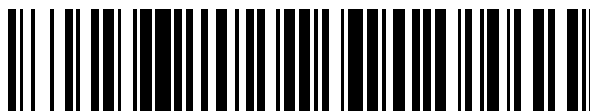


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 650 541**

51 Int. Cl.:

**H04S 5/00** (2006.01)

**H04S 7/00** (2006.01)

**H04S 3/00** (2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **10.03.2014 PCT/US2014/022793**

87 Fecha y número de publicación internacional: **02.10.2014 WO14159272**

96 Fecha de presentación y número de la solicitud europea: **10.03.2014 E 14714882 (9)**

97 Fecha y número de publicación de la concesión europea: **18.10.2017 EP 2926571**

54 Título: **Renderización de objetos de audio de tamaño aparente en distribuciones de altavoces arbitrarias**

30 Prioridad:

**28.03.2013 ES 201330461**

**11.06.2013 US 201361833581 P**

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

**19.01.2018**

73 Titular/es:

**DOLBY LABORATORIES LICENSING CORPORATION (50.0%)**

**100 Potrero Avenue**

**San Francisco, CA 94103-4813, US y**

**DOLBY INTERNATIONAL AB (50.0%)**

72 Inventor/es:

**MATEOS SOLE, ANTONIO y**

**TSINGOS, NICOLAS R.**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

**ES 2 650 541 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**

Renderización de objetos de audio de tamaño aparente en distribuciones de altavoces arbitrarias

5 Campo técnico

Esta divulgación se refiere a la creación y renderización de datos de reproducción de audio. En particular, esta divulgación se refiere a la creación y renderización de datos de reproducción de audio para entornos de reproducción, tales como sistemas de reproducción de sonido en salas de cine.

10

Antecedentes

Desde la introducción del cine sonoro en 1927 ha habido una constante evolución en la tecnología usada para capturar la intención artística de la banda sonora de una película y reproducirla en salas de cine. En la década de 1930, el sonido sincronizado en un disco dio lugar a un sonido de área variable en las películas, lo cual se mejoró aún más en la década de 1940 al introducirse las consideraciones acústicas de las salas de cine y un diseño mejorado de los altavoces, junto con la temprana introducción de la grabación multipista y la reproducción dirigible (que usa tonos de control para mover los sonidos). En la década de 1950 y de 1960, el uso de la banda magnética permitió una reproducción multicanal en las salas de cine, e introdujo canales envolventes y hasta cinco canales de pantalla en salas de teatro de gran calidad.

15

20

25

30

En la década de 1970, Dolby introdujo la reducción de ruido, tanto en la posproducción como en la película, junto con un medio económico de codificación y distribución de mezclas con 3 canales de pantalla y un canal envolvente mono. La calidad del sonido en las salas de cine se mejoró aún más en la década de 1980 con la reducción de ruido de la grabación espectral (SR) Dolby y programas de certificación tales como THX. Dolby introdujo el sonido digital en el cine durante la década de 1990 con un formato de 5.1 canales que proporciona un canal de pantalla izquierdo, uno central y otro derecho, una disposición envolvente izquierda y otra derecha, así como un canal de altavoz de graves para efectos de baja frecuencia. Dolby Surround 7.1 apareció en 2010 y aumentó el número de canales envolventes al dividir los canales envolventes izquierdo y derecho existentes en cuatro "zonas".

35

40

45

A medida que el número de canales aumenta y la distribución de los altavoces pasa de una disposición bidimensional (2D) plana a una disposición tridimensional (3D) que incluye elevación, las tareas de crear y renderizar sonidos son cada vez más complejas. Se necesitan procedimientos y dispositivos mejorados.

La publicación de solicitud de patente estadounidense n.º US 2010/092014 A1 da a conocer un aparato para generar una pluralidad de señales de altavoz para una disposición de altavoces que define un espacio de reproducción. El aparato incluye una unidad de etapa previa configurada para generar una pluralidad de señales de audio de salida mediante la utilización de una o más señales de audio asociadas a una o más posiciones virtuales, donde cada señal de audio de salida está asociada a una posición de altavoz de manera que la pluralidad de señales de audio de salida duplican conjuntamente una reproducción de la(s) señal(es) de audio de entrada en la(s) posición(es) virtual(es), y una pluralidad de señales de audio de salida son más pequeñas que una pluralidad de señales de altavoz. El aparato incluye además una etapa principal configurada para obtener la pluralidad de señales de audio de salida y, además, para obtener, como una posición virtual para cada señal de audio de salida, las posiciones de los altavoces, y para generar el número de señales de altavoz para la disposición de altavoces de manera que las posiciones de los altavoces se duplican como una fuente virtual mediante la disposición de altavoces.

Resumen

50

55

Algunos aspectos del contenido descrito en esta divulgación pueden implementarse en herramientas para renderizar datos de reproducción de audio que incluyen objetos de audio creados sin referencia a ningún entorno de reproducción particular. Tal y como se usa en el presente documento, el término "objeto de audio" puede referirse a un flujo de señales de audio y a metadatos asociados. Los metadatos indican al menos la posición y el tamaño aparente del objeto de audio. Sin embargo, los metadatos también pueden indicar datos de restricción de renderización, datos de tipo de contenido (por ejemplo, diálogos, efectos, etc.), datos de ganancia, datos de trayectoria, etc. Algunos objetos de audio pueden ser estáticos, mientras que otros pueden tener metadatos que varían en el tiempo: tales objetos de audio pueden moverse, pueden cambiar de tamaño y/o pueden tener otras propiedades que cambien en el tiempo.

60

Cuando los objetos de audio se supervisan o se reproducen en un entorno de reproducción, los objetos de audio pueden renderizarse según al menos los metadatos de posición y tamaño. El proceso de renderización puede implicar calcular un conjunto de valores de ganancia de objeto de audio para cada canal de un conjunto de canales de salida. Cada canal de salida puede corresponder a uno o más altavoces de reproducción del entorno de reproducción.

65

Algunas implementaciones descritas en el presente documento implican un proceso de "configuración" que puede tener lugar antes de renderizar cualquier objeto de audio particular. El proceso de configuración, que también puede

denominarse en el presente documento primera etapa o Etapa 1, puede implicar definir múltiples ubicaciones de fuente virtual en un volumen dentro del cual los objetos de audio pueden moverse. Tal y como se usa en el presente documento, una "ubicación de fuente virtual" es una ubicación de una fuente puntual estática. Según tales implementaciones, el proceso de configuración puede implicar recibir datos de ubicación de altavoz de reproducción y precalcular valores de ganancia de fuente virtual para cada una de las fuentes virtuales según los datos de ubicación de altavoz de reproducción y la ubicación de las fuentes virtuales. Tal y como se usa en el presente documento, el término "datos de ubicación de altavoz" puede incluir datos de ubicación que indican las posiciones de algunos o todos los altavoces del entorno de reproducción. Los datos de ubicación pueden proporcionarse como coordenadas absolutas de las ubicaciones de altavoz de reproducción, por ejemplo coordenadas cartesianas, coordenadas esféricas, etc. Como alternativa, o adicionalmente, los datos de ubicación pueden proporcionarse como coordenadas (por ejemplo, coordenadas cartesianas o coordenadas angulares) con respecto a otras ubicaciones del entorno de reproducción, tales como "puntos de sonido óptimo" acústicos del entorno de reproducción.

En algunas implementaciones, los valores de ganancia de fuente virtual pueden almacenarse y usarse durante el "tiempo de ejecución", durante el cual los datos de reproducción de audio se renderizan para los altavoces del entorno de reproducción. Durante el tiempo de ejecución, para cada objeto de audio, pueden calcularse valores de ganancia de fuente virtual a partir de ubicaciones de fuente virtual dentro de un área o volumen definido por los datos de posición de objeto de audio y los datos de tamaño de objeto de audio. El proceso de calcular valores de ganancia de fuente virtual de ubicaciones de fuente virtual puede implicar calcular un promedio ponderado de múltiples valores precalculados de ganancia de fuente virtual, determinados durante el proceso de configuración, para ubicaciones de fuente virtual que están dentro de un área o volumen de objeto de audio definido por el tamaño y la ubicación del objeto de audio. Un conjunto de valores de ganancia de objeto de audio para cada canal de salida del entorno de reproducción puede calcularse en función de, al menos en parte, las contribuciones calculadas de ganancia de fuente virtual. Cada canal de salida puede corresponder a al menos un altavoz de reproducción del entorno de reproducción.

Por consiguiente, algunos procedimientos descritos en el presente documento implican recibir datos de reproducción de audio que incluyen uno o más objetos de audio. Los objetos de audio pueden incluir señales de audio y metadatos asociados. Los metadatos pueden incluir al menos datos de posición de objeto de audio y datos de tamaño de objeto de audio. Los procedimientos pueden implicar calcular contribuciones de fuentes virtuales dentro de un área o volumen de objeto de audio definido por los datos de posición de objeto de audio y los datos de tamaño de objeto de audio. Los procedimientos pueden implicar calcular un conjunto de valores de ganancia de objeto de audio para cada uno de la pluralidad de canales de salida en función de, al menos en parte, los valores calculados de ganancia de fuente virtual. Cada canal de salida puede corresponder a al menos un altavoz de reproducción de un entorno de reproducción. Por ejemplo, el entorno de reproducción puede ser un entorno de sistema de sonido de una sala de cine.

El proceso de calcular ganancias de fuentes virtuales puede implicar calcular un promedio ponderado de valores de ganancia de fuente virtual de las fuentes virtuales dentro del área o volumen de objeto de audio. Los pesos para el promedio ponderado pueden depender de la posición del objeto de audio, del tamaño del objeto de audio y/o de cada ubicación de fuente virtual dentro del área o volumen de objeto de audio.

Los procedimientos también pueden implicar recibir datos de entorno de reproducción que incluyen datos de ubicación de altavoz de reproducción. Los procedimientos también pueden implicar definir una pluralidad de ubicaciones de fuente virtual según los datos de entorno de reproducción y calcular, para cada una de las ubicaciones de fuente virtual, un valor de ganancia de fuente virtual para cada una de la pluralidad de canales de salida. En algunas implementaciones, cada una de las ubicaciones de fuente virtual puede corresponder a una ubicación dentro del entorno de reproducción. Sin embargo, en algunas implementaciones, al menos algunas de las ubicaciones de fuente virtual pueden corresponder a ubicaciones fuera del entorno de reproducción.

En algunas implementaciones, las ubicaciones de fuente virtual pueden estar separadas de manera uniforme a lo largo de los ejes x, y y z. Sin embargo, en algunas implementaciones, la separación puede no ser la misma en todas las direcciones. Por ejemplo, las ubicaciones de fuente virtual pueden tener una primera separación uniforme a lo largo de los ejes x e y y una segunda separación uniforme a lo largo de un eje z. El proceso de calcular el conjunto de valores de ganancia de objeto de audio para cada uno de la pluralidad de canales de salida puede implicar cálculos independientes de contribuciones de ganancia de fuentes virtuales a lo largo de los ejes x, y y z. En implementaciones alternativas, las ubicaciones de fuente virtual pueden estar separadas de manera no uniforme.

En algunas implementaciones, el proceso de calcular el valor de ganancia de objeto de audio para cada uno de la pluralidad de canales de salida puede implicar determinar un valor de ganancia ( $g_i(x_o, y_o, z_o, s)$ ) para un objeto de audio de tamaño (s) que va a renderizarse en la ubicación  $x_o, y_o, z_o$ . Por ejemplo, el valor de ganancia de objeto de audio ( $g_i(x_o, y_o, z_o, s)$ ) puede expresarse como:

$$\left[ \sum_{x_{vs}, y_{vs}, z_{vs}} \left[ w(x_{vs}, y_{vs}, z_{vs}; x_o, y_o, z_o; s) g_l(x_{vs}, y_{vs}, z_{vs}) \right]^p \right]^{1/p},$$

5 donde  $(x_{vs}, y_{vs}, z_{vs})$  representa una ubicación de fuente virtual,  $g_l(x_{vs}, y_{vs}, z_{vs})$  representa un valor de ganancia para el canal  $l$  para la ubicación de fuente virtual  $x_{vs}, y_{vs}, z_{vs}$ , y  $w(x_{vs}, y_{vs}, z_{vs}; x_o, y_o, z_o; s)$  representa una o más funciones de ponderación para  $g_l(x_{vs}, y_{vs}, z_{vs})$  determinadas, al menos en parte, en función de la ubicación  $(x_o, y_o, z_o)$  del objeto de audio, el tamaño  $(s)$  del objeto de audio y la ubicación de fuente virtual  $(x_{vs}, y_{vs}, z_{vs})$ .

10 Según algunas de estas implementaciones,  $g_l(x_{vs}, y_{vs}, z_{vs}) = g_l(x_{vs})g_l(y_{vs})g_l(z_{vs})$ , donde  $g_l(x_{vs})$ ,  $g_l(y_{vs})$  y  $g_l(z_{vs})$  representan funciones de ganancia independientes de  $x$ ,  $y$  y  $z$ . En algunas de tales implementaciones, las funciones de ponderación pueden factorizarse como:

$$w(x_{vs}, y_{vs}, z_{vs}; x_o, y_o, z_o; s) = w_x(x_{vs}; x_o; s)w_y(y_{vs}; y_o; s)w_z(z_{vs}; z_o; s),$$

15 donde  $w_x(x_{vs}; x_o; s)$ ,  $w_y(y_{vs}; y_o; s)$  y  $w_z(z_{vs}; z_o; s)$  representan funciones de ponderación independientes de  $x_{vs}$ ,  $y_{vs}$  y  $z_{vs}$ . Según algunas de estas implementaciones,  $p$  puede ser una función del tamaño de objeto de audio  $(s)$ .

20 Algunos de tales procedimientos pueden implicar almacenar en un sistema de memoria valores calculados de ganancia de fuente virtual. El proceso de calcular contribuciones de ganancia de fuentes virtuales dentro del área o volumen de objeto de audio puede implicar recuperar, del sistema de memoria, valores calculados de ganancia de fuente virtual correspondientes a una posición y tamaño de objeto de audio e interpolar entre los valores calculados de ganancia de fuente virtual. El proceso de interpolación entre los valores calculados de ganancia de fuente virtual puede implicar: determinar una pluralidad de ubicaciones de fuente virtual próximas cerca de la posición de objeto de audio; determinar valores calculados de ganancia de fuente virtual para cada una de las ubicaciones de fuente virtual próximas; determinar una pluralidad de distancias entre la posición de objeto de audio y cada una de las  
25 ubicaciones de fuente virtual próximas; e interpolar entre los valores calculados de ganancia de fuente virtual según la pluralidad de distancias.

30 En algunas implementaciones, los datos de entorno de reproducción pueden incluir datos de límite de entorno de reproducción. El procedimiento puede implicar determinar que un área o volumen de objeto de audio incluye un área o volumen externo fuera de un límite del entorno de reproducción y aplicar un factor de desvanecimiento en función de, al menos en parte, el área o volumen externo. Algunos procedimientos pueden implicar determinar que un objeto de audio puede estar dentro de una distancia umbral desde un límite del entorno de reproducción y no proporcionar ninguna señal de alimentación de altavoz a altavoces de reproducción en un límite opuesto del entorno de reproducción. En algunas implementaciones, un área o volumen de objeto de audio puede ser un rectángulo, un  
35 prisma rectangular, un círculo, una esfera, una elipse y/o un elipsoide.

40 Algunos procedimientos pueden implicar descorrelacionar al menos algunos de los datos de reproducción de audio. Por ejemplo, los procedimientos pueden implicar descorrelacionar los datos de reproducción de audio para objetos de audio que tienen un tamaño de objeto de audio que supera un valor umbral.

45 En el presente documento se describen procedimientos alternativos. Algunos de tales procedimientos implican recibir datos de entorno de reproducción que incluyen datos de ubicación de altavoz de reproducción y datos de límite de entorno de reproducción, y recibir datos de reproducción de audio que incluyen uno o más objetos de audio y metadatos asociados. Los metadatos pueden incluir datos de posición de objeto de audio y datos de tamaño de objeto de audio. Los procedimientos pueden implicar determinar que un área o volumen de objeto de audio, definido por los datos de posición de objeto de audio y los datos de tamaño de objeto de audio, incluye un área o volumen externo fuera de un límite del entorno de reproducción y determinar un factor de desvanecimiento en función de, al menos en parte, el área o volumen externo. Los procedimientos pueden implicar calcular un conjunto de valores de ganancia para cada uno de una pluralidad de canales de salida en función de, al menos en parte, los metadatos  
50 asociados y el factor de desvanecimiento. Cada canal de salida puede corresponder a al menos un altavoz de reproducción del entorno de reproducción. El factor de desvanecimiento puede ser proporcional al área externo.

55 Los procedimientos también pueden implicar determinar que un objeto de audio puede estar dentro de una distancia umbral con respecto a un límite del entorno de reproducción y no proporcionar ninguna señal de alimentación de altavoz a altavoces de reproducción en un límite opuesto del entorno de reproducción.

Los procedimientos también pueden implicar calcular valores de ganancia de fuente virtual dentro del área o volumen de objeto de audio. Los procedimientos pueden implicar definir una pluralidad de ubicaciones de fuente virtual según los datos de entorno de reproducción y calcular, para cada una de las ubicaciones de fuente virtual, una ganancia de fuente virtual para cada una de una pluralidad de canales de salida. Las ubicaciones de fuente virtual pueden estar, o no, separadas de manera uniforme, dependiendo de la implementación particular.

Algunas implementaciones pueden manifestarse en uno o más medios no transitorios que incluyen software almacenado en los mismos. El software puede incluir instrucciones para controlar uno o más dispositivos para recibir datos de reproducción de audio que incluyen uno o más objetos de audio. Los objetos de audio pueden incluir señales de audio y metadatos asociados. Los metadatos pueden incluir al menos datos de posición de objeto de audio y datos de tamaño de objeto de audio. El software puede incluir instrucciones para calcular, para un objeto de audio del uno o más objetos de audio, contribuciones de ganancia de fuentes virtuales dentro de un área o volumen definido por los datos de posición de objeto de audio y los datos de tamaño de objeto de audio, y para calcular un conjunto de valores de ganancia de objeto de audio para cada uno de una pluralidad de canales de salida en función de, al menos en parte, las contribuciones calculadas. Cada canal de salida puede corresponder a al menos un altavoz de reproducción de un entorno de reproducción.

En algunas implementaciones, el proceso de calcular contribuciones de ganancia de fuentes virtuales puede implicar calcular un promedio ponderado de valores de ganancia de fuente virtual de las fuentes virtuales dentro del área o volumen de objeto de audio. Los pesos para el promedio ponderado pueden depender de la posición del objeto de audio, del tamaño del objeto de audio y/o de cada ubicación de fuente virtual dentro del área o volumen del objeto de audio.

El software puede incluir instrucciones para recibir datos del entorno de reproducción que incluyen datos de ubicación de altavoz de reproducción. El software puede incluir instrucciones para definir una pluralidad de ubicaciones de fuente virtual según los datos de entorno de reproducción y calcular, para cada una de las ubicaciones de fuente virtual, un valor de ganancia de fuente virtual para cada una de la pluralidad de canales de salida. Cada una de las ubicaciones de fuente virtual puede corresponder a una ubicación dentro del entorno de reproducción. En algunas implementaciones, al menos algunas de las ubicaciones de fuente virtual pueden corresponder a ubicaciones fuera del entorno de reproducción.

