar una respuesta al impulso binaural de una sala (BRIR) a cada canal del conjunto (por ejemplo, convolucionando cada canal del conjunto con una BRIR correspondiente a dicho canal, en los subsistemas 100 y 200 de la Fig. 3, o en los subsistemas 12,..., 14, y 15 de la Fig. 2), generando de este modo señales filtradas (por ejemplo, las salidas de los subsistemas 100 y 200 de la Fig. 3, o las salidas de los subsistemas 12,..., 14, y 15 de la Fig. 2), incluyendo mediante el uso al menos una red de retardo de retroalimentación (por ejemplo, las FDN 203, 204,..., 205 de la Fig. 3) para aplicar una reverberación tardía común a una mezcla descendente (por ejemplo, una mezcla descendente monofónica) de los canales del conjunto; y (b) combinar las señales filtradas (por ejemplo, en el subsistema 210 de la Fig. 3, o el subsistema que comprende los elementos 16 y 18 de la Fig. 2) para generar la señal binaural. Normalmente, se usa un banco de FDN para aplicar la reverberación tardía común a la mezcla descendente (por ejemplo con cada FDN aplicando reverberación tardía a una banda de frecuencias diferente). Normalmente, el paso (a) incluye el paso de aplicación a cada canal del conjunto una parte de "respuesta directa y reflexión temprana" de una BRIR de canal único para el canal (por ejemplo, en el subsistema 100 de la Fig. 3 o los subsistemas 12,..., 14 de la Fig. 2), y la reverberación tardía común se ha generado para emular los macro atributos colectivos de las partes de reverberación tardía de al menos algunas (por ejemplo, todas) de las BRIR de canal único.

En ejemplos típicos en la primera clase, cada una de las FDN se implementa en el dominio del filtro espejo en cuadratura complejo híbrido (HCQMF) o en el dominio del filtro espejo en cuadratura (QMF), y en algunos de dichos ejemplos, los atributos acústicos espaciales dependientes de la frecuencia de la señal binaural se controlan (por

ejemplo, usando el subsistema 209 de control de la Fig. 3) controlando la configuración de cada FDN empleada para aplicar la reverberación tardía. Normalmente, se usa una mezcla descendente monofónica de los canales (por ejemplo, la mezcla descendente generada por el subsistema 201 de la Fig. 3) como la entrada de las FDN para un procesamiento binaural eficiente del contenido de audio de la señal multi canal. Normalmente, el proceso de mezcla descendente está controlado en base a la distancia a la fuente para cada canal (esto es, la distancia entre la supuesta fuente del contenido de audio y la supuesta posición del usuario) y depende del manejo de las respuestas directas correspondientes a las distancias a la fuente para conservar la estructura temporal y de nivel de cada BRIR (esto es, cada BRIR determinada mediante las partes de respuesta directa y reflexión temprana de una BRIR de canal único para un canal, junto con la reverberación tardía para una mezcla descendente que incluye el canal). Aunque los canales a mezcla de manera descendente se pueden alinear en el tiempo y escalar de diferentes maneras durante la mezcla descendente, la relación temporal y de nivel entre las partes de respuesta directa, de reflexión temprana, y de reverberación tardía común de la BRIR para cada canal se deberían mantener. En las realizaciones en las que se usa un único banco de FDN para generar la parte de reverberación tardía común para todos los canales que se mezclan de manera descendente (para generar una mezcla descendente), es necesario que se aplique una ganancia y un retardo apropiados (a cada canal que se mezcla de manera descendente) durante la generación de la mezcla descendente.

Las realizaciones típicas en esta clase incluyen un paso de ajuste (por ejemplo, usando el subsistema 209 de control de la Fig. 3) de los coeficientes de la FDN correspondientes a los atributos dependientes de la frecuencia (por ejemplo, el tiempo de decaimiento de la reverberación, la coherencia interaural, la densidad modal y la relación directa a tardía). Esto permite una mejor coincidencia de los entornos acústicos y unas salidas de sonido más naturales.

En una clase adicional de ejemplos que no están cubiertos por las reivindicaciones, la invención es un método para generar una señal binaural en respuesta a una señal de entrada de audio multi canal, aplicando una respuesta al impulso binaural de una sala (BRIR) a cada canal (por ejemplo, convolucionando cada canal con la BRIR correspondiente) de un conjunto de canales de la señal de entrada (por ejemplo, cada uno de los canales de la señal de entrada o cada uno de los canales de rango completo de frecuencias de la señal de entrada), incluyendo: procesar cada canal del conjunto en una primera ruta de procesamiento (por ejemplo, implementada por el subsistema 100 de la Fig. 3 o los subsistemas 12, ..., 14 de la Fig. 2) que se configura para modelar, y aplicar a dicho cada canal, una parte de respuesta directa y reflexión temprana (por ejemplo, la EBRIR aplicada por el subsistema 12, 14, o 15 de la Fig. 2) de un BRIR de canal único para el canal; y procesar una mezcla descendente (por ejemplo, una mezcla descendente monofónica) de los canales del conjunto en una segunda ruta de procesamiento (por ejemplo, implementada por el subsistema 200 de la Fig. 3 o el subsistema 15 de la Fig. 2), en paralelo con la primera ruta de procesamiento. La segunda ruta de procesamiento se configura para modelar, y aplicar a la mezcla descendente, una reverberación tardía común (por ejemplo, la LBRIR aplicada por el subsistema 15 de la Fig. 2). Normalmente, la reverberación tardía común emula los macro atributos colectivos de las partes de reverberación tardía de al menos algunas (por ejemplo, todas) de las BRIR de canal único. Normalmente la segunda ruta de procesamiento incluye al menos una FDN (por ejemplo, una FDN para cada una de las múltiples bandas de frecuencias). Normalmente, se usa una mezcla descendente mono como entrada a todos los tanques de reverberación para cada FDN implementada por la segunda ruta de procesamiento. Normalmente, se proporcionan los mecanismos (por ejemplo el subsistema 209 de control de la Fig. 3) para el control sistemático de los macro atributos de cada FDN para simular mejor los entornos acústicos y producir una virtualización binaural más natural del sonido. Ya que la mayoría de dichos atributos son dependientes de la frecuencia, cada FDN se implementa normalmente en el dominio del filtro espejo en cuadratura complejo híbrido (HCQMF), el dominio de la frecuencia, el dominio, u en otro dominio del banco de filtros, y se usa una FDN diferente para cada banda de frecuencias. Un beneficio principal de implementar las FDN en el dominio del banco de filtros es permitir la aplicación de reverberación con las propiedades de la reverberación dependientes de la frecuencia. En diversas realizaciones, las FDN se implementan en cualquiera de entre una amplia variedad de dominios de banco de filtros, usando cualquiera de una variedad de bancos de filtros, que incluye, pero no se limita a los filtros espejo en cuadratura (QMF), los filtros de respuesta al impulso finita (filtros FIR), los filtros de respuesta al impulso infinita (filtros IIR), o los filtros de cruce.

Algunos ejemplos de la primera clase (y de la segunda clase) implementan una o más de las siguientes características:

1. una implementación de una FDN (por ejemplo, la implementación de la FDN de la Fig. 4) en el dominio del banco de filtros (por ejemplo, en el dominio del filtro espejo en cuadratura complejo híbrido), o una implementación de una FDN en el dominio del banco de filtros híbridos y una implementación del filtro de reverberación tardía en el dominio del tiempo (por ejemplo, la estructura descrita con referencia a la Fig. 8), que normalmente permite un ajuste independiente de los parámetros y/o los ajustes de la FDN para cada banda de frecuencias (que permite un control simple y flexible de los atributos acústicos dependientes de la frecuencia), por ejemplo, proporcionando la capacidad de variar los retardos del tanque de reverberación en diferentes bandas para cambiar la densidad modal como función de la frecuencia;

2. El proceso de mezcla descendente específico, empleado para generar (desde la señal de audio de entrada multi canal) la señal mezclada de manera descendente (por ejemplo, mezcla descendente monofónica) procesada en la segunda ruta de procesamiento, depende de la distancia a la fuente de cada canal y el manejo de la respuesta directa para mantener el nivel apropiado y la relación de temporización entre las respuestas directa y tardía;
3. Un filtro paso todo (por ejemplo, el APF 301 de la Fig. 4) se aplica en la segunda ruta de procesamiento (por ejemplo, en la entrada o la salida de un banco de FDN) para introducir diversidad de fase y aumentar la densidad de eco sin cambiar el espectro y/o el timbre de la reverberación resultante;
4. Se implementan retardos fraccionales en la ruta de retroalimentación de cada FDN en una estructura multi tasa, con un valor complejo, para superar los problemas relacionados con los retardos cuantificados a la cuadrícula de factor de disminución de muestras;
5. En las FDN, las salidas del tanque de reverberación se mezclan de manera descendente de manera lineal directamente en los canales binaurales (por ejemplo, mediante la matriz 312 de la Fig. 4) usando los coeficientes de mezcla de salida que se establecen en base a la coherencia interaural deseada en cada banda de frecuencias. De manera opcional, la correspondencia de los tanques de reverberación con los canales de salida binaural está alternando entre bandas de frecuencias para alcanzar un retardo equilibrado entre los canales binaurales. También de manera opcional, se aplican factores de normalización a la salidas del tanque de reverberación para igualar sus niveles mientras que se conserva el retardo fraccional y la energía general;
6. El tiempo de decaimiento de la reverberación dependiente de la frecuencia se controla (por ejemplo, usando el subsistema 209 de control de la Fig. 3) estableciendo una combinación apropiada de ganancias y retardos del tanque de reverberación en cada banda de frecuencias para simular salas reales;
7. se aplica un factor de escala (por ejemplo, mediante los elementos 306 y 309 de la Fig. 4) por banda de frecuencias (por ejemplo, en ya sea la entrada o la salida de la ruta de procesamiento relevante), para:
  - controlar la relación directa a tardía dependiente de la frecuencia (DLR) que coincide con la de una sala real (se puede usar un modelo simple para calcular el factor de escala requerido en base a la DLR y el tiempo de decaimiento de la reverberación objetivos, por ejemplo, T60);
  - proporcionar una atenuación de la baja frecuencia para mitigar el exceso de artefactos de peinado; y/o aplicar un conformado espectral de campo difuso a las respuestas de la FDN;
8. Se implementan modelos paramétricos simples (por ejemplo, mediante el subsistema 209 de control de la Fig. 3) para controlar los atributos dependientes de la frecuencia esenciales de la reverberación tardía, tales como el tiempo de decaimiento de la reverberación, la coherencia interaural, y/o la relación directa a tardía.

En algunas realizaciones (por ejemplo, para las aplicaciones en las que la latencia del sistema es crítica y el retardo provocado por los bancos de filtros de análisis y de síntesis es prohibitivo), las estructuras FDN en el dominio del banco de filtros de las realizaciones típicas del sistema inventivo (por ejemplo, la FDN de la Fig. 4 en cada banda de frecuencias) son reemplazadas por estructuras FDN implementadas en el dominio del tiempo (por ejemplo, la FDN 220 de la Fig. 10, que se puede implementar como se muestra en la Fig. 9). En las realizaciones en el dominio del tiempo del sistema inventivo, los subsistemas de las realizaciones en el dominio del banco de filtros que aplican un factor ( $G_{\text{entrada}}$ ) de ganancia de entrada, las ganancias ( $g_i$ ) del tanque de reverberación, y las ganancias ( $1/|g_i|$ ) de normalización son reemplazadas por filtros (y/o elementos de ganancia) en el dominio del tiempo para permitir controles dependientes de la frecuencia. La matriz de mezcla descendente de salida de una implementación en el dominio del banco de filtros (por ejemplo, la matriz 312 de mezcla de salida de la Fig. 4) es reemplazada (en las realizaciones en el dominio del tiempo típicas) por un conjunto de salida de filtros en el dominio del tiempo (por ejemplo, los elementos 500-503 de la implementación de la Fig. 11 del elemento 424 de la Fig. 9). A diferencia de los otros filtros de las realizaciones en el dominio del tiempo típicas, la respuesta de fase de este conjunto de salida de filtros normalmente es crítica (ya que la conservación de la energía y la coherencia interaural se podrían ver afectados por la respuesta de fase). En algunas realizaciones en el dominio del tiempo, los retardos del tanque de reverberación varían (por ejemplo, varían ligeramente) desde sus valores en una implementación correspondiente en el dominio del banco de filtros (por ejemplo, para evitar compartir el paso del banco de filtros como un factor común).

