

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 020 342**

51 Int. Cl.:

H04S 3/00 (2006.01)

H04S 7/00 (2006.01)

H04R 25/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **27.06.2012 E 22196385 (3)**

97 Fecha y número de publicación de la concesión europea: **26.03.2025 EP 4132011**

54 Título: **Aparato para renderizado de objetos de audio según restricciones impuestas de zona de altavoz, método correspondiente y producto de programa informático**

30 Prioridad:

01.07.2011 US 201161504005 P

20.04.2012 US 201261636102 P

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

22.05.2025

73 Titular/es:

**DOLBY LABORATORIES LICENSING
CORPORATION (100.00%)
1275 Market Street
San Francisco, CA 94103, US**

72 Inventor/es:

**TSINGOS, NICOLAS R.;
ROBINSON, CHARLES Q. y
SCHARPF, JURGEN**

74 Agente/Representante:

ELZABURU, S.L.P

ES 3 020 342 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Aparato para renderizado de objetos de audio según restricciones impuestas de zona de altavoz, método correspondiente y producto de programa informático

5

Referencia cruzada a solicitudes relacionadas

Esta solicitud es una solicitud divisional europea de la solicitud de patente europea EP 21179211.4 (referencia: D12070EP02), para la que se presentó el Formulario EPO 1001 el 14 de junio de 2021.

10

Esta solicitud reivindica prioridad de la solicitud provisional de EE. UU. N.º 61/504.005 presentada el 1 de julio de 2011 y la solicitud provisional EE. UU. N.º 61/636.102 presentada el 20 de abril de 2012.

Campo técnico

15

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

20

Antecedentes

Desde la introducción del sonido con la película en 1927, ha habido una evolución constante de la tecnología usada para capturar la intención artística de la banda sonora de la película y reproducirla en un entorno de cine. En la década de 1930, el sonido sincronizado en el disco dio paso al sonido de área variable en la película, que se mejoró aún más en la década de 1940 con consideraciones acústicas teatrales y un diseño de altavoz mejorado, junto con la introducción temprana de la grabación multipista y la reproducción orientable (usando tonos de control para sonidos de movimiento). En las décadas de 1950 y 1960, la banda magnética de la película permitió la reproducción multicanal en el cine, introduciendo canales envolventes y hasta cinco canales de pantalla en los cines premium.

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 rentable de codificación y distribución de mezclas con 3 canales de pantalla y un canal de sonido envolvente mono. La calidad del sonido de cine se mejoró aún más en la década de 1980 con programas de certificación y reducción de ruido Dolby Spectral Recording (SR) como THX. Dolby llevó el sonido digital al cine durante la década de 1990 con un formato de 5.1 canales que proporciona canales discretos de pantalla izquierdo, central y derecho, conjuntos de sonido envolvente izquierdo y derecho y un canal de subwoofer para efectos de baja frecuencia. Dolby Surround 7.1, introducido en 2010, aumentó la cantidad de canales de sonido envolvente al dividir los canales de sonido envolvente izquierdo y derecho existentes en cuatro "zonas".

35

40

A medida que aumenta la cantidad de canales y la disposición de los altavoces cambia de un conjunto bidimensional (2D) plano a un conjunto tridimensional (3D) que incluye la elevación, la tarea de posicionar y reproducir sonidos se vuelve cada vez más difícil. Sería deseable mejorar los métodos de creación y renderizado de audio.

45

El documento US2006109988 divulga un sistema y un método para grabar y reproducir eventos de sonido tridimensionales usando un volumen de sonido macro-micro integrado, discretizado, para reproducir una matriz acústica 3D que reproduce sonido incluyendo propagación natural y reverberación. El sistema y el método incluyen modelado y síntesis de sonido que puede permitir reproducir sonido como matriz volumétrica. La matriz volumétrica puede ser capturada, transferida, reproducida o procesada de otro modo, como espectros espaciales de acontecimientos sonoros reproducido discretamente con macro-micro relaciones controlables.

50

El documento US2006133628 divulga un sistema y un método para formar y renderizar mensajes MIDI 3D. El documento JP 2012049967 divulga un dispositivo de conversión de señales acústicas 1 que comprende medios de determinación de canal de reproducción 42 para determinar un altavoz de reproducción que incluye la dirección del altavoz original en una región de dirección identificada por direcciones de tres altavoces de reproducción; medios de cálculo de coeficientes de ponderación 43 para calcular como coeficiente de ponderación la ratio de distribución de la señal acústica original para cada altavoz de reproducción donde concuerdan la cantidad física acústica en un punto recibido de la señal acústica original y la cantidad física acústica en un punto recibido de una señal acústica de reproducción correspondiente a cada altavoz de reproducción determinado, el cálculo se realiza en función de posiciones de altavoces originales y la posición de cada altavoz de reproducción determinado; y medios de distribución de señales acústicas 50 para distribuir la señal acústica original en función de un coeficiente de ponderación, generando de ese modo señales acústicas de reproducción para el número de canales que coinciden con los altavoces de reproducción.

55

60

65

El documento US US5636283 divulga un sistema para mezclar sonido por cinco canales que rodea un plano de audio.

5 "Report ITU-R BS.2159-3, Multichannel sound technology in home and broadcasting applications, BS Series Broadcasting service (sound)", 26 de enero de 2012, BS.2159-3 divulga un sistema de sonido multicanal 22.2 que tiene nueve canales en la capa superior, diez canales en la capa media, tres canales en la capa inferior y dos canales de efectos de frecuencia baja (LFE). Este sistema es idóneo para pantallas anchas tales como una pantalla FPD de 100 pulgadas (254 cm), porque puede localizar bidimensionalmente una imagen acústica sobre la pantalla entera usando tres canales inferiores, cinco canales medios y tres canales superiores alrededor de la pantalla.

10 El documento WO2011119401 divulga percepción de audio en proximidad local para pistas visuales.

15 El documento JP2011066868 divulga que un divisor de espacio tridimensional 11 tiene como salida información de plano e información de correlación de canales, en función de posiciones de una pluralidad de altavoces dispuestos estereoscópicamente en el espacio tridimensional para tener como salida señales de audio de una pluralidad de canales, y en función de una dirección de división para dividir el espacio tridimensional en una pluralidad de planos. Los codificadores de plano 12 a 14 generan elementos de codificación como resultado de codificar como grupo de programas para cada plano bidimensional en función de la información de plano y la información de correlación de canal, y además genera y tiene como salida información posicional de plano. Una sección integradora de flujo 15 integra todos los elementos de codificación y la información posicional de plano para generar y tener como salida un flujo de codificación.

20 El documento US2011144783 divulga un aparato para controlar un renderizador de síntesis de campo de onda con objetos sonoros incluye un proveedor para proporcionar una descripción de escena.

25 **Compendio**

30 Algunos aspectos de la materia de asunto descrita en esta divulgación pueden implementarse en herramientas para crear y renderizar datos de reproducción de audio. Algunas de estas herramientas de creación permiten generalizar los datos de reproducción de audio para una amplia variedad de entornos de reproducción. Según algunas implementaciones de este tipo, los datos de reproducción de audio pueden crearse mediante la creación de metadatos para objetos sonoros. Los metadatos pueden crearse con referencia a las zonas de altavoz. Durante el proceso de renderizado, los datos de reproducción de audio pueden reproducirse según la disposición de los altavoces de reproducción de un entorno de reproducción particular.

35 Según la invención, se proporciona un aparato según la reivindicación 1, un método según la reivindicación 7 y un medio no transitorio que tiene software almacenado en el mismo según la reivindicación 8.

40 Los detalles de una o más implementaciones de la materia de asunto descrita en esta memoria descriptiva se exponen en los dibujos adjuntos y en la descripción a continuación. Otras características, aspectos y ventajas serán evidentes a partir de la descripción, los dibujos y las reivindicaciones. Cabe señalar que las dimensiones relativas de las siguientes figuras pueden no estar dibujadas a escala.

Breve descripción de los dibujos

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

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

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

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

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

60 Las Figuras 5A-5C muestran ejemplos de respuestas del altavoz correspondientes a un objeto sonoro que tiene una posición restringida a una superficie bidimensional de un espacio tridimensional.

Las Figuras 5D y 5E muestran ejemplos de superficies bidimensionales a las que se puede restringir un objeto sonoro.

65 La Figura 6A es un diagrama de flujo que esboza un ejemplo de un proceso de restricción de posiciones de un objeto sonoro a una superficie bidimensional.

- La Figura 6B es un diagrama de flujo que esboza un ejemplo de un proceso de correlación de una posición de objeto sonoro a una única ubicación de altavoz o una única zona de altavoz.
- 5 La Figura 7 es un diagrama de flujo que esboza un proceso de establecimiento y uso de altavoces virtuales.
- Las Figuras 8A-8C muestran ejemplos de altavoces virtuales correlacionados a puntos finales de línea y respuestas de los altavoces correspondientes.
- 10 Las Figuras 9A-9C muestran ejemplos del uso de un amarre virtual para mover un objeto sonoro.
- La Figura 10A es un diagrama de flujo que esboza un proceso de uso de un amarre virtual para mover un objeto sonoro.
- 15 La Figura 10B es un diagrama de flujo que esboza un proceso alternativo de uso de un amarre virtual para mover un objeto sonoro.
- Las Figuras 10C-10E muestran ejemplos del proceso descrito en la Figura 10B.
- 20 La Figura 11 muestra un ejemplo de aplicación de la restricción de zona de altavoz en un entorno de reproducción virtual.
- La Figura 12 es un diagrama de flujo que esboza algunos ejemplos de la aplicación de reglas de restricción de zona de altavoz.
- 25 Las Figuras 13A y 13B muestran un ejemplo de una GUI que puede cambiar entre una vista bidimensional y una vista tridimensional de un entorno de reproducción virtual.
- Las Figuras 13C-13E muestran combinaciones de representaciones bidimensionales y tridimensionales de entornos de reproducción.
- 30 La Figura 14A es un diagrama de flujo que esboza un proceso de control de un aparato para presentar las GUI como las que se muestran en las Figuras 13C-13E.
- 35 La Figura 14B es un diagrama de flujo que esboza un proceso de renderizado de objetos sonoros para un entorno de reproducción.
- La Figura 15A muestra un ejemplo de un objeto sonoro y la anchura de objeto sonoro asociado en un entorno de reproducción virtual.
- 40 La Figura 15B muestra un ejemplo de un perfil de dispersión correspondiente a la anchura de objeto sonoro que se muestra en la Figura 15A.
- La Figura 16 es un diagrama de flujo que esboza un proceso de manejo en blobs (objeto binario grande) de objetos sonoros.
- 45 Las Figuras 17A y 17B muestran ejemplos de un objeto sonoro posicionado en un entorno de reproducción virtual tridimensional.
- 50 La Figura 18 muestra ejemplos de zonas que se corresponden con los modos de panoramización.
- Las Figuras 19A-19D muestran ejemplos de aplicación de técnicas de panoramización de campo cercano y campo lejano a objetos sonoros en diferentes ubicaciones.
- 55 La Figura 20 indica las zonas de altavoz de un entorno de reproducción que pueden usarse en un proceso de control de polarización de pantalla a sala.
- La Figura 21 es un diagrama de bloques que proporciona ejemplos de componentes de un aparato de creación y/o renderizado.
- 60 La Figura 22A es un diagrama de bloques que representa algunos componentes que pueden usarse para la creación de contenido de audio.
- La Figura 22B es un diagrama de bloques que representa algunos componentes que pueden usarse para la reproducción de audio en un entorno de reproducción.
- 65

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

Descripción de realizaciones de ejemplo

5 La siguiente descripción se dirige a ciertas implementaciones con el fin de describir algunos aspectos innovadores de esta divulgación, así como ejemplos de contextos en los que estos aspectos innovadores pueden implementarse. Sin embargo, las enseñanzas en esta memoria se pueden aplicar de diversas maneras diferentes.

10 Por ejemplo, aunque se han descrito diversas implementaciones en términos de entornos de reproducción particulares, las enseñanzas en esta memoria son ampliamente aplicables a otros entornos de reproducción conocidos, así como entornos de reproducción que puedan introducirse en el futuro. De manera similar, mientras que en esta memoria se presentan ejemplos de interfaces gráficas de usuario (GUI), algunas de las cuales proporcionan ejemplos de ubicaciones de altavoces, zonas de altavoz, etc., los inventores contemplan otras implementaciones. Además, las implementaciones descritas pueden implementarse en diversas herramientas de creación y/o renderizado, que pueden implementarse en una variedad de hardware, software, firmware, etc. En consecuencia, no se pretende que las enseñanzas de esta divulgación se limiten a las implementaciones mostradas en las figuras y/o descritas en esta memoria, sino que tienen una amplia aplicabilidad. La siguiente descripción es útil para ilustrar la invención que se define en las reivindicaciones adjuntas.

20 La Figura 1 muestra un ejemplo de un entorno de reproducción que tiene una configuración Dolby Surround 5.1. Dolby Surround 5.1 se desarrolló en la década de 1990, pero esta configuración aún se despliega ampliamente en entornos de sistemas de sonido de cine. Un proyector 105 puede configurarse para proyectar imágenes de vídeo, por ejemplo, para una película, en la pantalla 150.

25 Los datos de reproducción de audio pueden sincronizarse con las imágenes de vídeo y procesarse mediante el procesador 110 de sonido. Los amplificadores de potencia 115 pueden proporcionar señales de alimentación de altavoz a los altavoces del entorno de reproducción 100.

30 La configuración de Dolby Surround 5.1 incluye el conjunto de sonido envolvente izquierdo 120, el conjunto de sonido envolvente derecho 125, cada uno de los cuales se controla en grupo por un solo canal. La configuración de Dolby Surround 5.1 también incluye canales separados para el canal de pantalla izquierdo 130, el canal de pantalla central 135 y el canal de pantalla derecho 140. Se proporciona un canal separado para el subwoofer 145 para efectos de baja frecuencia (LFE).

35 En 2010, Dolby proporcionó mejoras al sonido del cine digital al presentar 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 vídeo digital y proyectar imágenes de vídeo en la pantalla 150. Los datos de reproducción de audio pueden ser procesados 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.

40 La configuración de Dolby Surround 7.1 incluye el conjunto de sonido envolvente de lado izquierdo 220 y el conjunto de sonido envolvente de lado derecho 225, cada uno de los cuales puede ser accionado por un solo canal. Al igual que Dolby Surround 5.1, la configuración de Dolby Surround 7.1 incluye canales separados para el canal de pantalla izquierdo 230, el canal de pantalla central 235, el canal de pantalla derecho 240 y el subwoofer 245. Sin embargo, Dolby Surround 7.1 aumenta el número de canales de sonido envolvente al dividir los canales de sonido envolvente izquierdo y derecho de Dolby Surround 5.1 en cuatro zonas: además del conjunto de sonido envolvente de lado izquierdo 220 y el conjunto de sonido envolvente de lado derecho 225, se incluyen canales separados para los altavoces de sonido envolvente traseros izquierdos 224 y los altavoces de sonido envolvente traseros derechos 226. Aumentar el número de zonas de sonido envolvente dentro del entorno 200 de reproducción puede mejorar significativamente la localización del sonido.