Según algunas implementaciones, las ubicaciones de fuente virtual pueden estar separadas de manera uniforme. En algunas implementaciones, las ubicaciones de fuente virtual pueden tener una primera separación uniforme a lo largo de los ejes  $x$  e  $y$  y una segunda separación uniforme a lo largo de un eje  $z$ . El proceso de calcular el conjunto de valores de ganancia de objeto de audio para cada uno de la pluralidad de canales de salida puede implicar cálculos independientes de contribuciones de ganancia de fuentes virtuales a lo largo de los ejes  $x$ ,  $y$  y  $z$ .

En el presente documento se describen varios dispositivos y aparatos. Algunos de estos aparatos pueden incluir un sistema de interfaz y un sistema lógico. El sistema de interfaz puede incluir una interfaz de red. En algunas implementaciones, el aparato puede incluir un dispositivo de memoria. El sistema de interfaz puede incluir una interfaz entre el sistema lógico y el dispositivo de memoria.

El sistema lógico puede adaptarse para recibir, desde el sistema de interfaz, datos de reproducción de audio que incluyen uno o más objetos de audio. Los objetos de audio pueden incluir señales de audio y metadatos asociados. Los metadatos pueden incluir al menos datos de posición de objeto de audio y datos de tamaño de objeto de audio. El sistema lógico puede estar adaptado para calcular, para un objeto de audio del uno o más objetos de audio, ganancias de fuentes virtuales dentro de un área o volumen de objeto de audio definido por los datos de posición de objeto de audio y los datos de tamaño de objeto de audio. El sistema lógico puede estar adaptado para calcular un conjunto de valores de ganancia de objeto de audio para cada uno de una pluralidad de canales de salida en función de, al menos en parte, las contribuciones de ganancia calculadas. Cada canal de salida puede corresponder a al menos un altavoz de reproducción de un entorno de reproducción.

El proceso de calcular valores de ganancia de fuentes virtuales puede implicar calcular un promedio ponderado de valores de ganancia de fuente virtual de las fuentes virtuales dentro del área o volumen de objeto de audio. Los pesos para el promedio ponderado pueden depender de la posición del objeto de audio, del tamaño del objeto de audio y de cada ubicación de fuente virtual dentro del área o volumen de objeto de audio. El sistema lógico puede estar adaptado para recibir, desde el sistema de interfaz, datos de entorno de reproducción que incluyen datos de ubicación de altavoz de reproducción.

El sistema lógico puede estar adaptado para definir una pluralidad de ubicaciones de fuente virtual según los datos de entorno de reproducción y calcular, para cada una de las ubicaciones de fuente virtual, un valor de ganancia de fuente virtual para cada una de la pluralidad de canales de salida. Cada una de las ubicaciones de fuente virtual puede corresponder a una ubicación dentro del entorno de reproducción. Sin embargo, en algunas implementaciones, al menos algunas de las ubicaciones de fuente virtual pueden corresponder a ubicaciones fuera del entorno de reproducción. Las ubicaciones de fuente virtual pueden estar, o no, separadas de manera uniforme, dependiendo de la implementación. En algunas implementaciones, las ubicaciones de fuente virtual pueden tener

una primera separación uniforme a lo largo de los ejes x e y y una segunda separación uniforme a lo largo de un eje z. El proceso de calcular el conjunto de valores de ganancia de objeto de audio para cada uno de la pluralidad de canales de salida puede implicar cálculos independientes de contribuciones de ganancia de fuentes virtuales a lo largo de los ejes x, y y z.

5 El aparato también puede incluir una interfaz de usuario. El sistema lógico puede estar adaptado para recibir entradas de usuario, tales como datos de tamaño de objeto de audio, a través de la interfaz de usuario. En algunas implementaciones, el sistema lógico puede estar adaptado para escalar los datos de tamaño de objeto de audio de entrada.

10 Los detalles de una o más implementaciones del contenido descrito en esta especificación se exponen en los dibujos adjuntos y en la siguiente descripción. Otras características, aspectos y ventajas resultarán evidentes a partir de la descripción, los dibujos y las reivindicaciones. Debe observarse que las dimensiones relativas de las siguientes figuras pueden no estar dibujadas a escala.

15 Breve descripción de los dibujos

La Figura 1 muestra un ejemplo de un entorno de reproducción que tiene una configuración Dolby Surround 5.1.

La Figura 2 muestra un ejemplo de un entorno de reproducción que tiene una configuración Dolby Surround 7.1.

20 La Figura 3 muestra un ejemplo de un entorno de reproducción que tiene una configuración de sonido envolvente Hamasaki 22.2.

La Figura 4A muestra un ejemplo de una interfaz gráfica de usuario (GUI) que muestra zonas de altavoz a diferentes alturas en un entorno de reproducción virtual.

La Figura 4B muestra un ejemplo de otro entorno de reproducción.

25 La Figura 5A es un diagrama de flujo que proporciona una visión general de un procedimiento de procesamiento de audio.

La Figura 5B es un diagrama de flujo que proporciona un ejemplo de un proceso de configuración.

30 La Figura 5C es un diagrama de flujo que proporciona un ejemplo de un proceso en tiempo de ejecución que calcula valores de ganancia para objetos de audio recibidos según valores de ganancia precalculados para ubicaciones de fuente virtual.

La Figura 6A muestra un ejemplo de ubicaciones de fuente virtual con respecto a un entorno de reproducción.

La Figura 6B muestra un ejemplo alternativo de ubicaciones de fuente virtual con respecto a un entorno de reproducción.

35 Las Figuras 6C a 6F muestran ejemplos que aplican técnicas de panoramización de campo cercano y de campo lejano en objetos de audio situados en diferentes ubicaciones.

La Figura 6G ilustra un ejemplo de un entorno de reproducción que tiene un altavoz en cada esquina de un cuadrado que tiene una longitud de borde igual a 1.

La Figura 7 muestra un ejemplo de contribuciones de fuentes virtuales dentro de un área definida por datos de posición de objeto de audio y datos de tamaño de objeto de audio.

40 Las Figuras 8A y 8B muestran un objeto de audio en dos posiciones en un entorno de reproducción.

La Figura 9 es un diagrama de flujo que muestra un procedimiento para determinar un factor de desvanecimiento en función de, al menos en parte, qué cantidad de un área o volumen de un objeto de audio se extiende fuera de un límite de un entorno de reproducción.

45 La Figura 10 es un diagrama de bloques que proporciona ejemplos de componentes de un aparato de creación y/o renderización.

La Figura 11A es un diagrama de bloques que representa algunos componentes que pueden usarse en la creación de contenido de audio.

La Figura 11B es un diagrama de bloques que representa algunos componentes que pueden usarse en la reproducción de audio en un entorno de reproducción.

50 Números de referencia y designaciones similares en los diversos dibujos indican elementos similares.

## Descripción de formas de realización de ejemplo

La siguiente descripción está dirigida a ciertas implementaciones con el fin de describir algunos aspectos innovadores de esta divulgación, así como ejemplos de contextos en los que pueden implementarse estos aspectos innovadores. Sin embargo, las enseñanzas del presente documento pueden aplicarse de varias maneras diferentes. Por ejemplo, mientras que varias implementaciones se describen en lo que respecta a entornos de reproducción particulares, las enseñanzas del presente documento pueden aplicarse de manera generalizada a otros entornos de reproducción conocidos, así como a entornos de reproducción que puedan introducirse en el futuro. Además, las implementaciones descritas pueden implementarse en varias herramientas de creación y/o renderización, las cuales pueden implementarse en hardware, software, firmware, etc., diverso. Por consiguiente, las enseñanzas de esta divulgación no están limitadas a las implementaciones mostradas en las figuras y/o descritas en el presente documento, sino que tienen una amplia aplicabilidad.

La Figura 1 muestra un ejemplo de un entorno de reproducción que tiene una configuración Dolby Surround 5.1. El sistema Dolby Surround 5.1 se desarrolló en la década de 1990, pero esta configuración sigue utilizándose ampliamente en entornos de sistemas de sonido en salas de cine. Un proyector 105 puede estar configurado para proyectar imágenes de vídeo, por ejemplo, de una película, en la pantalla 150. Datos de reproducción de audio pueden sincronizarse con las imágenes de vídeo y procesarse por el procesador de sonido 110. Los amplificadores de potencia 115 pueden proporcionar señales de alimentación de altavoz a los altavoces del entorno de reproducción 100.

La configuración Dolby Surround 5.1 incluye una disposición envolvente izquierda 120 y una disposición envolvente derecha 125, cada una de las cuales incluye un grupo de altavoces que se activan en grupo mediante un único canal. La configuración Dolby Surround 5.1 incluye además canales diferentes para el canal de pantalla izquierdo 130, el canal de pantalla central 135 y el canal de pantalla derecho 140. Un canal aparte para el altavoz de graves 145 se proporciona para efectos de baja frecuencia (LFE).

En la década de 2010, Dolby ofreció mejoras en el sonido digital de salas de cine con la introducción del sistema Dolby Surround 7.1. La Figura 2 muestra un ejemplo de un entorno de reproducción que tiene una configuración Dolby Surround 7.1. Un proyector digital 205 puede configurarse para recibir datos de vídeo digital y para proyectar imágenes de vídeo en la pantalla 150. Los datos de reproducción de audio pueden procesarse por el procesador de sonido 210. Los amplificadores de potencia 215 pueden proporcionar señales de alimentación de altavoz a los altavoces del entorno de reproducción 200.

La configuración Dolby Surround 7.1 incluye la disposición envolvente lateral izquierda 220 y la disposición envolvente lateral derecha 225, cada una de las cuales puede activarse mediante un único canal. Al igual que la configuración Dolby Surround 5.1, la configuración Dolby Surround 7.1 incluye canales diferentes para el canal de pantalla izquierdo 230, el canal de pantalla central 235, el canal de pantalla derecho 240 y el altavoz de graves 245. Sin embargo, la configuración Dolby Surround 7.1 aumenta el número de canales envolventes al dividir los canales envolventes izquierdo y derecho de la configuración Dolby Surround 5.1 en cuatro zonas: además de la disposición envolvente lateral izquierda 220 y la disposición envolvente lateral derecha 225, se incluyen otros canales para los altavoces envolventes traseros izquierdos 224 y los altavoces envolventes traseros derechos 226. Aumentar el número de zonas envolventes dentro del entorno de reproducción 200 puede mejorar considerablemente la localización del sonido.

En un esfuerzo por crear un entorno más inmersivo, algunos entornos de reproducción pueden configurarse con un mayor número de altavoces, activados por un mayor número de canales. Además, algunos entornos de reproducción pueden incluir altavoces dispuestos a diferente altura, algunos de los cuales pueden estar situados por encima de la zona de asientos del entorno de reproducción.

La Figura 3 muestra un ejemplo de un entorno de reproducción que tiene una configuración de sonido envolvente Hamasaki 22.2. La configuración Hamasaki 22.2 se desarrolló en los laboratorios NHK Science & Technology Research Laboratories de Japón como el componente de sonido envolvente de la televisión de definición ultraalta. La configuración Hamasaki 22.2 proporciona 24 canales de altavoz, que pueden usarse para activar altavoces dispuestos en tres capas. La capa de altavoces superior 310 del entorno de reproducción 300 puede activarse mediante 9 canales. La capa de altavoces central 320 puede activarse mediante 10 canales. La capa de altavoces inferior 330 puede activarse mediante 5 canales, dos de los cuales son para los altavoces de graves 345a y 345b.

Por consiguiente, la tendencia actual es incluir no solamente más altavoces y más canales, sino también incluir altavoces a diferente altura. A medida que el número de canales aumenta y la distribución de los altavoces pasa de una disposición 2D a una disposición 3D, las tareas de situar y renderizar sonidos son cada vez más complejas. Por consiguiente, el presente cesionario ha desarrollado varias herramientas, así como interfaces de usuario relacionadas, que aumentan la funcionalidad y/o reducen la complejidad de creación en un sistema de sonido de audio 3D. Algunas de estas herramientas se describen en detalle con referencia a las Figuras 5A a 19D de la solicitud de patente provisional estadounidense n.º 61/636.102, presentada el 20 de abril de 2012 y titulada "System and Tools for Enhanced 3D Audio Authoring and Rendering" (la "solicitud de creación y renderización").

La Figura 4A muestra un ejemplo de una interfaz gráfica de usuario (GUI) que muestra zonas de altavoz a diferentes alturas en un entorno de reproducción virtual. Por ejemplo, la GUI 400 puede mostrarse en un dispositivo de visualización según las instrucciones de un sistema lógico, según las señales recibidas desde dispositivos de entrada de usuario, etc. Algunos de tales dispositivos se describen posteriormente con referencia a la Figura 10.

Tal y como se usa en el presente documento con referencia a entornos de reproducción virtual tales como el entorno de reproducción virtual 404, el término "zona de altavoz" se refiere generalmente a una construcción lógica que puede tener, o no, una correspondencia de uno a uno con un altavoz de reproducción de un entorno de reproducción real. Por ejemplo, una "ubicación de zona de altavoz" puede corresponder, o no, a una ubicación de altavoz de reproducción particular de un entorno de reproducción de una sala de cine. En cambio, el término "ubicación de zona de altavoz" puede referirse normalmente a una zona de un entorno de reproducción virtual. En algunas implementaciones, un zona de altavoz de un entorno de reproducción virtual puede corresponder a un altavoz virtual, por ejemplo, a través del uso de tecnología de virtualización tal como Dolby Headphone™ (denominada en ocasiones Mobile Surround™), que crea un entorno de sonido envolvente virtual en tiempo real usando un conjunto de auriculares estéreo de dos canales. En la GUI 400 hay siete zonas de altavoz 402a a una primera altura y dos zonas de altavoz 402b a una segunda altura, lo que hace un total de nueve zonas de altavoz en el entorno de reproducción virtual 404. En este ejemplo, las zonas de altavoz 1 a 3 están en la zona delantera 405 del entorno de reproducción virtual 404. El área delantera 405 puede corresponder, por ejemplo, a un área de un entorno de reproducción de sala de cine en la que está dispuesta una pantalla 150, a una zona de una casa en la que está ubicada una pantalla de televisión, etc.

En este caso, la zona de altavoz 4 corresponde, en general, a altavoces del área izquierda 410 y la zona de altavoz 5 corresponde a altavoces del área derecha 415 del entorno de reproducción virtual 404. La zona de altavoz 6 corresponde un área trasera izquierda 412 y la zona de altavoz 7 corresponde a un área trasera derecha 414 del entorno de reproducción virtual 404. La zona de altavoz 8 corresponde a altavoces de un área superior 420a y la zona de altavoz 9 corresponde a altavoces de un área superior 420b, que puede ser un área de techo virtual. Por consiguiente, y como se describe en mayor detalle en la solicitud de creación y renderización, las ubicaciones de las zonas de altavoz 1 a 9 mostradas en la Figura 4A pueden corresponder, o no, a las ubicaciones de altavoces de reproducción de un entorno de reproducción real. Además, otras implementaciones pueden incluir más o menos zonas de altavoz y/o alturas.

En varias implementaciones descritas en la solicitud de creación y renderización, una interfaz de usuario tal como la GUI 400 puede usarse como parte de una herramienta de creación y/o una herramienta de renderización. En algunas implementaciones, la herramienta de creación y/o la herramienta de renderización pueden implementarse a través de software almacenado en uno o más medios no transitorios. La herramienta de creación y/o la herramienta de renderización pueden implementarse (al menos en parte) mediante hardware, firmware, etc., tal como el sistema lógico y otros dispositivos descritos posteriormente con referencia a la Figura 10. En algunas implementaciones de creación, una herramienta de creación asociada puede usarse para crear metadatos para datos de audio asociados. Los metadatos pueden incluir, por ejemplo, datos que indican la posición y/o la trayectoria de un objeto de audio en un espacio tridimensional, datos de restricción de zona de altavoz, etc. Los metadatos pueden crearse con respecto a las zonas de altavoz 402 del entorno de reproducción virtual 404, en lugar de con respecto a una distribución de altavoces particular de un entorno de reproducción real. Una herramienta de renderización puede recibir datos de audio y metadatos asociados, y puede calcular ganancias de audio y señales de alimentación de altavoz para un entorno de reproducción. Tales ganancias de audio y señales de alimentación de altavoz pueden calcularse según un proceso de panoramización en amplitud, que puede crear la percepción de que el sonido procede de una posición  $P$  del entorno de reproducción. Por ejemplo, las señales de alimentación de altavoz pueden proporcionarse a altavoces de reproducción 1 a  $N$  del entorno de reproducción según la siguiente ecuación:

$$x_i(t) = g_i x(t), \quad i = 1, \dots, N \quad (\text{Ecuación 1})$$

En la Ecuación 1,  $x_i(t)$  representa la señal de alimentación de altavoz que va a aplicarse al altavoz  $i$ ,  $g_i$  representa el factor de ganancia del canal correspondiente,  $x(t)$  representa la señal de audio y  $t$  representa el tiempo. Los factores de ganancia pueden determinarse, por ejemplo, según los procedimientos de panoramización en amplitud descritos en la sección 2, páginas 3 y 4 del documento de V. Pulkki, *Compensating Displacement of Amplitude-Panned Virtual Sources* (conferencia internacional de la Sociedad de Ingeniería de Audio (AES) acerca de audio virtual, sintético y de entretenimiento). En algunas implementaciones, las ganancias pueden depender de la frecuencia. En algunas implementaciones puede introducirse un retardo de tiempo sustituyendo  $x(t)$  por  $x(t-\Delta t)$ .

En algunas implementaciones de renderización, los datos de reproducción de audio creados con referencia a las zonas de altavoz 402 pueden correlacionarse con ubicaciones de altavoz de una amplia gama de entornos de reproducción, que pueden estar en una configuración Dolby Surround 5.1, una configuración Dolby Surround 7.1, una configuración Hamasaki 22.2 u otra configuración. Por ejemplo, haciendo referencia a la Figura 2, una herramienta de renderización puede correlacionar datos de reproducción de audio para las zonas de altavoz 4 y 5 con la disposición envolvente lateral izquierda 220 y la disposición envolvente lateral derecha 225 de un entorno de



reproducción que tiene una configuración Dolby Surround 7.1. Los datos de reproducción de audio para las zonas de altavoz 1, 2 y 3 pueden correlacionarse con el canal de pantalla izquierdo 230, con el canal de pantalla derecho 240 y con el canal de pantalla central 235, respectivamente. Los datos de reproducción de audio para las zonas de altavoz 6 y 7 pueden correlacionarse con los altavoces envolventes traseros izquierdos 224 y con los altavoces envolventes traseros derechos 226.

La Figura 4B muestra un ejemplo de otro entorno de reproducción. En algunas implementaciones, una herramienta de renderización puede correlacionar datos de reproducción de audio para las zonas de altavoz 1, 2 y 3 con altavoces de pantalla 455 correspondientes del entorno de reproducción 450. Una herramienta de renderización puede correlacionar datos de reproducción de audio para las zonas de altavoz 4 y 5 con la disposición envolvente lateral izquierda 460 y la disposición envolvente lateral derecha 465 y puede correlacionar datos de reproducción de audio para las zonas de altavoz 8 y 9 con altavoces izquierdos 470a situados por encima de la cabeza y altavoces derechos 470b situados por encima de la cabeza. Los datos de reproducción de audio para las zonas de altavoz 6 y 7 pueden correlacionarse con los altavoces envolventes traseros izquierdos 480a y con los altavoces envolventes traseros derechos 480b.