La FIG. 10 es un diagrama de bloques de una realización del sistema de virtualización de auriculares inventivo similar al de la Fig. 3, excepto en que los elementos 202-207 del sistema de la Fig. 3 se han reemplazado en el sistema de la Fig. 10 por una FDN 220 única que se implementa en el dominio del tiempo (por ejemplo, la FDN 220 de la Fig. 10 se puede implementar como la FDN de la Fig. 9). En la Fig. 10, se emiten dos señales (canal izquierdo y derecho) en el dominio del tiempo desde el subsistema 100 de respuesta directa y reflexión temprana, y se emiten dos señales (canal izquierdo y derecho) en el dominio del tiempo desde el subsistema 221 de procesamiento de la reverberación tardía. El elemento 210 de adición se acopla a las salidas de los subsistemas 100 y 200. El elemento

210 se configura para combinar (mezclar) las salidas del canal izquierdo de los subsistemas 100 y 221 para generar el canal izquierdo, L, de la salida de la señal de audio binaural del virtualizador de la Fig. 10, y para combinar (mezclar) las salidas del canal derecho de los subsistemas 100 y 221 para generar el canal derecho, R, de la salida de la señal de audio binaural del virtualizador de la Fig. 10. Se puede implementar el elemento 210 para simplemente sumar las muestras correspondientes del canal izquierdo emitidas desde los subsistemas 100 y 221 para generar el canal izquierdo de la señal de salida binaural, y para simplemente las muestras correspondientes del canal derecho emitidas desde los subsistemas 100 y 221 para generar el canal derecho de la señal de salida binaural, suponiendo que se implementan los ajustes de nivel y los alineamientos de tiempo apropiados en los subsistemas 100 y 221.

En el sistema de la Fig. 10, la señal de entrada de audio multi canal (que tiene los canales,  $X_i$ ) se dirige hacia, y experimenta el procesamiento en, dos rutas de procesamiento paralelas: una a través del subsistema 100 de procesamiento de la respuesta directa y la reflexión temprana; la otra a través del subsistema 221 de procesamiento de la reverberación tardía. El sistema de la Fig. 10 se configura para aplicar una BRIR, a cada canal  $X_i$ . Cada  $BRIR_i$  se puede descomponer en dos partes: una parte de respuesta directa y reflexión temprana (aplicada por el subsistema 100), y una parte de reverberación tardía (aplicada por el subsistema 221). En funcionamiento, el subsistema 100 de respuesta directa y reflexión temprana genera por tanto la respuesta directa y las partes de las reflexiones tempranas de la señal de audio binaural que se emite desde el virtualizador, y el subsistema 221 de procesamiento de la reverberación tardía (“generador de la reverberación tardía”) genera por tanto la parte de reverberación tardía de la señal de audio binaural que se emite desde el virtualizador. Las salidas de los subsistemas 100 y 221 son mezcladas (mediante el subsistema 210) para generar la señal de audio binaural, que se impone normalmente desde el subsistema 210 a un sistema de procesamiento (no mostrado) en el que ésta experimenta un procesamiento binaural para su reproducción por los auriculares.

El subsistema 201 de mezcla descendente (del subsistema 221 de procesamiento de la reverberación) se configura para mezclar de manera descendente los canales de la señal de entrada multi canal en una mezcla descendente mono (que es una señal en el dominio del tiempo), y la FDN 220 se configura para aplicar la parte de la reverberación tardía a la mezcla descendente mono.

Con referencia a la Fig. 9, describiremos a continuación un ejemplo de una FDN en el dominio del tiempo que se puede emplear como la FDN 220 del virtualizador de la Fig. 10. La FDN de la Fig. 9 incluye el filtro 400 de entrada que se acopla para recibir una mezcla descendente mono (por ejemplo, generada por el subsistema 201 del sistema de la Fig. 10) de todos los canales de una señal de entrada de audio multi canal. La FDN de la Fig. 9 incluye también un filtro 401 paso todo (APF) (que corresponde al APF 301 de la Fig. 4) acoplado a la salida del filtro 400, el elemento 401A de ganancia de entrada acoplado a la salida del filtro 401, los elementos 402, 403, 404, y 405 de adición (que corresponden a los elementos 302, 303, 304, y 305 de la Fig. 4) acoplados a la salida del elemento 401A, y cuatro tanques de reverberación. Cada tanque de reverberación se acopla a la salida de uno diferente de los elementos 402, 403, 404, y 405, y comprende uno de entre los filtros 406, y 406A, 407 y 407A, 408 y 408A, y 409 y 409A de reverberación, una de las líneas 410, 411, 412, y 413 de retardo (correspondientes a las líneas 307 de retardo de la Fig. 4) acoplada a éste, y uno de los elementos 417, 418, 419, y 420 de ganancia acoplado a la salida de una de las líneas de retardo.

La matriz 415 unitaria (correspondiente a la matriz 308 unitaria de la Fig. 4, y normalmente implementada para ser idéntica a la matriz 308) se acopla a las salidas de las líneas 410, 411, 412, y 413 de retardo. La matriz 415 se configura para imponer una salida de retroalimentación a una segunda entrada de cada uno de los elementos 402, 403, 404, y 405.

Cuando el retardo ( $n_1$ ) aplicado por la línea 410 es menor que el aplicado ( $n_2$ ) por la línea 411, el retardo aplicado por la línea 411 es menor que el aplicado ( $n_3$ ) por la línea 412, y el retardo aplicado por la línea 412 es menor que el aplicado ( $n_4$ ) por la línea 413, las salidas 417 y 419 de los elementos de ganancia (del primer y tercer banco de reverberación) se imponen a las entradas del elemento 422 de adición, y las salidas 418 y 420 de los elementos de ganancia (del segundo y cuarto banco de reverberación) se imponen a las entradas del elemento 423 de adición. La salida del elemento 422 se impone en una entrada del filtro 424 de IACC y mezclado, y la salida del elemento 423 se impone en la otra entrada de la etapa 424 de filtrado y mezcla IACC.

Ejemplos de las implementaciones de los elementos 417 - 420 de ganancia y los elementos 422, 423, y 424 de la Fig. 9 se describirán con referencia a la implementación típica de los elementos 310 y 311 y la matriz 312 de mezcla de salida de la Fig. 4. La matriz 312 de mezcla de salida de la Fig. 4 (también identificada como  $M_{\text{salida}}$ ) es una matriz  $2 \times 2$  configurada para mezclar los canales binaurales no mezclados (las salidas de los elementos 310 y 311, respectivamente) desde la distribución inicial para generar los canales de salida binaural izquierdo y derecho (las señales del oído izquierdo, “L”, y del oído derecho, “R”, impuestas en la salida de la matriz 312) que tengan la coherencia interaural deseada. Esta distribución inicial es implementada por los elementos 310 y 311, cada uno de los cuales combina dos salidas del tanque de reverberación para generar uno de los canales binaurales no mezclados, con la salida del tanque de reverberación teniendo el menor retardo que se impone a la entrada del elemento 310 y la salida del tanque de reverberación teniendo el segundo menor retardo impuesto a la entrada del elemento 311. Los elementos 422 y 423 de la realización de la Fig. 9 realizan el mismo tipo de distribución inicial

(sobre las señales en el dominio del tiempo impuestas a sus entradas) que los elementos 310 y 311 (en cada banda de frecuencias) de la realización de la Fig. 4 realizan sobre los flujos de componentes en el dominio del banco de filtros (en la banda de frecuencias relevante) impuestas en sus entradas.

5 Los canales binaurales no mezclados (salida de los elementos 310 y 311 de la Fig. 4, o de los elementos 422 y 423 de la Fig. 9), que están cerca de no estar correlados ya que no están compuestos de ninguna salida del tanque de reverberación común, se pueden mezclar (por la matriz 312 de la Fig. 4 o la etapa 424 de la Fig. 9) para implementar un patrón de distribución que alcance una coherencia interaural deseada para los canales de salida binaurales izquierdo y derecho. Sin embargo, ya que los retardos del tanque de reverberación son diferentes en cada FDN (esto es, la FDN de la Fig. 9, o la FDN implementada para cada banda de frecuencias diferente en la Fig. 4), un canal binaural no mezclado (la salida de uno de los elementos 310 y 311, o 422 y 423) guía constantemente al otro canal binaural no mezclado (la salida del otro de los elementos 310 y 311, o 422 y 423).

10 Por tanto, en la realización de la Fig. 4, si la combinación de los retardos del tanque de reverberación y el patrón de distribución es idéntico a lo largo de todas las bandas de frecuencias, resultaría en un sesgo de la imagen de sonido. Este sesgo se puede mitigar si al patrón de distribución se alterna a lo largo de las bandas de frecuencias de manera tal que los canales de salida binaural mezclados se guíen y arrastren el uno al otro en las bandas de frecuencias alternas. Por ejemplo, si la coherencia interaural deseada es Coh, donde  $|Coh| \leq 1$ , se puede implementar la matriz 312 de mezcla descendente de salida en las bandas impares para multiplicar las dos entradas impuestas a ésta por una matriz que tenga la siguiente forma:

$$M_{\text{salida}} = \begin{bmatrix} \cos \beta & \text{sen} \beta \\ \text{sen} \beta & \cos \beta \end{bmatrix}, \text{ donde } \beta = \arcsen(Coh)/2,$$

20 y se puede implementar la matriz 312 de mezcla descendente de salida en las bandas de frecuencias pares para multiplicar las dos entradas impuestas a ésta por una matriz que tenga la siguiente forma:

$$M_{\text{sal, alt}} = \begin{bmatrix} \text{sen} \beta & \cos \beta \\ \cos \beta & \text{sen} \beta \end{bmatrix}$$

donde  $\beta = \arcsen(Coh)/2$ .

25 De manera alternativa, el sesgo de la imagen de sonido anteriormente indicado en los canales de salida binaurales puede ser mitigado mediante la implementación de la matriz 312 para que sea idéntica en las FDN para todas las bandas de frecuencias, si el orden de las entradas del canal se conmuta para alternar algunas de las bandas de frecuencias (por ejemplo el elemento 310 de salida se puede imponer a la primera entrada de la matriz 312 y la salida del elemento 311 se puede imponer a la segunda entrada de la matriz 312 en las bandas de frecuencias impares, y la salida del elemento 311 se puede imponer a la primera entrada de la matriz 312 y la salida del elemento 310 se puede imponer a la segunda entrada de la matriz 312 en las bandas de frecuencias pares).

30 En la realización de la Fig. 9 (y otras realizaciones en el dominio del tiempo de una FDN del sistema inventivo), no es trivial alternar la distribución basada en frecuencia para abarcar el sesgo de la imagen de sonido que de otra manera resultaría cuando la salida del canal binaural no mezclado del elemento 422 guíe (o retrase) constantemente la salida del canal binaural no mezclado del elemento 423. Este sesgo de la imagen de sonido es abarcado en una realización en el dominio del tiempo típica de una FDN del sistema inventivo de una manera diferente que la abarcada normalmente en la realización en el dominio del banco de filtros de una FDN del sistema inventivo. Específicamente, en la realización de la Fig. 9 (y algunas otras realizaciones en el dominio del tiempo de una FDN del sistema inventivo), las ganancias relativas de los canales binaurales no mezclados (por ejemplo, la salida de los elementos 422 y 423 de la Fig. 9) son determinadas por los elementos de ganancia (por ejemplo, los elementos 417, 418, 419, y 420 de la Fig.9) para compensar el sesgo de la imagen de sonido que de otra manera resultaría debido a la temporización no equilibrada observada. Mediante la implementación de un elemento de ganancia (por ejemplo, el elemento 417) para atenuar la señal de llegada más temprana (que se ha distribuido a un lado, por ejemplo, mediante el elemento 422) y la implementación del elemento de ganancia (por ejemplo, el elemento 418) para aumentar la siguiente señal más temprana (que se ha distribuido al otro lado, por ejemplo, mediante el elemento 423), la imagen estéreo se vuelve a centrar. Por tanto, el tanque de reverberación que incluye el elemento 417 de ganancia aplica una primera ganancia a la salida del elemento 417, y el tanque de reverberación que incluye el elemento 418 de ganancia aplica una segunda ganancia (diferente de la primera ganancia) a la salida del elemento 418. para que la primera ganancia y la segunda ganancia atenúen el primer canal binaural no mezclado (salida del elemento 422) en relación al segundo canal binaural no mezclado (salida del elemento 423).