45 En un esfuerzo por crear un entorno más inmersivo, algunos entornos de reproducción pueden configurarse con una mayor cantidad de altavoces, accionados por una mayor cantidad de canales. Además, algunos entornos de reproducción pueden incluir altavoces desplegados a diversas alturas, algunos de los cuales pueden estar por encima de un área de asientos del entorno de reproducción.

50 La Figura 3 muestra un ejemplo de un entorno de reproducción que tiene una configuración de sonido envolvente Hamasaki 22.2. Hamasaki 22.2 se desarrolló en los laboratorios de investigación científica y tecnológica de NHK en Japón como el componente de sonido envolvente de la televisión de definición ultraalta. Hamasaki 22.2 proporciona 24 canales de altavoces, que se pueden usar para accionar los altavoces dispuestos en tres capas. La capa superior de altavoces 310 del entorno de reproducción 300 puede ser accionada por 9 canales. La capa intermedia de altavoces 320 puede ser accionada por 10 canales. La capa inferior de altavoces 330 puede ser accionada por 5 canales, dos de los cuales son para los subwoofers 345a y 345b.

65 En consecuencia, la tendencia moderna es incluir no solo más altavoces y más canales, sino también incluir

altavoces a diferentes alturas. A medida que aumenta la cantidad de canales y la disposición de los altavoces cambia de un conjunto 2D a un conjunto 3D, las tareas de posicionamiento y renderizado de sonidos se vuelven cada vez más difíciles.

5 Esta divulgación proporciona diversas herramientas, así como interfaces de usuario relacionadas, que aumentan la funcionalidad y/o reducen la complejidad de creación para un sistema de sonido de audio 3D.

10 La Figura 4A muestra un ejemplo de una interfaz gráfica de usuario (GUI) que representa las zonas de altavoz a diferentes alturas en un entorno de reproducción virtual. La GUI 400 puede, por ejemplo, mostrarse en un dispositivo de visualización según las instrucciones de un sistema lógico, según las señales recibidas de los dispositivos de entrada del usuario, etc. Algunos de estos dispositivos se describen a continuación con referencia a la Figura 21.

15 Como se usa en esta memoria con referencia a entornos de reproducción virtual como el entorno de reproducción virtual 404, el término "zona de altavoz" generalmente se refiere a una construcción lógica que puede tener o no una correspondencia 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 altavoces de reproducción particular de un entorno de reproducción de cine. En cambio, el término "ubicación de zona de altavoz" puede referirse en general a una zona de un entorno de reproducción virtual. En algunas implementaciones, una zona de altavoz de un entorno de reproducción virtual puede corresponder a un altavoz virtual, por ejemplo, mediante el uso de tecnología de virtualización como Dolby Headphone™ (a veces denominada Mobile Surround™), que crea un entorno de sonido envolvente virtual en tiempo real usando un conjunto de auriculares estéreo de dos canales.

25 En la GUI 400, hay siete zonas de altavoz 402a en una primera elevación y dos zonas de altavoz 402b en una segunda elevación, haciendo un total de nueve zonas de altavoz en el entorno de reproducción virtual 404. En este ejemplo, las zonas de altavoz 1-3 están en el área frontal 405 del entorno de reproducción virtual 404. El área frontal 405 puede corresponder, por ejemplo, a una zona de un entorno de reproducción de cine en el que se ubica una pantalla 150, a una zona de una vivienda en la que se sitúa una pantalla de televisión, etc.

30 Aquí, la zona de altavoz 4 corresponde generalmente a los altavoces del área izquierda 410 y la zona de altavoz 5 corresponde a los altavoces del área derecha 415 del entorno de reproducción virtual 404. La zona de altavoz 6 corresponde a 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 en un área superior 420a y la zona de altavoz 9 corresponde a altavoces en un área superior 420b, que puede ser un área de techo virtual como un área del techo virtual 520 que se muestra en las Figuras 5D y 5E. En consecuencia, y como se describe con más detalle a continuación, las ubicaciones de las zonas de altavoz 1-9 que se muestran en la Figura 4A pueden corresponder o no a las ubicaciones de los altavoces de reproducción de un entorno de reproducción real. Además, otras implementaciones pueden incluir más o menos zonas de altavoz y/o elevaciones.

40 En diversas implementaciones descritas en esta memoria, una interfaz de usuario como la GUI 400 puede usarse como parte de una herramienta de creación y/o una herramienta de renderizado. En algunas implementaciones, la herramienta de creación y/o la herramienta de renderizado pueden implementarse mediante software almacenado en uno o más medios no transitorios. La herramienta de creación y/o la herramienta de renderizado puede implementarse (al menos en parte) mediante hardware, firmware, etc., como el sistema lógico y otros dispositivos que se describen a continuación con referencia a la Figura 21. En algunas implementaciones de creación, se puede usar una herramienta de creación asociada para crear metadatos para datos de audio asociados. Los metadatos pueden, por ejemplo, incluir datos que indiquen la posición y/o trayectoria de un objeto sonoro 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 disposición de los altavoces particular de un entorno de reproducción real. Una herramienta de renderizado puede recibir datos de audio y metadatos asociados, y puede calcular las ganancias de audio y las 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 de amplitud, que puede crear la percepción de que un sonido proviene de una posición P en el entorno de reproducción. Por ejemplo, las señales de alimentación de altavoz pueden proporcionarse a los altavoces de reproducción 1 a N del entorno de reproducción según la siguiente ecuación:

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

60 En la ecuación 1, $x_i(t)$ representa la señal de alimentación de altavoz que se aplicará 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 métodos de panoramización de amplitud descritos en la sección 2, páginas 3-4 del documento de V. Pulkki, *Compensating Displacement of Amplitude-Panned Virtual Sources* (Audio Engineering Society (AES) Conferencia internacional sobre audio virtual, sintético y de ocio).

En algunas implementaciones, las ganancias pueden depender de la frecuencia.

En algunas implementaciones, se puede introducir un retraso de tiempo al sustituir $x(t)$ por $x(t-\Delta t)$.

5 En algunas implementaciones de renderizado, los datos de reproducción de audio creados con referencia a las zonas de altavoz 402 son correlacionados a ubicaciones de altavoces de una amplia gama de entornos de reproducción, que pueden estar en una configuración Dolby Surround 5.1, una configuración Dolby Surround 10 7.1, una configuración Hamasaki 22.2, u otra configuración. Por ejemplo, con referencia a la Figura 2, una herramienta de renderizado puede correlacionar datos de reproducción de audio para las zonas de altavoz 4 y 5 al conjunto de sonido envolvente de lado izquierdo 220 y al conjunto de sonido envolvente de lado derecho 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 al canal de pantalla izquierdo 230, al canal de pantalla derecho 240 y al canal de pantalla central 235, respectivamente. Los datos de reproducción de 15 audio para las zonas de altavoz 6 y 7 pueden correlacionarse a los altavoces de sonido envolvente traseros izquierdos 224 y a los altavoces de sonido envolvente traseros derechos 226.

La Figura 4B muestra un ejemplo de otro entorno de reproducción. En algunas implementaciones, una herramienta de renderizado puede correlacionar datos de reproducción de audio para las zonas de altavoz 1, 20 2 y 3 a los altavoces de pantalla 455 correspondientes del entorno de reproducción 450. Una herramienta de renderizado puede correlacionar datos de reproducción de audio para las zonas de altavoz 4 y 5 al conjunto de sonido envolvente de lado izquierdo 460 y al conjunto de sonido envolvente de lado derecho 465 y puede correlacionar datos de reproducción de audio para las zonas de altavoz 8 y 9 a los altavoces superiores izquierdos 470a y los altavoces superiores derechos 470b. Los datos de reproducción de audio para las zonas 25 de altavoz 6 y 7 pueden correlacionarse a los altavoces de sonido envolvente traseros izquierdos 480a y a los altavoces de sonido envolvente traseros derechos 480b.

En algunas implementaciones de creación, se puede usar una herramienta de creación para crear metadatos para objetos sonoros. Como se usa en esta memoria, el término "objeto sonoro" puede referirse a un flujo de 30 datos de audio y metadatos asociados. Los metadatos indican típicamente la posición 3D del objeto, las restricciones de renderizado y el tipo de contenido (por ejemplo, diálogo, efectos, etc.). Dependiendo de la implementación, los metadatos pueden incluir otros tipos de datos, como datos de anchura, datos de ganancia, datos de trayectoria, etc. Algunos objetos sonoros pueden estar estáticos, mientras que otros pueden moverse. Los detalles de objeto sonoro se pueden crear o renderizar según los metadatos asociados que, entre otras 35 cosas, pueden indicar la posición del objeto sonoro en un espacio tridimensional en un momento determinado. Cuando los objetos sonoros se monitorizan o reproducen en un entorno de reproducción, los objetos sonoros pueden renderizarse según los metadatos posicionales usando los altavoces de reproducción que están presentes en el entorno de reproducción, en lugar de enviarse a un canal físico predeterminado, como es el caso con sistemas tradicionales basados en canales como Dolby 5.1 y Dolby 7.1.

40 En esta memoria, se describen diversas herramientas de creación y renderizado con referencia a una GUI que es sustancialmente la misma que la GUI 400. Sin embargo, otras diversas interfaces de usuario, incluidas, entre otras, las GUI, se pueden usar en asociación con estas herramientas de creación y renderizado. Algunas de estas herramientas pueden simplificar el proceso de creación mediante la aplicación de diversos tipos de 45 restricciones. Algunas implementaciones se describirán ahora con referencia a las Figuras 5A y siguientes.

Las Figuras 5A-5C muestran ejemplos de respuestas de los altavoces correspondientes a un objeto sonoro que tiene una posición restringida a una superficie bidimensional de un espacio tridimensional, que es un hemisferio en este ejemplo. En estos ejemplos, las respuestas de los altavoces han sido calculadas por un 50 renderizador asumiendo una configuración de 9 altavoces, con cada altavoz correspondiente a una de las zonas 1-9 de altavoz. Sin embargo, como se indica en otra parte en esta memoria, generalmente no puede haber una correlación uno a uno entre las zonas de altavoz de un entorno de reproducción virtual y los altavoces de reproducción en un entorno de reproducción. Con referencia primero a la Figura 5A, el objeto sonoro 505 se muestra en una ubicación en la parte frontal izquierda del entorno de reproducción virtual 404. En consecuencia, el altavoz correspondiente a la zona de altavoz 1 indica una ganancia sustancial y los altavoces 55 correspondientes a las zonas de altavoz 3 y 4 indican ganancias moderadas.

En este ejemplo, la ubicación del objeto sonoro 505 se puede cambiar colocando un cursor 510 en el objeto sonoro 505 y "arrastrando" el objeto sonoro 505 a una ubicación deseada en el plano x, y del entorno de 60 reproducción virtual 404. A medida que el objeto es arrastrado hacia el centro del entorno de reproducción, también se correlaciona en la superficie de un hemisferio y su elevación aumenta. Aquí, los aumentos en la elevación del objeto sonoro 505 se indican mediante un aumento en el diámetro del círculo que representa el objeto sonoro 505: como se muestra en las Figuras 5B y 5C, a medida que el objeto sonoro 505 se arrastra al centro superior del entorno de reproducción virtual 404, el objeto sonoro 505 parece cada vez más grande. 65 Alternativa o adicionalmente, la elevación del objeto sonoro 505 puede indicarse mediante cambios de color, brillo, una indicación de elevación numérica, etc. Cuando el objeto sonoro 505 se posiciona en el centro superior

del entorno de reproducción virtual 404, como se muestra en la Figura 5C, los altavoces correspondientes a las zonas de altavoz 8 y 9 indican ganancias sustanciales y los otros altavoces indican poca o ninguna ganancia.

5 En esta implementación, la posición del objeto sonoro 505 se restringe a una superficie bidimensional, como una superficie esférica, una superficie elíptica, una superficie cónica, una superficie cilíndrica, una cuña, etc. Las Figuras 5D y 5E muestran ejemplos de superficies bidimensionales a las que se puede restringir un objeto sonoro. Las Figuras 5D y 5E son vistas en corte transversal a través del entorno de reproducción virtual 404, con el área frontal 405 mostrada a la izquierda. En las Figuras 5D y 5E, los valores y del eje y-z aumentan en la dirección del área frontal 405 del entorno de reproducción virtual 404, para mantener la coherencia con las orientaciones de los ejes x-y que se muestran en las Figuras 5A-5C.

15 En el ejemplo que se muestra en la Figura 5D, la superficie bidimensional 515a es una sección de un elipsoide. En el ejemplo que se muestra en la Figura 5E, la superficie bidimensional 515b es una sección de una cuña. Sin embargo, las formas, orientaciones y posiciones de las superficies bidimensionales 515 mostradas en las Figuras 5D y 5E son meros ejemplos. En implementaciones alternativas, al menos una parte de la superficie bidimensional 515 puede extenderse fuera del entorno de reproducción virtual 404. En algunas de tales implementaciones, la superficie bidimensional 515 puede extenderse por encima del techo virtual 520. En consecuencia, el espacio tridimensional dentro del que se extiende la superficie bidimensional 515 no tiene necesariamente la misma extensión que el volumen del entorno de reproducción virtual 404. En otras implementaciones, un objeto sonoro puede restringirse a características unidimensionales como curvas, líneas rectas, etc.

25 La Figura 6A es un diagrama de flujo que esboza un ejemplo de un proceso de restricción de posiciones de un objeto sonoro a una superficie bidimensional. Al igual que con otros diagramas de flujo que se proporcionan en esta memoria, las operaciones del proceso 600 no se realizan necesariamente en el orden mostrado. Además, el proceso 600 (y otros procesos proporcionados en esta memoria) pueden incluir más o menos operaciones que las que se indican en los dibujos y/o se describen. En este ejemplo, los bloques 605 a 622 los realiza una herramienta de creación y los bloques 624 a 630 los realiza una herramienta de renderizado. La herramienta de creación y la herramienta de renderizado pueden implementarse en un solo aparato o en más de un aparato. Aun que la Figura 6A (y otros diagramas de flujo proporcionados en esta memoria) pueden crear la impresión de que los procesos de creación y renderizado se realizan de manera secuencial, en muchas implementaciones los procesos de creación y renderizado se realizan sustancialmente al mismo tiempo. Los procesos de creación y los procesos de renderizado pueden ser interactivos. Por ejemplo, los resultados de una operación de creación pueden enviarse a la herramienta de renderizado, los resultados correspondientes de la herramienta de renderizado pueden ser evaluados por un usuario, que puede realizar una creación adicional basada en estos resultados, etc.

40 En el bloque 605, se recibe una indicación de que la posición de un objeto sonoro debe restringirse a una superficie bidimensional. La indicación puede, por ejemplo, ser recibida por un sistema lógico de un aparato que se configura para proporcionar herramientas de creación y/o renderizado. Al igual que con otras implementaciones descritas en esta memoria, el sistema lógico puede estar funcionando según las instrucciones del software almacenadas en un medio no transitorio, según el firmware, etc. La indicación puede ser una señal de un dispositivo de entrada del usuario (como una pantalla táctil, un ratón, una bola de seguimiento, un dispositivo de reconocimiento de gestos, etc.) en respuesta a la entrada de un usuario.