En algunas implementaciones de creación, una herramienta de creación puede usarse para crear metadatos para objetos de audio. Tal y como se ha mencionado anteriormente, el término "objeto de audio" puede referirse a un flujo de señales de datos de audio y a metadatos asociados. Los metadatos pueden indicar la posición 3D del objeto de audio, el tamaño aparente del objeto de audio, restricciones de renderización, el tipo de contenido (por ejemplo, diálogo, efectos), etc. Dependiendo de la implementación, los metadatos pueden incluir otros tipos de datos, tales como datos de ganancia, datos de trayectoria, etc. Algunos objetos de audio pueden ser estáticos, mientras que otros pueden moverse. Los detalles de un objeto de audio pueden crearse o renderizarse según los metadatos asociados que, entre otras cosas, pueden indicar la posición del objeto de audio en un espacio tridimensional en un instante de tiempo dado. Cuando los objetos de audio se supervisan o reproducen en un entorno de reproducción, los objetos de audio pueden renderizarse según su posición y los metadatos de tamaño según la distribución de altavoces de reproducción del entorno de reproducción.

La Figura 5A es un diagrama de flujo que proporciona una visión general de un procedimiento de procesamiento de audio. Ejemplos más detallados se describen posteriormente con referencia a la Figura 5B y siguientes. Estos procedimientos pueden incluir más o menos bloques que los mostrados y descritos en el presente documento y no se llevan a cabo necesariamente en el orden mostrado en el presente documento. Estos procedimientos pueden llevarse a cabo, al menos en parte, mediante un aparato tal como los mostrados en las Figuras 10 y 11B y descritos posteriormente. En algunas formas de realización, estos procedimientos pueden implementarse, al menos en parte, mediante software almacenado en uno o más medios no transitorios. El software puede incluir instrucciones para controlar uno o más dispositivos para llevar a cabo los procedimientos descritos en el presente documento.

En el ejemplo mostrado en la Figura 5A, el procedimiento 500 comienza con un proceso de configuración que determina valores de ganancia de fuente virtual para ubicaciones de fuente virtual con respecto a un entorno de reproducción particular (bloque 505). La Figura 6A muestra un ejemplo de ubicaciones de fuente virtual con respecto a un entorno de reproducción. Por ejemplo, el bloque 505 puede implicar determinar valores de ganancia de fuente virtual de las ubicaciones de fuente virtual 605 con respecto a las ubicaciones de altavoz de reproducción 625 del entorno de reproducción 600a. Las ubicaciones de fuente virtual 605 y las ubicaciones de altavoz de reproducción 625 son simplemente ejemplos. En el ejemplo mostrado en la Figura 6A, las ubicaciones de fuente virtual 605 están separadas de manera uniforme a lo largo de los ejes x, y y z. Sin embargo, en implementaciones alternativas, las ubicaciones de fuente virtual 605 pueden estar separadas de diferente manera. Por ejemplo, en algunas implementaciones, las ubicaciones de fuente virtual 605 pueden tener una primera separación uniforme a lo largo de los ejes x e y y una segunda separación uniforme a lo largo del eje z. En otras implementaciones, las ubicaciones de fuente virtual 605 pueden estar separadas de manera no uniforme.

En el ejemplo mostrado en la Figura 6A, el entorno de reproducción 600a y el volumen de fuente virtual 602a son coextensivos, de manera que cada una de las ubicaciones de fuente virtual 605 corresponde a una ubicación dentro del entorno de reproducción 600a. Sin embargo, en implementaciones alternativas, el entorno de reproducción 600 y el volumen de fuente virtual 602 pueden no ser coextensivos. Por ejemplo, al menos algunas de las ubicaciones de fuente virtual 605 pueden corresponder a ubicaciones fuera del entorno de reproducción 600.

La Figura 6B muestra un ejemplo alternativo de ubicaciones de fuente virtual con respecto a un entorno de reproducción. En este ejemplo, el volumen de fuente virtual 602b se extiende fuera del entorno de reproducción 600b.

Haciendo referencia a la Figura 5A, en este ejemplo, el proceso de configuración del bloque 505 tiene lugar antes de renderizar algún objeto de audio particular. En algunas implementaciones, los valores de ganancia de fuente virtual determinados en el bloque 505 pueden almacenarse en un sistema de almacenamiento. Los valores almacenados de ganancia de fuente virtual pueden usarse durante un proceso en "tiempo de ejecución" que calcula valores de ganancia de objeto de audio para objetos de audio recibidos según al menos algunos de los valores de ganancia de fuente virtual (bloque 510). Por ejemplo, el bloque 510 puede implicar calcular los valores de ganancia de objeto de

audio en función de, al menos en parte, valores de ganancia de fuente virtual correspondientes a ubicaciones de fuente virtual que están dentro de un área o volumen de objeto de audio.

5 En algunas implementaciones, el procedimiento 500 puede incluir un bloque opcional 515, que implica descorrelacionar datos de audio. El bloque 515 puede ser parte de un proceso en tiempo de ejecución. En algunas de tales implementaciones, el bloque 515 puede implicar la convolución en el dominio de frecuencia. Por ejemplo, el bloque 515 puede implicar aplicar un filtro de respuesta finita al impulso (FIR) para cada señal de alimentación de altavoz.

10 En algunas implementaciones, los procesos del bloque 515 pueden llevarse a cabo, o no, dependiendo de un tamaño de objeto de audio y/o de la intención artística del autor. Según algunas de estas implementaciones, una herramienta de creación puede vincular el tamaño de objeto de audio con la descorrelación indicando (por ejemplo, a través de una bandera de descorrelación incluida en metadatos asociados) que la descorrelación debe iniciarse cuando el tamaño de objeto de audio sea mayor que o igual a un valor umbral de tamaño y que la descorrelación debe interrumpirse si el tamaño de objeto de audio es menor que el valor umbral de tamaño. En algunas implementaciones, la descorrelación puede controlarse (por ejemplo, aumentarse, reducirse o inhabilitarse) según la entrada de usuario con respecto al valor umbral de tamaño y/u otros valores de entrada.

20 La Figura 5B es un diagrama de flujo que proporciona un ejemplo de un proceso de configuración. Por consiguiente, todos los bloques mostrados en la Figura 5B son ejemplos de procesos que pueden llevarse a cabo en bloque 505 de la Figura 5A. En este caso, el proceso de configuración empieza con la recepción de datos de entorno de reproducción (bloque 520). Los datos de entorno de reproducción pueden incluir datos de ubicación de altavoz de reproducción. Los datos de entorno de reproducción también pueden incluir datos que representan límites de un entorno de reproducción, tales como paredes, techos, etc. Si el entorno de reproducción es una sala de cine, los datos de entorno de reproducción también pueden incluir una indicación de la ubicación de la pantalla.

30 Los datos de entorno de reproducción también pueden incluir datos que indican una correlación de canales de salida con altavoces de reproducción de un entorno de reproducción. Por ejemplo, el entorno de reproducción puede tener una configuración Dolby Surround 7.1, tal como la mostrada en la Figura 2 y descrita anteriormente. Por consiguiente, los datos de entorno de reproducción también pueden incluir datos que indican una correlación entre un canal Lss y los altavoces envolventes laterales izquierdos 220, entre un canal Lrs y los altavoces envolventes traseros izquierdos 224, etc.

35 En este ejemplo, el bloque 525 implica definir ubicaciones de fuente virtual 605 según los datos de entorno de reproducción. Las ubicaciones de fuente virtual 605 pueden definirse dentro de un volumen de fuente virtual. En algunas implementaciones, el volumen de fuente virtual puede corresponder a un volumen dentro del cual pueden moverse objetos de audio. Como se muestra en las Figuras 6A y 6B, en algunas implementaciones, el volumen de fuente virtual 602 puede ser coextensivo con un volumen del entorno de reproducción 600, mientras que en otras implementaciones al menos algunas de las ubicaciones de fuente virtual 605 pueden corresponder a ubicaciones fuera del entorno de reproducción 600.

45 Además, las ubicaciones de fuente virtual 605 pueden estar, o no, separadas de manera uniforme dentro del volumen de fuente virtual 602, dependiendo de la implementación particular. En algunas implementaciones, las ubicaciones de fuente virtual 605 pueden estar separadas de manera uniforme en todas direcciones. Por ejemplo, las ubicaciones de fuente virtual 605 pueden formar una cuadrícula rectangular de  $N_x$  por  $N_y$  por  $N_z$  ubicaciones de fuente virtual 605. En algunas implementaciones, el valor de  $N$  puede estar en el intervalo comprendido entre 5 y 100. El valor de  $N$  puede depender, al menos en parte, del número de altavoces de reproducción del entorno de reproducción: puede ser deseable incluir dos o más ubicaciones de fuente virtual 605 entre cada ubicación de altavoz de reproducción.

50 En otras implementaciones, las ubicaciones de fuente virtual 605 pueden tener una primera separación uniforme a lo largo de los ejes  $x$  e  $y$  y una segunda separación uniforme a lo largo de un eje  $z$ . Las ubicaciones de fuente virtual 605 pueden formar una cuadrícula rectangular de  $N_x$  por  $N_y$  por  $M_z$  ubicaciones de fuente virtual 605. Por ejemplo, en algunas implementaciones puede haber menos ubicaciones de fuente virtual 605 a lo largo del eje  $z$  que a lo largo de los ejes  $x$  o  $y$ . En algunas de tales implementaciones, el valor de  $N$  puede estar en el intervalo comprendido entre 10 y 100, mientras que el valor de  $M$  puede estar en el intervalo comprendido entre 5 y 10.

60 En este ejemplo, el bloque 530 implica calcular valores de ganancia de fuente virtual para cada una de las ubicaciones de fuente virtual 605. En algunas implementaciones, el bloque 530 implica calcular, para cada una de las ubicaciones de fuente virtual 605, valores de ganancia de fuente virtual para cada canal de una pluralidad de canales de salida del entorno de reproducción. En algunas implementaciones, el bloque 530 puede implicar aplicar un algoritmo de panoramización en amplitud basada en vectores (VBAP), un algoritmo de panoramización basado en pares o un algoritmo similar para calcular valores de ganancia para fuentes puntuales ubicadas en cada una de las ubicaciones de fuente virtual 605. En otras implementaciones, el bloque 530 puede implicar aplicar un algoritmo separable para calcular valores de ganancia para fuentes puntuales ubicadas en cada una de las ubicaciones de fuente virtual 605. Tal y como se usa en el presente documento, un algoritmo "separable" es uno para el que la

ganancia de un altavoz dado puede expresarse como el producto de dos o más factores que pueden calcularse por separado para cada una de las coordenadas de la ubicación de fuente virtual. Ejemplos incluyen algoritmos implementados en varios panoramizadores de consola de mezclas existentes, que incluyen, sin limitarse a, el software Pro Tools™ y panoramizadores implementados en consolas de películas digitales proporcionadas por AMS Neve. A continuación se proporcionan algunos ejemplos bidimensionales.

Las Figuras 6C a 6F muestran ejemplos que aplican técnicas de panoramización de campo cercano y de campo lejano en objetos de audio situados en diferentes ubicaciones. Haciendo referencia primero a la Figura 6C, el objeto de audio está sustancialmente fuera del entorno de reproducción virtual 400a. Por lo tanto, en este caso, se aplicarán uno o más procedimientos de panoramización de campo lejano. En algunas implementaciones, los procedimientos de panoramización de campo lejano pueden basarse en ecuaciones de panoramización en amplitud basada en vectores (VBAP) conocidas por los expertos en la técnica. Por ejemplo, los procedimientos de panoramización de campo lejano pueden basarse en las ecuaciones VBAP descritas en la sección 2.3, página 4 del documento de V. Pulkki, *Compensating Displacement of Amplitude-Panned Virtual Sources* (Conferencia internacional de la AES acerca de audio virtual, sintético y de entretenimiento). En implementaciones alternativas pueden usarse otros procedimientos para panoramizar objetos de audio de campo lejano y de campo cercano, por ejemplo procedimientos que implican la síntesis de planos acústicos correspondientes o de una onda esférica. El documento de D. de Vries, *Wave Field Synthesis* (Monografía de la AES, 1999), describe procedimientos pertinentes.

Haciendo referencia a continuación a la Figura 6D, el objeto de audio 610 está dentro del entorno de reproducción virtual 400a. Por lo tanto, en este caso, se aplicarán uno o más procedimientos de panoramización de campo cercano. Algunos de tales procedimientos de panoramización de campo cercano usarán una pluralidad de zonas de altavoz que incluyen el objeto de audio 610 en el entorno de reproducción virtual 400a.

La Figura 6G ilustra un ejemplo de un entorno de reproducción que tiene un altavoz en cada esquina de un cuadrado que tiene una longitud de borde igual a 1. En este ejemplo, el origen (0,0) del eje x-y coincide con el altavoz de pantalla izquierdo (L) 130. Por consiguiente, el altavoz de pantalla derecho (R) 140 tiene las coordenadas (1, 0), el altavoz envolvente izquierdo (Ls) 120 tiene las coordenadas (0, 1) y el altavoz envolvente derecho (Rs) 125 tiene las coordenadas (1, 1). La posición de objeto de audio 615 (x, y) está x unidades a la derecha del altavoz L e y unidades desde la pantalla 150. En este ejemplo, cada uno de los cuatro altavoces recibe un factor cos/sen proporcional a su distancia a lo largo del eje x y del eje y. Según algunas implementaciones, las ganancias pueden calcularse de la siguiente manera:

$$\begin{aligned} G_L(x) &= \cos(\pi/2 * x) \quad \text{si } l=L, Ls \\ G_L(x) &= \sin(\pi/2 * x) \quad \text{si } l=R, Rs \\ G_L(y) &= \cos(\pi/2 * y) \quad \text{si } l=L, R \\ G_L(y) &= \sin(\pi/2 * y) \quad \text{si } l=Ls, Rs \end{aligned}$$

La ganancia global es el producto:  $G_L(x,y) = G_L(x) G_L(y)$ . En general, estas funciones dependen de todas las coordenadas de todos los altavoces. Sin embargo,  $G_L(x)$  no depende de la posición y de la fuente, y  $G_L(y)$  no depende de su posición x. Para ilustrar un cálculo simple, supóngase que la posición de objeto de audio 615 es (0, 0), la ubicación del altavoz L.  $G_L(x) = \cos(0) = 1$ .  $G_L(y) = \cos(0) = 1$ . La ganancia global es el producto:  $G_L(x,y) = G_L(x) G_L(y) = 1$ . Cálculos similares dan lugar a  $G_{Ls} = G_{Rs} = G_R = 0$ .

Puede ser deseable mezclar diferentes modos de panoramización cuando un objeto de audio entra o sale del entorno de reproducción virtual 400a. Por ejemplo, una mezcla de ganancias calculadas según procedimientos de panoramización de campo cercano y procedimientos de panoramización de campo lejano puede aplicarse cuando el objeto de audio 610 se mueve desde la ubicación de objeto de audio 615 mostrada en la Figura 6C hacia la ubicación de objeto de audio 615 mostrada en la Figura 6D, o viceversa. En algunas implementaciones, puede usarse una ley de panoramización basada en pares (por ejemplo, una ley de potencia o de seno de conservación de energía) para mezclar las ganancias calculadas según los procedimientos de panoramización de campo cercano y los procedimientos de panoramización de campo lejano. En implementaciones alternativas, la ley de panoramización basada en pares puede conservar la amplitud en lugar de conservar la energía, de manera que la suma es igual a uno en lugar de que la suma de los cuadrados sea igual a uno. También es posible mezclar las señales procesadas resultantes, por ejemplo para procesar la señal de audio usando de manera independiente ambos procedimientos de panoramización y para atenuar de manera cruzada las dos señales de audio resultantes.

Haciendo de nuevo referencia a la Figura 5B, independientemente del algoritmo usado en el bloque 530, los valores de ganancia resultantes pueden almacenarse en un sistema de memoria (bloque 535) para usarse en operaciones en tiempo de ejecución.

La Figura 5C es un diagrama de flujo que proporciona un ejemplo de un proceso en tiempo de ejecución que calcula valores de ganancia para objetos de audio recibidos según valores de ganancia precalculados para ubicaciones de

fuentes virtuales. Todos los bloques mostrados en la Figura 5C son ejemplos de procesos que pueden llevarse a cabo en el bloque 510 de la Figura 5A.

5 En este ejemplo, el proceso en tiempo de ejecución comienza con la recepción de datos de reproducción de audio que incluyen uno o más objetos de audio (bloque 540). Los objetos de audio incluyen señales de audio y metadatos asociados, que incluyen, al menos, datos de posición de objeto de audio y datos de tamaño de objeto de audio en este ejemplo. Haciendo referencia a la Figura 6A, por ejemplo, el objeto de audio 610 se define, al menos en parte, mediante una posición de objeto de audio 615 y un volumen de objeto de audio 620a. En este ejemplo, los datos recibidos de tamaño de objeto de audio indican que el volumen de objeto de audio 620a corresponde al de un prisma rectangular. En el ejemplo, mostrado en la Figura 6B, los datos recibidos de tamaño de objeto de audio indican, sin embargo, que el volumen de objeto de audio 620b corresponde al de una esfera. Estos tamaños y formas son simplemente ejemplos; en implementaciones alternativas, los objetos de audio pueden tener otros tamaños y/o formas diversos. En algunos ejemplos alternativos, el área o volumen de un objeto de audio puede ser un rectángulo, un círculo, una elipse, un elipsoide o un sector esférico.

15 En esta implementación, el bloque 545 implica calcular contribuciones de ganancia de fuentes virtuales dentro de un área o volumen definido por los datos de posición de objeto de audio y los datos de tamaño de objetos de audio. En los ejemplos mostrados en las Figuras 6A y 6B, el bloque 545 puede implicar calcular contribuciones de ganancia de las fuentes virtuales en las ubicaciones de fuente virtual 605 que están dentro del volumen de objeto de audio 620a o del volumen de objeto de audio 620b. Si los metadatos del objeto de audio cambian en el tiempo, el bloque 545 puede llevarse a cabo de nuevo según los nuevos valores de metadatos. Por ejemplo, si el tamaño de objeto de audio y/o la posición de objeto de audio cambia, diferentes ubicaciones de fuente virtual 605 pueden estar dentro del volumen de objeto de audio 620 y/o las ubicaciones de fuente virtual 605 usadas en un cálculo anterior pueden estar a una distancia diferente de la posición de objeto de audio 615. En el bloque 545, las contribuciones correspondientes de ganancia de fuente virtual se calcularán según el nuevo tamaño y/o posición de objeto de audio.

30 En algunos ejemplos, el bloque 545 puede implicar recuperar, de un sistema de memoria, valores calculados de ganancia de fuente virtual para ubicaciones de fuente virtual correspondientes a una posición y tamaño de objeto de audio, e interpolar entre los valores calculados de ganancia de fuente virtual. El proceso de interpolación entre los valores calculados de ganancia de fuente virtual puede implicar determinar una pluralidad de ubicaciones de fuente virtual próximas cerca de la posición de objeto de audio, determinar valores calculados de ganancia de fuente virtual para cada una de las ubicaciones de fuente virtual próximas, determinar una pluralidad de distancias entre la posición de objeto de audio y cada una de las ubicaciones de fuente virtual próximas e interpolar entre los valores calculados de ganancia de fuente virtual según la pluralidad de distancias.