50 Más específicamente, en una implementación típica de la FDN de la Fig. 9, las cuatro líneas 410, 411, 412, y 413 de retardo tienen una longitud aumentada, con valores  $n_1$ ,  $n_2$ ,  $n_3$ , y  $n_4$ , respectivamente aumentados. En esta

implementación, el filtro 417 aplica una ganancia de  $g_1$ . Por tanto, la salida del filtro 417 es una versión retardada de la entrada de la línea 410 de retardo a la que se ha aplicado la ganancia  $g_1$ . De manera similar, el filtro 418 aplica una ganancia de  $g_2$ , el filtro 419 aplica una ganancia de  $g_3$ , y el filtro 420 aplica una ganancia de  $g_4$ . Por tanto, la salida del filtro 418 es una versión retardada de la entrada a la línea 411 de retardo a la que se ha aplicado una ganancia de  $g_2$ , y la salida del filtro 419 es una versión retardada de la entrada a la línea 412 de retardo a la que se le ha aplicado una ganancia  $g_3$ , y la salida del filtro 420 es una versión retardada de la entrada a la línea 413 de retardo a la que se ha aplicado una ganancia  $g_4$ .

En esta implementación, la elección de los siguientes valores de ganancia puede resultar en un sesgo indeseable de la imagen de sonido de salida (indicado por la salida de los canales binaurales del elemento 424) a un lado (esto es, al canal de la izquierda o de la derecha):  $g_1 = 0,5$ ,  $g_2 = 0,5$ ,  $g_3 = 0,5$  y  $g_4 = 0,5$ . De acuerdo con una realización de la invención, los valores  $g_1$ ,  $g_2$ ,  $g_3$ , y  $g_4$  de ganancia (aplicados por los elementos 417, 418, 419, y 420, respectivamente) se eligen como sigue para centrar la imagen de sonido:  $g_1 = 0,38$ ,  $g_2 = 0,6$ ,  $g_3 = 0,5$  y  $g_4 = 0,5$ . Por tanto, la imagen estéreo de salida se vuelve a centrar de acuerdo con una realización de la invención mediante la atenuación de la señal de llegada más temprana (que se ha distribuido a una lado, mediante el elemento 422 en el ejemplo) en relación a la segunda señal de llegada más temprana (esto es, eligiendo  $g_1 < g_3$ ), y aumentando la segunda señal más temprana (que se ha distribuido al otro lado, mediante el elemento 423 en el ejemplo), en relación a la señal de llegada más tardía (esto es, eligiendo  $g_4 < g_2$ ).

Las implementaciones típicas de la FDN en el dominio del tiempo de la Fig. 9 tienen las siguientes diferencias y similitudes a la FDN en el dominio del banco de filtros (dominio CQMF) de la Fig. 4:

la misma matriz de retroalimentación unitaria, A (matriz 308 de la Fig. 4 y matriz 415 de la Fig. 9);

retardos del tanque de reverberación similares,  $n_i$  (esto es, los retardos en la implementación CQMF de la Fig. 4 pueden ser  $n_1 = 17 \cdot 64 T_s = 1088 T_s$ ,  $n_2 = 21 \cdot 64 T_s = 1344 T_s$ ,  $n_3 = 26 \cdot 64 T_s = 1664 T_s$ , y  $n_4 = 29 \cdot 64 T_s = 1856 T_s$ , donde  $1/T_s$  es la tasa de muestreo ( $1/T_s$  normalmente es igual a 48 KHz), en donde los retardos en la implementación en el dominio del tiempo pueden ser:  $n_1 = 1089 T_s$ ,  $n_2 = 1345 T_s$ ,  $n_3 = 1663 T_s$ , y  $n_4 = 185 T_s$ . Observe que en las implementaciones CQMF típicas existe la limitación práctica de que cada retardo es un múltiplo entero de la duración de un bloque de 64 muestras (la tasa de muestreo es normalmente de 48 KHz), pero en el dominio del tiempo no existe más flexibilidad en cuanto a la elección de cada retardo y por tanto más flexibilidad en cuanto a la elección del retardo para cada tanque de reverberación);

implementaciones del filtro paso todo similares (esto es, implementaciones similares del filtro 301 de la Fig. 4 y del filtro 401 de la Fig. 9). Por ejemplo, el filtro paso todo puede ser implementado poniendo en cascada varios (por ejemplo, 3) filtros paso todo. Por ejemplo, cada filtro paso todo puesto en cascada puede ser de la forma

$$\frac{g - z^{-n_i}}{1 - g \cdot z^{-n_i}}$$
 donde  $g = 0,6$ . El filtro 301 paso todo de la Fig. 4 puede ser implementado mediante tres filtros paso todo puestos en cascada con unos retardos adecuados de los bloques de muestreo (por ejemplo,  $n_1 = 64 T_s$ ,  $n_2 = 128 T_s$ , y  $n_3 = 196 T_s$ ), donde todos los filtros 401 paso todo de la Fig. 9 (los filtros paso todo en el dominio del tiempo) pueden ser implementados por tres filtros paso todo en cascada con retardos similares (por ejemplo,  $n_1 = 61 T_s$ ,  $n_2 = 127 T_s$ , y  $n_3 = 191 T_s$ ).

En algunas implementaciones de la FDN en el dominio del tiempo de la Fig. 9, el filtro 400 de entrada se implementa para que provoque que la relación directa a tardía (DLR) de la BRIR sea aplicada por el sistema de la Fig. 9 para hacer coincidir (al menos de manera sustancial) una DLR objetivo, y para que la DLR de la BRIR a ser aplicada por un virtualizador que incluya el sistema de la Fig. 9 (por ejemplo, el virtualizador de la Fig. 10) pueda ser cambiada por el filtro 400 de reemplazo (o controlar una configuración del filtro 400). Por ejemplo, en algunas realizaciones, el filtro 400 se implementa como una cascada de filtros (por ejemplo, un primer filtro 400A y un segundo filtro 400B, acoplados como se muestra en la Fig. 9A) para implementar la DLR objetivo y de manera opcional también para implementar el control de la DLR deseada. Por ejemplo, los filtros de la cascada son filtros IIR (por ejemplo el filtro 400A es un filtro paso alto Butterworth de primer orden (un filtro IIR) configurado para hacer coincidir las características de baja frecuencia objetivo, y el filtro 400B es un filtro IIR paso bajo configurado para hacer coincidir las características de alta frecuencia objetivo). Como otro ejemplo, los filtros de la cascada son filtros IIR y FIR (por ejemplo, el filtro 400A es un filtro paso alto Butterworth de segundo orden (un filtro IIR) configurado para hacer coincidir las características de baja frecuencia, y el filtro 400B es un filtro FIR de orden 14 configurado para hacer coincidir las características de alta frecuencia objetivo). Normalmente, la señal directa es fija, y el filtro 400 modifica la señal tardía para alcanzar la DLR objetivo. El filtro 401 paso todo (APF) se implementa preferiblemente para realizar la misma función que hace el APF 301 de la Fig. 4, principalmente para introducir diversidad de fase y aumentar la densidad de eco para generar una salida de la FDN con un sonido más natural. El APF 401 normalmente controla la respuesta de fase mientras que el filtro 400 de entrada controla la respuesta de amplitud.

En la Fig. 9, el filtro 406 y el elemento 406A de ganancia juntos implementan un filtro de reverberación, el filtro 407 y el elemento 407A de ganancia juntos implementan otro filtro de reverberación, el filtro 408 y el elemento 408A de ganancia juntos implementan otro filtro de reverberación, y el filtro 409 y el elemento 409A de ganancia juntos

implementan otro filtro de reverberación. Cada uno de los filtros 406, 407, 408, y 409 de la Fig. 9 se implementan preferiblemente como un filtro con un valor de ganancia máximo cercano a uno (ganancia unidad), y cada uno de los elementos 406A, 407A, 408A, y 409A de ganancia se configuran para aplicar una ganancia de decaimiento a la salida del correspondiente de los filtros 406, 407, 408, y 409 que hace coincidir el decaimiento deseado (después del retardo del tanque de reverberación relevante,  $n_i$ ). Específicamente, el elemento 406A de ganancia se configura para aplicar una ganancia de decaimiento (gananciadecaimiento<sub>1</sub>) a la salida del filtro 406 para provocar que la salida del elemento 406A tenga una ganancia tal que la salida de la línea 410 de retardo (después del retardo del tanque de reverberación,  $n_1$ ) tenga una primera ganancia de decaimiento objetivo, el elemento 407A de ganancia se configura para aplicar una ganancia de decaimiento (gananciadecaimiento<sub>2</sub>) a la salida del filtro 407 para provocar que la salida del elemento 407A tenga una ganancia tal que la salida de la línea 411 de retardo (después del retardo del tanque de reverberación,  $n_2$ ) tenga una segunda ganancia de decaimiento objetivo, el elemento 408A de ganancia se configura para aplicar una ganancia de decaimiento (gananciadecaimiento<sub>3</sub>) a la salida del filtro 408 para provocar que la salida del elemento 408A tenga una ganancia tal que la salida de la línea 412 de retardo (después del retardo del tanque de reverberación,  $n_3$ ) tenga una tercera ganancia de decaimiento objetivo y se configura el elemento 409A de ganancia se configura para aplicar una ganancia de decaimiento (gananciadecaimiento<sub>4</sub>) a la salida del filtro 409 para provocar que la salida del elemento 409A tenga una ganancia tal que la salida de la línea 413 de retardo (después del retardo del tanque de reverberación,  $n_4$ ) tenga una cuarta ganancia de decaimiento objetivo.

Cada uno de los filtros 406, 407, 408, y 409, y cada uno de los elemento 406A, 407A, 408A, y 409A del sistema de la Fig. 9 se implementan de manera preferible (con cada uno de los filtros 406, 407, 408, y 409 implementados preferiblemente como un filtro IIR, por ejemplo un filtro limitador o una cascada de filtros limitadores) para alcanzar una característica T60 objetivo de la BRIR a ser aplicada mediante un virtualizador que incluye el sistema de la Fig. 9 (por ejemplo el virtualizador de la Fig. 10), donde "T60" denota el tiempo ( $T_{60}$ ) de decaimiento de reverberación. Por ejemplo, en algunas realizaciones cada uno de los filtros 406, 407, 408, y 409 se implementan como un filtro limitador (por ejemplo, un filtro limitador que tenga  $Q = 0,3$  y una frecuencia de corte de 500 Hz, para alcanzar la característica  $T_{60}$  mostrada en la Fig. 13, en la que  $T_{60}$  tiene unidades de segundos) o como una cascada de dos filtros atenuadores IIR (por ejemplo, que tengan frecuencias de corte de 100 Hz y 1000Hz, para alcanzar la característica  $T_{60}$  mostrada en la Fig. 14, en la que  $T_{60}$  tiene unidades de segundos). La forma de cada uno de los filtros atenuadores se determina para hacer que coincida la curva de cambio deseada desde la baja frecuencia a la alta frecuencia. Cuando el filtro 406 se implementa como un filtro atenuador (o una cascada de filtros atenuadores), el filtro de reverberación que comprende el filtro 406 y el elemento 406A de ganancia es también un filtro atenuador (o una cascada de filtros atenuadores). De la misma manera, cuando cada uno de los filtros 407, 408, y 409 se implementa como un filtro atenuador (o una cascada de filtros atenuadores), cada filtro de reverberación que comprende el filtro 407 (o 408 o 409) y el correspondiente elemento de ganancia (407A, 408A, o 409A) es también un filtro atenuador (o una cascada de filtros atenuadores).