45 En el bloque opcional 607, se reciben datos de audio. El bloque 607 es opcional en este ejemplo, ya que los datos de audio también pueden ir directamente a un renderizador desde otra fuente (por ejemplo, una mesa de mezclas) que está sincronizada en el tiempo con la herramienta de creación de metadatos. En algunas de estas implementaciones, puede existir un mecanismo implícito para vincular cada flujo de audio a un flujo de metadatos entrante correspondiente para formar un objeto sonoro. Por ejemplo, el flujo de metadatos puede contener un identificador para el objeto sonoro que representa, por ejemplo, un valor numérico de 1 a N. Si el aparato de reproducción se configura con entradas de audio que también están numeradas de 1 a N, la herramienta de renderizado puede automáticamente suponer que un objeto sonoro se forma por el flujo de metadatos identificado con un valor numérico (por ejemplo, 1) y los datos de audio recibidos en la primera entrada de audio. De manera similar, cualquier flujo de metadatos identificado con el número 2 puede formar un objeto con el audio recibido en el segundo canal de entrada de audio. En algunas implementaciones, el audio y los metadatos pueden ser empaquetados previamente por la herramienta de creación para formar objetos sonoros y los objetos sonoros pueden proporcionarse a la herramienta de renderizado, por ejemplo, enviados a través de una red como paquetes TCP/IP.

60 En implementaciones alternativas, la herramienta de creación puede enviar solo los metadatos en la red y la herramienta de renderizado puede recibir audio de otra fuente (por ejemplo, a través de un flujo de modulación de código de pulso (PCM), a través de audio analógico, etc.). En dichas implementaciones, la herramienta de renderizado puede configurarse para agrupar los datos y metadatos de audio para formar los objetos sonoros. Los datos de audio pueden, por ejemplo, ser recibidos por el sistema lógico a través de una interfaz. La interfaz puede ser, por ejemplo, una interfaz de red, una interfaz de audio (por ejemplo, una interfaz configurada para la comunicación a través del estándar AES3 desarrollado por Audio Engineering Society y la Unión Europea

de Radiodifusión, también conocida como AES/EBU, a través del protocolo de interfaz digital de audio multicanal (MADI), a través de señales analógicas, etc.) o una interfaz entre el sistema lógico y un dispositivo de memoria. En este ejemplo, los datos recibidos por el renderizador incluyen al menos un objeto sonoro.

5 En el bloque 610, se reciben las coordenadas (x, y) o (x, y, z) de una posición de objeto sonoro. El bloque 610 puede, por ejemplo, implicar recibir una posición inicial del objeto sonoro. El bloque 610 también puede implicar recibir una indicación de que un usuario ha posicionado o reposicionado el objeto sonoro, por ejemplo, como se describe anteriormente con referencia a las Figuras 5A-5C. Las coordenadas del objeto sonoro se correlacionan a una superficie bidimensional en el bloque 615. La superficie bidimensional puede ser similar a
10 una de las descritas anteriormente con referencia a las Figuras 5D y 5E, o puede ser una superficie bidimensional diferente. En este ejemplo, cada punto del plano x-y se correlacionará a un solo valor z, por lo que el bloque 615 implica correlacionar las coordenadas x e y recibidas en el bloque 610 a un valor de z. En otras implementaciones, se pueden usar diferentes procesos de correlación y/o sistemas de coordenadas. El objeto sonoro puede mostrarse (bloque 620) en la ubicación (x, y, z) que se determina en el bloque 615. Los datos y metadatos de audio, incluida la ubicación correlacionada (x, y, z) que se determina en el bloque 615, pueden almacenarse en el bloque 621. Los datos de audio y los metadatos pueden enviarse a una herramienta de renderizado (bloque 622). En algunas implementaciones, los metadatos pueden enviarse continuamente mientras se realizan algunas operaciones de creación, por ejemplo, mientras el objeto sonoro se posiciona, restringe, visualiza en la GUI 400, etc.

20 En el bloque 623, se determina si continuará el proceso de creación. Por ejemplo, el proceso de creación puede finalizar (bloque 625) al recibir una entrada desde una interfaz de usuario que indica que un usuario ya no desea restringir las posiciones de los objetos sonoros a una superficie bidimensional. De lo contrario, el proceso de creación puede continuar, por ejemplo, volviendo al bloque 607 o al bloque 610. En algunas implementaciones, las operaciones de renderizado pueden continuar tanto si continúa el proceso de creación como si no. En algunas implementaciones, los objetos sonoros pueden grabarse en un disco en la plataforma de creación y luego reproducirse desde un procesador de sonido dedicado o un servidor de cine conectado a un procesador de sonido, por ejemplo, un procesador de sonido similar al procesador de sonido 210 de la Figura 2, con fines de exhibición.

30 En algunas implementaciones, la herramienta de renderizado puede ser un software que se ejecuta en un aparato que se configura para proporcionar la funcionalidad de creación. En otras implementaciones, la herramienta de renderizado se puede proporcionar en otro dispositivo. El tipo de protocolo de comunicación usado para la comunicación entre la herramienta de creación y la herramienta de renderizado puede variar según si ambas herramientas se ejecutan en el mismo dispositivo o si se comunican a través de una red.

35 En el bloque 626, los datos y metadatos de audio (incluidas las posiciones (x, y, z) determinadas en el bloque 615) son recibidos por la herramienta de renderizado. En implementaciones alternativas, los datos y metadatos de audio pueden recibirse por separado e interpretarse mediante la herramienta de renderizado como un objeto sonoro a través de un mecanismo implícito. Como se ha indicado anteriormente, por ejemplo, un flujo de metadatos puede contener un código de identificación de objeto sonoro (por ejemplo, 1, 2, 3, etc.) y puede adjuntarse respectivamente con las entradas de audio primera, segunda y tercera (es decir, conexión de audio digital o analógica) en el sistema de renderizado para formar un objeto sonoro que se puede renderizar en los altavoces. Durante las operaciones de renderizado del proceso 600 (y otras operaciones de renderizado descritas en esta memoria), las ecuaciones de ganancia de panoramización pueden aplicarse según la disposición de los altavoces de reproducción de un entorno de reproducción particular. En consecuencia, el sistema lógico de la herramienta de renderizado puede recibir datos de entorno de reproducción que comprenden una indicación de un número de altavoces de reproducción en el entorno de reproducción y una indicación de la ubicación de cada altavoz de reproducción dentro del entorno de reproducción. Estos datos
40 pueden recibirse, por ejemplo, accediendo a una estructura de datos que se almacena en una memoria accesible por el sistema lógico o recibida a través de un sistema de interfaz.

45 En este ejemplo, se aplican ecuaciones de ganancia de panoramización para la posición o posiciones (x, y, z) para determinar los valores de ganancia (bloque 628) para aplicar a los datos de audio (bloque 630).

55 En algunas implementaciones, los datos de audio cuyo nivel se ha ajustado en respuesta a los valores de ganancia pueden ser reproducidos por altavoces de reproducción, por ejemplo, por altavoces de auriculares (u otros altavoces) que se configuran para comunicarse con un sistema lógico de la herramienta de renderizado. En algunas implementaciones, las ubicaciones de los altavoces de reproducción pueden corresponder a las ubicaciones de las zonas de altavoz de un entorno de reproducción virtual, como el entorno de reproducción virtual 404 descrito anteriormente. Las respuestas correspondientes del altavoz pueden mostrarse en un dispositivo de visualización, por ejemplo, como se muestra en las Figuras 5A-5C.

60 En el bloque 635, se determina si el proceso continuará. Por ejemplo, el proceso puede finalizar (bloque 640) al recibir una entrada desde una interfaz de usuario que indica que un usuario ya no desea continuar con el proceso de renderizado. De lo contrario, el proceso puede continuar, por ejemplo, volviendo al bloque 626. Si

el sistema lógico recibe una indicación de que el usuario desea volver al proceso de creación correspondiente, el proceso 600 puede volver al bloque 607 o al bloque 610.

5 Otras implementaciones pueden implicar la imposición de otros diversos tipos de restricciones y la creación de otros tipos de metadatos de restricciones para objetos sonoros. La Figura 6B es un diagrama de flujo que esboza un ejemplo de un proceso de correlación de una posición de objeto sonoro a una ubicación de un solo altavoz. Este proceso también puede denominarse en esta memoria "ajuste". En el bloque 655, se recibe una indicación de que la posición de un objeto sonoro puede ajustarse a una sola ubicación de altavoz o una sola zona de altavoz. En este ejemplo, la indicación es que la posición del objeto sonoro se ajustará a una sola ubicación de altavoz, cuando corresponda. La indicación puede, por ejemplo, ser recibida por un sistema lógico de un aparato que se configura para proporcionar herramientas de creación. La indicación puede corresponder con la entrada recibida desde un dispositivo de entrada del usuario. Sin embargo, la indicación también puede corresponder con una categoría del objeto sonoro (por ejemplo, como un sonido de bala, una vocalización, etc.) y/o una anchura de objeto sonoro. La información sobre la categoría y/o la anchura puede, por ejemplo, recibirse como metadatos para el objeto sonoro. En tales implementaciones, el bloque 657 puede ocurrir antes del bloque 655.

20 En el bloque 656, se reciben datos de audio. Las coordenadas de la posición de un objeto sonoro se reciben en el bloque 657. En este ejemplo, se muestra la posición del objeto sonoro (bloque 658) según las coordenadas recibidas en el bloque 657. Los metadatos, incluidas las coordenadas del objeto sonoro y un indicador de ajuste, que indica la funcionalidad de ajuste, se guardan en el bloque 659. Los datos y metadatos de audio son enviados por la herramienta de creación a una herramienta de renderizado (bloque 660).

25 En el bloque 662, se determina si continuará el proceso de creación. Por ejemplo, el proceso de creación puede finalizar (bloque 663) al recibir una entrada desde una interfaz de usuario que indica que un usuario ya no desea ajustar las posiciones de los objetos sonoros a la ubicación de un altavoz. De lo contrario, el proceso de creación puede continuar, por ejemplo, volviendo al bloque 665. En algunas implementaciones, las operaciones de renderizado pueden continuar tanto si continúa el proceso de creación como si no.

30 Los datos de audio y los metadatos enviados por la herramienta de creación son recibidos por la herramienta de renderizado en el bloque 664. En el bloque 665, se determina (por ejemplo, por el sistema lógico) si ajustar la posición del objeto sonoro a la ubicación de altavoz. Esta determinación puede basarse, al menos en parte, en la distancia entre la posición del objeto sonoro y la ubicación de altavoz de reproducción más cercana de un entorno de reproducción.

35 En este ejemplo, si se determina en el bloque 665 ajustar la posición del objeto sonoro a la ubicación de un altavoz, la posición del objeto sonoro se correlacionará a una ubicación de altavoz en el bloque 670, generalmente la más cercana a la posición deseada (x, y, z) recibida para el objeto sonoro. En este caso, la ganancia de los datos de audio reproducidos por esta ubicación de altavoz será 1,0, mientras que la ganancia de los datos de audio reproducidos por otros altavoces será cero. En implementaciones alternativas, la posición del objeto sonoro puede correlacionarse a un grupo de ubicaciones de altavoces en el bloque 670.

45 Por ejemplo, con referencia nuevamente a la Figura 4B, el bloque 670 puede implicar ajustar la posición del objeto sonoro a uno de los altavoces superiores izquierdos 470a. Alternativamente, el bloque 670 puede involucrar ajustar la posición del objeto sonoro a un solo altavoz y altavoces vecinos, por ejemplo, 1 o 2 altavoces vecinos. En consecuencia, los metadatos correspondientes pueden aplicarse a un pequeño grupo de altavoces de reproducción y/o a un altavoz de reproducción individual.

50 Sin embargo, si se determina en el bloque 665 que la posición del objeto sonoro no se ajustará a la ubicación de un altavoz, por ejemplo, si esto diese como resultado una gran discrepancia en la posición en relación con la posición original prevista recibida para el objeto, las reglas de panoramización serán aplicadas (bloque 675). Las reglas de panoramización se pueden aplicar según la posición del objeto sonoro, así como otras características del objeto sonoro (como anchura, volumen, etc.).

55 Los datos de ganancia determinados en el bloque 675 pueden aplicarse a los datos de audio en el bloque 681 y el resultado puede guardarse. En algunas implementaciones, los datos de audio resultantes pueden ser reproducidos por altavoces que se configuran para comunicarse con el sistema lógico.

60 Si se determina en el bloque 685 que el proceso 650 continuará, el proceso 650 puede volver al bloque 664 para continuar con las operaciones de renderizado. Alternativamente, el proceso 650 puede volver al bloque 655 para reanudar las operaciones de creación.

65 El proceso 650 puede implicar diversos tipos de operaciones de suavizado. Por ejemplo, el sistema lógico puede configurarse para suavizar las transiciones en las ganancias aplicadas a los datos de audio cuando se pasa de correlacionar una posición de objeto sonoro desde una primera ubicación de un solo altavoz a una segunda ubicación de un solo altavoz. Con referencia nuevamente a la Figura 4B, si la posición del objeto

sonoro fue correlacionada inicialmente a uno de los altavoces superiores izquierdos 470a y luego fue correlacionada a uno de los altavoces envolventes traseros derechos 480b, el sistema lógico puede configurarse para suavizar la transición entre los altavoces para que el objeto sonoro no parezca "saltar" repentinamente de un altavoz (o zona de altavoz) a otro. En algunas implementaciones, el suavizado puede implementarse según un parámetro de velocidad de desvanecimiento cruzado.

En algunas implementaciones, el sistema lógico puede configurarse para suavizar las transiciones en las ganancias aplicadas a los datos de audio cuando se realiza la transición entre la correlación de una posición de objeto sonoro a una ubicación de un solo altavoz y la aplicación de reglas de panoramización para la posición del objeto sonoro. Por ejemplo, si se determinara posteriormente en el bloque 665 que la posición del objeto sonoro se había movido a una posición que se determinó que estaba demasiado lejos del altavoz más cercano, se pueden aplicar reglas de panoramización para la posición del objeto sonoro en el bloque 675. Sin embargo, cuando se pasa de ajuste a panoramización (o viceversa), el sistema lógico puede configurarse para suavizar las transiciones en las ganancias aplicadas a los datos de audio. El proceso puede terminar en el bloque 690, por ejemplo, al recibir la entrada correspondiente desde una interfaz de usuario.

Algunas implementaciones alternativas pueden implicar la creación de restricciones lógicas. En algunos casos, por ejemplo, un mezclador de sonido puede desear un control más explícito sobre el conjunto de altavoces que se usa durante una operación de panoramización particular. Algunas implementaciones permiten a un usuario generar "correlaciones lógicas" de una o dos dimensiones entre conjuntos de altavoces y una interfaz de panoramización.