40 El proceso de calcular contribuciones de valores de ganancia de fuentes virtuales puede implicar calcular un promedio ponderado de valores calculados de ganancia de fuente virtual para ubicaciones de fuente virtual dentro de un área o volumen definido por el tamaño del objeto de audio. Los pesos para el promedio ponderado pueden depender, por ejemplo, de la posición del objeto de audio, del tamaño del objeto de audio y de cada ubicación de fuente virtual dentro del área o volumen.

45 La Figura 7 muestra un ejemplo de contribuciones de fuentes virtuales dentro de un área definida por datos de posición de objeto de audio y datos de tamaño de objeto de audio. La Figura 7 ilustra una sección cruzada de un entorno de audio 200a, tomada de manera perpendicular al eje z. Por consiguiente, la Figura 7 está dibujada desde la perspectiva de una persona que mira hacia abajo en el entorno de audio 200a, a lo largo del eje z. En este ejemplo, el entorno de audio 200a es el entorno de un sistema de sonido de una sala de cine que tiene una configuración Dolby Surround 7.1, tal como la mostrada en la Figura 2 y descrita anteriormente. Por consiguiente, el entorno de reproducción 200a incluye los altavoces envolventes laterales izquierdos 220, los altavoces envolventes traseros izquierdos 224, los altavoces envolventes laterales derechos 225, los altavoces envolventes traseros derechos 226, el canal de pantalla izquierdo 230, el canal de pantalla central 235, el canal de pantalla derecho 240 y el altavoz de graves 245.

55 El objeto de audio 610 tiene un tamaño indicado por el volumen de objeto de audio 620b, cuya área de sección transversal rectangular se muestra en la Figura 7. Dada la posición de objeto de audio 615 en el instante de tiempo ilustrado en la Figura 7, 12 ubicaciones de fuente virtual 605 están incluidas en el área abarcada por el volumen de objeto de audio 620b en el plano x-y. Dependiendo de la extensión del volumen de objeto de audio 620b en la dirección z y la separación de las ubicaciones de fuente virtual 605 a lo largo del eje z, ubicaciones de fuente virtual adicionales 605s pueden estar incluidas, o no, dentro del volumen de objeto de audio 620b.

60 La Figura 7 indica contribuciones de ganancia de las ubicaciones de fuente virtual 605 dentro del área o volumen definido por el tamaño del objeto de audio 610. En este ejemplo, el diámetro del círculo usado para ilustrar cada una de las ubicaciones de fuente virtual 605 corresponde a la contribución de la ubicación de fuente virtual 605 correspondiente. Las ubicaciones de fuente virtual 605a que están más cerca de la posición de objeto de audio 615 se muestran como las más grandes, que indican la mayor contribución de las fuentes virtuales correspondientes. Las segundas contribuciones de ganancia más grandes son de fuentes virtuales en las ubicaciones de fuente virtual

605b, que son las segundas más cercanas a la posición de objeto de audio 615. Contribuciones más pequeñas son las realizadas por las ubicaciones de fuente virtual 605c, que están más alejadas de la posición de objeto de audio 615 pero aún dentro del volumen de objeto de audio 620b. Las ubicaciones de fuente virtual 605d que están fuera del volumen de objeto de audio 620b se muestran como las más pequeñas, lo que indica que, en este ejemplo, las fuentes virtuales correspondientes no realizan ninguna contribución.

Haciendo referencia a la Figura 5C, en este ejemplo, el bloque 550 implica calcular un conjunto de valores de ganancia de objeto de audio para cada uno de una pluralidad de canales de salida en función de, al menos en parte, los valores calculados de ganancia de fuente virtual. Cada canal de salida puede corresponder a al menos un altavoz de reproducción del entorno de reproducción. El bloque 550 puede implicar normalizar los valores de ganancia de objeto de audio resultantes. Para la implementación mostrada en la Figura 7, por ejemplo, cada canal de salida puede corresponder a un único altavoz o a un grupo de altavoces.

El proceso de calcular el valor de ganancia de objeto de audio para cada uno de la pluralidad de canales de salida puede implicar determinar un valor de ganancia ( $g_l^{\text{tamaño}}(x_o, y_o, z_o, s)$ ) para un objeto de audio de tamaño ( $s$ ) que va a renderizarse en la ubicación  $x_o, y_o, z_o$ . Este valor de ganancia de objeto de audio puede denominarse algunas veces en el presente documento "contribución de tamaño de objeto de audio". Según algunas implementaciones, el valor de ganancia de objeto de audio ( $g_l^{\text{tamaño}}(x_o, y_o, z_o, s)$ ) puede expresarse como:

$$\left[ \sum_{x_{vs}, y_{vs}, z_{vs}} [w(x_{vs}, y_{vs}, z_{vs}; x_o, y_o, z_o; s) g_l(x_{vs}, y_{vs}, z_{vs})]^p \right]^{1/p} \quad \text{(Ecuación 2)}$$

En la Ecuación 2,  $(x_{vs}, y_{vs}, z_{vs})$  representa una ubicación de fuente virtual,  $g_l(x_{vs}, y_{vs}, z_{vs})$  representa un valor de ganancia para el canal  $l$  para la ubicación de fuente virtual  $x_{vs}, y_{vs}, z_{vs}$ , y  $w(x_{vs}, y_{vs}, z_{vs}; x_o, y_o, z_o; s)$  representa un peso para  $g_l(x_{vs}, y_{vs}, z_{vs})$  determinado, al menos en parte, en función de la ubicación  $(x_o, y_o, z_o)$  del objeto de audio, el tamaño ( $s$ ) del objeto de audio y la ubicación de fuente virtual  $(x_{vs}, y_{vs}, z_{vs})$ .

En algunos ejemplos, el exponente  $p$  puede tener un valor comprendido entre 1 y 10. En algunas implementaciones,  $p$  puede ser una función del tamaño de objeto de audio  $s$ . Por ejemplo, si  $s$  es relativamente mayor, en algunas implementaciones  $p$  puede ser relativamente menor. Según algunas de tales implementaciones,  $p$  puede determinarse de la siguiente manera:

$$p = 6, \quad \text{si } s \leq 0.5$$

$$p = 6 + (-4)(s - 0.5) / (s_{\text{max}} - 0.5), \quad \text{si } s > 0.5,$$

donde  $s_{\text{max}}$  corresponde al valor máximo de un tamaño interno escalado de manera ascendente  $s_{\text{interno}}$  (descrito posteriormente) y donde un tamaño de objeto de audio  $s = 1$  puede corresponder a un objeto de audio que tiene un tamaño (por ejemplo, un diámetro) igual a una longitud de uno de los límites del entorno de reproducción (por ejemplo, igual a la longitud de una pared del entorno de reproducción).

Dependiendo, en parte, del/de los algoritmo(s) usado(s) para calcular los valores de ganancia de fuente virtual, puede ser posible simplificar la Ecuación 2 si las ubicaciones de fuente virtual están distribuidas de manera uniforme a lo largo de un eje y si las funciones de ponderación y las funciones de ganancia pueden separarse, por ejemplo como se ha descrito anteriormente. Si estas condiciones se cumplen, entonces  $g_l(x_{vs}, y_{vs}, z_{vs})$  puede expresarse como  $g_{lx}(x_{vs})g_{ly}(y_{vs})g_{lz}(z_{vs})$ , donde  $g_{lx}(x_{vs})$ ,  $g_{ly}(y_{vs})$  y  $g_{lz}(z_{vs})$  representan funciones de ganancia independientes de coordenadas  $x$ ,  $y$  y  $z$  para una ubicación de fuente virtual.

Asimismo,  $w(x_{vs}, y_{vs}, z_{vs}; x_o, y_o, z_o; s)$  puede factorizarse como  $w_x(x_{vs}; x_o; s)w_y(y_{vs}; y_o; s)w_z(z_{vs}; z_o; s)$ , donde  $w_x(x_{vs}; x_o; s)$ ,  $w_y(y_{vs}; y_o; s)$  y  $w_z(z_{vs}; z_o; s)$  representan funciones de ponderación independientes de coordenadas  $x$ ,  $y$  y  $z$  para una ubicación de fuente virtual. Un ejemplo de este tipo se muestra en la Figura 7. En este ejemplo, la función de ponderación 710, expresada como  $w_x(x_{vs}; x_o; s)$ , puede calcularse de manera independiente a partir de la función de ponderación 720, expresada como  $w_y(y_{vs}; y_o; s)$ . En algunas implementaciones, las funciones de ponderación 710 y 720 pueden ser funciones gaussianas, donde la función de ponderación  $w_z(z_{vs}; z_o; s)$  puede ser un producto de funciones gaussianas y de coseno.

Si  $w(x_{vs}, y_{vs}, z_{vs}; x_o, y_o, z_o; s)$  puede factorizarse como  $w_x(x_{vs}; x_o; s)w_y(y_{vs}; y_o; s)w_z(z_{vs}; z_o; s)$ , entonces la Ecuación 2 se simplifica pasando a ser:

$$\left[ f_l^x(x_o; s) f_l^y(y_o; s) f_l^z(z_o; s) \right]^{1/p},$$

donde

$$f_l^x(x_o; s) = \sum_{x_s} [g_l(x_s) w(x_s; x_o; s)]^p,$$

$$f_l^y(y_o; s) = \sum_{y_s} [g_l(y_s) w(y_s; y_o; s)]^p$$

y

$$f_l^z(z_o; s) = \sum_{z_s} [g_l(z_s) w(z_s; z_o; s)]^p.$$

5 Las funciones  $f$  pueden contener toda la información requerida según las fuentes virtuales. Si las posibles posiciones de objeto se vuelven discretas a lo largo de cada eje, puede expresarse cada función  $f$  como una matriz. Cada función  $f$  puede precalcularse durante el proceso de configuración del bloque 505 (véase la Figura 5A) y almacenarse en un sistema de memoria, por ejemplo como una matriz o como una tabla de consulta. En tiempo de ejecución (bloque 510), las matrices o tablas de consulta pueden recuperarse del sistema de memoria. El proceso en tiempo de ejecución puede implicar interpolar, dados un tamaño y posición de objeto de audio, entre los valores correspondientes más próximos de estas matrices. En algunas implementaciones, la interpolación puede ser lineal.

15 En algunas implementaciones, la contribución de tamaño de objeto de audio  $g_l^{tamaño}$  puede combinarse con el resultado de "casi ganancia de objeto de audio" para la posición de objeto de audio. Como se usa en el presente documento, la "casi ganancia de objeto de audio" es una ganancia calculada que está basada en la posición de objeto de audio 615. El cálculo de ganancias puede realizarse usando el mismo algoritmo usado para calcular cada uno de los valores de ganancia de fuente virtual. Según algunas de tales implementaciones, un cálculo de desvanecimiento cruzado puede realizarse entre la contribución de tamaño de objeto de audio y el resultado de casi ganancia de objeto de audio, por ejemplo en función del tamaño de objeto de audio. Tales implementaciones pueden proporcionar una panoramización gradual y un crecimiento gradual de objetos de audio, y puede permitir una transición gradual entre el tamaño de objeto de audio más pequeño y el más grande. En una de estas implementaciones,

$$25 \quad g_l^{total}(x_o, y_o, z_o; s) = \alpha(s) g_l^{casi ganancia}(x_o, y_o, z_o; s) + \beta(s) \tilde{g}_l^{tamaño}(x_o, y_o, z_o; s),$$

donde

$$30 \quad \begin{aligned} s < s_{desv_x}, & \quad \alpha = \cos((s / s_{desv_x})(\pi / 2)), & \quad \beta = \text{sen}((s / s_{desv_x})(\pi / 2)) \\ s \geq s_{desv_x}, & \quad \alpha = 0, \beta = 1, \end{aligned}$$

y donde  $\tilde{g}_l^{tamaño}$  representa la versión normalizada de  $g_l^{tamaño}$  antes calculada. En algunas de estas implementaciones,  $s_{desv_x} = 0,2$ . Sin embargo, en implementaciones alternativas,  $s_{desv_x}$  puede tener otros valores.

35 Según algunas implementaciones, el valor de tamaño de objeto de audio puede escalarse de manera ascendente en la parte más grande de su intervalo de posibles valores. En algunas implementaciones de creación, por ejemplo, un usuario puede estar expuesto a valores de tamaño de objeto de audio  $s_{usuario} \in [0,1]$  que están correlacionados, en el tamaño real usado por el algoritmo, con un intervalo mayor, por ejemplo el intervalo  $[0, s_{max}]$ , donde  $s_{max} > 1$ . Esta correlación puede garantizar que cuando el tamaño se ajuste a un valor máximo por el usuario, las ganancias se vuelven completamente independientes de la posición del objeto. Según algunas de estas implementaciones, tales correlaciones pueden realizarse según una función lineal orientada a elementos que conecta pares de puntos ( $s_{usuario}, s_{interno}$ ), donde  $s_{usuario}$  representa un tamaño de objeto de audio seleccionado por el usuario y  $s_{interno}$  representa un tamaño de objeto de audio correspondiente determinado por el algoritmo. Según algunas de estas implementaciones, la correlación puede realizarse según una función lineal orientada a elementos que conecta pares de puntos (0, 0), (0,2, 0,3), (0,5, 0,9), (0,75, 1,5) y (1,  $s_{max}$ ). En una de estas implementaciones,  $s_{max} = 2,8$ .

50 Las Figuras 8A y 8B muestran un objeto de audio en dos posiciones dentro de un entorno de reproducción. En estos ejemplos, el volumen de objeto de audio 620b es una esfera que tiene un radio inferior a la mitad de la longitud o anchura del entorno de reproducción 200a. El entorno de reproducción 200a está configurado según la configuración Dolby 7.1. En el instante de tiempo ilustrado en la Figura 8A, la posición de objeto de audio 615 está relativamente más cerca de la mitad del entorno de reproducción 200a. En el instante de tiempo ilustrado en la Figura 8B, la posición de objeto de audio 615 se ha acercado a un límite del entorno de reproducción 200a. En este ejemplo, el límite es la pared izquierda de una sala de cine y coincide con las ubicaciones de los altavoces envolventes laterales izquierdos 220.

55 Por razones estéticas, puede ser deseable modificar los cálculos de ganancia de objeto de audio para objetos de audio que se aproximan a un límite de un entorno de reproducción. En las Figuras 8A y 8B, por ejemplo, no se proporciona ninguna señal de alimentación a los altavoces en un límite opuesto del entorno de reproducción (en este

caso, los altavoces envolventes laterales derechos 225) cuando la posición de objeto de audio 615 está dentro de una distancia umbral desde el límite izquierdo 805 del entorno de reproducción. En el ejemplo mostrado en la Figura 8B, ninguna señal de alimentación de altavoz se proporciona a altavoces correspondientes al canal de pantalla izquierdo 230, al canal de pantalla central 235, al canal de pantalla derecho 240 o al altavoz de graves 245 cuando la posición de objeto de audio 615 está dentro de una distancia de umbral (que puede ser una distancia umbral diferente) del límite izquierdo 805 del entorno de reproducción, si la posición de objeto de audio 615 está también más lejos que una distancia umbral con respecto a la pantalla.

En el ejemplo mostrado en la Figura 8B, el volumen de objeto de audio 620b incluye un área o volumen externo del límite izquierdo 805. Según algunas implementaciones, un factor de desvanecimiento para cálculos de ganancia puede basarse, al menos en parte, en la cantidad del límite izquierdo 805 que está dentro del volumen de objeto de audio 620b y/o en la cantidad del área o volumen de un objeto de audio que se extiende fuera de tal límite.

La Figura 9 es un diagrama de flujo que muestra un procedimiento para determinar un factor de desvanecimiento en función de, al menos en parte, qué cantidad de un área o volumen de un objeto de audio se extiende fuera de un límite de un entorno de reproducción. En el bloque 905 se reciben datos de entorno de reproducción. En este ejemplo, los datos de entorno de reproducción incluyen datos de ubicación de altavoz de reproducción y datos de límite de entorno de reproducción. El bloque 910 implica recibir datos de reproducción de audio que incluyen uno o más objetos de audio y metadatos asociados. Los metadatos incluyen, en este ejemplo, al menos datos de posición de objeto de audio y datos de tamaño de objeto de audio.

En esta implementación, el bloque 915 implica determinar que un área o volumen de objeto de audio, definido por los datos de posición de objeto de audio y los datos de tamaño de objeto de audio, incluye un área o volumen externo fuera de un límite de entorno de reproducción. El bloque 915 también puede implicar determinar qué proporción del área o volumen de objeto de audio está fuera del límite de entorno de reproducción.

En el bloque 920, se determina un factor de desvanecimiento. En este ejemplo, el factor de desvanecimiento puede basarse, al menos en parte, en el área externa. Por ejemplo, el factor de desvanecimiento puede ser proporcional al área externa.

En el bloque 925, un conjunto de valores de ganancia de objeto de audio puede calcularse para cada uno de una pluralidad de canales de salida en función de, al menos en parte, los metadatos asociados (en este ejemplo, los datos de posición de objeto de audio y los datos de tamaño de objeto de audio) y el factor de desvanecimiento. Cada canal de salida puede corresponder a al menos un altavoz de reproducción del entorno de reproducción.

En algunas implementaciones, los cálculos de ganancia de objeto de audio pueden implicar calcular contribuciones de ganancia de fuentes virtuales dentro de un área o volumen de objeto de audio. Las fuentes virtuales pueden corresponder a una pluralidad de ubicaciones de fuente virtual que pueden definirse con referencia a los datos de entorno de reproducción. Las ubicaciones de fuente virtual pueden estar, o no, separadas de manera uniforme. Para cada una de las ubicaciones de fuente virtual, un valor de ganancia de fuente virtual puede calcularse para cada uno de la pluralidad de canales de salida. Como se ha descrito anteriormente, en algunas implementaciones, estos valores de ganancia de fuente virtual pueden calcularse y almacenarse durante un proceso de configuración, y después recuperarse para usarse durante operaciones en tiempo de ejecución.

En algunas implementaciones, el factor de desvanecimiento puede aplicarse a todos los valores de ganancia de fuente virtual correspondientes a ubicaciones de fuente de virtual dentro de un entorno de reproducción. En algunas implementaciones,  $g_l^{tamaño}$  puede modificarse de la siguiente manera:

$$g_l^{tamaño} = \left[ g_l^{límite} + (\text{factor\_desvanecimiento}) \times g_l^{dentro} \right]^{1/p},$$

donde

$$\begin{aligned} \text{factor\_desvanecimiento} &= 1, & \text{si } d_{límite} &\geq s, \\ \text{factor\_desvanecimiento} &= d_{límite} / s, & \text{si } d_{límite} &< s, \end{aligned}$$

donde  $d_{límite}$  representa la distancia mínima entre una ubicación de objeto de audio y un límite del entorno de reproducción, y  $g_l^{límite}$  representa la contribución de fuentes virtuales a lo largo de un límite. Por ejemplo, haciendo referencia a la Figura 8B,  $g_l^{límite}$  puede representar la contribución de fuentes virtuales dentro del volumen de objeto de audio 620b y adyacentes al límite 805. En este ejemplo, como el de la Figura 6A, no hay fuentes virtuales situadas fuera del entorno de reproducción.