La Fig. 9B es un ejemplo de filtro 406 implementado como una cascada de un primer filtro 406B atenuador y un segundo filtro 406C atenuador, acoplados como se muestra en la Fig. 9B. Cada uno de los filtros 407, 408, y 409 se puede implementar como la implementación de la Fig. 9B del filtro 406.

En algunas realizaciones, las ganancias de decaimiento (gananciadecaimiento<sub>i</sub>) aplicada por los elementos 406A, 407A, 408A, y 409A se determina como sigue:

$$\text{gananciadecaimiento}_i = 10^{((-60 \cdot (n_i / F_s) / T) / 20)}$$

donde  $i$  es el índice del tanque de reverberación (esto es, el elemento 406A aplica la gananciadecaimiento<sub>1</sub>, el elemento 407A aplica la gananciadecaimiento<sub>2</sub>, y así sucesivamente),  $n_i$  es el retardo del  $i$ -ésimo tanque de reverberación (por ejemplo,  $n_1$  es el retardo aplicado por la línea 410 de retardo).  $F_s$  es la tasa de muestreo,  $T$  es el tiempo ( $T_{60}$ ) de decaimiento de reverberación deseado a una predeterminada baja frecuencia.

La Fig. 11 es un diagrama de bloques de una realización de los siguientes elementos de la Fig. 9: los elementos 422 y 423, y la etapa 424 de filtrado y de mezcla IACC (coeficiente de correlación cruzada interaural). El elemento 422 se acopla y configura para sumar las salidas de los filtros 417 y 419 (de la Fig. 9) y para imponer la señal sumada a la entrada del filtro 500 atenuador de frecuencias bajas, y el elemento 423 se acopla y configura para sumar las salidas de los filtros 418 y 420 (de la Fig. 9) y para imponer la señal sumada a la entrada del filtro 501 paso alto. Las salidas de los filtros 500 y 501 se suman (se mezclan) en el elemento 502 para generar la señal de salida del oído izquierdo binaural, y las salidas de los filtros 500 y 501 se mezclan en el elemento 503 para generar la señal de salida del oído derecho binaural. Los elementos 502 y 503 mezclan (suman y restan) las salidas filtradas de los filtros 500 y 501 para generar las señales de salida binaurales que alcancen (dentro de una precisión aceptable) la característica IACC objetivo. En la realización de la Fig. 11, cada uno de entre el filtro 500 atenuador de frecuencias bajas y el filtro 501 paso alto se implementa normalmente como un filtro IIR de primer orden. En un ejemplo en el que los filtros 500 y 501 tienen dicha implementación, la realización de la Fig. 11 puede alcanzar la característica IACC ejemplar señalada como curva "I" en la Fig. 12, que es una buena coincidencia con la característica IACC objetivo señalada como "I<sub>T</sub>" en la Fig. 12.

La Fig. 11A es una gráfica de la respuesta (R1) en frecuencia de una implementación típica del filtro 500 de la Fig. 11, la respuesta (R2) en frecuencia de una implementación típica del filtro 501 de la Fig. 11, y la respuesta de los filtros 500 y 501 conectados en paralelo. Es evidente a partir de la FIG. 11A, que la respuesta combinada es de manera deseable plana a lo largo del rango 100 Hz - 10.000 Hz.

5 Por tanto, en una clase de ejemplos que no están cubiertos por las reivindicaciones, la invención es un sistema (por ejemplo, el de la Fig. 10) y el método para generar una señal binaural (por ejemplo, la salida del elemento 210 de la Fig. 10) en respuesta a un conjunto de canales de una señal de entrada de audio multi canal, que incluye mediante la aplicación de una respuesta al impulso binaural de una sala (BRIR) a cada canal del conjunto, generando de este modo señales filtradas, incluyendo el uso de una red de retardo de retroalimentación única (FDN) para aplicar una reverberación tardía común a una mezcla descendente de los canales del conjunto; y combinando las señales filtradas para generar la señal binaural. La FDN se implementa en el dominio del tiempo. En algunos de dichos ejemplos, la FDN en el dominio del tiempo (por ejemplo, la FDN 220 de la Fig. 10, configurada como en la Fig. 9) incluye:

10 un filtro de entrada (por ejemplo, el filtro 400 de la Fig. 9) que tiene una entrada acoplada para recibir la mezcla descendente, en donde el filtro de entrada se configura para generar una primera mezcla descendente filtrada en respuesta a la mezcla descendente;

un filtro paso todo (por ejemplo, el filtro 401 de la Fig. 9) que tiene una entrada acoplada para recibir la mezcla descendente, en donde el filtro de entrada se configura para generar una primera mezcla descendente filtrada en respuesta a la mezcla descendente;

20 un subsistema de aplicación de la reverberación (por ejemplo, todos los elementos de la Fig. 9 distintos de los elementos 400, 401, y 424), que tienen una primera salida (por ejemplo, la salida del elemento 422) y una segunda salida (por ejemplo, la salida del elemento 423), en donde el subsistema de aplicación de la reverberación comprende un conjunto de tanques de reverberación, teniendo cada uno de los tanques de reverberación un retardo diferente, y en donde el subsistema de aplicación de la reverberación se acopla y configura para generar un primer canal binaural no mezclado y un segundo canal binaural no mezclado en respuesta a la segunda mezcla descendente filtrada, para imponer el primer canal binaural no mezclado en la primera salida, y para imponer el segundo canal binaural no mezclado en la segunda salida; y

25 una etapa de filtrado y de mezcla de coeficiente de correlación cruzada interaural (IACC) (por ejemplo la etapa 424 de la Fig. 9, que se puede implementar como los elementos 500, 501, 502, y 503 de la Fig. 11) acoplada al subsistema de aplicación de la reverberación y configurada para generar un primer canal binaural no mezclado y un segundo canal binaural no mezclado en respuesta al primer canal binaural no mezclado y al segundo canal binaural no mezclado.

30 El filtro de entrada se puede implementar para generar (preferiblemente como una cascada de dos filtros configurados para generar) la primera mezcla descendente filtrada de manera tal que cada BRIR tiene una relación directa a tardía (DLR) que coincide, al menos de manera sustancial, con una DLR objetivo.

35 Cada tanque de reverberación se puede configurar para generar la señal retardada, y puede incluir un filtro de reverberación (por ejemplo, implementado como un filtro atenuador o una cascada de filtros atenuadores) acoplado y configurado para aplicar una ganancia a una señal que se propaga en dicho cada uno de los tanques de reverberación, para provocar que la señal retardada tenga una ganancia que coincida, al menos de manera sustancial, con la ganancia de decaimiento objetivo para dicha señal retardada, en un esfuerzo para alcanzar una característica de tiempo de decaimiento de reverberación objetivo (por ejemplo la característica  $T_{60}$ ) de cada BRIR.

40 En algunas realizaciones, el primer canal binaural no mezclado guía al segundo canal binaural no mezclado, los tanques de reverberación incluyen un primer tanque de reverberación (por ejemplo, el tanque de reverberación de la Fig. 9 que incluye la línea 410 de retardo) configurado para generar una primera señal retardada que tiene el menor retardo y un segundo tanque de reverberación (por ejemplo, el tanque de reverberación de la Fig. 9 que incluye la línea 411 de retardo) configurado para generar una segunda señal retardada que tiene el segundo menor retardo, en donde el primer tanque de reverberación se configura para aplicar una primera ganancia a la primera señal retardada, el segundo tanque de reverberación se configura para aplicar una segunda ganancia a la segunda señal retardada, la segunda ganancia es diferente de la primera ganancia, la segunda ganancia es diferente de la primera ganancia, y la aplicación de la primera ganancia y de la segunda ganancia resulta en la atenuación del primer canal binaural no mezclado en relación al segundo canal binaural no mezclado. Normalmente, el primer canal binaural no mezclado y el segundo canal binaural no mezclado son indicativos de una imagen en estéreo que se vuelve a centrar. En algunas realizaciones, la etapa de filtrado y de mezcla IACC se configura para generar el primer canal binaural no mezclado y el segundo canal binaural no mezclado de manera tal que dicho primer canal binaural no mezclado y dicho segundo canal binaural no mezclado tienen una característica IACC que al menos de manera sustancial coincide con la característica IACC objetivo.

Los aspectos de la invención incluyen los métodos y sistemas (por ejemplo, el sistema 20 de la Fig. 2, o el sistema de la Fig. 3, o de la Fig.10) que realizan (o se configuran para realizar, o soportar la realización de) la virtualización

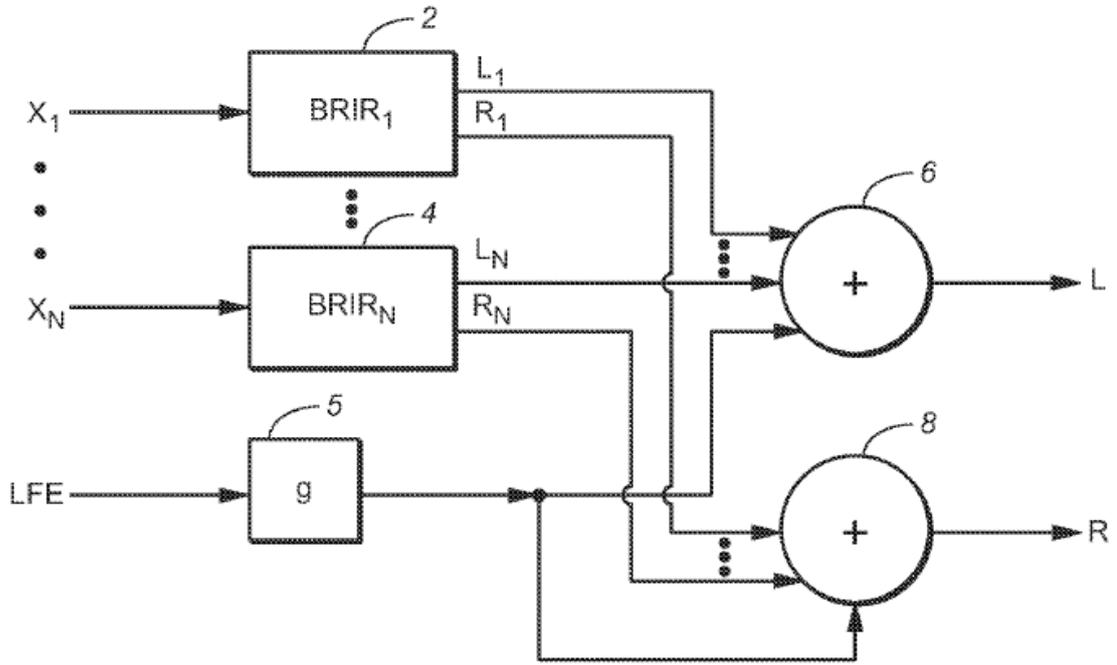
binaural de las señales de audio (por ejemplo, las señales de audio cuyo contenido de audio consiste de canales de altavoces, y/o señales de audio basadas en objetos).