La Figura 7 es un diagrama de flujo que esboza un proceso de establecimiento y uso de altavoces virtuales. Las Figuras 8A-8C muestran ejemplos de altavoces virtuales correlacionados a puntos finales de línea y respuestas de zona de altavoz correspondientes. Con referencia primero al proceso 700 de la Figura 7, se recibe una indicación en el bloque 705 para crear altavoces virtuales. La indicación puede ser recibida, por ejemplo, por un sistema lógico de un aparato de creación y puede corresponder con la entrada recibida desde un dispositivo de entrada del usuario.

En el bloque 710, se recibe una indicación de la ubicación de un altavoz virtual. Por ejemplo, con referencia a la Figura 8A, un usuario puede usar un dispositivo de entrada de usuario para colocar el cursor 510 en la posición del altavoz virtual 805a y seleccionar esa ubicación, por ejemplo, mediante un clic del ratón. En el bloque 715, se determina (por ejemplo, según la entrada del usuario) que se seleccionarán altavoces virtuales adicionales en este ejemplo. El proceso vuelve al bloque 710 y el usuario selecciona la posición del altavoz virtual 805b, que se muestra en la Figura 8A, en este ejemplo.

En este caso, el usuario solo desea establecer dos ubicaciones de altavoces virtuales. Por lo tanto, en el bloque 715, se determina (por ejemplo, según la entrada del usuario) que no se seleccionarán altavoces virtuales adicionales. Se puede mostrar una polilínea 810, como se muestra en la Figura 8A, conectando las posiciones del altavoz virtual 805a y 805b. En algunas implementaciones, la posición del objeto sonoro 505 estará restringida a la polilínea 810. En algunas implementaciones, la posición del objeto sonoro 505 puede estar restringida a una curva paramétrica. Por ejemplo, se puede proporcionar un conjunto de puntos de control según la entrada del usuario y se puede usar un algoritmo de ajuste de curvas, como un spline, para determinar la curva paramétrica. En el bloque 725, se recibe una indicación de la posición de un objeto sonoro a lo largo de la polilínea 810. En algunas de estas implementaciones, la posición se indicará como un valor escalar entre cero y uno. En el bloque 725, pueden visualizarse las coordenadas (x, y, z) del objeto sonoro y la polilínea definida por los altavoces virtuales. Se pueden mostrar los datos de audio y los metadatos asociados, incluida la posición escalar obtenida y las coordenadas (x, y, z) de los altavoces virtuales. (Bloque 727). Aquí, los datos de audio y los metadatos pueden enviarse a una herramienta de renderizado a través de un protocolo de comunicación apropiado en el bloque 728.

En el bloque 729, se determina si continuará el proceso de creación. Si no, el proceso 700 puede terminar (bloque 730) o puede continuar con las operaciones de renderizado, según la entrada del usuario. Sin embargo, como se ha indicado anteriormente, en muchas implementaciones, al menos algunas operaciones de renderizado pueden realizarse simultáneamente con las operaciones de creación.

En el bloque 732, la herramienta de renderizado recibe los datos y metadatos de audio. En el bloque 735, las ganancias a aplicar a los datos de audio se calculan para cada posición de altavoz virtual. La Figura 8B muestra las respuestas del altavoz para la posición del altavoz virtual 805a. La Figura 8C muestra las respuestas del altavoz para la posición del altavoz virtual 805b. En este ejemplo, como en muchos otros ejemplos descritos en esta memoria, las respuestas de los altavoces indicadas son para altavoces de reproducción que tienen ubicaciones correspondientes a las ubicaciones que se muestran para las zonas de altavoz de la GUI 400. Aquí, los altavoces virtuales 805a y 805b y la línea 810 se han posicionado en un plano que no está cerca de los altavoces de reproducción que tienen ubicaciones correspondientes a las zonas de altavoz 8 y 9. Por lo tanto, no se indica ninguna ganancia para estos altavoces en las Figuras 8B o 8C.

5 Cuando el usuario mueve el objeto sonoro 505 a otras posiciones a lo largo de la línea 810, el sistema lógico calculará el desvanecimiento cruzado que corresponde a estas posiciones (bloque 740), por ejemplo, según el parámetro de posición escalar del objeto sonoro. En algunas implementaciones, se puede usar una ley de panoramización por pares (por ejemplo, una ley de seno o potencia que conserva la energía) para combinar las ganancias que se aplicarán a los datos de audio para la posición del altavoz virtual 805a y las ganancias que se aplicarán a los datos de audio para la posición del altavoz virtual 805b.

10 En el bloque 742, se puede determinar (por ejemplo, según la entrada del usuario) si se continúa con el proceso 700. A un usuario se le puede presentar, por ejemplo, (por ejemplo, a través de una GUI) la opción de continuar con las operaciones de renderizado o de volver a las operaciones de creación. Si se determina que el proceso 700 no continuará, el proceso finaliza. (Bloque 745).

15 Al realizar una panorámica de objetos sonoros que se mueven rápidamente (por ejemplo, objetos sonoros que corresponden a automóviles, aviones, etc.), puede ser difícil crear una trayectoria suave si el usuario selecciona las posiciones de objetos sonoros punto por punto. La falta de suavidad en la trayectoria del objeto sonoro puede influir en la imagen acústica percibida. En consecuencia, algunas implementaciones de creación proporcionadas en esta memoria aplican un filtro de paso bajo a la posición de un objeto sonoro para suavizar las ganancias de panoramización resultantes. Las implementaciones de creación alternativas aplican un filtro de paso bajo a la ganancia aplicada a los datos de audio.

20 Otras implementaciones de creación pueden permitir que un usuario simule agarrar, tirar, lanzar o interactuar de manera similar con objetos sonoros. Algunas de estas implementaciones pueden implicar la aplicación de leyes físicas simuladas, como conjuntos de reglas que se usan para describir la velocidad, la aceleración, el impulso, la energía cinética, la aplicación de fuerzas, etc.

25 Las Figuras 9A-9C muestran ejemplos del uso de un amarre virtual para arrastrar un objeto sonoro. En la Figura 9A, se ha formado un amarre virtual 905 entre el objeto sonoro 505 y el cursor 510. En este ejemplo, el amarre virtual 905 tiene una constante de resorte virtual.

30 En algunas de tales implementaciones, la constante de resorte virtual puede seleccionarse según la entrada del usuario.

35 La Figura 9B muestra el objeto sonoro 505 y el cursor 510 en un momento posterior, después del cual el usuario ha movido el cursor 510 hacia la zona de altavoz 3. El usuario puede haber movido el cursor 510 usando un ratón, un joystick, una bola de seguimiento, un aparato de detección de gestos u otro tipo de dispositivo de entrada de usuario. El amarre virtual 905 se ha estirado y el objeto sonoro 505 se ha movido cerca de la zona de altavoz 8. El objeto sonoro 505 tiene aproximadamente el mismo tamaño en las Figuras 9A y 9B, lo que indica (en este ejemplo) que la elevación del objeto sonoro 505 no ha cambiado sustancialmente.

40 La Figura 9C muestra el objeto sonoro 505 y el cursor 510 en un momento posterior, después de lo cual el usuario ha movido el cursor alrededor de la zona de altavoz 9. El amarre virtual 905 se ha estirado aún más. El objeto sonoro 505 se ha movido hacia abajo, como lo indica la disminución del tamaño del objeto sonoro 505. El objeto sonoro 505 se ha movido en un arco suave. Este ejemplo ilustra un beneficio potencial de tales implementaciones, que es que el objeto sonoro 505 se puede mover en una trayectoria más suave que si un usuario simplemente selecciona posiciones para el objeto sonoro 505 punto por punto.

50 La Figura 10A es un diagrama de flujo que esboza un proceso de uso de un amarre virtual para mover un objeto sonoro. El proceso 1000 comienza con el bloque 1005, en el que se reciben los datos de audio. En el bloque 1007, se recibe una indicación para unir un amarre virtual entre un objeto sonoro y un cursor. La indicación puede ser recibida por un sistema lógico de un aparato de creación y puede corresponder con la entrada recibida desde un dispositivo de entrada del usuario. Con referencia a la Figura 9A, por ejemplo, un usuario puede colocar el cursor 510 sobre el objeto sonoro 505 y luego indicar, a través de un dispositivo de entrada de usuario o una GUI, que el amarre virtual 905 debe formarse entre el cursor 510 y el objeto sonoro 505. Se pueden recibir datos de posición del cursor y del objeto. (Bloque 1010).

55 En este ejemplo, los datos de velocidad y/o aceleración del cursor pueden ser calculados por el sistema lógico según los datos de posición del cursor, a medida que se mueve el cursor 510. (Bloque 1015). Los datos de posición y/o datos de trayectoria para el objeto sonoro 505 se pueden calcular según la constante de resorte virtual de el amarre virtual 905 y los datos de posición, velocidad y aceleración del cursor. Algunas implementaciones de este tipo pueden implicar la asignación de una masa virtual al objeto sonoro 505. (Bloque 1020). Por ejemplo, si el cursor 510 se mueve a una velocidad relativamente constante, el amarre virtual 905 puede no estirarse y el objeto sonoro 505 puede ser arrastrado a una velocidad relativamente constante. Si el cursor 510 acelera, el amarre virtual 905 puede estirarse y el amarre virtual 905 puede aplicar una fuerza correspondiente al objeto sonoro 505. Puede haber un retraso de tiempo entre la aceleración del cursor 510 y la fuerza aplicada por el amarre virtual 905. En implementaciones alternativas, la posición y/o la trayectoria del objeto sonoro 505 se pueden determinar de una manera diferente, por ejemplo, sin asignar una constante de

resorte virtual a el amarre virtual 905, aplicando reglas de fricción y/o inercia al objeto sonoro 505, etc.

Pueden visualizarse posiciones discretas y/o la trayectoria del objeto sonoro 505 y el cursor 510 (bloque 1025). En este ejemplo, el sistema lógico muestrea posiciones de objetos sonoros en un intervalo de tiempo (bloque 1030). En algunas de tales implementaciones, el usuario puede determinar el intervalo de tiempo para el muestreo. Se pueden guardar la ubicación del objeto sonoro y/o los metadatos de la trayectoria, etc. (Bloque 1034).

En el bloque 1036 se determina si continuará este modo de creación. El proceso puede continuar si el usuario así lo desea, por ejemplo, volviendo al bloque 1005 o al bloque 1010. De lo contrario, el proceso 1000 puede terminar (bloque 1040).

La Figura 10B es un diagrama de flujo que esboza un proceso alternativo de uso de un amarre virtual para mover un objeto sonoro. Las Figuras 10C-10E muestran ejemplos del proceso descrito en la Figura 10B. Con referencia primero a la Figura 10B, el proceso 1050 comienza con el bloque 1055, en el que se reciben los datos de audio. En el bloque 1057, se recibe una indicación para unir un amarre virtual entre un objeto sonoro y un cursor. La indicación puede ser recibida por un sistema lógico de un aparato de creación y puede corresponder con la entrada recibida desde un dispositivo de entrada del usuario. Con referencia a la Figura 10C, por ejemplo, un usuario puede colocar el cursor 510 sobre el objeto sonoro 505 y luego indicar, a través de un dispositivo de entrada de usuario o una GUI, que el amarre virtual 905 debe formarse entre el cursor 510 y el objeto sonoro 505.

Los datos de posición del cursor y del objeto sonoro pueden recibirse en el bloque 1060. En el bloque 1062, el sistema lógico puede recibir una indicación (a través de un dispositivo de entrada de usuario o una GUI, por ejemplo), de que el objeto sonoro 505 debe mantenerse en una posición indicada, por ejemplo, una posición indicada por el cursor 510. En el bloque 1065, el dispositivo lógico recibe una indicación de que el cursor 510 se ha movido a una nueva posición, que puede mostrarse junto con la posición del objeto sonoro 505 (bloque 1067). Con referencia a la Figura 10D, por ejemplo, el cursor 510 se ha movido del lado izquierdo al lado derecho del entorno de reproducción virtual 404. Sin embargo, el objeto sonoro 510 aún se mantiene en la misma posición indicada en la Figura 10C. Como resultado, el amarre virtual 905 se ha estirado sustancialmente.

En el bloque 1069, el sistema lógico recibe una indicación (a través de un dispositivo de entrada de usuario o una GUI, por ejemplo) de que el objeto sonoro 505 se va a liberar. El sistema lógico puede calcular los datos de la posición del objeto sonoro y/o la trayectoria resultantes, que pueden visualizarse (bloque 1075). La pantalla resultante puede ser similar a la que se muestra en la Figura 10E, que muestra el objeto sonoro 505 moviéndose suave y rápidamente a través del entorno de reproducción virtual 404. El sistema lógico puede guardar la ubicación del objeto sonoro y/o los metadatos de la trayectoria en un sistema de memoria (bloque 1080).

En el bloque 1085, se determina si continuará el proceso de creación 1050. El proceso puede continuar si el sistema lógico recibe una indicación de que el usuario desea hacerlo. Por ejemplo, el proceso 1050 puede continuar volviendo al bloque 1055 o al bloque 1060. De lo contrario, la herramienta de creación puede enviar los datos y metadatos de audio a una herramienta de renderizado (bloque 1090), después de lo cual el proceso 1050 puede finalizar (bloque 1095).

Para optimizar la verosimilitud del movimiento percibido de un objeto sonoro, puede ser conveniente permitir que el usuario de una herramienta de creación (o una herramienta de renderizado) seleccione un subconjunto de altavoces en un entorno de reproducción y limitar el conjunto de altavoces activos al subconjunto elegido. En algunas implementaciones, las zonas de altavoz y/o los grupos de zonas de altavoz pueden designarse como activos o inactivos durante una operación de creación o renderizado. Por ejemplo, con referencia a la Figura 4A, las zonas de altavoz del área frontal 405, el área izquierda 410, el área derecha 415 y/o el área superior 420 pueden controlarse como un grupo. Las zonas de altavoz de un área posterior que incluye las zonas de altavoz 6 y 7 (y, en otras implementaciones, una o más zonas de altavoz ubicadas entre las zonas de altavoz 6 y 7) también pueden controlarse como un grupo. Se puede proporcionar una interfaz de usuario para habilitar o deshabilitar dinámicamente todos los altavoces que corresponden a una zona de altavoz particular o a un área que incluye una pluralidad de zonas de altavoz.

En algunas implementaciones, el sistema lógico de un dispositivo de creación (o un dispositivo de renderizado) puede configurarse para crear metadatos de restricción de zona de altavoz según la entrada del usuario recibida a través de un sistema de entrada del usuario. Los metadatos de restricción de zona de altavoz pueden incluir datos para deshabilitar zonas de altavoz seleccionadas. Algunas de tales implementaciones se describirán ahora con referencia a las Figuras 11 y 12.