En implementaciones alternativas,  $g_l^{tamaño}$  puede modificarse de la siguiente manera:

$$g_l^{\text{tamaño}} = \left[ g_l^{\text{fuera}} + (\text{factor\_desvanecimiento}) \times g_l^{\text{dentro}} \right]^{1/p},$$

donde  $g_l^{\text{fuera}}$  representa ganancias de objeto de audio basadas en fuentes virtuales ubicadas fuera de un entorno de reproducción pero dentro de un área o volumen de objeto de audio. Por ejemplo, haciendo referencia a la Figura 8B,  $g_l^{\text{fuera}}$  puede representar la contribución de fuentes virtuales dentro del volumen de objeto de audio 620b y fuera del límite 805. En este ejemplo, al igual que en la Figura 6B, hay fuentes virtuales situadas dentro y fuera del entorno de reproducción.

La Figura 10 es un diagrama de bloques que proporciona ejemplos de componentes de un aparato de creación y/o renderización. En este ejemplo, el dispositivo 1000 incluye un sistema de interfaz 1005. El sistema de interfaz 1005 puede incluir una interfaz de red, tal como una interfaz de red inalámbrica. Como alternativa, o adicionalmente, el sistema de interfaz 1005 puede incluir una interfaz de bus serie universal (USB) u otra interfaz de este tipo.

El dispositivo 1000 incluye un sistema lógico 1010. El sistema lógico 1010 puede incluir un procesador, tal como un procesador de propósito general de un solo o varios chips. El sistema lógico 1010 puede incluir un procesador de señales digitales (DSP), un circuito integrado de aplicación específica (ASIC), una matriz de puertas programables *in situ* (FPGA) u otro dispositivo de lógica programable, lógica de transistor o puertas discretas, componentes de hardware discretos, o combinaciones de los mismos. El sistema lógico 1010 puede estar configurado para controlar los otros componentes del dispositivo 1000. Aunque no se muestra ninguna interfaz entre los componentes del dispositivo 1000 en la Figura 10, el sistema lógico 1010 puede estar configurado con interfaces para la comunicación con otros componentes. Los otros componentes pueden estar configurados, o no, para una comunicación recíproca, según sea apropiado.

El sistema lógico 1010 puede estar configurado para llevar a cabo una funcionalidad de creación y/o renderización de audio que incluye, pero sin limitarse a, los tipos de funcionalidad de creación y/o renderización de audio descritos en el presente documento. En algunas de estas implementaciones, el sistema lógico 1010 puede estar configurado para funcionar, al menos en parte, según el software almacenado en uno o más medios no transitorios. Los medios no transitorios pueden incluir una memoria asociada al sistema lógico 1010, tal como una memoria de acceso aleatorio (RAM) y/o una memoria de solo lectura (ROM). Los medios no transitorios pueden incluir una memoria del sistema de memoria 1015. El sistema de memoria 1015 puede incluir uno o más tipos adecuados de medios de almacenamiento no transitorios, tales como memoria flash, un disco duro, etc.

El sistema de visualización 1030 puede incluir uno o más tipos de visualización adecuados, dependiendo de la implementación de dispositivo 1000. Por ejemplo, el sistema de visualización 1030 puede incluir una pantalla de cristal líquido, una pantalla de plasma, una pantalla biestable, etc.

El sistema de entrada de usuario 1035 puede incluir uno o más dispositivos configurados para aceptar entradas de un usuario. En algunas implementaciones, el sistema de entrada de usuario 1035 puede incluir una pantalla táctil superpuesta a una pantalla del sistema de visualización 1030. El sistema de entrada de usuario 1035 puede incluir un ratón, una bola de seguimiento, un sistema de detección de gestos, una palanca de control, una o más GUI y/o menús presentados en el sistema de visualización 1030, botones, un teclado, conmutadores, etc. En algunas implementaciones, el sistema de entrada de usuario 1035 puede incluir el micrófono 1025: un usuario puede proporcionar comandos de voz para el dispositivo 1000 a través del micrófono 1025. El sistema lógico puede estar configurado para el reconocimiento de voz y para controlar al menos algunas operaciones del dispositivo 1000 según tales comandos de voz.

El sistema de suministro de energía 1040 puede incluir uno o más dispositivos de almacenamiento de energía adecuados, tales como una batería de níquel-cadmio o una batería de iones de litio. El sistema de suministro de energía 1040 puede estar configurado para recibir energía de una toma de corriente.

La Figura 11A es un diagrama de bloques que representa algunos componentes que pueden usarse en la creación de contenido de audio. El sistema 1100 puede usarse, por ejemplo, para la creación de contenido de audio en estudios de mezcla y/o etapas de doblaje. En este ejemplo, el sistema 1100 incluye una herramienta de creación de audio y metadatos 1105 y una herramienta de renderización 1110. En esta implementación, la herramienta de creación de audio y metadatos 1105 y la herramienta de renderización 1110 incluyen interfaces de conexión de audio 1107 y 1112, respectivamente, que pueden configurarse para la comunicación a través de AES/EBU, MADI, medios analógicos, etc. La herramienta de creación de audio y metadatos 1105 y la herramienta de renderización 1110 incluyen interfaces de red 1109 y 1117, respectivamente, que pueden configurarse para enviar y recibir metadatos mediante TCP/IP o cualquier otro protocolo adecuado. La interfaz 1120 está configurada para proporcionar datos de audio a los altavoces.

El sistema 1100 puede incluir, por ejemplo, un sistema de creación existente, tal como el sistema Pro Tools™, que ejecute una herramienta de creación de metadatos (es decir, un panoramizador como el descrito en el presente documento) como accesorio. El panoramizador también puede ejecutarse en un sistema autónomo (por ejemplo, un



- PC o una consola de mezclas) conectado a la herramienta de renderización 1110, o puede ejecutarse en el mismo dispositivo físico que la herramienta de renderización 1110. En el segundo caso, el panoramizador y el renderizador pueden usar una conexión local, por ejemplo mediante una memoria compartida. La GUI del panoramizador también puede estar presente en un dispositivo de tipo tableta, un ordenador portátil, etc. La herramienta de renderización
- 5 1110 puede comprender un sistema de renderización que incluye un procesador de sonido que está configurado para ejecutar procedimientos de renderización como los descritos en las Fig. 5A a C y en la Fig. 9. El sistema de renderización puede incluir, por ejemplo, un ordenador personal, un ordenador portátil, etc., que incluya interfaces para la entrada/salida de audio y un sistema lógico apropiado.
- 10 La Figura 11B es un diagrama de bloques que representa algunos componentes que pueden usarse en la reproducción de audio en un entorno de reproducción (por ejemplo, en una sala de cine). El sistema 1150 incluye un servidor de sala de cine 1155 y un sistema de renderización 1160 en este ejemplo. El servidor de sala de cine 1155 y el sistema de renderización 1160 incluyen interfaces de red 1157 y 1162, respectivamente, que pueden configurarse para enviar y recibir objetos de audio a través de TCP/IP o cualquier otro protocolo adecuado. La
- 15 interfaz 1164 está configurada para proporcionar datos de audio a los altavoces.
- Varias modificaciones en las implementaciones descritas en esta divulgación pueden resultar fácilmente evidentes a los expertos en la técnica. Los principios generales definidos en el presente documento pueden aplicarse a otras implementaciones sin apartarse del espíritu o el alcance de esta divulgación. Por tanto, las reivindicaciones no están
- 20 limitadas a las implementaciones mostradas en el presente documento, sino que se les concede el alcance más amplio coherente con esta divulgación, los principios y características novedosas descritos en el presente documento.

**REIVINDICACIONES**

1. Un procedimiento, que comprende:

- 5 recibir (540) datos de reproducción de audio que comprenden uno o más objetos de audio, comprendiendo los objetos de audio señales de audio y metadatos asociados, incluyendo los metadatos al menos datos de posición de objeto de audio y datos de tamaño de objeto de audio;
- calcular (545), para un objeto de audio de los uno o más objetos de audio, valores de ganancia de fuente virtual de fuentes virtuales en ubicaciones de fuente virtual respectivas dentro de un área o volumen de objeto de audio definido por los datos de posición de objeto de audio y los datos de tamaño de objeto de audio; y
- 10 calcular (550) un conjunto de valores de ganancia de objeto de audio para cada uno de una pluralidad de canales de salida en función de, al menos en parte, los valores calculados de ganancia de fuente virtual, donde cada canal de salida corresponde a al menos un altavoz de reproducción de un entorno de reproducción y cada una de dichas ubicaciones de fuente virtual corresponde a una ubicación estática respectiva dentro del entorno de reproducción.
- 15
2. El procedimiento según la reivindicación 1, en el que el proceso de calcular el conjunto de valores de ganancia de objeto de audio implica calcular un promedio ponderado de valores de ganancia de fuente virtual de las fuentes virtuales dentro del área o volumen de objeto de audio.
- 20
3. El procedimiento según la reivindicación 2, en el que los pesos para el promedio ponderado dependen de la posición del objeto de audio, del tamaño del objeto de audio y de cada ubicación de fuente virtual dentro del área o volumen del objeto de audio.
- 25
4. El procedimiento según la reivindicación 1, que comprende además:
- recibir datos de entorno de reproducción que incluyen datos de ubicación de altavoz de reproducción.
- 30
5. El procedimiento según la reivindicación 4, que comprende además:
- definir una pluralidad de ubicaciones de fuente virtual según los datos de entorno de reproducción; y calcular, para cada una de las ubicaciones de fuentes virtual, un valor de ganancia de fuente virtual para cada uno de la pluralidad de canales de salida.
- 35
6. El procedimiento según la reivindicación 5, en el que al menos algunas de las ubicaciones de fuente virtual corresponden a ubicaciones fuera del entorno de reproducción.
- 40
7. El procedimiento según la reivindicación 5, en el que las ubicaciones de fuente virtual están separadas de manera uniforme a lo largo de los ejes x, y y z.
- 45
8. El procedimiento según la reivindicación 5, en el que las ubicaciones de fuente virtual tienen una primera separación uniforme a lo largo de los ejes x e y y una segunda separación uniforme a lo largo de un eje z.
9. El procedimiento según la reivindicación 7 o la reivindicación 8, en el que el proceso de calcular el conjunto de valores de ganancia de objeto de audio para cada uno de la pluralidad de canales de salida implica cálculos independientes de valores de ganancia de fuente virtual de fuentes virtuales a lo largo de los ejes x, y y z.
- 50
10. El procedimiento según la reivindicación 5, en el que las ubicaciones de fuente virtual están separadas de manera no uniforme.
- 55
11. El procedimiento según la reivindicación 4, que comprende además almacenar en un sistema de memoria valores calculados de ganancia de fuente virtual.
12. El procedimiento según la reivindicación 11, en el que el proceso de calcular el conjunto de valores de ganancia de objeto de audio implica:
- recuperar, del sistema de memoria, valores calculados de ganancia de fuente virtual correspondientes a una posición y tamaño de objeto de audio; e
- interpolarse entre los valores calculados de ganancia de fuente virtual.
- 60
13. El procedimiento según la reivindicación 1, en el que el área o volumen de objeto de audio es al menos uno de entre un rectángulo, un prisma rectangular, un círculo, una esfera, una elipse o un elipsoide.

14. El procedimiento según la reivindicación 1, en el que el entorno de reproducción comprende un entorno de sistema de sonido de una sala de cine.
- 5 15. El procedimiento según la reivindicación 1, que comprende además descorrelacionar al menos algunos de los datos de reproducción de audio.
16. El procedimiento según la reivindicación 1, que comprende además descorrelacionar los datos de reproducción de audio para objetos de audio que tienen un tamaño de objeto de audio que supera un valor umbral.
- 10 17. El procedimiento según la reivindicación 1, en el que los datos de entorno de reproducción incluyen datos de límite de entorno de reproducción, que comprende además:
- determinar que el área o volumen de objeto de audio incluye un área o volumen externo fuera de un límite de entorno de reproducción; y
- 15 aplicar un factor de desvanecimiento en función de, al menos en parte, el área o volumen externo.
18. El procedimiento según la reivindicación 1, que comprende además:
- recibir datos de entorno de reproducción que incluyen datos de ubicación de altavoz de reproducción y datos de límite de entorno de reproducción;
- 20 determinar que el área o volumen de objeto de audio incluye un área o volumen externo fuera de un límite de entorno de reproducción; y
- determinar un factor de desvanecimiento en función de, al menos en parte, el área o volumen externo, donde calcular el conjunto de valores de ganancia de la pluralidad de canales de salida también se basa en los metadatos asociados y el factor de desvanecimiento.
- 25 19. El procedimiento según la reivindicación 18, en el que el factor de desvanecimiento es proporcional al área externa.
- 30 20. El procedimiento según la reivindicación 18, que comprende además:
- determinar que un objeto de audio está dentro de una distancia umbral con respecto a un límite de entorno de reproducción; y
- no proporcionar ninguna señal de alimentación de altavoz a altavoces de reproducción en un límite opuesto del entorno de reproducción.
- 35 21. El procedimiento según la reivindicación 18, que comprende además:
- definir una pluralidad de ubicaciones de fuente virtual según los datos de entorno de reproducción; y
- 40 calcular, para cada una de las ubicaciones de fuente virtual, una ganancia de fuente virtual para cada uno de una pluralidad de canales de salida.
22. Un medio no transitorio que tiene software almacenado en el mismo, incluyendo el software instrucciones para controlar al menos un aparato para llevar a cabo el procedimiento según una cualquiera de las reivindicaciones anteriores.
- 45 23. Un aparato (1000), que comprende:
- un sistema de interfaz (1005); y
- 50 un sistema lógico (1010) adaptado para controlar el aparato para llevar a cabo el procedimiento según una cualquiera de las reivindicaciones anteriores.