- 5 En algunas realizaciones, el virtualizador inventivo es o incluye un procesador de propósito general acoplado para recibir o generar los datos de entrada indicativos de la señal de audio multi canal, y programados con software (o firmware) y/o de otra manera configurados (por ejemplo, en respuesta a los datos de control) para realizar cualquiera de entre una variedad de operaciones en los datos de entrada, incluyendo una realización del método inventivo. Dicho procesador de propósito general se acoplaría normalmente a un dispositivo de entrada (por ejemplo, un ratón y/o un teclado), una memoria, y un dispositivo de presentación. Por ejemplo, el sistema de la Fig. 3 (o el sistema 20 de la Fig. 2, o el sistema virtualizador que comprende los elementos 12,..., 14, 15, 16 y 18 del sistema 20) se podría
- 10 implementar en un procesador de propósito general, siendo las entradas datos de audio indicativos de N canales de la señal de entrada de audio, y siendo las salidas los datos de audio indicativos de los dos canales de una señal de audio binaural. Un conversor digital a analógico convencional (DAC) podría operar sobre los datos de salida para generar las versiones análogas de los canales de la señal binaural para su reproducción por los altavoces (por ejemplo, un par de auriculares).
- 15 Aunque las realizaciones específicas de las presente invención y aplicaciones de la invención se han descrito en la presente memoria, será evidente para aquellos de habilidad ordinaria en la técnica que son posibles muchas variaciones sobre las realizaciones y aplicaciones descritas en la presente memoria sin salir del alcance de la invención descrita y reivindicada en la presente memoria.

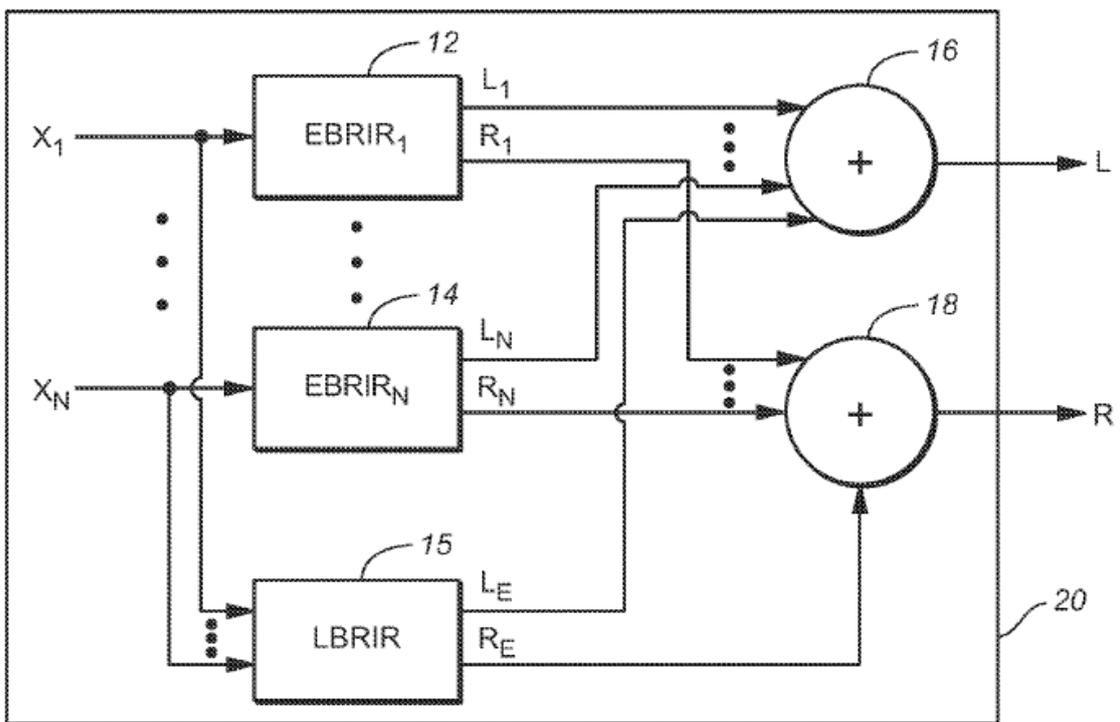
**REIVINDICACIONES**

1. Un método para generar una señal binaural en respuesta a un conjunto de canales de una señal de entrada de audio multi canal, que incluye los pasos de:
- 5 (a) aplicar una respuesta al impulso binaural de una sala, BRIR, a cada canal del conjunto, generando de este modo las señales filtradas, que incluye el uso de al menos una red (203, 204, 205, 220) de retardo de retroalimentación para aplicar una reverberación tardía común a una mezcla descendente de los canales del conjunto; y
- (b) combinar las señales filtradas para generar la señal binaural,
- en donde en el paso (a), la parte de reverberación tardía común emula los macro atributos colectivos de las partes de reverberación tardías de al menos algunas BRIR de canal único.
- 10 incluyendo el método también un paso de imposición de valores de control a la red (203, 204, 205) de retardo de retroalimentación para establecer al menos una ganancia de entrada, unas ganancias del tanque de reverberación, unos retardos del tanque de reverberación, o unos parámetros de la matriz de salida que establezcan la coherencia interaural para dicha red (203, 204, 205) de retardo de retroalimentación, en donde los valores de control se imponen de tal manera que la parte de reverberación tardía común emula los atributos macro colectivos de las
- 15 partes de reverberación tardías de dichas al menos algunas BRIR de canal único.
2. El método de la reivindicación 1, en donde el paso (a) incluye un paso de generación de la mezcla descendente de una manera que depende de la distancia a la fuente para cada uno de los canales que se mezcla de manera descendente para generar dicha mezcla descendente, y sobre el manejo de la parte de respuesta directa de la BRIR para dicho cada uno de los canales que se mezclan de manera descendente para generar dicha mezcla
- 20 descendente, para mantener el nivel apropiado y la relación de temporización entre la parte de respuesta directa de dicha BRIR y la reverberación tardía común.
3. Un sistema configurado para generar una señal binaural en respuesta a un conjunto de canales de una señal de entrada de audio multicanal, incluyendo dicho sistema:
- 25 un subsistema de filtrado acoplado y configurado para aplicar una respuesta al impulso binaural de una sala, BRIR, a cada canal del conjunto, generando de este modo las señales filtradas, incluyendo generar una mezcla descendente de los canales del conjunto y procesar dicha mezcla descendente en al menos una red (203, 204, 205, 220) de retardo de retroalimentación para aplicar una reverberación tardía común a dicha mezcla descendente; y
- un subsistema (210) de combinación de señal, acoplado al subsistema de filtrado, y configurado para generar la señal binaural mediante la combinación de las señales filtradas,
- 30 en donde la reverberación tardía común emula los atributos macro colectivos de las partes de reverberación tardías de al menos una BRIR de canal único,
- incluyendo el sistema también un subsistema (209) de control acoplado al subsistema de filtrado y configurado para imponer los valores de control a la red (203, 204, 205) de retardo de retroalimentación para establecer al menos uno de entre la ganancia de entrada, las ganancias del tanque de reverberación, los retardos del tanque de
- 35 reverberación, o los parámetros de la matriz de salida que establezcan la coherencia interaural para dicha red (203, 204, 205) de retardo de retroalimentación, en donde los valores de control se imponen de tal manera que la parte de reverberación tardía común emula los atributos macro colectivos de las partes de reverberación tardías de dichas al menos algunas BRIR de canal único.
4. El sistema de la reivindicación 3, en donde el subsistema de filtrado se configura para aplicar a cada canal del conjunto una parte de respuesta directa y de reverberación temprana de la BRIR de canal único para el canal.
- 40 5. El sistema de la reivindicación 3 o de la reivindicación 4,
- en donde el subsistema de filtrado incluye un banco de redes (203, 204, 205) de retardo de retroalimentación configurado para aplicar la reverberación tardía común a la mezcla descendente, con cada red (203, 204, 205) de retardo de retroalimentación del banco aplicando la reverberación tardía a una banda de frecuencias distinta de la
- 45 mezcla descendente.
6. El sistema de cualquiera de las reivindicaciones 3 – 5, en donde la mezcla descendente de los canales del conjunto es una mezcla descendente monofónica de dichos canales del conjunto.
7. El sistema de la reivindicación 3 o de la reivindicación 4,
- 50 en donde la red (220) de retardo de retroalimentación se implementa en el dominio del tiempo, y el subsistema de filtrado se configura para procesar la mezcla descendente en el dominio del tiempo en dicha red (220) de retardo de retroalimentación para aplicar la reverberación tardía común a dicha mezcla descendente.

8. El sistema de la reivindicación 7, en donde la red (220) de retardo de retroalimentación incluye:  
 un filtro (400) de entrada que tiene una entrada acoplada para recibir la mezcla descendente, en donde el filtro (400) de entrada se configura para generar una primera mezcla descendente filtrada en respuesta a la mezcla descendente;
- 5 un filtro (401) paso todo, acoplado y configurado para generar una segunda mezcla descendente en respuesta a la primera mezcla descendente filtrada;
- 10 un subsistema de aplicación de la reverberación, que tiene una primera salida y una segunda salida, en donde el subsistema de aplicación de la reverberación comprende el conjunto de tanques de reverberación, teniendo cada uno de los tanques de reverberación un retardo diferente, y en donde el subsistema de aplicación de la reverberación se acopla y configura para generar un primer canal binaural no mezclado y un segundo canal binaural no mezclado en la primera salida, y para imponer el segundo canal binaural no mezclado en la segunda salida; y
- una etapa (424) de filtrado y mezcla de coeficiente de correlación cruzada interaural, IACC, acoplada al subsistema de aplicación de la reverberación y configurado para genera un primer canal binaural mezclado y un segundo canal binaural mezclado en respuesta al primer canal binaural no mezclado y al segundo canal binaural no mezclado.
- 15 9. El sistema de la reivindicación 8, en donde el filtro (400) de entrada se implementa como una cascada de dos filtros configurados para generar la primera mezcla descendente filtrada de manera tal que cada BRIR tenga una relación directa a tardía, DLR, que coincida, con la DLR objetivo.
10. El sistema de la reivindicación 8 o la reivindicación 9,
- 20 en donde cada uno de los tanques de reverberación se configura para generar una señal retardada, e incluye un filtro (406, 406A, 407, 407A, 408, 408A, 409, 409A) de reverberación acoplado y configurado para aplicar una ganancia a una señal que se propaga en dicho cada uno de los tanques de reverberación, para provocar que la señal retardada tenga una ganancia que coincida con una ganancia de decaimiento objetivo para dicha señal retardada, para alcanzar un tiempo de decaimiento de reverberación objetivo característico de cada una de dichas BRIR.
- 25 11. El sistema de cualquiera de las reivindicaciones 8 – 10,
- en donde el primer canal binaural no mezclado guía al segundo canal binaural no mezclado, los tanques de reverberación incluyen un primer tanque de reverberación configurado para generar una primera señal retardada que tiene el menor retardo y una segundo tanque de reverberación configurado para generar una segunda señal retardada que tenga el segundo menor retardo, en donde el primer tanque de reverberación se configura para aplicar una primera ganancia a la primera señal retardada, el segundo tanque de reverberación se configura para aplicar una segunda ganancia a la segunda señal retardada, la segunda ganancia es diferente que la primera ganancia, y la aplicación de la primera ganancia y de la segunda ganancia resulta en la atenuación del primer canal binaural no mezclado en relación al segundo canal binaural no mezclado.
- 30 12. El sistema de cualquiera de las reivindicaciones 8 -11,
- 35 en donde el primer canal binaural mezclado y el segundo canal binaural mezclado son indicativos de una imagen estéreo que se vuelve a centrar; y/o
- en donde la etapa (424) de filtrado y mezclado IACC se configura para generar el primer canal binaural mezclado y el segundo canal binaural mezclado de manera tal que dicho primer canal binaural mezclado y dicho canal binaural mezclado tengan una característica IACC que coincida con la característica IACC objetivo.
- 40