La Figura 11 muestra un ejemplo de aplicación de una restricción de zona de altavoz en un entorno de reproducción virtual. En algunas implementaciones de este tipo, un usuario puede seleccionar zonas de altavoz haciendo clic en sus representaciones en una GUI, como la GUI 400, usando un dispositivo de entrada de usuario como un ratón. Aquí, un usuario ha desactivado las zonas de altavoz 4 y 5, a los lados del entorno de

reproducción virtual 404. Las zonas de altavoz 4 y 5 pueden corresponder a la mayoría (o todos) de los altavoces en un entorno de reproducción física, como un entorno de sistema de sonido de cine. En este ejemplo, el usuario también ha restringido las posiciones del objeto sonoro 505 a posiciones a lo largo de la línea 1105. Con la mayoría o todos los altavoces a lo largo de las paredes laterales deshabilitadas, una panorámica desde la pantalla 150 hasta la parte posterior del entorno de reproducción virtual 404 estaría limitada a no usar los altavoces laterales. Esto puede crear un movimiento percibido mejorado de adelante atrás para un área de audiencia amplia, particularmente para los miembros de la audiencia que se sientan cerca de los altavoces de reproducción correspondientes a las zonas de altavoz 4 y 5.

En algunas implementaciones, las restricciones de la zona de altavoz pueden llevarse a cabo a través de todos los modos de renderizado. Por ejemplo, las restricciones de la zona de altavoz pueden llevarse a cabo en situaciones en las que hay menos zonas disponibles para renderizar, por ejemplo, cuando se renderiza para una configuración Dolby Surround 7.1 o 5.1 que expone solo 7 o 5 zonas. Las restricciones de la zona de altavoz también se pueden llevar a cabo cuando hay más zonas disponibles para renderizar. Como tal, las restricciones de la zona de altavoz también se pueden ver como una forma de guiar el renderizado, proporcionando una solución no ciega al proceso tradicional de "mezcla ascendente/mezcla descendente".

La Figura 12 es un diagrama de flujo que esboza algunos ejemplos de la aplicación de reglas de restricción de zona de altavoz. El proceso 1200 comienza con el bloque 1205, en el que se reciben una o más indicaciones para aplicar reglas de restricción de zona de altavoz. La indicación o indicaciones pueden ser recibidas por un sistema lógico de un aparato de creación o renderizado y puede corresponder con la entrada recibida desde un dispositivo de entrada de usuario. Por ejemplo, las indicaciones pueden corresponder a la selección de un usuario de una o más zonas de altavoz para desactivar. En algunas implementaciones, el bloque 1205 puede implicar recibir una indicación de qué tipo de reglas de restricción de zona de altavoz deben aplicarse, por ejemplo, como se describe a continuación.

En el bloque 1207, una herramienta de creación recibe los datos de audio. Los datos de posición de objeto sonoro pueden recibirse (bloque 1210), por ejemplo, según la entrada de un usuario de la herramienta de creación, y visualizarse (bloque 1215). Los datos de posición son coordenadas (x, y, z) en este ejemplo. Aquí, las zonas de altavoz activas e inactivas para las reglas de restricción de zonas de altavoz seleccionadas también se muestran en el bloque 1215. En el bloque 1220, se guardan los datos de audio y los metadatos asociados. En este ejemplo, los metadatos incluyen la posición del objeto sonoro y los metadatos de restricción de zona de altavoz, que pueden incluir un indicador de identificación de la zona de altavoz.

En algunas implementaciones, los metadatos de restricción de zona de altavoz pueden indicar que una herramienta de renderizado debe aplicar ecuaciones de panoramización para calcular las ganancias de forma binaria, por ejemplo, considerando que todos los altavoces de las zonas de altavoz seleccionadas (deshabilitadas) están "apagados" y todas las demás zonas de altavoz están "activadas". El sistema lógico puede configurarse para crear metadatos de restricción de zonas de altavoz que incluyen datos para deshabilitar las zonas de altavoz seleccionadas.

En implementaciones alternativas, los metadatos de restricción de zona de altavoz pueden indicar que la herramienta de renderizado aplicará ecuaciones de panoramización para calcular las ganancias de una manera combinada que incluye cierto grado de contribución de los altavoces de las zonas de altavoz deshabilitadas. Por ejemplo, el sistema lógico puede configurarse para crear metadatos de restricción de zonas de altavoz que indiquen que la herramienta de renderizado debe atenuar las zonas de altavoz seleccionadas realizando las siguientes operaciones: calcular las primeras ganancias que incluyen las contribuciones de las zonas de altavoz seleccionadas (desactivadas); calcular las segundas ganancias que no incluyen las contribuciones de las zonas de altavoz seleccionadas; y mezclar las primeras ganancias con las segundas ganancias. En algunas implementaciones, se puede aplicar una polarización a las primeras ganancias y/o las segundas ganancias (por ejemplo, desde un valor mínimo seleccionado hasta un valor máximo seleccionado) para permitir un abanico de contribuciones potenciales desde zonas de altavoz seleccionadas.

En este ejemplo, la herramienta de creación envía los datos y metadatos de audio a una herramienta de renderizado en el bloque 1225. El sistema lógico puede entonces determinar si el proceso de creación continuará (bloque 1227). El proceso de creación puede continuar si el sistema lógico recibe una indicación de que el usuario desea hacerlo. De lo contrario, el proceso de creación puede finalizar (bloque 1229). En algunas implementaciones, las operaciones de renderizado pueden continuar, según la entrada del usuario.

Los objetos sonoros, incluidos los datos de audio y los metadatos creados por la herramienta de creación, son recibidos por la herramienta de renderizado en el bloque 1230. Los datos de posición para un objeto sonoro particular se reciben en el bloque 1235 en este ejemplo. El sistema lógico de la herramienta de renderizado puede aplicar ecuaciones de panoramización para calcular las ganancias de los datos de posición de objeto sonoro, según las reglas de restricción de zona de altavoz.

En el bloque 1245, las ganancias calculadas se aplican a los datos de audio. El sistema lógico puede guardar

la ganancia, la ubicación del objeto sonoro y los metadatos de limitación de la zona de altavoz en un sistema de memoria. En algunas implementaciones, los datos de audio pueden ser reproducidos por un sistema de altavoces. Las respuestas correspondientes del altavoz pueden mostrarse en una pantalla en algunas implementaciones.

5

En el bloque 1248, se determina si el proceso 1200 continuará. El proceso puede continuar si el sistema lógico recibe una indicación de que el usuario desea hacerlo. Por ejemplo, el proceso de renderizado puede continuar volviendo al bloque 1230 o al bloque 1235. Si se recibe una indicación de que un usuario desea volver al proceso de creación correspondiente, el proceso puede volver al bloque 1207 o al bloque 1210. De lo contrario, el proceso 1200 puede terminar (bloque 1250).

10

Las tareas de posicionamiento y renderizado de objetos sonoros en un entorno de reproducción virtual tridimensional se están volviendo cada vez más difíciles. Parte de la dificultad se relaciona con los desafíos de representar el entorno de reproducción virtual en una GUI. Algunas implementaciones de creación y renderizado proporcionadas en esta memoria permiten que un usuario cambie entre una panoramización bidimensional del espacio de la pantalla y una panoramización tridimensional del espacio de la sala. Tal funcionalidad puede ayudar a preservar la precisión del posicionamiento del objeto sonoro mientras proporciona una GUI que es conveniente para el usuario.

15

Las Figuras 13A y 13B muestran un ejemplo de una GUI que puede cambiar entre una vista bidimensional y una vista tridimensional de un entorno de reproducción virtual. Con referencia primero a la Figura 13A, la GUI 400 representa una imagen 1305 en la pantalla. En este ejemplo, la imagen 1305 es la de un tigre dientes de sable. En esta vista superior del entorno de reproducción virtual 404, un usuario puede observar fácilmente que el objeto sonoro 505 está cerca de la zona de altavoz 1. La elevación puede inferirse, por ejemplo, por el tamaño, el color o algún otro atributo del objeto sonoro 505. Sin embargo, la relación de la posición con la de la imagen 1305 puede ser difícil de determinar en esta vista.

20

25

En este ejemplo, puede parecer que la GUI 400 gira dinámicamente alrededor de un eje, como el eje 1310. La Figura 13B muestra la GUI 1300 después del proceso de rotación. En esta vista, un usuario puede ver más claramente la imagen 1305 y puede usar la información de la imagen 1305 para posicionar el objeto sonoro 505 con mayor precisión. En este ejemplo, el objeto sonoro corresponde a un sonido hacia el que mira el tigre dientes de sable. Ser capaz de cambiar entre la vista superior y una vista de pantalla del entorno de reproducción virtual 404 permite al usuario seleccionar de forma rápida y precisa la elevación adecuada para el objeto sonoro 505, usando información del material en pantalla.

30

35

En esta memoria, se proporcionan diversas otras GUI convenientes para la creación y/o el renderizado. Las Figuras 13C-13E muestran combinaciones de representaciones bidimensionales y tridimensionales de entornos de reproducción. Con referencia primero a la Figura 13C, se representa una vista superior del entorno de reproducción virtual 404 en un área izquierda de la GUI 1310. La GUI 1310 también incluye una representación tridimensional 1345 de un entorno de reproducción virtual (o real). El área 1350 de la representación tridimensional 1345 se corresponde con la pantalla 150 de la GUI 400. La posición del objeto sonoro 505, particularmente su elevación, puede verse claramente en la representación tridimensional 1345. En este ejemplo, la anchura de objeto sonoro 505 también se muestra en la representación tridimensional 1345.

40

La disposición 1320 de los altavoces representa las ubicaciones de los altavoces 1324 a 1340, cada una de las cuales puede indicar una ganancia correspondiente a la posición del objeto sonoro 505 en el entorno de reproducción virtual 404. En algunas implementaciones, la disposición 1320 de los altavoces puede, por ejemplo, representar ubicaciones de altavoces de reproducción de un entorno de reproducción real, como una configuración Dolby Surround 5.1, una configuración Dolby Surround 7.1, una configuración Dolby 7.1 aumentada con altavoces superiores, etc. Cuando un sistema lógico recibe una indicación de una posición del objeto sonoro 505 en el entorno de reproducción virtual 404, el sistema lógico puede configurarse para correlacionar esta posición a ganancias para las ubicaciones 1324 a 1340 de los altavoces de la disposición 1320 de los altavoces, por ejemplo, mediante el proceso de panoramización de amplitud descrito anteriormente. Por ejemplo, en la Figura 13C, las ubicaciones 1325, 1335 y 1337 de los altavoces tienen cada una un cambio de color que indica las ganancias correspondientes a la posición del objeto sonoro 505.

45

50

55

Con referencia ahora a la Figura 13D, el objeto sonoro se ha movido a una posición detrás de la pantalla 150. Por ejemplo, un usuario puede haber movido el objeto sonoro 505 colocando un cursor sobre el objeto sonoro 505 en la GUI 400 y arrastrándolo a una nueva posición. Esta nueva posición también se muestra en la representación tridimensional 1345, que se ha girado a una nueva orientación. Las respuestas de la disposición de altavoces 1320 pueden parecer sustancialmente iguales en las Figuras 13C y 13D. Sin embargo, en una GUI real, las ubicaciones 1325, 1335 y 1337 de los altavoces pueden tener una apariencia diferente (como un brillo o color diferente) para indicar las diferencias de ganancia correspondientes causadas por la nueva posición del objeto sonoro 505.

60

65

Con referencia ahora a la Figura 13E, el objeto sonoro 505 se ha movido rápidamente a una posición en la

parte trasera derecha del entorno de reproducción virtual 404. En el momento representado en la Figura 13E, la ubicación 1326 del altavoz está respondiendo a la posición actual del objeto sonoro 505 y las ubicaciones 1325 y 1337 de los altavoces siguen respondiendo a la posición anterior del objeto sonoro 505.

5 La Figura 14A es un diagrama de flujo que esboza un proceso de control de un aparato para presentar las GUI como las que se muestran en las Figuras 13C-13E. El proceso 1400 comienza con el bloque 1405, en el que se reciben una o más indicaciones para visualizar ubicaciones de objetos sonoros, ubicaciones de zonas de altavoz y ubicaciones de altavoces de reproducción para un entorno de reproducción. Las ubicaciones de las zonas de altavoz pueden corresponder a un entorno de reproducción virtual y/o un entorno de reproducción
10 real, por ejemplo, como se muestra en las Figuras 13C-13E. La indicación o indicaciones pueden ser recibidas por un sistema lógico de un aparato de reproducción y/o creación y puede corresponder con la entrada recibida desde un dispositivo de entrada del usuario. Por ejemplo, las indicaciones pueden corresponder a la selección de un usuario de una configuración de entorno de reproducción.

15 En el bloque 1407, se reciben datos de audio. Los datos de posición y anchura de objeto sonoro se reciben en el bloque 1410, por ejemplo, según la entrada del usuario. En el bloque 1415, se muestran el objeto sonoro, las ubicaciones de las zonas de altavoz y las ubicaciones de los altavoces de reproducción. La posición del objeto sonoro puede mostrarse en vistas bidimensionales y/o tridimensionales, por ejemplo, como se muestra en las Figuras 13C-13E. Los datos de anchura pueden usarse no solo para el renderizado de objetos sonoros,
20 sino que también pueden afectar cómo se muestra el objeto sonoro (véase la representación del objeto sonoro 505 en la representación tridimensional 1345 de las Figuras 13C-13E).

Los datos de audio y los metadatos asociados pueden grabarse. (Bloque 1420).

25 En el bloque 1425, la herramienta de creación envía los datos y metadatos de audio a una herramienta de renderizado. El sistema lógico puede entonces determinar (bloque 1427) si continuará el proceso de creación. El proceso de creación puede continuar (por ejemplo, volviendo al bloque 1405) si el sistema lógico recibe una indicación de que el usuario desea hacerlo. De lo contrario, el proceso de creación puede finalizar. (Bloque 1429).

30 Los objetos sonoros, incluidos los datos de audio y los metadatos creados por la herramienta de creación, son recibidos por la herramienta de renderizado en el bloque 1430. Los datos de posición para un objeto sonoro particular se reciben en el bloque 1435 en este ejemplo. El sistema lógico de la herramienta de renderizado puede aplicar ecuaciones de panoramización para calcular las ganancias de los datos de posición de objeto sonoro, según los metadatos de anchura.

35 En algunas implementaciones de renderizado, el sistema lógico puede correlacionar las zonas de altavoz a los altavoces de reproducción del entorno de reproducción. Por ejemplo, el sistema lógico puede acceder a una estructura de datos que incluye zonas de altavoz y ubicaciones de altavoces de reproducción correspondientes. A continuación se describen más detalles y ejemplos con referencia a la Figura 14B.

40 En algunas implementaciones, pueden aplicarse ecuaciones de panoramización, por ejemplo, mediante un sistema lógico, según la posición del objeto sonoro, la anchura y/u otra información, como las ubicaciones de los altavoces del entorno de reproducción (bloque 1440). En el bloque 1445, los datos de audio se procesan según las ganancias que se obtienen en el bloque 1440. Al menos algunos de los datos de audio resultantes
45 pueden almacenarse, si así se desea, junto con los datos de posición de objeto sonoro correspondiente y otros metadatos recibidos de la herramienta de creación. Los datos de audio pueden ser reproducidos por altavoces.

50 El sistema lógico puede entonces determinar (bloque 1448) si el proceso 1400 continuará. El proceso 1400 puede continuar si, por ejemplo, el sistema lógico recibe una indicación de que el usuario desea hacerlo. De lo contrario, el proceso 1400 puede finalizar (bloque 1449).