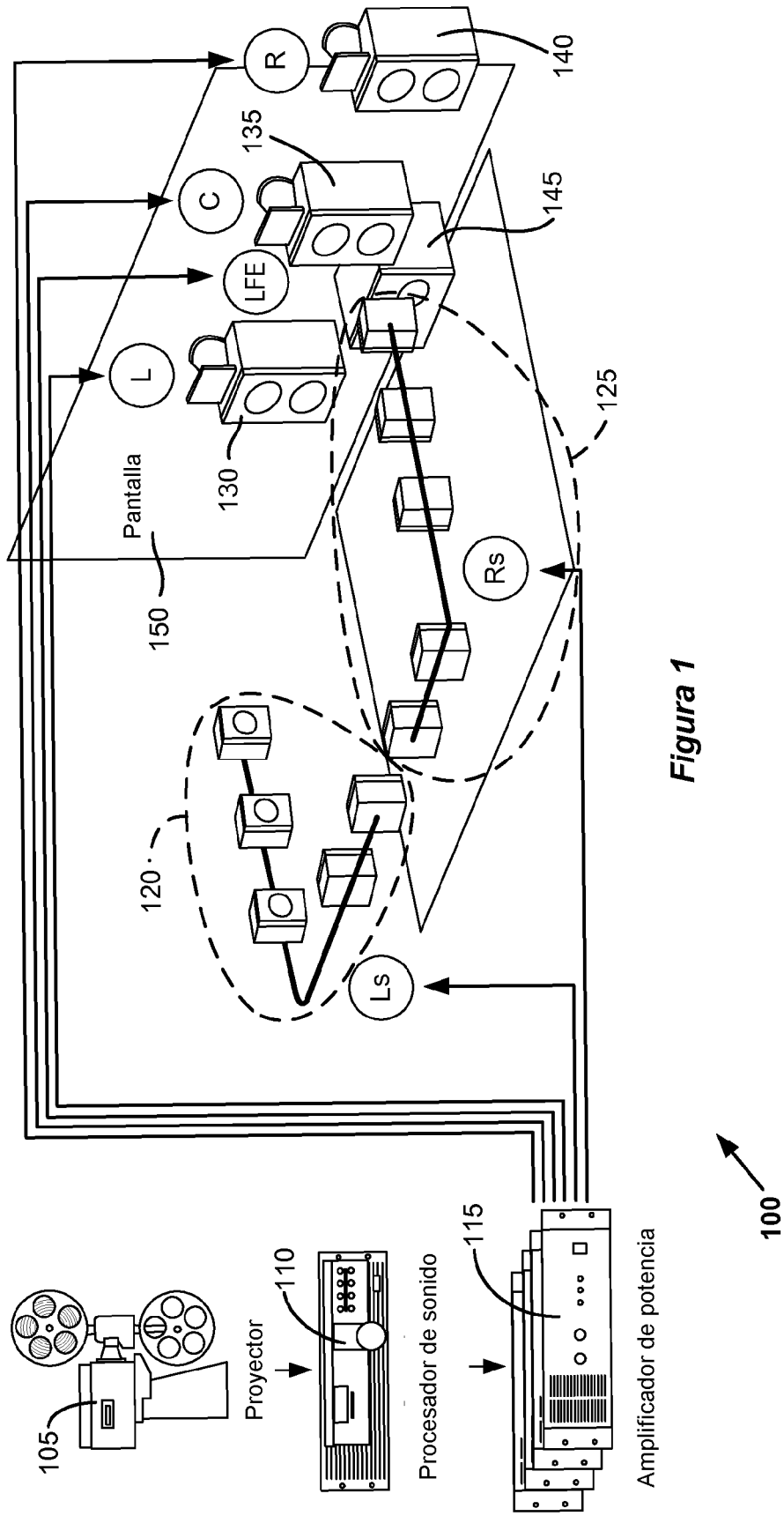


Figura 1

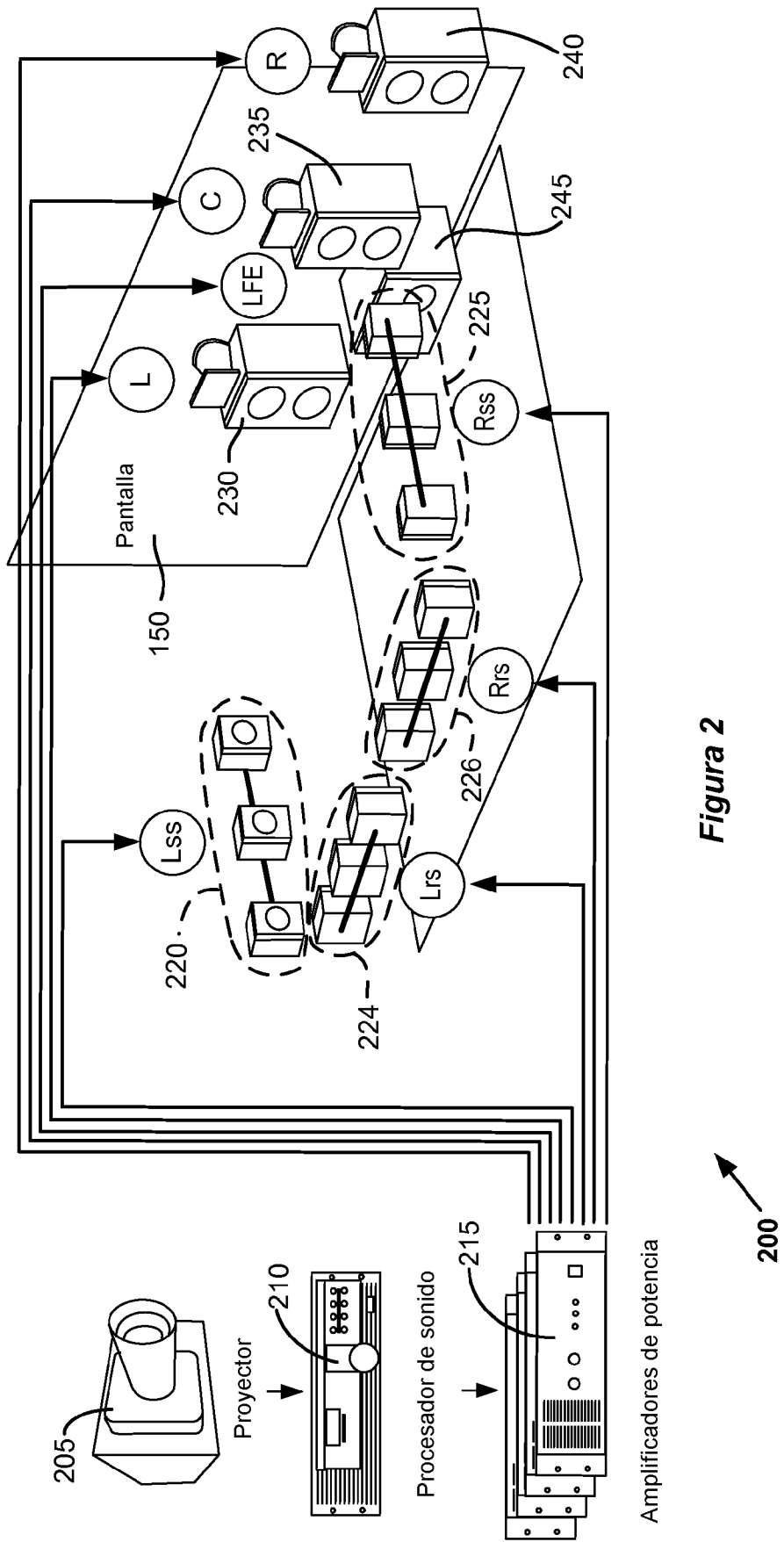


Figura 2

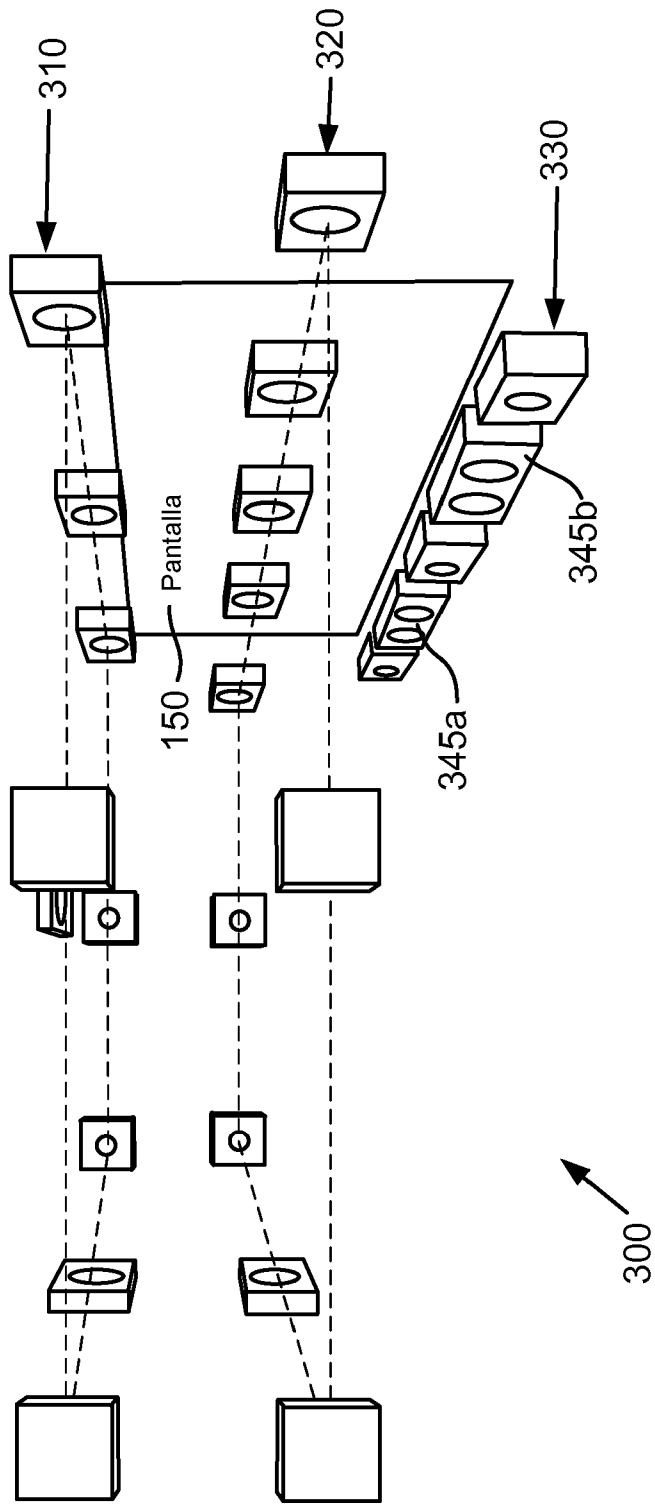


Figura 3

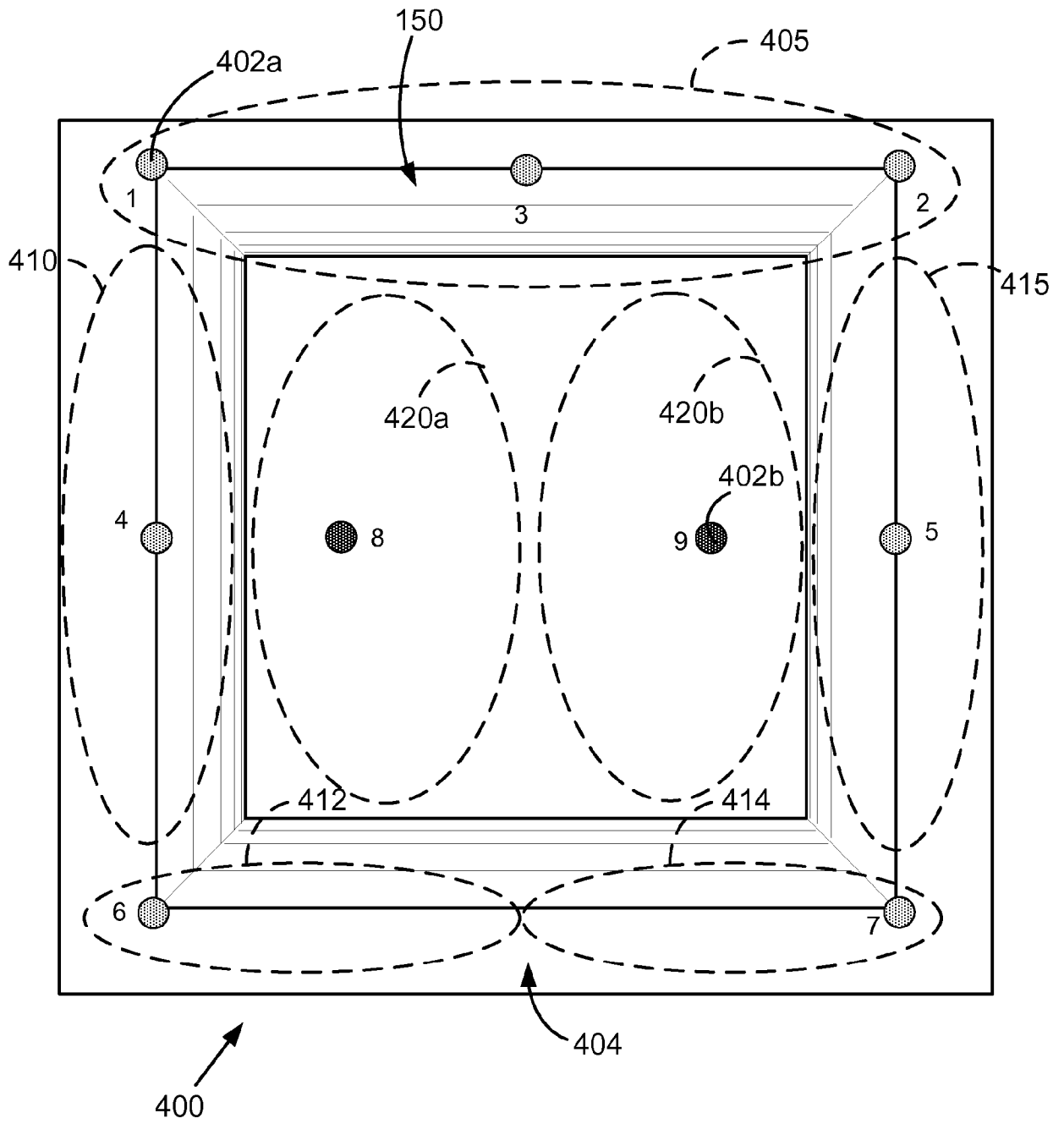


Figura 4a

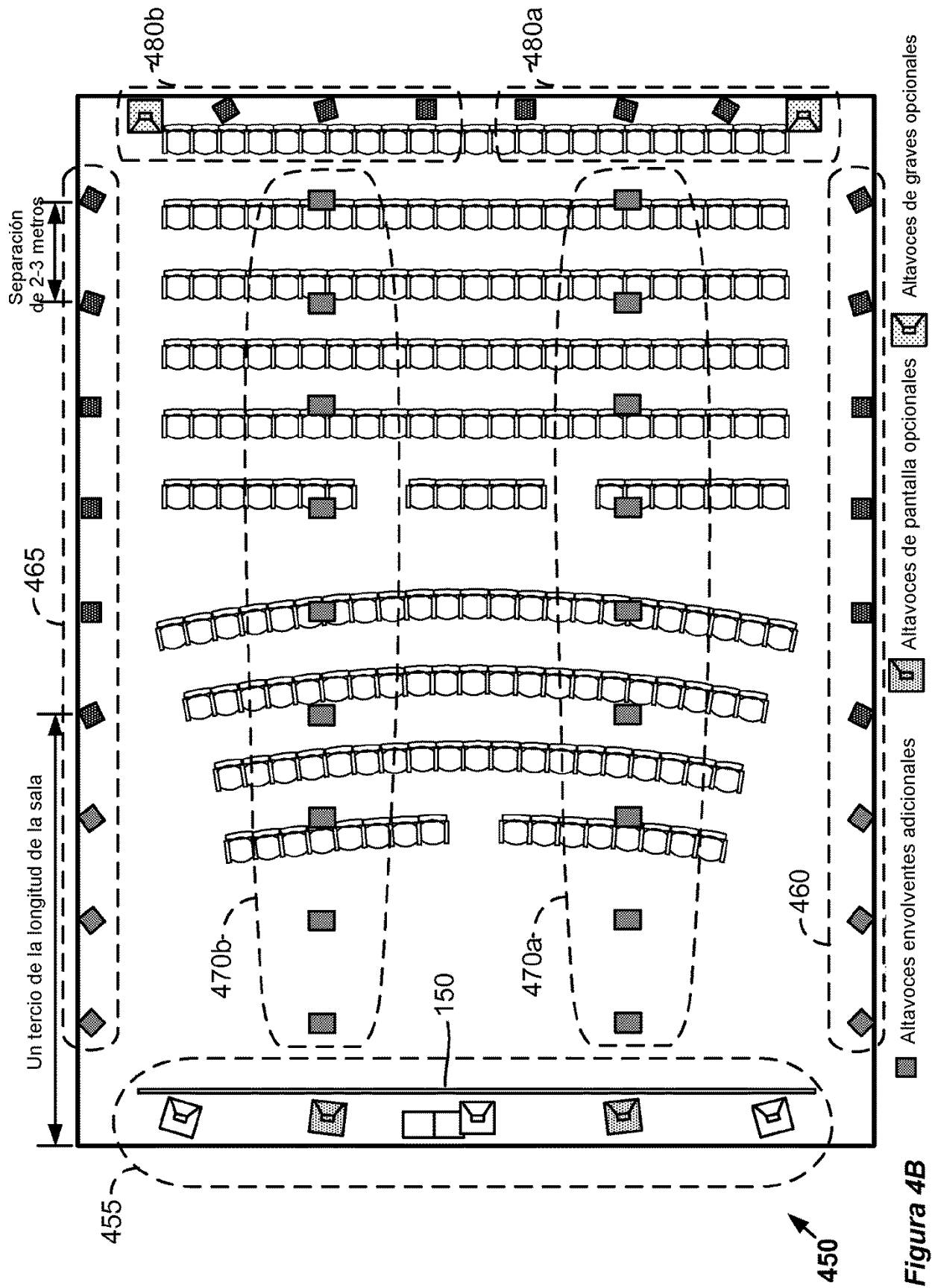
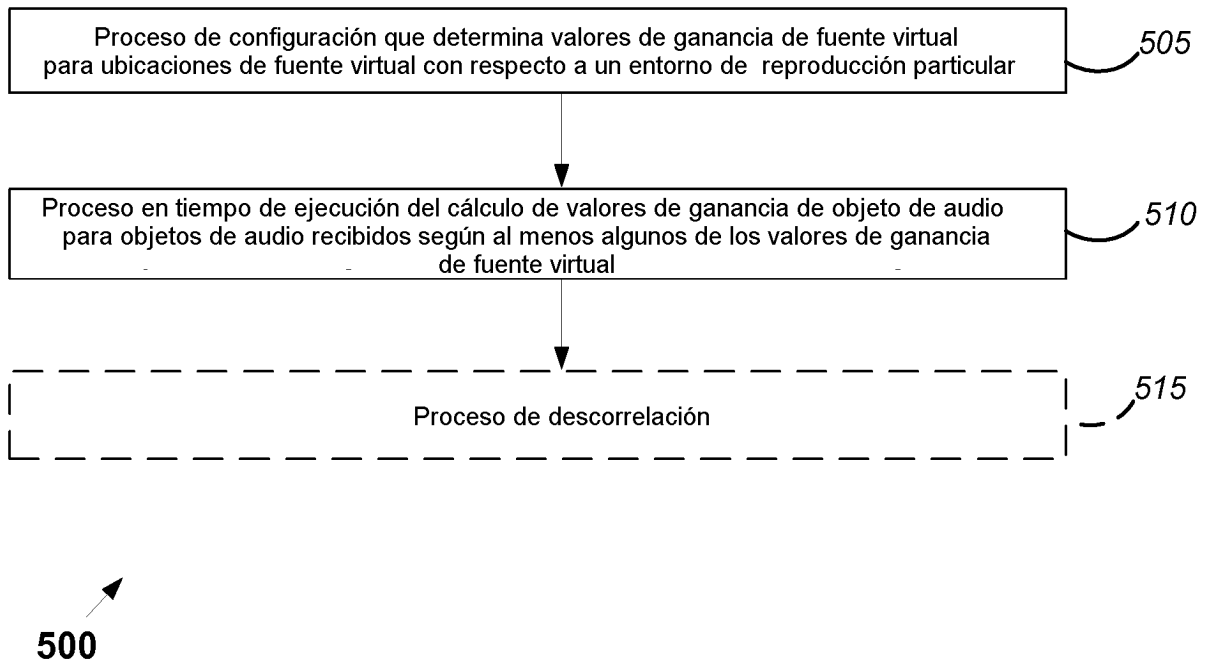
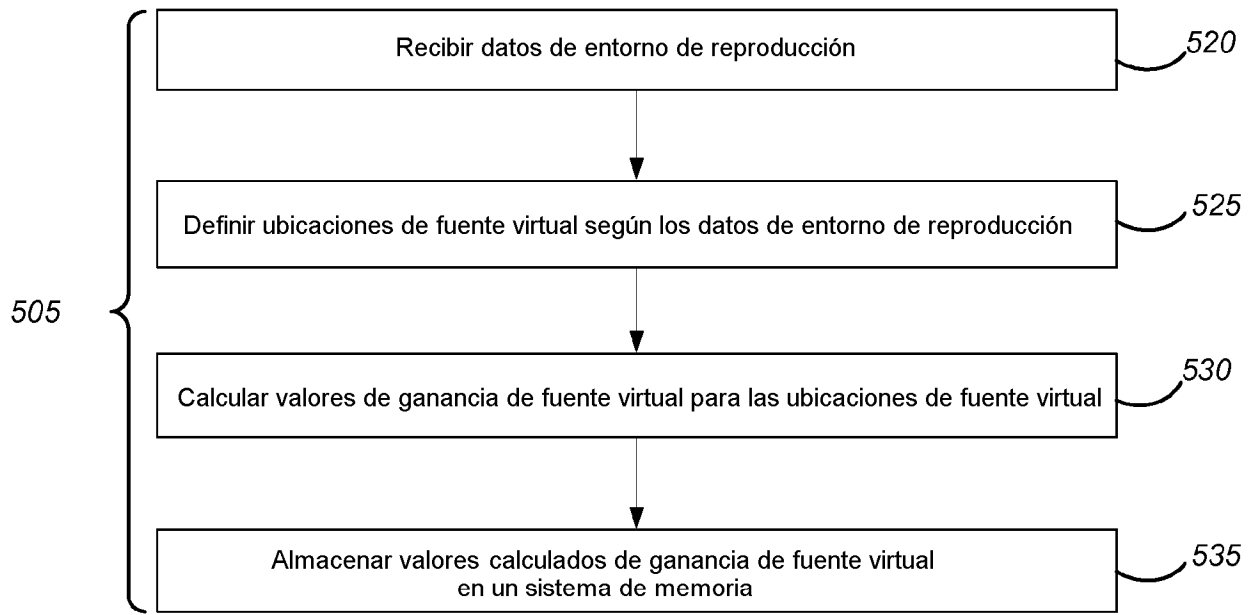