**FIG. 1**  
(TÉCNICA ANTERIOR)



**FIG. 2**

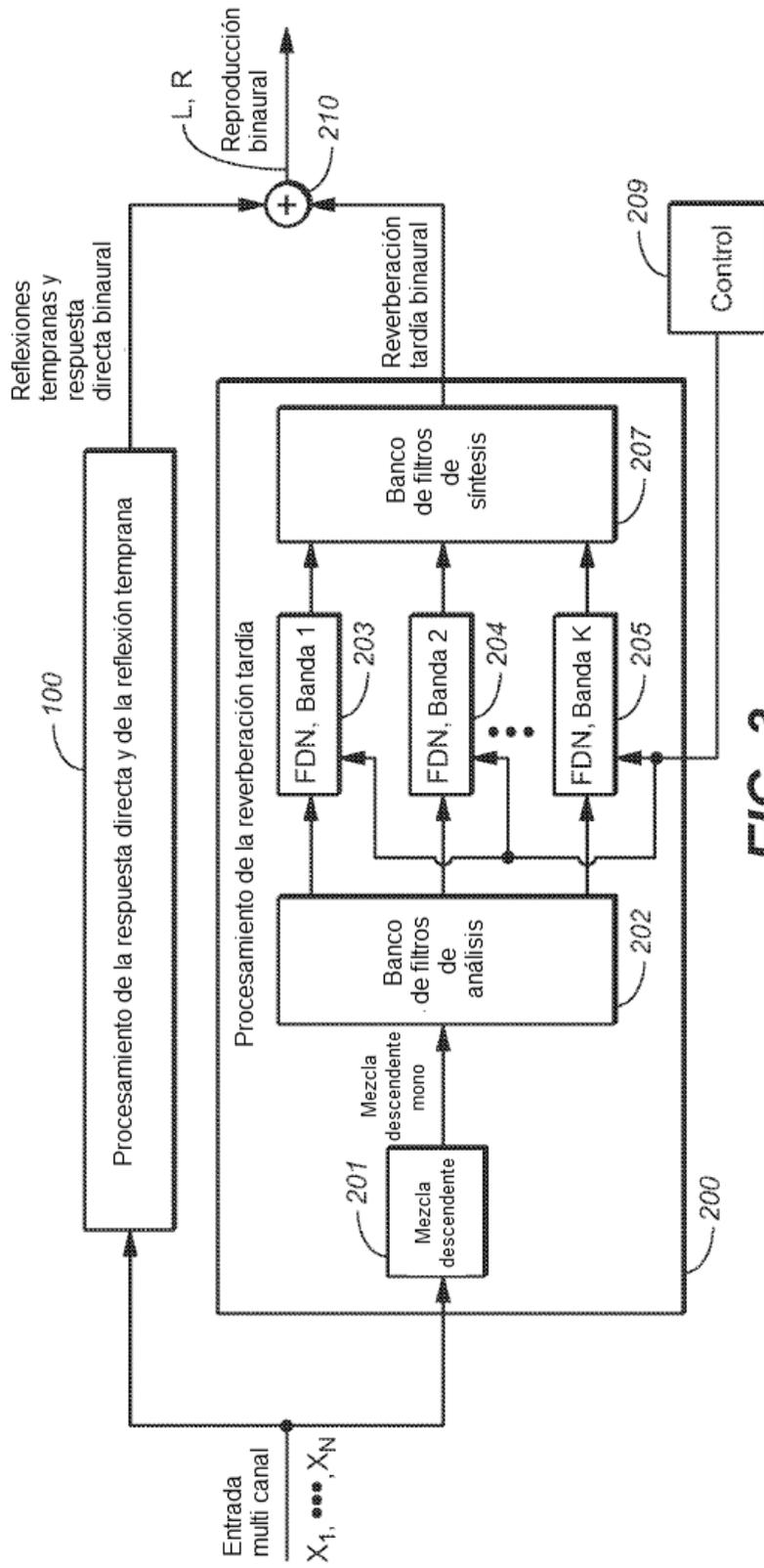
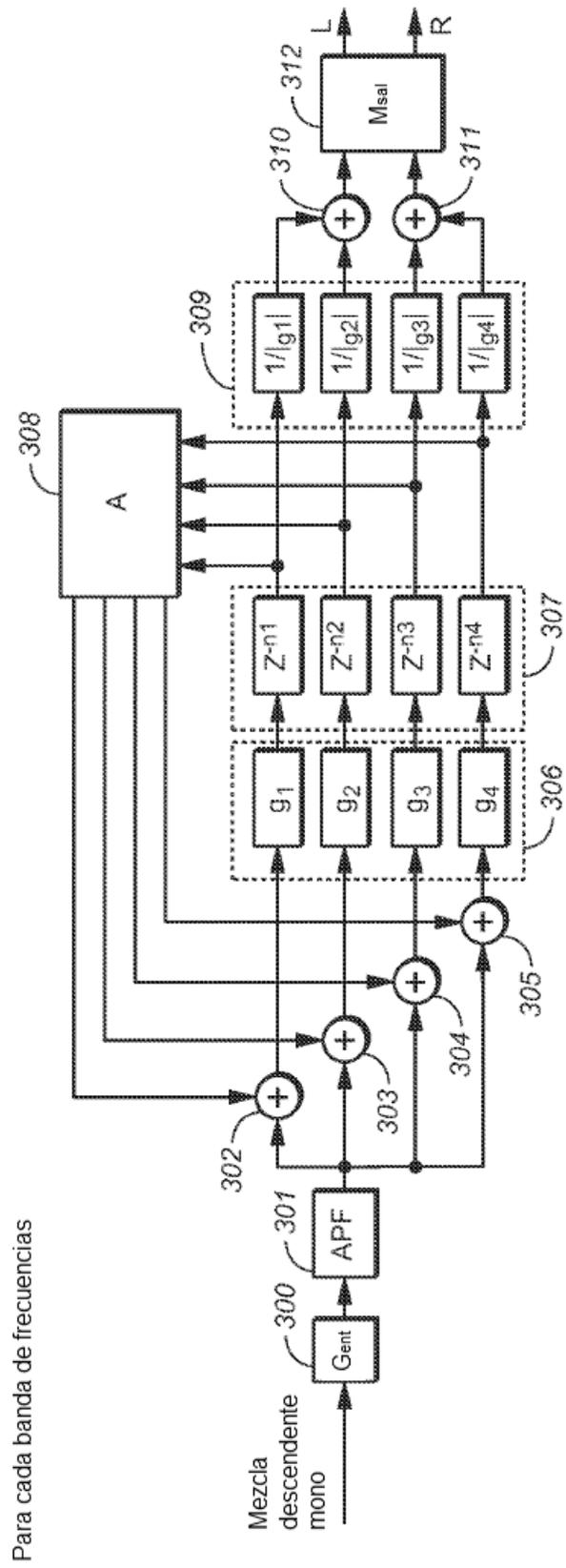
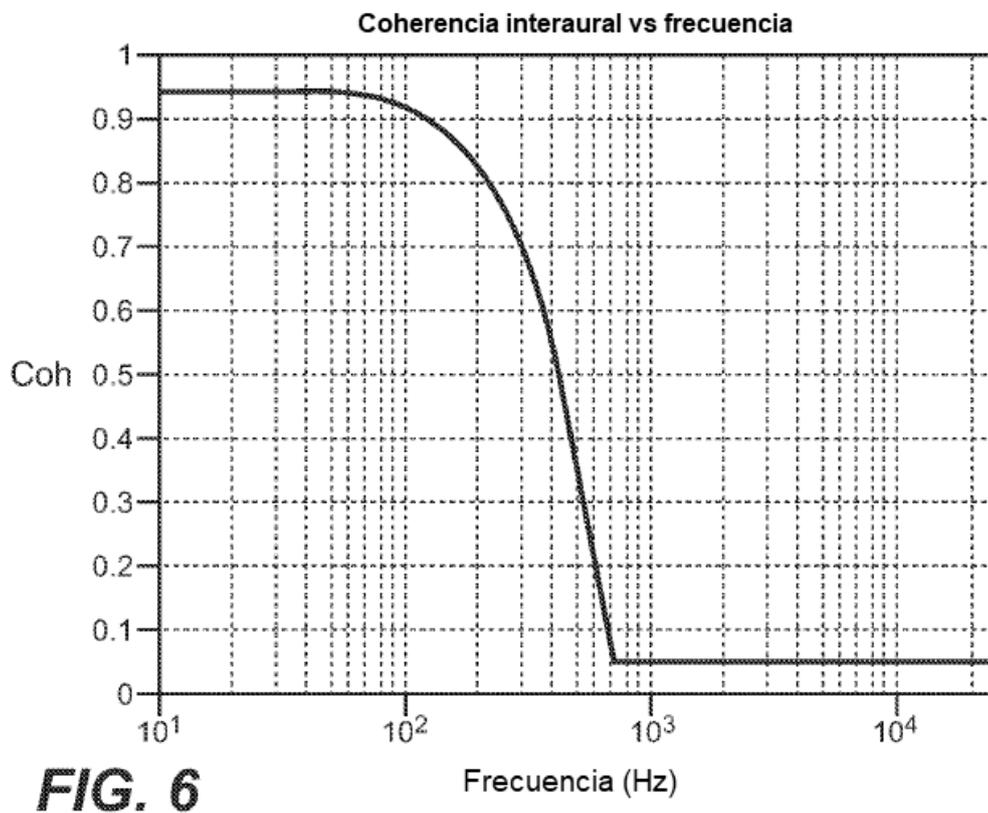
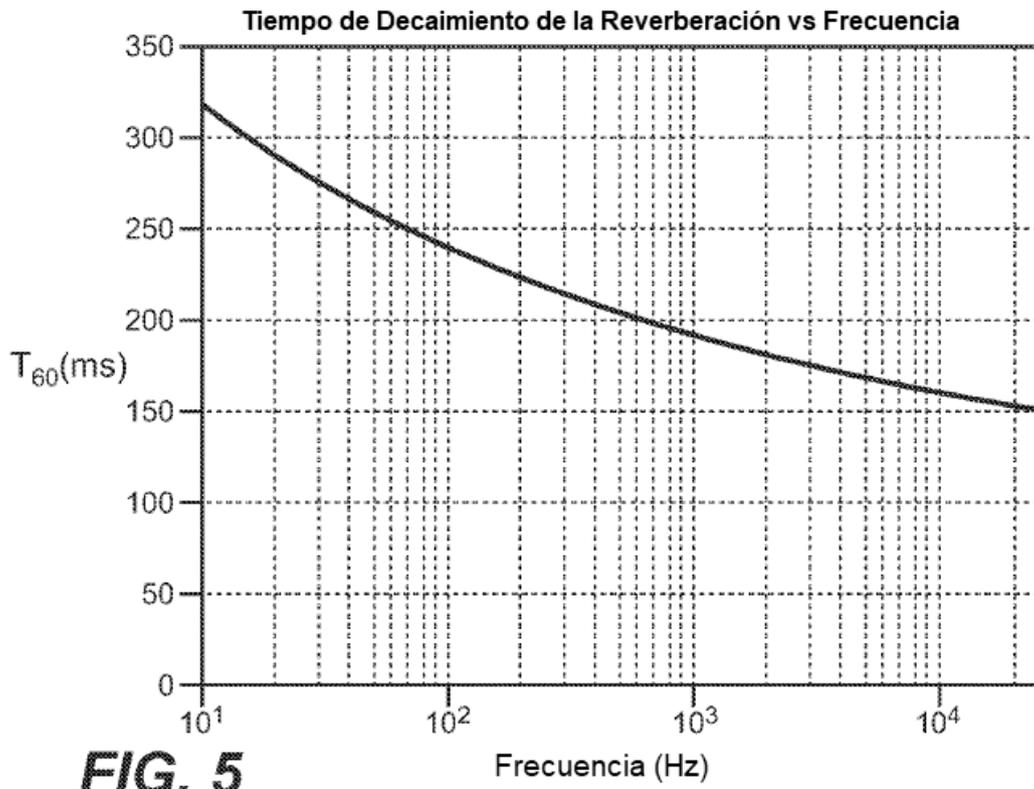
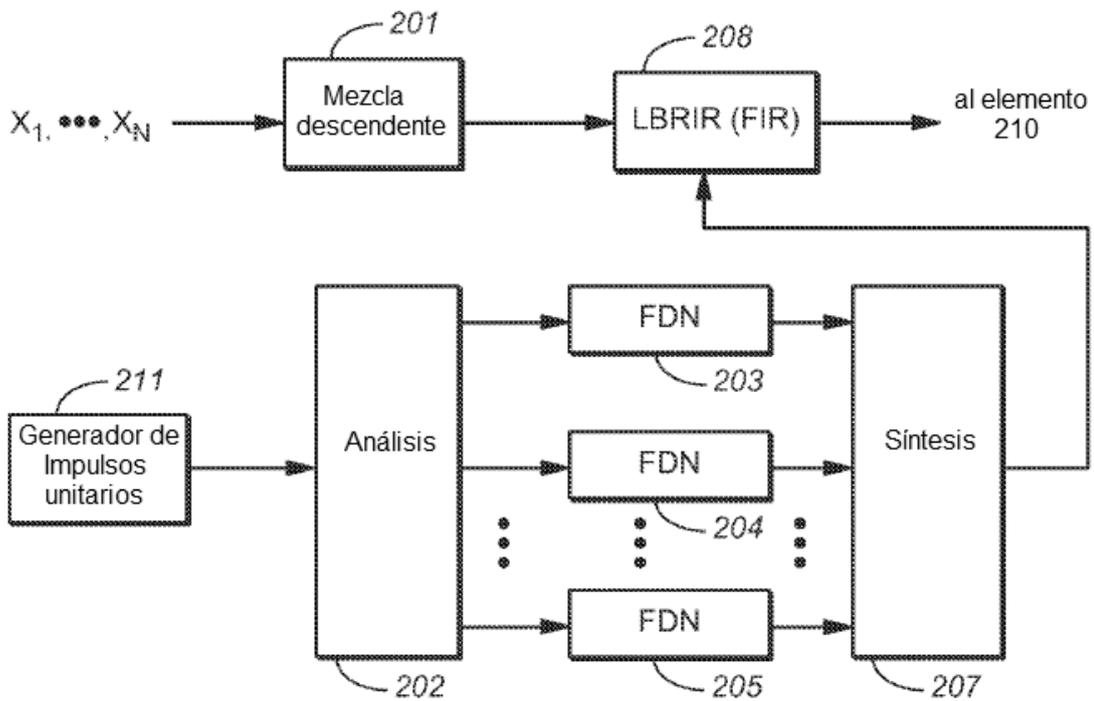
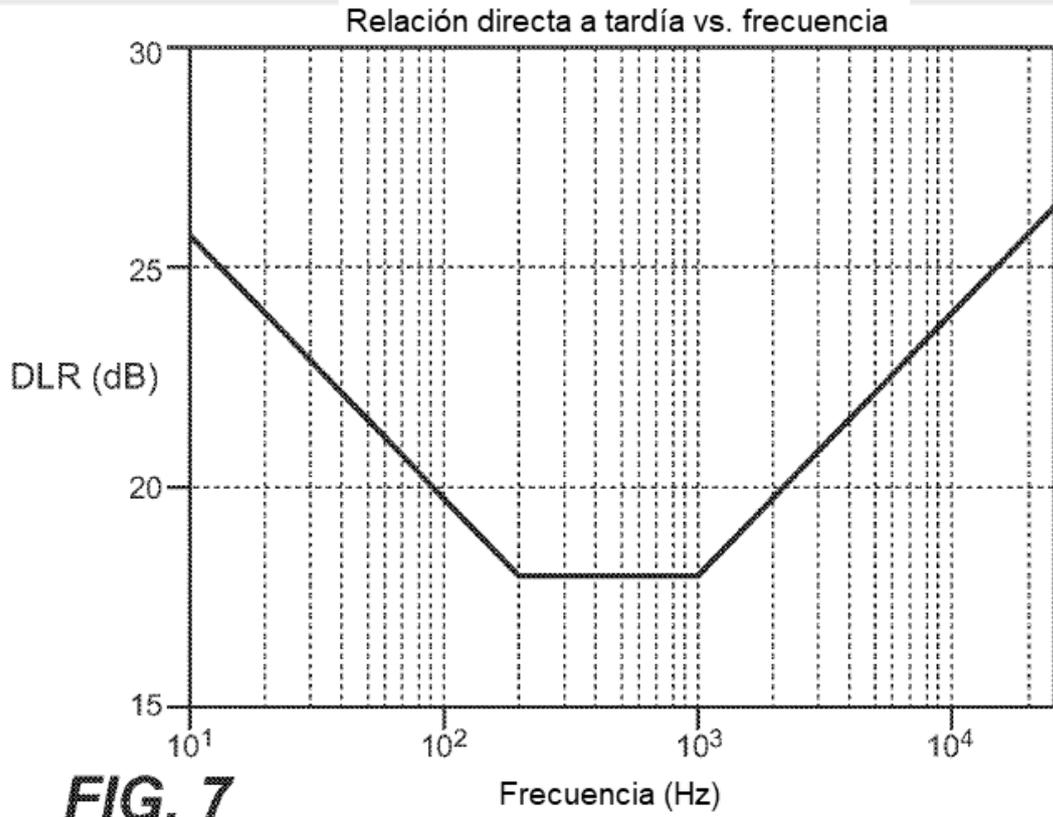