La Figura 14B es un diagrama de flujo que esboza un proceso de renderizado de objetos sonoros para un entorno de reproducción. El proceso 1450 comienza con el bloque 1455, en el que se reciben una o más
55 indicaciones para generar objetos sonoros para un entorno de reproducción. La indicación o indicaciones pueden ser recibidas por un sistema lógico de un aparato de renderizado y puede corresponder con la entrada recibida desde un dispositivo de entrada del usuario. Por ejemplo, las indicaciones pueden corresponder a la selección de un usuario de una configuración de entorno de reproducción.

60 En el bloque 1457, se reciben datos de reproducción de audio (incluyendo uno o más objetos sonoros y metadatos asociados). Los datos de entorno de reproducción pueden recibirse en el bloque 1460. Los datos de entorno de reproducción pueden incluir una indicación de un número de altavoces de reproducción en el entorno de reproducción y una indicación de la ubicación de cada altavoz de reproducción dentro del entorno de reproducción. El entorno de reproducción puede ser un entorno de sistema de sonido de cine, un entorno de cine en casa, etc. En algunas implementaciones, los datos de entorno de reproducción pueden incluir datos
65 de disposición de zona de altavoz de reproducción que indican zonas de altavoz de reproducción y ubicaciones de altavoces de reproducción que se corresponden con las zonas de altavoz.

El entorno de reproducción puede visualizarse en el bloque 1465. En algunas implementaciones, el entorno de reproducción puede mostrarse de manera similar a la disposición 1320 de los altavoces que se muestra en las Figuras 13C-13E.

5

En el bloque 1470, los objetos sonoros pueden renderizarse en una o más señales de alimentación de altavoz para el entorno de reproducción. En algunas implementaciones, los metadatos asociados con los objetos sonoros pueden haber sido creados de una manera como la descrita anteriormente, de modo que los metadatos pueden incluir datos de ganancia correspondientes a las zonas de altavoz (por ejemplo, correspondientes a las zonas 1-9 de altavoz de la GUI 400). El sistema lógico puede correlacionar las zonas de altavoz a los altavoces de reproducción del entorno de reproducción. Por ejemplo, el sistema lógico puede acceder a una estructura de datos, almacenada en una memoria, que incluye zonas de altavoz y ubicaciones de altavoces de reproducción correspondientes. El dispositivo de renderizado puede tener una variedad de tales estructuras de datos, cada una de las cuales corresponde a una configuración de altavoz diferente. En algunas implementaciones, un aparato de renderizado puede tener tales estructuras de datos para una variedad de configuraciones de entorno de reproducción estándar, como una configuración Dolby Surround 5.1, una configuración Dolby Surround 7.1 y/o una configuración de sonido envolvente Hamasaki 22.2.

10

15

En algunas implementaciones, los metadatos de los objetos sonoros pueden incluir otra información del proceso de creación. Por ejemplo, los metadatos pueden incluir datos de restricción de altavoz. Los metadatos pueden incluir información para correlacionar una posición de objeto sonoro a una ubicación de altavoz de reproducción única o una zona de altavoz de reproducción única. Los metadatos pueden incluir datos que restringen la posición de un objeto sonoro a una curva unidimensional o una superficie bidimensional. Los metadatos pueden incluir datos de trayectoria para un objeto sonoro. Los metadatos pueden incluir un identificador para el tipo de contenido (por ejemplo, diálogo, música o efectos).

20

25

En consecuencia, el proceso de renderizado puede implicar el uso de los metadatos, por ejemplo, para imponer restricciones de zona de altavoz. En algunas de tales implementaciones, el aparato de reproducción puede proporcionar al usuario la opción de modificar las restricciones indicadas por los metadatos, por ejemplo, modificar las restricciones del altavoz y volver a procesar en consecuencia. El renderizado puede implicar la creación de una ganancia agregada basada en una o más de una posición de objeto sonoro deseada, una distancia desde la posición de objeto sonoro deseada a una posición de referencia, una velocidad de un objeto sonoro o un tipo de contenido de objeto sonoro. Es posible que se visualicen las respuestas correspondientes de los altavoces de reproducción. (Bloque 1475). En algunas implementaciones, el sistema lógico puede controlar los altavoces para reproducir el sonido correspondiente a los resultados del proceso de renderizado.

30

35

En el bloque 1480, el sistema lógico puede determinar si el proceso 1450 continuará. El proceso 1450 puede continuar si, por ejemplo, el sistema lógico recibe una indicación de que el usuario desea hacerlo. Por ejemplo, el proceso 1450 puede continuar volviendo al bloque 1457 o al bloque 1460. De lo contrario, el proceso 1450 puede terminar (bloque 1485).

40

El control de anchura de fuente aparente y la dispersión son características de algunos sistemas de creación/renderización de sonido envolvente existentes. En esta divulgación, el término "difusión" se refiere a distribuir la misma señal sobre múltiples altavoces para difuminar la imagen acústica. El término "anchura" se refiere a la decorrelación de las señales de salida a cada canal para el control de anchura aparente. La anchura puede ser un valor escalar adicional que controla la cantidad de decorrelación aplicada a cada señal de alimentación de altavoz.

45

Algunas implementaciones descritas en esta memoria proporcionan un control de dispersión orientado al eje 3D. Una implementación de este tipo se describirá ahora con referencia a las Figuras 15A y 15B. La Figura 15A muestra un ejemplo de un objeto sonoro y la anchura de objeto sonoro asociado en un entorno de reproducción virtual. Aquí, la GUI 400 indica un elipsoide 1505 que se extiende alrededor del objeto sonoro 505, indicando la anchura de objeto sonoro. La anchura de objeto sonoro puede indicarse mediante metadatos del objeto sonoro y/o recibirse según la entrada del usuario. En este ejemplo, las dimensiones x e y del elipsoide 1505 son diferentes, pero en otras implementaciones estas dimensiones pueden ser las mismas. Las dimensiones z del elipsoide 1505 no se muestran en la Figura 15A.

50

55

La Figura 15B muestra un ejemplo de un perfil de dispersión correspondiente a la anchura de objeto sonoro que se muestra en la Figura 15A. La dispersión se puede representar como un parámetro vectorial tridimensional. En este ejemplo, el perfil de dispersión 1507 se puede controlar de forma independiente a lo largo de 3 dimensiones, por ejemplo, según la entrada del usuario. Las ganancias a lo largo de los ejes x e y se representan en la Figura 15B por la altura respectiva de las curvas 1510 y 1520. La ganancia para cada muestra 1512 también se indica mediante el tamaño de los círculos correspondientes 1515 dentro del perfil de dispersión 1507. Las respuestas de los altavoces 1510 se indican mediante sombreado gris en la Figura 15B.

60

65

En algunas implementaciones, el perfil de dispersión 1507 puede implementarse mediante una integral

separable para cada eje. Según algunas implementaciones, se puede establecer automáticamente un valor de dispersión mínimo en función de la colocación de altavoces para evitar discrepancias tímbricas durante la panoramización. Alternativa o adicionalmente, se puede establecer automáticamente un valor mínimo de dispersión en función de la velocidad del objeto sonoro panoramizado, de modo que a medida que aumenta la velocidad del objeto sonoro, un objeto se vuelve más disperso espacialmente, de manera similar a la rapidez con la que se mueven las imágenes en una película parecen desdibujarse.

Cuando se usan implementaciones de reproducción de audio basadas en objetos sonoros como las descritas en esta memoria, es posible que se entreguen sin mezclar al entorno de reproducción una cantidad potencialmente grande de pistas de audio y los metadatos que las acompañan (incluidos, entre otros, los metadatos que indican las posiciones de los objetos sonoros en el espacio tridimensional). Una herramienta de renderizado en tiempo real puede usar dichos metadatos e información sobre el entorno de reproducción para calcular las señales de alimentación de altavoz para optimizar la reproducción de cada objeto sonoro.

Cuando se mezcla una gran cantidad de objetos sonoros en las salidas de los altavoces, puede producirse una sobrecarga en el dominio digital (por ejemplo, la señal digital puede recortarse antes de la conversión analógica) o en el dominio analógico, cuando la señal analógica amplificada es reproducida por los altavoces de reproducción.

Ambos casos pueden dar como resultado una distorsión audible, lo cual no es deseable. La sobrecarga en el dominio analógico también podría dañar los altavoces de reproducción.

En consecuencia, algunas implementaciones descritas en esta memoria implican el "manejo en blobs" de objetos dinámicos en respuesta a la sobrecarga del altavoz de reproducción. Cuando los objetos sonoros se renderizan con un perfil de dispersión dado, en algunas implementaciones la energía puede dirigirse a un mayor número de altavoces de reproducción vecinos mientras se mantiene la energía global constante. Por ejemplo, si la energía para el objeto sonoro se distribuyera uniformemente entre N altavoces de reproducción, podría contribuir a cada salida de altavoz de reproducción con una ganancia de $1/\sqrt{N}$. Este enfoque proporciona un "margen superior" de mezcla adicional y puede aliviar o prevenir la distorsión del altavoz de reproducción, como el recorte.

Para usar un ejemplo numérico, supóngase que un altavoz se recortará si recibe una entrada mayor que 1,0. Supóngase que se indica que se mezclan dos objetos en el altavoz A, uno en el nivel 1,0 y el otro en el nivel 0,25. Si no se usara el manejo en blobs, el nivel de mezcla en el altavoz A totalizaría 1,25 y se produciría un recorte. Sin embargo, si el primer objeto se maneja en blobs con otro altavoz B, entonces (según algunas implementaciones) cada altavoz recibiría el objeto a 0,707, lo que daría como resultado un "margen superior" adicional en el altavoz A para mezclar objetos adicionales. El segundo objeto se puede mezclar de forma segura en el altavoz A sin recorte, ya que el nivel de mezcla para el altavoz A será $0,707 + 0,25 = 0,957$.

En algunas implementaciones, durante la fase de creación, cada objeto sonoro puede mezclarse con un subconjunto de las zonas de altavoz (o todas las zonas de altavoz) con una ganancia de mezcla determinada. Por lo tanto, se puede construir una lista dinámica de todos los objetos que contribuyen a cada altavoz. En algunas implementaciones, esta lista se puede ordenar por niveles de energía decrecientes, por ejemplo, usando el producto del nivel original de raíz cuadrática media (RMS) de la señal multiplicado por la ganancia de mezcla. En otras implementaciones, la lista puede clasificarse según otros criterios, como la importancia relativa asignada al objeto sonoro.

Durante el proceso de renderizado, si se detecta una sobrecarga para una salida de altavoz de reproducción determinada, la energía de los objetos sonoros puede dispersarse entre varios altavoces de reproducción. Por ejemplo, la energía de los objetos sonoros puede dispersarse usando una anchura o un factor de dispersión que sea proporcional a la cantidad de sobrecarga y a la contribución relativa de cada objeto sonoro al altavoz de reproducción dado. Si el mismo objeto sonoro contribuye a que varios altavoces de reproducción se sobrecarguen, su anchura o factor de dispersión puede, en algunas implementaciones, aumentarse de forma aditiva y aplicarse a la siguiente trama renderizada de datos de audio.

En general, un limitador duro recortará cualquier valor que exceda un umbral al valor del umbral. Como en el ejemplo anterior, si un altavoz recibe un objeto mixto en el nivel 1,25 y solo puede permitir un nivel máximo de 1,0, el objeto estará "limitado" a 1,0. Un limitador suave comenzará a aplicar la limitación antes de alcanzar el umbral absoluto para proporcionar un resultado más suave y agradable al oído. Los limitadores suaves también pueden usar una función de "mirar hacia adelante" para predecir cuándo puede ocurrir un recorte futuro para reducir suavemente la ganancia antes de que ocurra el recorte y así evitar el recorte.

Diversas implementaciones de "manejo en blobs" proporcionadas en esta memoria pueden usarse junto con un limitador duro o suave para limitar la distorsión audible mientras se evita la degradación de la precisión/nitidez espacial. A diferencia de una difusión global o el uso exclusivo de limitadores, las implementaciones de manejo en blobs pueden apuntar selectivamente a objetos ruidosos u objetos de un tipo

de contenido determinado. Tales implementaciones pueden ser controladas por el mezclador. Por ejemplo, si los metadatos de restricción de zona de altavoz para un objeto sonoro indican que no se debe usar un subconjunto de altavoces de reproducción, el aparato de reproducción puede aplicar las reglas de restricción de zona de altavoz correspondientes además de implementar un método de manejo en blobs.

5

La Figura 16 es un diagrama de flujo que esboza un proceso de manejar en blobs de objetos sonoros. El proceso 1600 comienza con el bloque 1605, en donde se reciben una o más indicaciones para activar la funcionalidad de manejo en blobs de objetos sonoros. La indicación o indicaciones pueden ser recibidas por un sistema lógico de un aparato de renderizado y puede corresponder con la entrada recibida desde un dispositivo de entrada del usuario. En algunas implementaciones, las indicaciones pueden incluir la selección de un usuario de una configuración de entorno de reproducción. En implementaciones alternativas, el usuario puede haber seleccionado previamente una configuración de entorno de reproducción.

10

En el bloque 1607, se reciben datos de reproducción de audio (incluyendo uno o más objetos sonoros y metadatos asociados). En algunas implementaciones, los metadatos pueden incluir metadatos de restricción de zona de altavoz, por ejemplo, como se describe anteriormente. En este ejemplo, la posición del objeto sonoro, el tiempo y los datos de dispersión se analizan a partir de los datos de reproducción de audio (o se reciben de otro modo, por ejemplo, a través de la entrada desde una interfaz de usuario) en el bloque 1610.

15

Las respuestas de los altavoces de reproducción se determinan para la configuración del entorno de reproducción aplicando ecuaciones de panoramización para los datos del objeto sonoro, por ejemplo, como se describe anteriormente (bloque 1612). En el bloque 1615, se visualizan la posición del objeto sonoro y las respuestas del altavoz de reproducción (bloque 1615). Las respuestas de los altavoces de reproducción también pueden reproducirse a través de altavoces que se configuran para comunicarse con el sistema lógico.

20

En el bloque 1620, el sistema lógico determina si se detecta una sobrecarga para cualquier altavoz de reproducción del entorno de reproducción. Si es así, pueden aplicarse reglas de manejo como blobs de objetos sonoros como las descritas anteriormente hasta que no se detecte sobrecarga (bloque 1625). La salida de datos de audio en el bloque 1630 se puede guardar, si así se desea, y se puede enviar a los altavoces de reproducción.

25

En el bloque 1635, el sistema lógico puede determinar si el proceso 1600 continuará. El proceso 1600 puede continuar si, por ejemplo, el sistema lógico recibe una indicación de que el usuario desea hacerlo. Por ejemplo, el proceso 1600 puede continuar volviendo al bloque 1607 o al bloque 1610. De lo contrario, el proceso 1600 puede terminar (bloque 1640).