Figura 4B

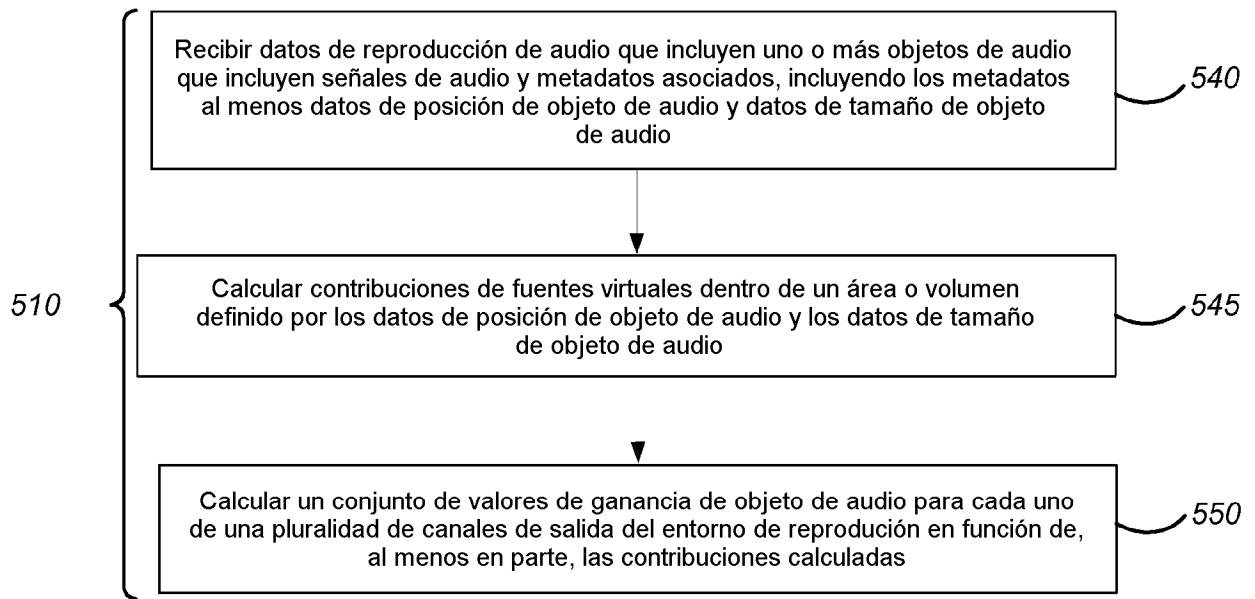




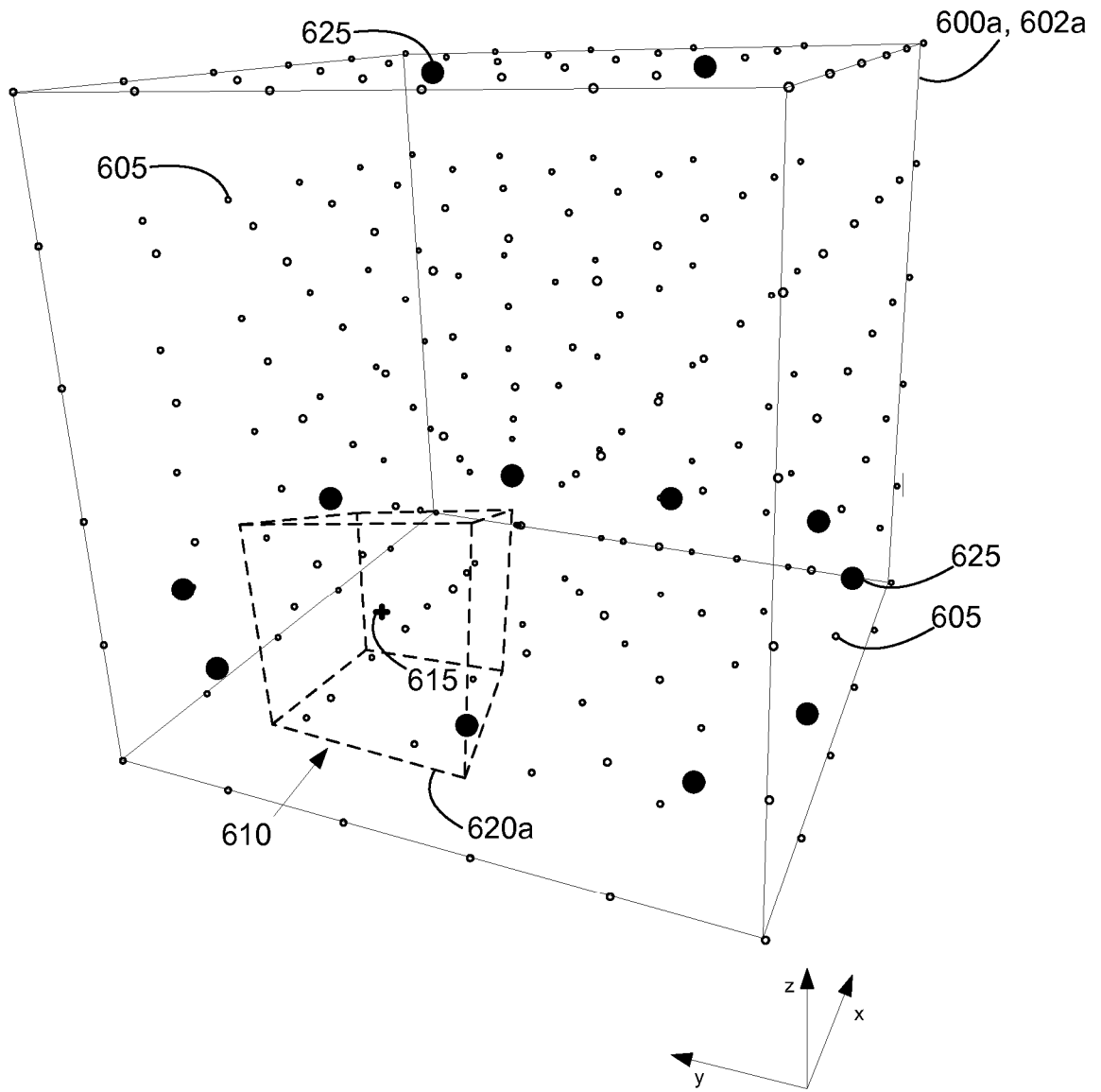
**Figura 5A**



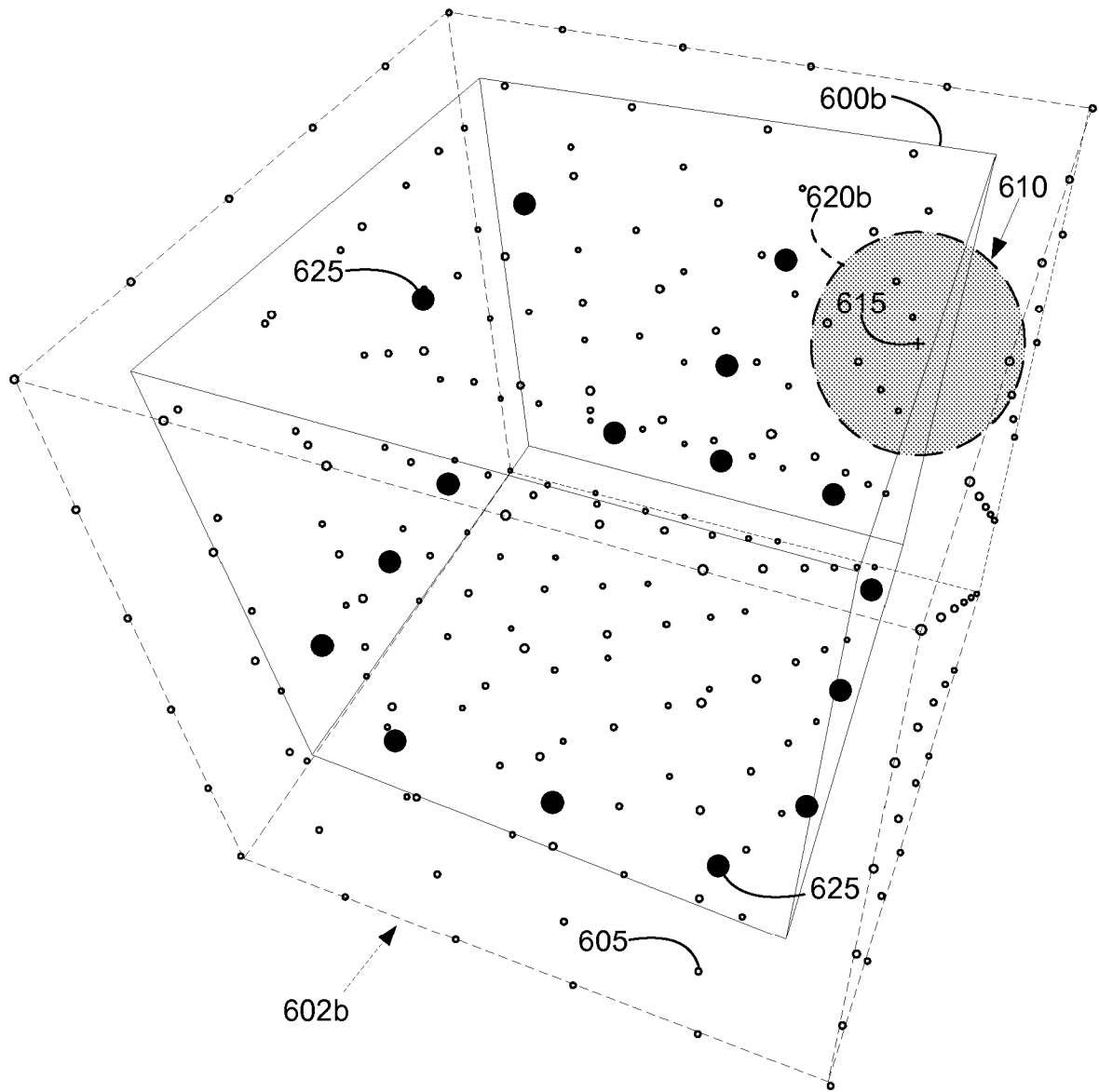
**Figura 5B**



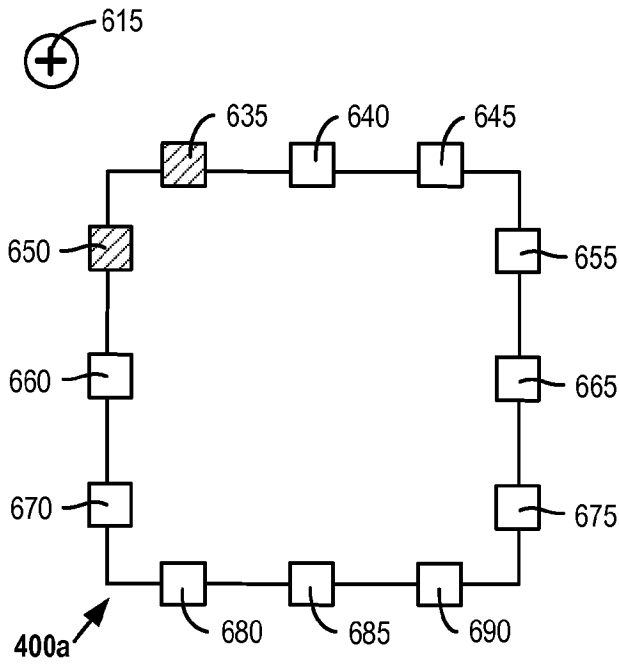
**Figura 5C**



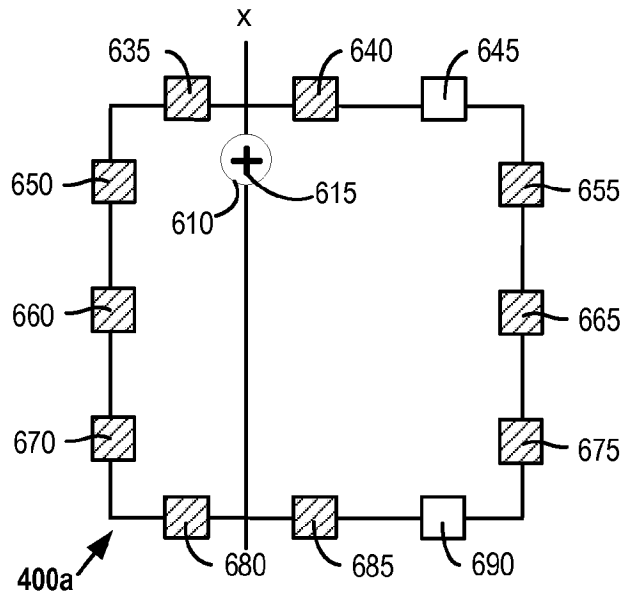
**Figura 6A**



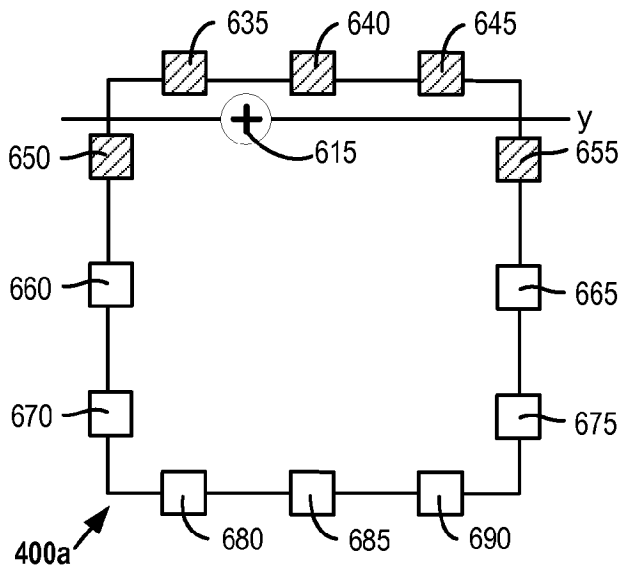
**Figura 6B**



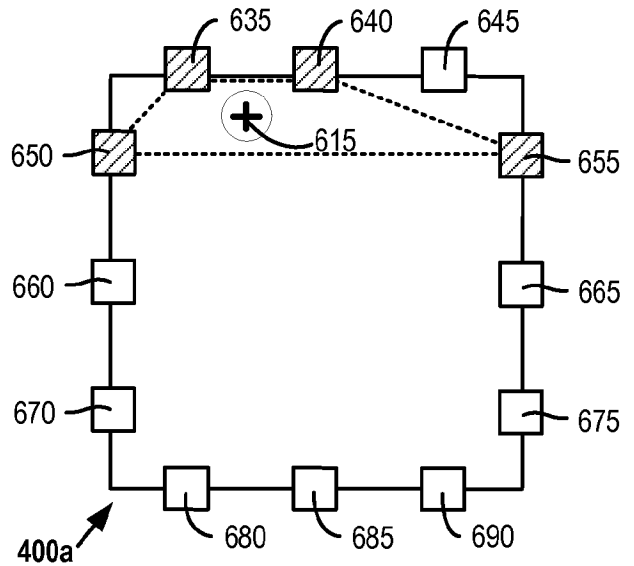
**Figura 6C**



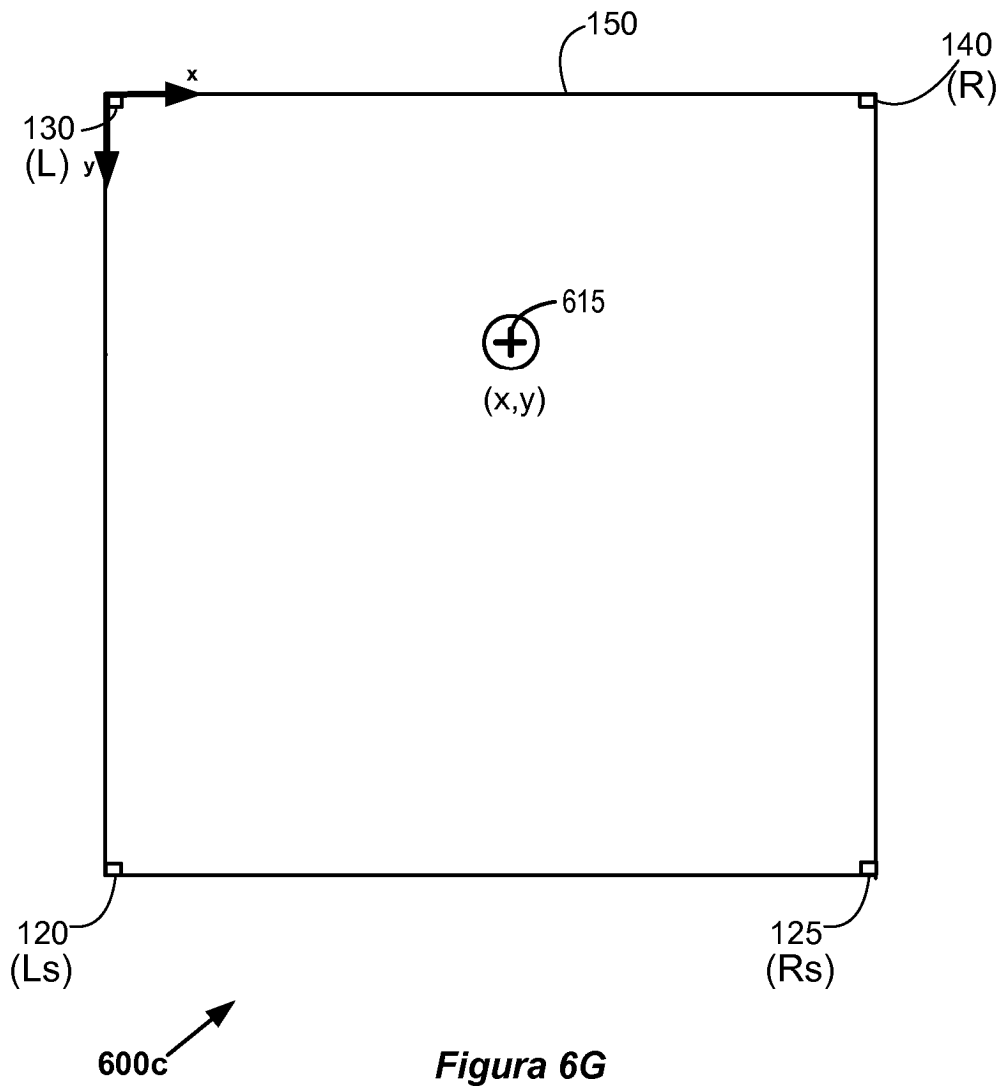