FIG. 3



**FIG. 4**





**FIG. 8**

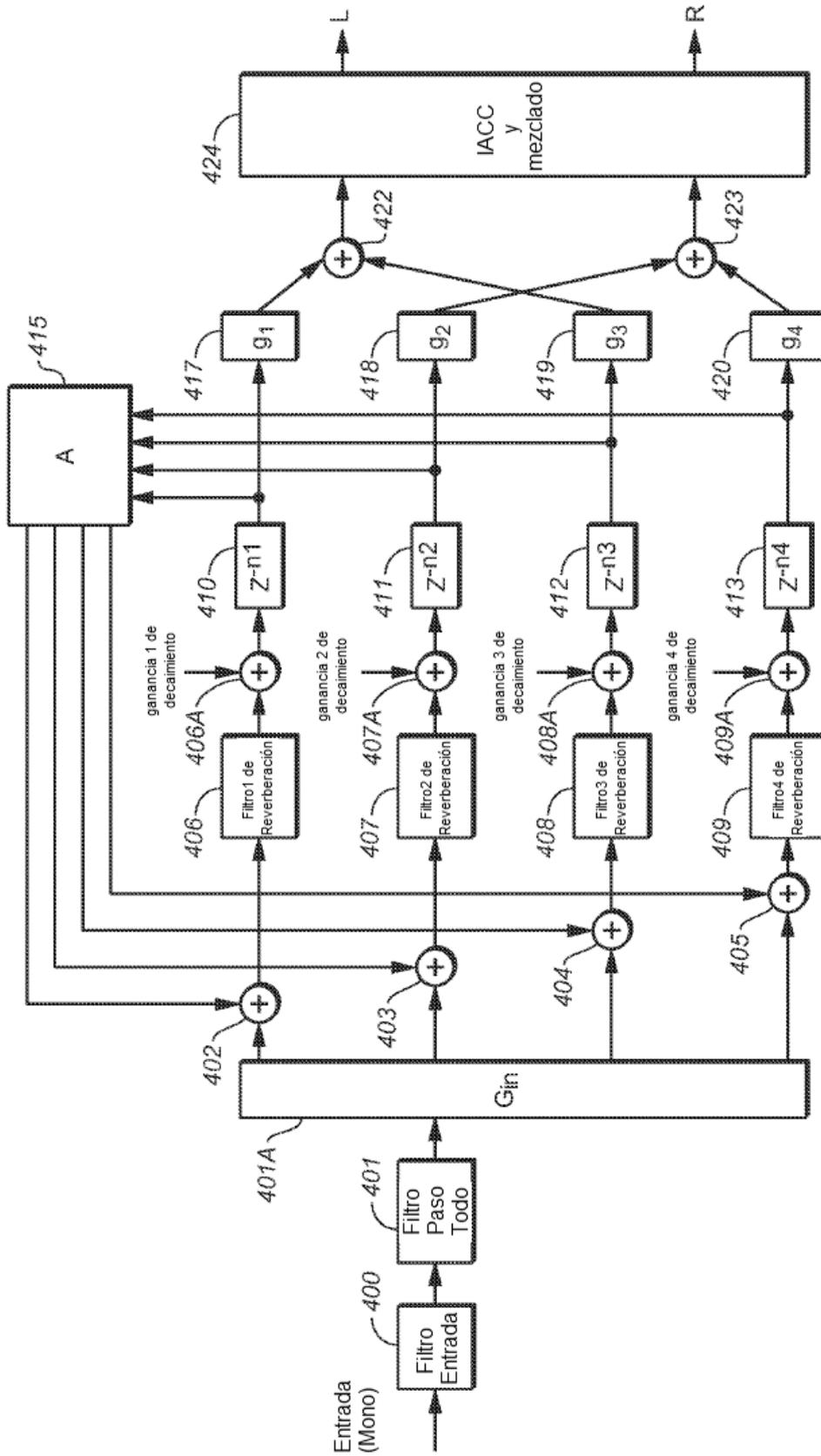
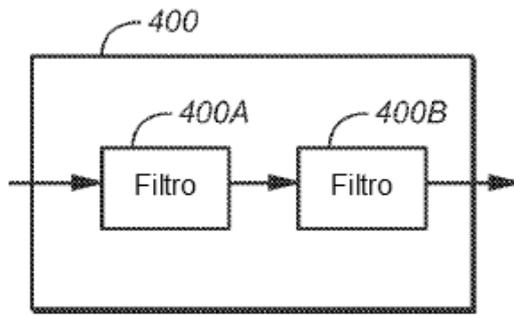
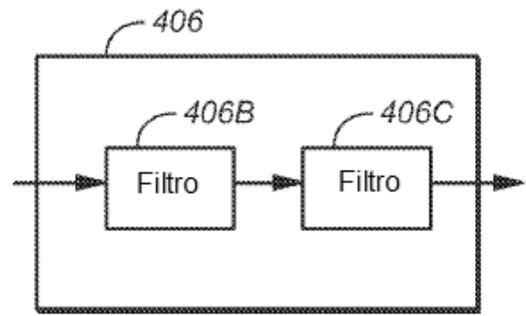