30

Algunas implementaciones proporcionan ecuaciones de ganancia de panoramización extendidas que se pueden usar para generar una imagen de la posición de un objeto sonoro en un espacio tridimensional. A continuación se describirán algunos ejemplos con referencia a las Figuras 17A y 17B. Las Figuras 17A y 17B muestran ejemplos de un objeto sonoro posicionado en un entorno de reproducción virtual tridimensional. Con referencia primero a la Figura 17A, la posición del objeto sonoro 505 puede verse dentro del entorno de reproducción virtual 404. En este ejemplo, las zonas 1-7 de altavoz se ubican en un plano y las zonas de altavoz 8 y 9 se ubican en otro plano, como se muestra en la Figura 17B. Sin embargo, los números de zonas de altavoz, planos, etc., se dan meramente a modo de ejemplo; los conceptos descritos en esta memoria pueden extenderse a diferentes números de zonas de altavoz (o altavoces individuales) y más de dos planos de elevación.

40

En este ejemplo, un parámetro de elevación "z", que puede variar de cero a 1, correlaciona la posición de un objeto sonoro a los planos de elevación. En este ejemplo, el valor $z = 0$ corresponde al plano base que incluye las zonas de altavoz 1-7, mientras que el valor $z = 1$ corresponde al plano superior que incluye las zonas de altavoz 8 y 9.

50

Los valores de z entre cero y 1 corresponden a una mezcla entre una imagen acústica generada usando solo los altavoces en el plano base y una imagen acústica generada usando solo los altavoces en el plano superior.

En el ejemplo que se muestra en la Figura 17B, el parámetro de elevación para el objeto sonoro 505 tiene un valor de 0,6. En consecuencia, en una implementación, se puede generar una primera imagen acústica usando ecuaciones de panoramización para el plano base, según las coordenadas (x, y) del objeto sonoro 505 en el plano base. Se puede generar una segunda imagen acústica usando ecuaciones de panoramización para el plano superior, según las coordenadas (x, y) del objeto sonoro 505 en el plano superior. Se puede producir una imagen acústica resultante combinando la primera imagen acústica con la segunda imagen acústica, según la proximidad del objeto sonoro 505 a cada plano. Se puede aplicar una función de conservación de energía o amplitud de la elevación z . Por ejemplo, suponiendo que z puede variar de cero a uno, los valores de ganancia de la primera imagen acústica pueden multiplicarse por $\cos(z \cdot \pi/2)$ y los valores de ganancia de la segunda imagen acústica pueden multiplicarse por $\sin(z \cdot \pi/2)$, de modo que la suma de sus cuadrados sea 1 (preservación de energía).

55

60

65

Otras implementaciones descritas en esta memoria pueden implicar el cálculo de ganancias basándose en dos o más técnicas de panoramización y la creación de una ganancia agregada basada en uno o más parámetros. Los parámetros pueden incluir uno o más de los siguientes: posición deseada del objeto sonoro; distancia desde la posición deseada del objeto sonoro hasta una posición de referencia; la velocidad o velocidad del objeto sonoro; o tipo de contenido de objeto sonoro.

Algunas implementaciones de este tipo se describirán ahora con referencia a las Figuras 18 y siguientes. La Figura 18 muestra ejemplos de zonas que se corresponden con diferentes modos de panoramización. Los tamaños, formas y extensión de estas zonas se dan simplemente a modo de ejemplo. En este ejemplo, los métodos de panoramización de campo cercano se aplican a los objetos sonoros ubicados dentro de la zona 1805 y los métodos de panoramización de campo lejano se aplican a los objetos sonoros ubicados en la zona 1815, fuera de la zona 1810.

Las Figuras 19A-19D muestran ejemplos de aplicación de técnicas de panoramización de campo cercano y campo lejano a objetos sonoros en diferentes ubicaciones. Con referencia primero a la Figura 19A, el objeto sonoro está sustancialmente fuera del entorno 1900 de reproducción virtual. Esta ubicación corresponde a la zona 1815 de la Figura 18. Por lo tanto, en este caso se aplicarán uno o más métodos de panoramización de campo lejano. En algunas implementaciones, los métodos de panoramización de campo lejano pueden basarse en ecuaciones de panoramización de amplitud basada en vectores (VBAP) que son conocidas por los expertos en la técnica. Por ejemplo, los métodos 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 (AES International Conference on Virtual, Synthetic and Entertainment Audio)*, que se incorpora en la presente por la referencia. En implementaciones alternativas, se pueden usar otros métodos para panoramizar objetos sonoros de campo lejano y de campo cercano, por ejemplo, métodos que involucran la síntesis de planos acústicos u ondas esféricas correspondientes. Documento de D. de Vries, *Wave Field Synthesis (AES Monográfico 1999)*.

Con referencia ahora a la Figura 19B, el objeto sonoro está dentro del entorno 1900 de reproducción virtual. Esta ubicación corresponde a la zona 1805 de la Figura 18. Por lo tanto, en este caso se aplicarán uno o más métodos de panoramización de campo cercano. Algunos de estos métodos de panoramización de campo cercano usarán una serie de zonas de altavoz que encierran el objeto sonoro 505 en el entorno 1900 de reproducción virtual.

En algunas implementaciones, el método de panoramización de campo cercano puede implicar una panoramización de "equilibrio dual" y la combinación de dos conjuntos de ganancias. En el ejemplo representado en la Figura 19B, el primer conjunto de ganancias corresponde a un equilibrio delante/detrás entre dos conjuntos de zonas de altavoz que encierran posiciones del objeto sonoro 505 a lo largo del eje y. Las respuestas correspondientes implican todas las zonas de altavoz del entorno de reproducción virtual 1900, excepto las zonas de altavoz 1915 y 1960.

En el ejemplo representado en la Figura 19C, el segundo conjunto de ganancias corresponde a un equilibrio izquierda/derecha entre dos conjuntos de zonas de altavoz que encierran posiciones del objeto sonoro 505 a lo largo del eje x. Las respuestas correspondientes involucran las zonas de altavoz 1905 a 1925. La Figura 19D indica el resultado de combinar las respuestas indicadas en las Figuras 19B y 19C.

Puede ser deseable combinar diferentes modos de panoramización cuando un objeto sonoro entra o sale del entorno de reproducción virtual 1900. En consecuencia, se aplica una combinación de ganancias calculadas según métodos de panoramización de campo cercano y métodos de panoramización de campo lejano para objetos sonoros ubicados en la zona 1810 (consúltese la Figura 18). En algunas implementaciones, se puede usar una ley de panoramización por pares (por ejemplo, una ley del seno o potencia que conserva la energía) para combinar las ganancias calculadas según los métodos de panoramización de campo cercano y los métodos de panoramización de campo lejano. En implementaciones alternativas, la ley de panoramización por pares puede conservar la amplitud en lugar de conservar la energía, de modo que la suma sea igual a uno en lugar de que la suma de los cuadrados sea igual a uno. También es posible combinar las señales procesadas resultantes, por ejemplo, para procesar la señal de audio usando ambos métodos de panoramización de forma independiente y para atenuar las dos señales de audio resultantes.

Puede ser deseable proporcionar un mecanismo que permita al creador de contenido y/o al reproductor de contenido ajustar fácilmente los diferentes rerenderizados para una trayectoria creada determinada. En el contexto de la mezcla para películas, el concepto de equilibrio de energía de pantalla a sala se considera importante. En algunos casos, una reproducción automática de una trayectoria de sonido determinada (o "panorámica") dará como resultado un equilibrio diferente de pantalla a sala, según la cantidad de altavoces de reproducción en el entorno de reproducción. Según algunas implementaciones, la polarización de pantalla a sala puede controlarse según los metadatos creados durante un proceso de creación. Según implementaciones alternativas, la polarización de pantalla a sala puede controlarse únicamente en el lado del renderizado (es decir, bajo el control del reproductor de contenido), y no en respuesta a los metadatos.

En consecuencia, algunas implementaciones descritas en esta memoria proporcionan una o más formas de control de polarización de pantalla a sala. En algunas implementaciones de este tipo, la polarización de pantalla a sala puede implementarse como una operación de escalado. Por ejemplo, la operación de escalado puede implicar la trayectoria prevista original de un objeto sonoro a lo largo de la dirección de adelante atrás y/o un escalado de las posiciones de los altavoces usadas en el renderizador para determinar las ganancias de panoramización. En algunas de tales implementaciones, el control de polarización de pantalla a sala puede ser un valor variable entre cero y un valor máximo (por ejemplo, uno). La variación puede, por ejemplo, ser controlable con una GUI, un control deslizante virtual o físico, un botón, etc.

Alternativa adicionalmente, el control de polarización de pantalla a sala puede implementarse usando alguna forma de restricción del área del altavoz. La Figura 20 indica las zonas de altavoz de un entorno de reproducción que pueden usarse en un proceso de control de polarización de pantalla a sala. En este ejemplo, se puede establecer el área de altavoz delantero 2005 y el área de altavoz trasero 2010 (o 2015). La polarización de pantalla a sala se puede ajustar en función de las áreas de altavoz seleccionadas. En algunas implementaciones de este tipo, se puede implementar una polarización de pantalla a sala como una operación de escala entre el área 2005 de altavoz delantero y el área 2010 (o 2015) de altavoz trasero. En implementaciones alternativas, la polarización de pantalla a sala puede implementarse de forma binaria, por ejemplo, permitiendo que un usuario seleccione una polarización del lado frontal, una polarización del lado posterior o ninguna polarización. Los ajustes de polarización para cada caso pueden corresponder con niveles de polarización predeterminados (y generalmente distintos de cero) para el área 2005 de altavoz delantero y el área 2010 (o 2015) de altavoz trasero. En esencia, tales implementaciones pueden proporcionar tres ajustes preestablecidos para el control de polarización de pantalla a sala en lugar de (o además de) una operación de escalado de valor continuo.

Según algunas implementaciones de este tipo, se pueden crear dos zonas de altavoz lógicas adicionales en una GUI de creación (por ejemplo, 400) dividiendo las paredes laterales en una pared lateral delantera y una pared lateral trasera. En algunas implementaciones, las dos zonas de altavoz lógicas adicionales corresponden a las áreas de pared izquierda/sonido envolvente izquierdo y pared derecha/sonido envolvente derecho del renderizador. Dependiendo de la selección de un usuario de cuál de estas dos zonas lógicas de altavoz están activas, la herramienta de renderizado podría aplicar factores de escala preestablecidos (por ejemplo, como se describe arriba) al renderizar en configuraciones Dolby 5.1 o Dolby 7.1. La herramienta de renderizado también puede aplicar dichos factores de escala preestablecidos al renderizar para entornos de reproducción que no admitan la definición de estas dos zonas lógicas adicionales, por ejemplo, porque sus configuraciones de altavoz físico no tienen más de un altavoz físico en la pared lateral.

La Figura 21 es un diagrama de bloques que proporciona ejemplos de componentes de un aparato de creación y/o renderizado. En este ejemplo, el dispositivo 2100 incluye un sistema de interfaz 2105. El sistema de interfaz 2105 puede incluir una interfaz de red, como una interfaz de red inalámbrica. Como alternativa, o adicionalmente, el sistema de interfaz 2105 puede incluir una interfaz de bus universal en serie (USB) u otra interfaz similar.

El dispositivo 2100 incluye un sistema lógico 2110. El sistema lógico 2110 puede incluir un procesador, tal como un procesador de chip único o multichip de uso general. El sistema lógico 2110 puede incluir un procesador de señal digital (DSP), un circuito integrado de aplicación específica (ASIC), un conjunto de puertas programables en campo (FPGA) u otro dispositivo lógico programable, lógica de puerta discreta o transistor, o componentes de hardware discretos, o combinaciones de los mismos. El sistema lógico 2110 puede configurarse para controlar los otros componentes del dispositivo 2100. Aunque en la Figura 21 no se muestran interfaces entre los componentes del dispositivo 2100, el sistema lógico 2110 puede configurarse con interfaces para la comunicación con los otros componentes. Los otros componentes pueden o no configurarse para comunicarse entre sí, según corresponda.

El sistema lógico 2110 puede configurarse para realizar funciones de creación y/o renderizado de audio, incluidos, entre otros, los tipos de funciones de creación y/o renderizado de audio descritos en esta memoria. En algunas de tales implementaciones, el sistema lógico 2110 puede configurarse para operar (al menos en parte) según el software almacenado en uno o más medios no transitorios. Los medios no transitorios pueden incluir memoria asociada con el sistema lógico 2110, como memoria de acceso aleatorio (RAM) y/o memoria de solo lectura (ROM). Los medios no transitorios pueden incluir la memoria del sistema de memoria 2115. El sistema de memoria 2115 puede incluir uno o más tipos adecuados de medios de almacenamiento no transitorios, como memoria flash, disco duro, etc.

El sistema de visualización 2130 puede incluir uno o más tipos de visualización adecuados, dependiendo de la manifestación del dispositivo 2100. Por ejemplo, el sistema de visualización 2130 puede incluir una pantalla de cristal líquido, una pantalla de plasma, una pantalla biestable, etc.

El sistema de entrada de usuario 2135 puede incluir uno o más dispositivos configurados para aceptar entradas

- de un usuario. En algunas implementaciones, el sistema de entrada de usuario 2135 puede incluir una pantalla táctil que superpone una pantalla del sistema de visualización 2130. El sistema de entrada de usuario 2135 puede incluir un ratón, una bola de seguimiento, un sistema de detección de gestos, un joystick, una o más GUI y/o menús presentados en el sistema de visualización 2130, botones, un teclado, interruptores, etc. En algunas implementaciones, el sistema de entrada de usuario 2135 puede incluir el micrófono 2125: un usuario puede proporcionar órdenes de voz para el dispositivo 2100 a través del micrófono 2125. El sistema lógico puede configurarse para reconocimiento de voz y para controlar al menos algunas operaciones del dispositivo 2100 según dichas órdenes de voz.
- El sistema de potencia 2140 puede incluir uno o más dispositivos de almacenamiento de energía adecuados, como una batería de níquel-cadmio o una batería de iones de litio. El sistema de potencia 2140 puede configurarse para recibir potencia de una toma de corriente.
- La Figura 22A es un diagrama de bloques que representa algunos componentes que pueden usarse para la creación de contenido de audio. El sistema 2200 puede, por ejemplo, usarse para la creación de contenido de audio en estudios de mezcla y/o escenarios de doblaje. En este ejemplo, el sistema 2200 incluye una herramienta de creación de audio y metadatos 2205 y una herramienta de renderizado 2210. En esta implementación, la herramienta de creación de audio y metadatos 2205 y la herramienta de renderizado 2210 incluyen interfaces de conexión de audio 2207 y 2212, respectivamente, que pueden configurarse para comunicación a través de AES/EBU, MADI, analógico, etc. La herramienta de creación de audio y metadatos 2205 y la herramienta de renderizado 2210 incluyen interfaces de red 2209 y 2217, respectivamente, que pueden configurarse para enviar y recibir metadatos a través de TCP/IP o cualquier otro protocolo adecuado. La interfaz 2220 se configura para enviar datos de audio a los altavoces.
- El sistema 2200 puede, por ejemplo, incluir un sistema de creación existente, tal como un sistema Pro Tools™, ejecutando una herramienta de creación de metadatos (es decir, un panoramizador como se describe en esta memoria) como complemento. El panoramizador también podría ejecutarse en un sistema independiente (por ejemplo, un PC o una mesa de mezclas) conectado a la herramienta de renderizado 2210, o podría ejecutarse en el mismo dispositivo físico que la herramienta de renderizado 2210. En el último caso, el panoramizador y el renderizador podrían usar una conexión local, por ejemplo, a través de la memoria compartida. La GUI del panoramizador también se puede controlar de forma remota en una tableta, un ordenador portátil, etc. La herramienta de renderizado 2210 puede comprender un sistema de renderizado que incluye un procesador de sonido que se configura para ejecutar software de renderizado. El sistema de reproducción puede incluir, por ejemplo, un ordenador personal, un ordenador portátil, etc., que incluya interfaces para entrada/salida de audio y un sistema lógico apropiado.
- La Figura 22B es un diagrama de bloques que representa algunos componentes que pueden usarse para la reproducción de audio en un entorno de reproducción (por ejemplo, una sala de cine). El sistema 2250 incluye un servidor de cine 2255 y un sistema de renderizado 2260 en este ejemplo. El servidor de cine 2255 y el sistema de renderizado 2260 incluyen interfaces de red 2257 y 2262, respectivamente, que pueden configurarse para enviar y recibir objetos sonoros a través de TCP/IP o cualquier otro protocolo adecuado. La interfaz 2264 se configura para enviar datos de audio a los altavoces.
- Diversas modificaciones a las implementaciones descritas en esta divulgación pueden ser fácilmente evidentes para los expertos en la técnica. Los principios generales definidos en esta memoria pueden aplicarse a otras implementaciones sin apartarse del alcance de esta divulgación. Por lo tanto, no se pretende que las reivindicaciones se limiten a las implementaciones que se muestran en esta memoria, sino que deben otorgarles el alcance más amplio compatible con esta divulgación, los principios y las características novedosas que se divulgan en esta memoria.