**Figura 6D**



**Figura 6E**



**Figura 6F**



**Figura 6G**

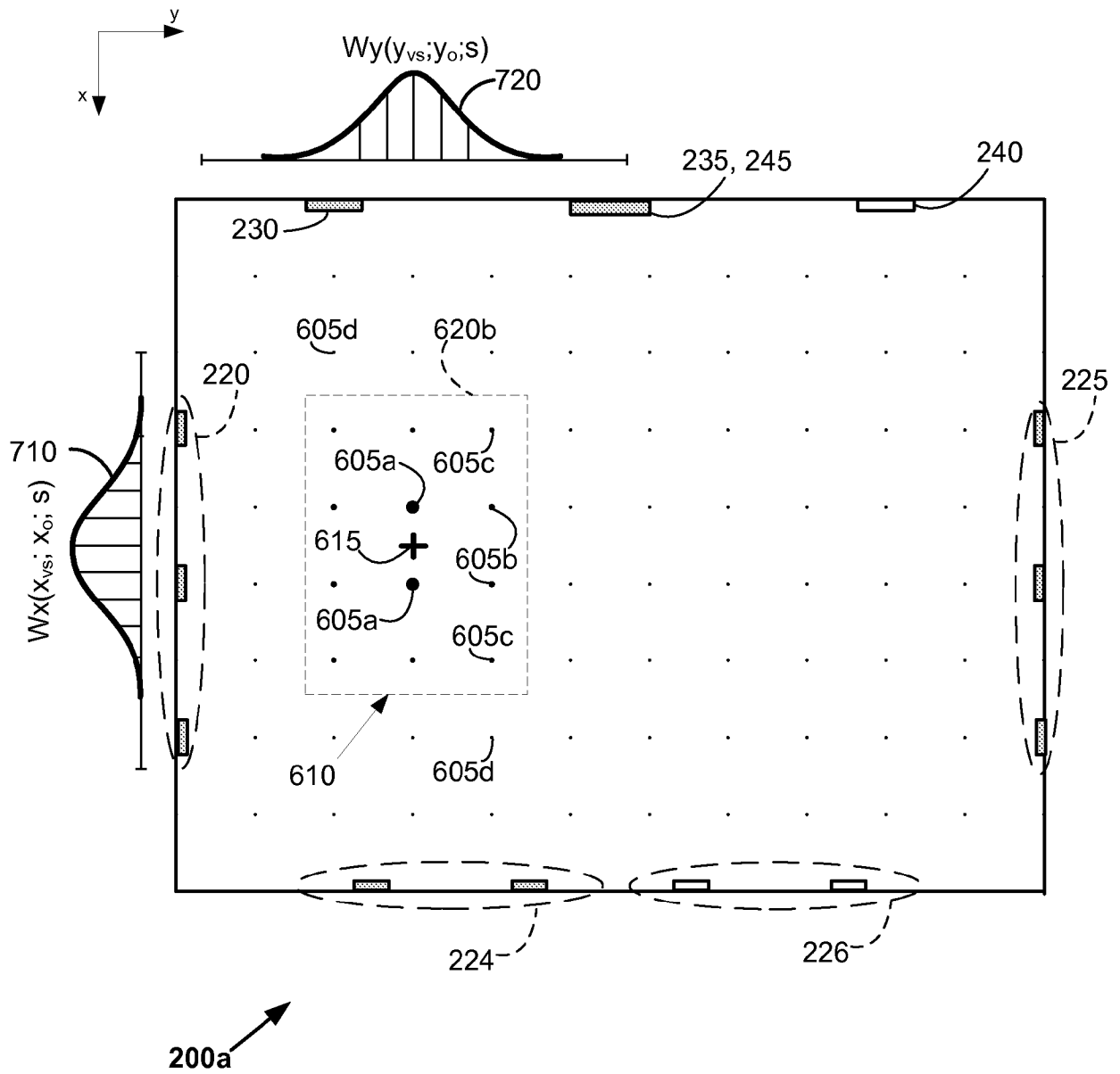
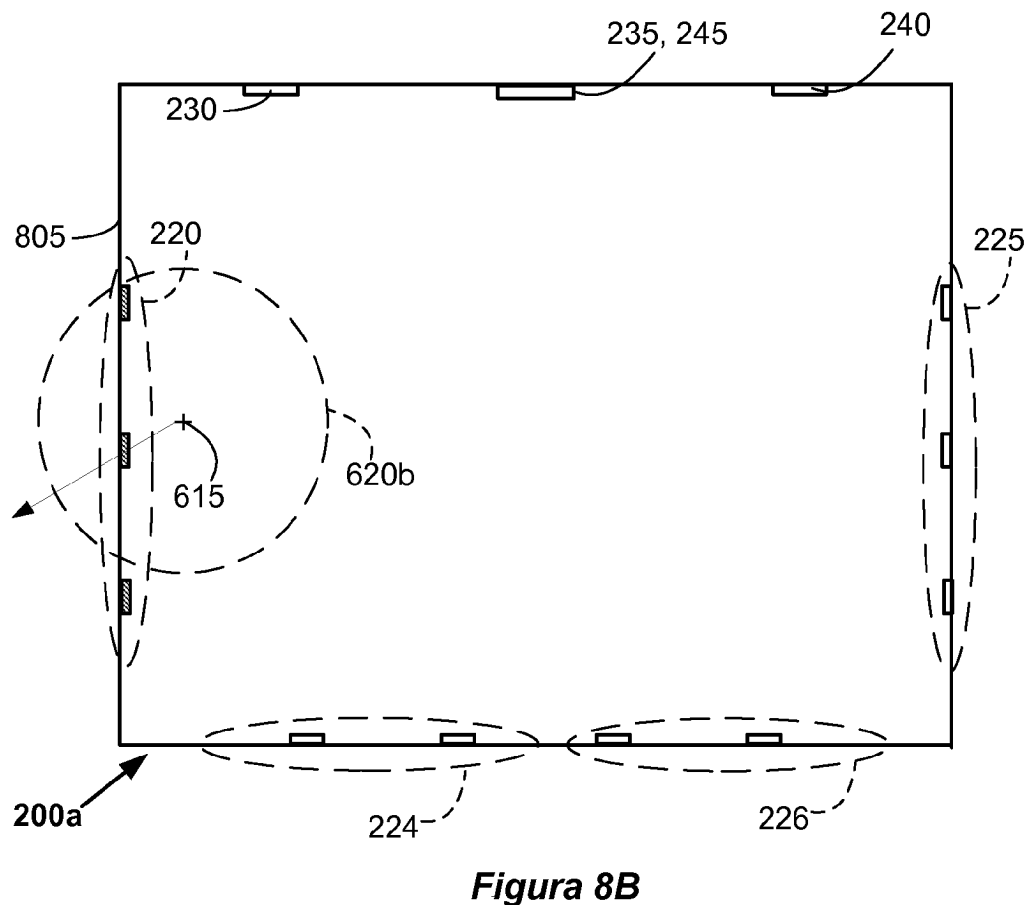
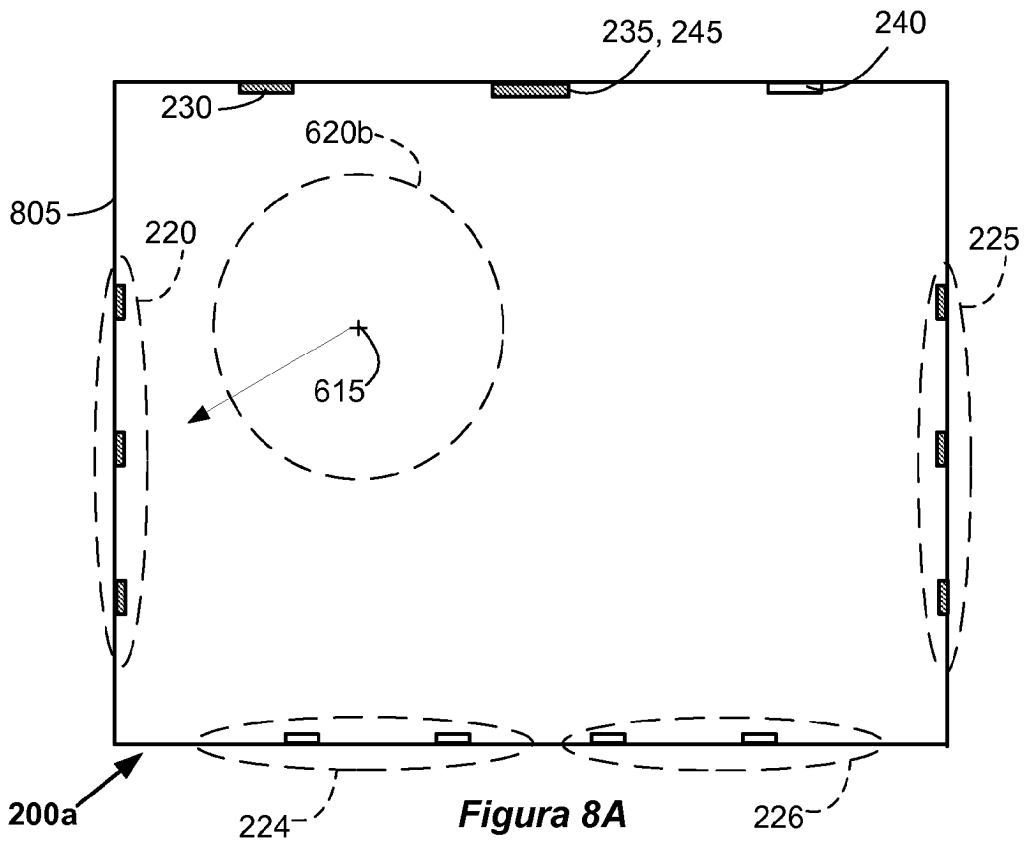
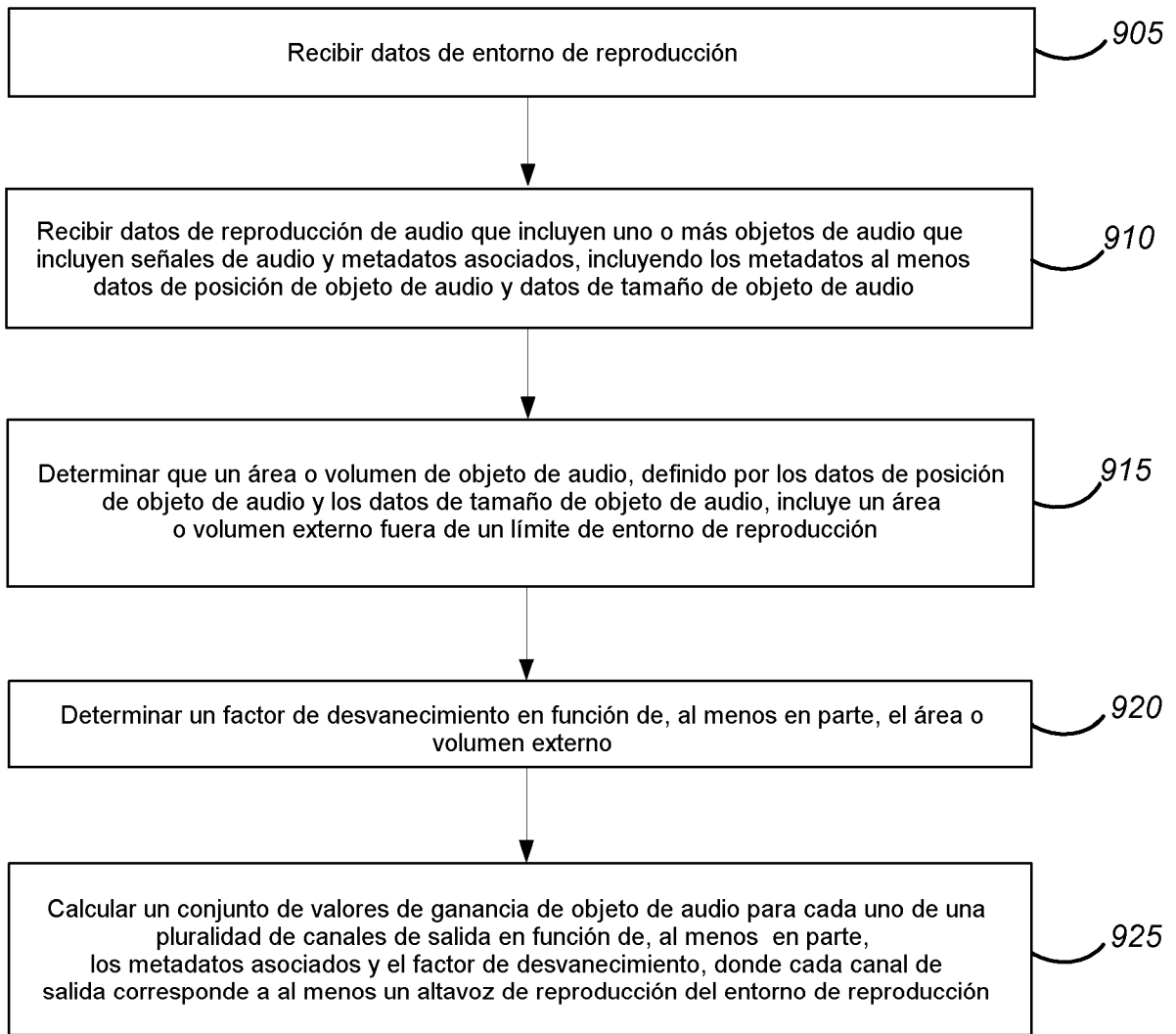


Figura 7

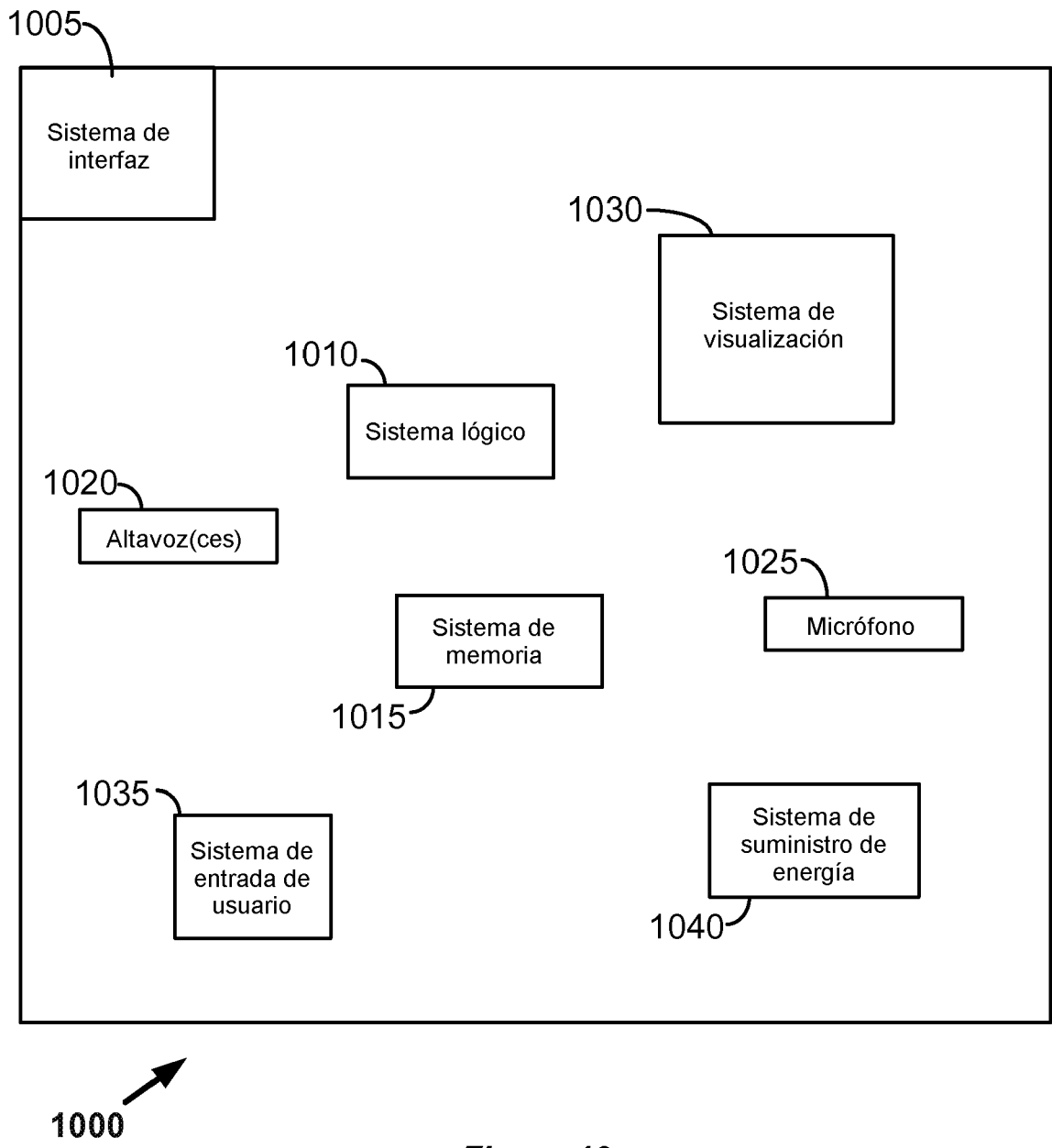




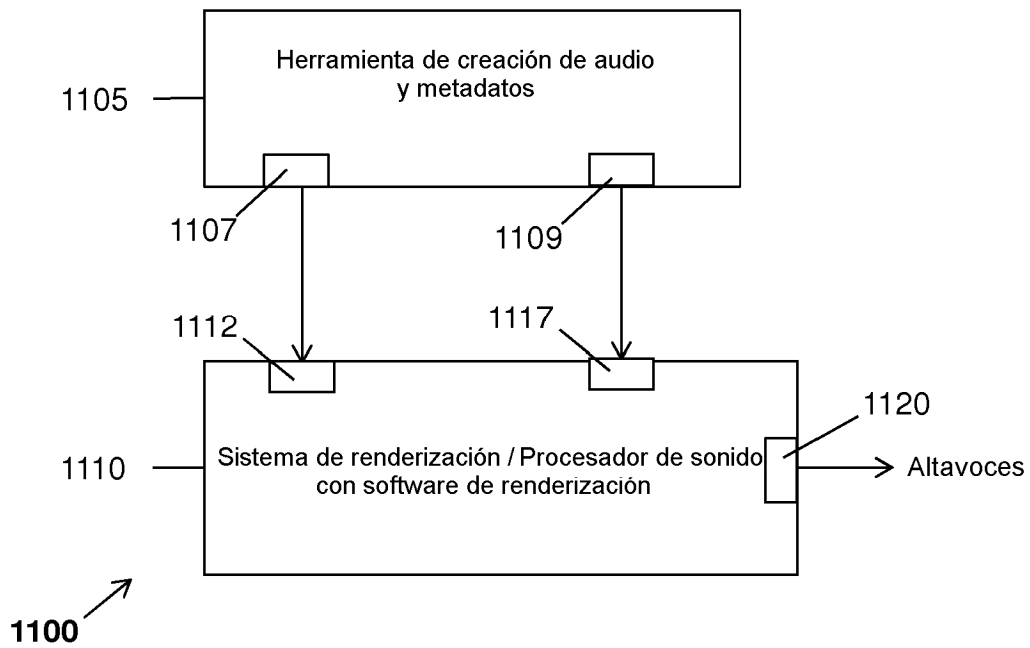


900 ↗

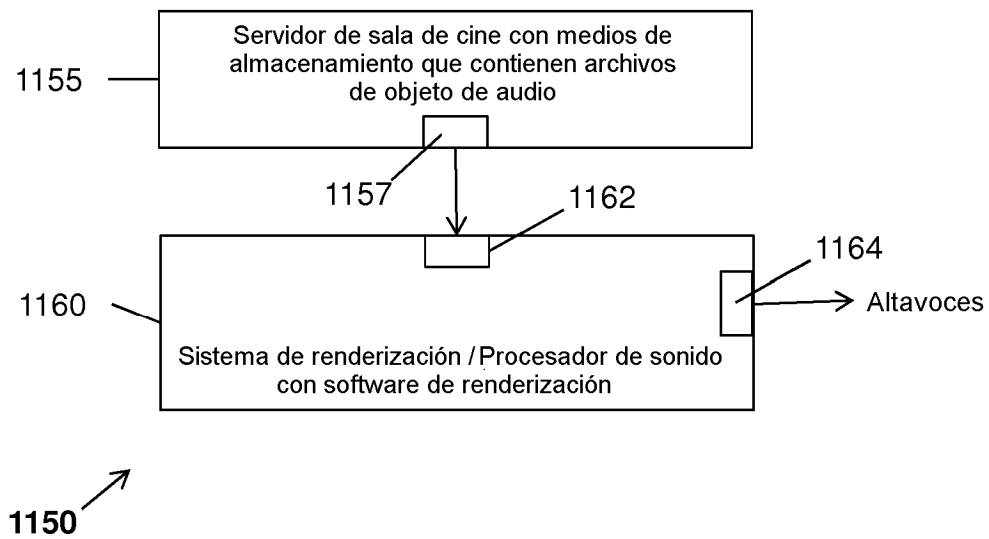
**Figura 9**



**Figura 10**



**Figura 11A**



**Figura 11B**