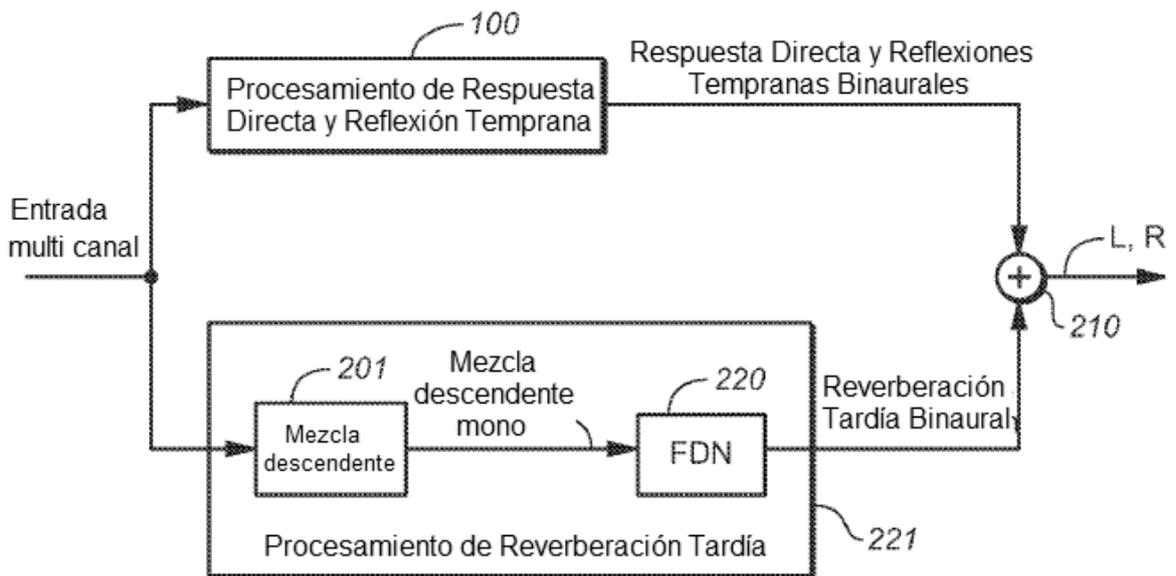
FIG. 9



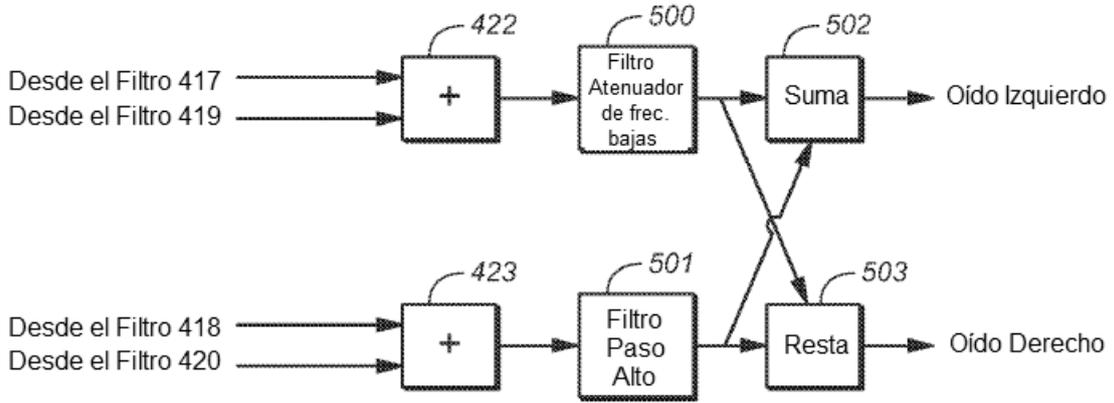
**FIG. 9A**



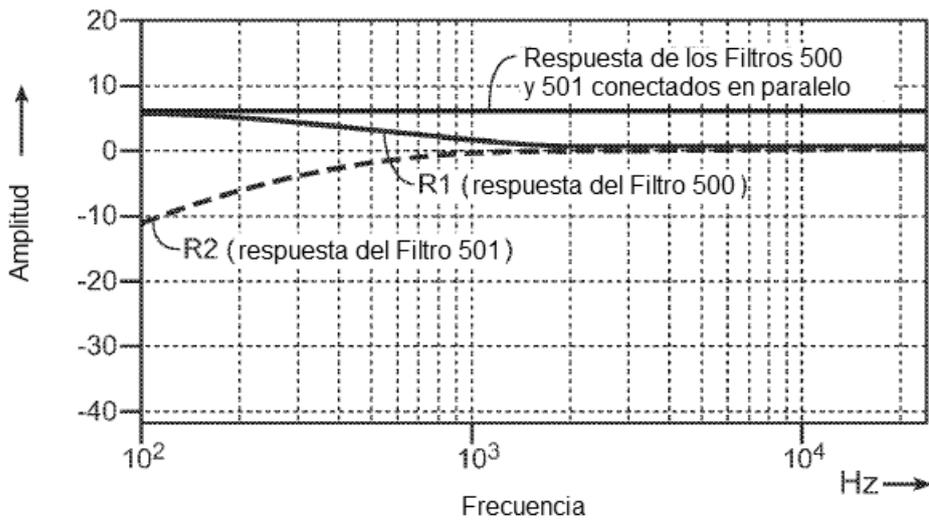
**FIG. 9B**



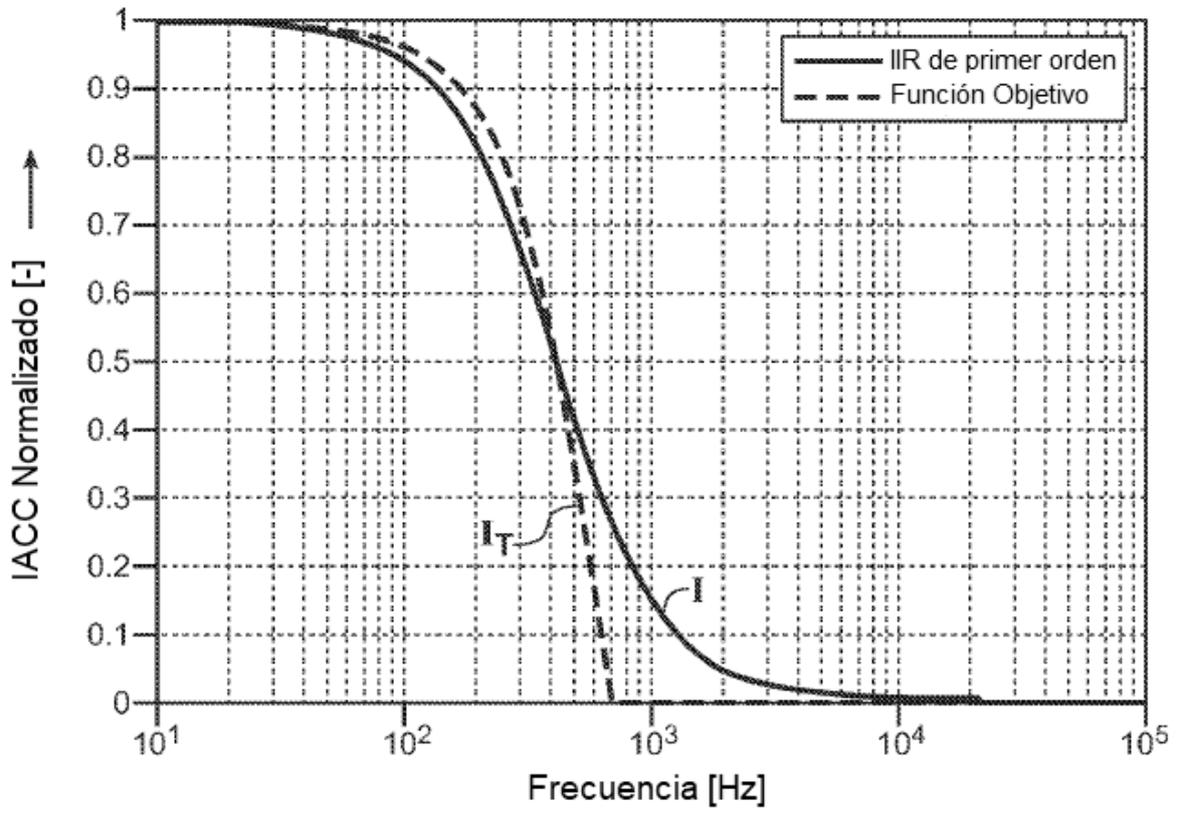
**FIG. 10**



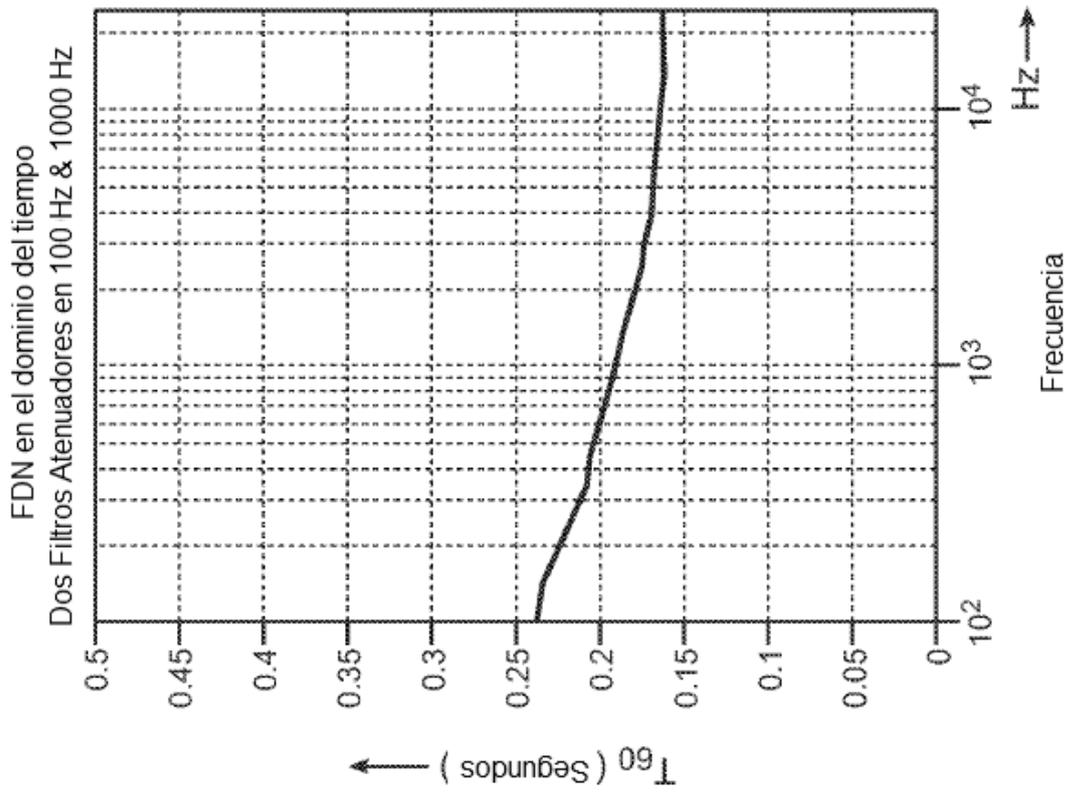
**FIG. 11**



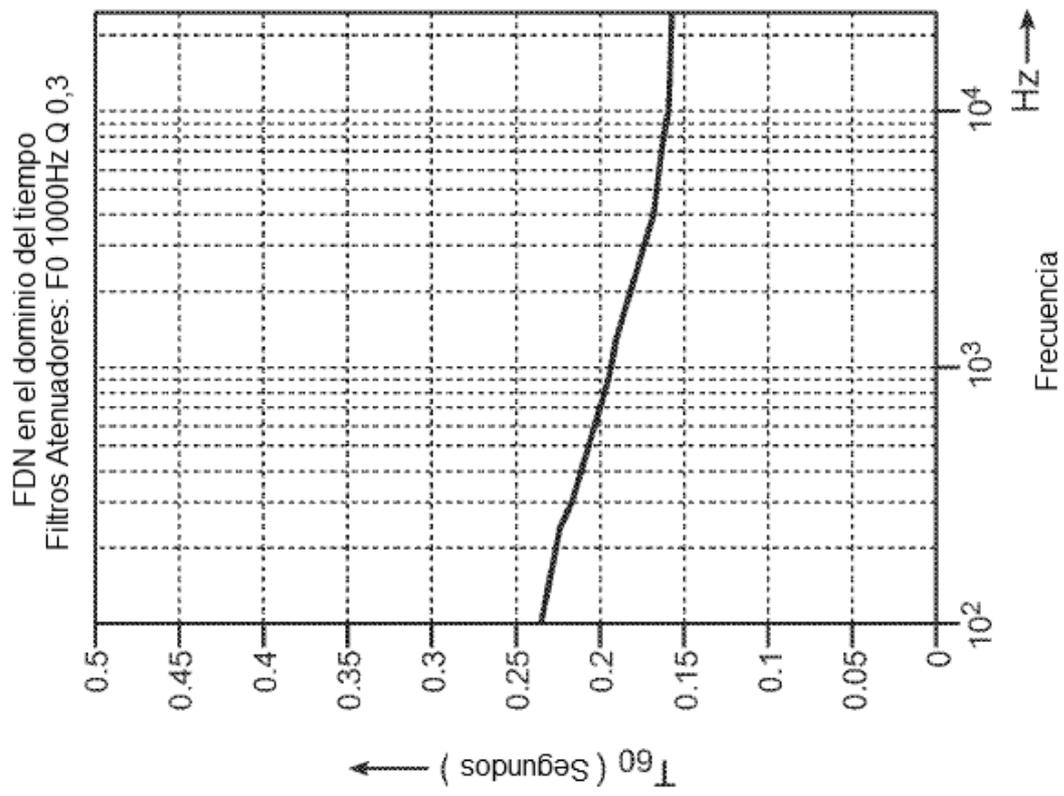
**FIG. 11A**



**FIG. 12**



**FIG. 14**



**FIG. 13**