REIVINDICACIONES

1. Un aparato, que comprende:
- 5 un sistema de interfaz (2105); y
- un sistema lógico (2110) configurado para:
- 10 recibir, por medio del sistema de interfaz (2105), datos de reproducción de audio que comprenden uno o más objetos sonoros y metadatos asociados;
- recibir, a través del sistema de interfaz (2105), datos de entorno de reproducción que comprenden una indicación de un número de altavoces de reproducción de un entorno de reproducción tridimensional
- 15 real y una indicación de la ubicación de cada altavoz de reproducción dentro del entorno de reproducción real; y
- renderizar uno o más objetos sonoros en una o más señales de alimentación de altavoz basándose, al menos en parte, en los metadatos asociados, en donde cada señal de alimentación de altavoz
- 20 corresponde a al menos uno de los altavoces de reproducción dentro del entorno de reproducción real,
- el aparato se caracteriza por que:
- los metadatos asociados con cada objeto sonoro incluyen metadatos de restricción de zona de
- 25 altavoz que indican si renderizar el objeto sonoro respectivo implica imponer restricciones de zona de altavoz, y
- en donde renderizar el uno o más objetos sonoros incluye renderizar el respectivo objeto sonoro al imponer restricciones de zona de altavoz en respuesta a los metadatos de restricción de zona
- 30 de altavoz, y
- en donde renderizar el uno o más objetos sonoros incluye además renderizar el respectivo objeto sonoro ya sea al aplicar reglas de panoramización para correlacionar el objeto sonoro a una pluralidad de altavoces de reproducción, o al correlacionar el objeto sonoro a un único altavoz de
- 35 reproducción.
2. El aparato de la reivindicación 1, en donde el renderizado implica crear una ganancia basada en uno o más de una posición de objeto sonoro deseada, una distancia desde la posición de objeto sonoro deseada a una posición de referencia, una velocidad de un objeto sonoro o un tipo de contenido de objeto sonoro.
- 40 3. El aparato de la reivindicación 1, en donde el renderizado implica el manejo como blobs dinámico de objetos en respuesta a la sobrecarga del altavoz, dirigiendo la energía de audio a un mayor número de altavoces de reproducción vecinos mientras se mantiene la energía global constante.
- 45 4. El aparato de la reivindicación 1, en donde el renderizado implica la correlación de posiciones de objetos sonoros a planos de conjuntos de altavoz del entorno de reproducción real.
5. El aparato de cualquiera de las reivindicaciones 1-4, en donde el sistema lógico se configura además para calcular las ganancias de los altavoces correspondientes a la pluralidad de zonas de altavoz.
- 50 6. El aparato de la reivindicación 5, en donde el sistema lógico se configura además para calcular las ganancias de los altavoces para posiciones de objetos sonoros a lo largo de una curva unidimensional entre posiciones de altavoces virtuales.
- 55 7. Un método, que comprende:
- recibir (1457) datos de reproducción de audio que comprenden uno o más objetos sonoros y metadatos asociados;
- 60 recibir (1460) datos de entorno de reproducción que comprenden una indicación de un número de altavoces de reproducción en un entorno de reproducción real y una indicación de la ubicación de cada altavoz de reproducción del entorno de reproducción real tridimensional; y
- renderizar (1470) uno o más objetos sonoros en una o más señales de alimentación de altavoz basándose, al menos en parte, en los metadatos asociados, en donde cada señal de alimentación de altavoz corresponde a al menos uno de los altavoces de reproducción dentro del entorno de reproducción
- 65

real,

el método se caracteriza por que:

5 los metadatos asociados con cada objeto sonoro incluyen metadatos de restricción de zona de altavoz que indican si renderizar el objeto sonoro respectivo implica imponer restricciones de zona de altavoz, y

10 en donde renderizar el uno o más objetos sonoros incluye renderizar el respectivo objeto sonoro al imponer restricciones de zona de altavoz en respuesta a los metadatos de restricción de zona de altavoz, y

15 en donde renderizar el uno o más objetos sonoros incluye además renderizar el respectivo objeto sonoro ya sea al aplicar reglas de panoramización para correlacionar el objeto sonoro a una pluralidad de altavoces de reproducción, o al correlacionar el objeto sonoro a un único altavoz de reproducción.

8. Un medio no transitorio que tiene software almacenado, el software incluye instrucciones que, cuando son ejecutadas por un ordenador, hacen que el ordenador lleve a cabo las siguientes operaciones:

20 recibir (1457) datos de reproducción de audio que comprenden uno o más objetos sonoros y metadatos asociados;

25 recibir (1460) datos de entorno de reproducción que comprenden una indicación de un número de altavoces de reproducción en un entorno de reproducción real y una indicación de la ubicación de cada altavoz de reproducción del entorno de reproducción real tridimensional; y

30 renderizar (1470) uno o más objetos sonoros en una o más señales de alimentación de altavoz basándose, al menos en parte, en los metadatos asociados, en donde cada señal de alimentación de altavoz corresponde a al menos uno de los altavoces de reproducción dentro del entorno de reproducción real,

caracterizado porque:

35 los metadatos asociados con cada objeto sonoro incluyen metadatos de restricción de zona de altavoz que indican si renderizar el objeto sonoro respectivo implica imponer restricciones de zona de altavoz, y

40 en donde renderizar el uno o más objetos sonoros incluye renderizar el respectivo objeto sonoro al imponer restricciones de zona de altavoz en respuesta a los metadatos de restricción de zona de altavoz, y

en donde renderizar el uno o más objetos sonoros incluye además renderizar el respectivo objeto sonoro ya sea al aplicar reglas de panoramización para correlacionar el objeto sonoro a una pluralidad de altavoces de reproducción, o al correlacionar el objeto sonoro a un único altavoz de reproducción.

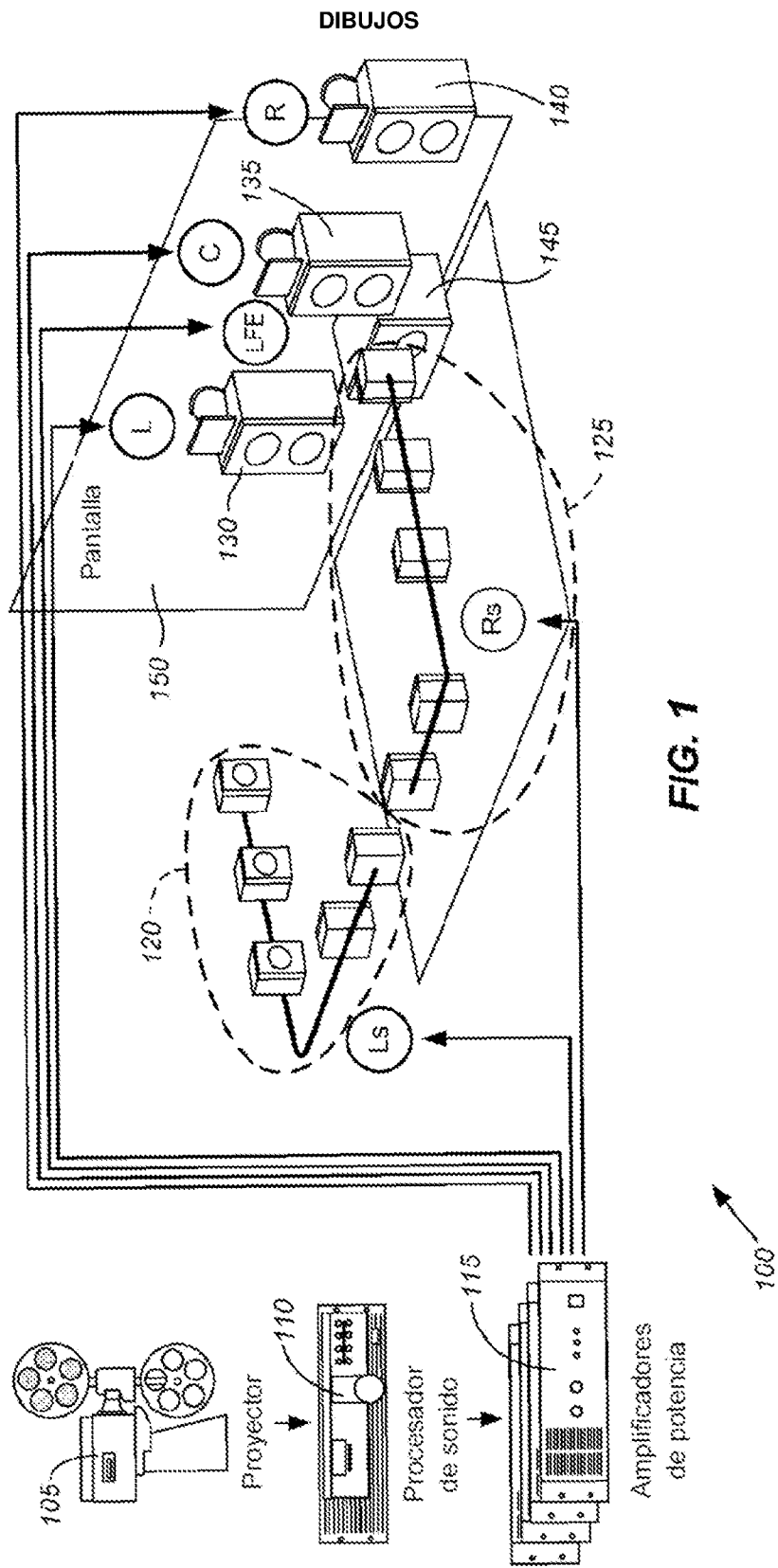


FIG. 1

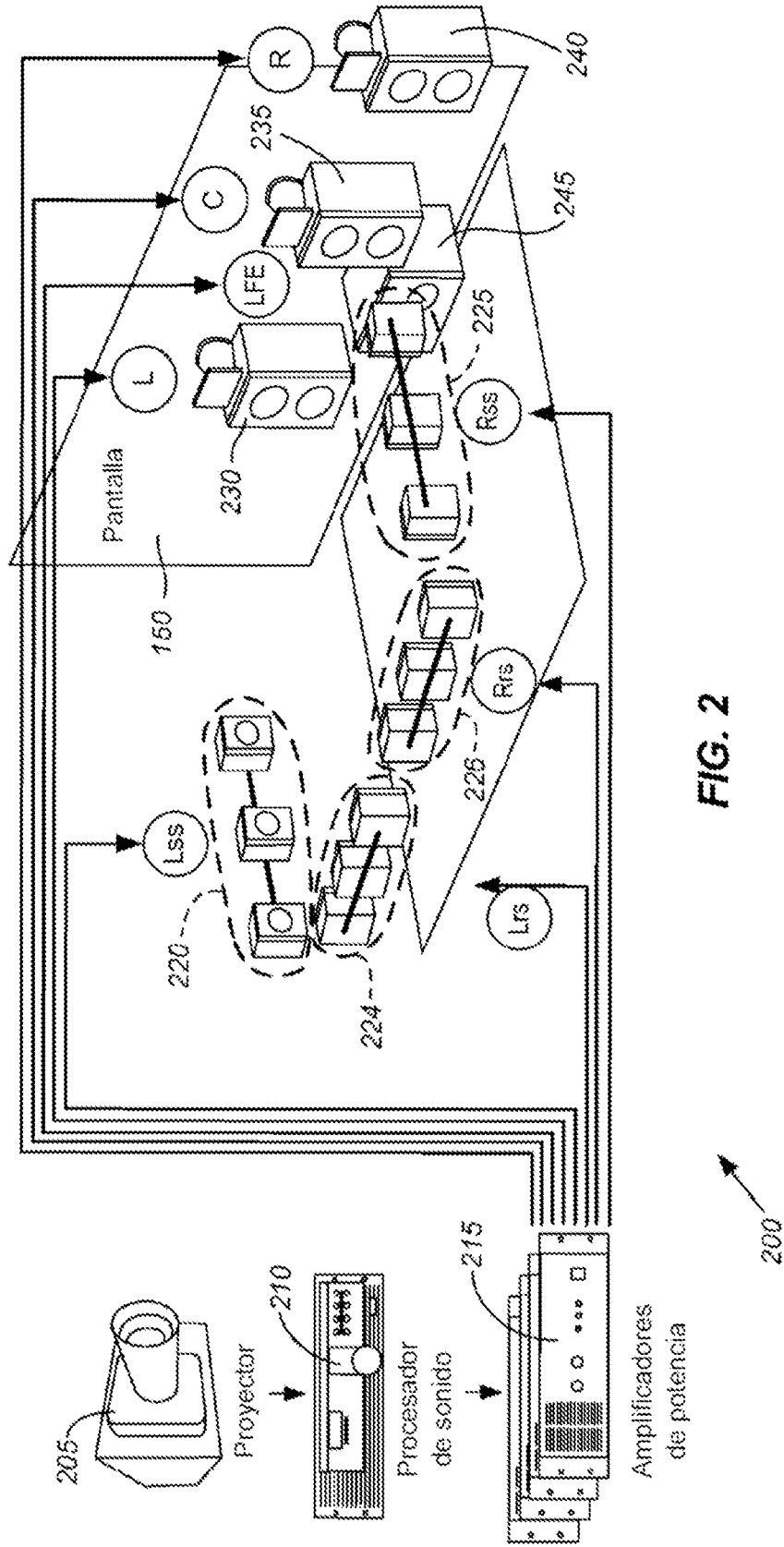


FIG. 2

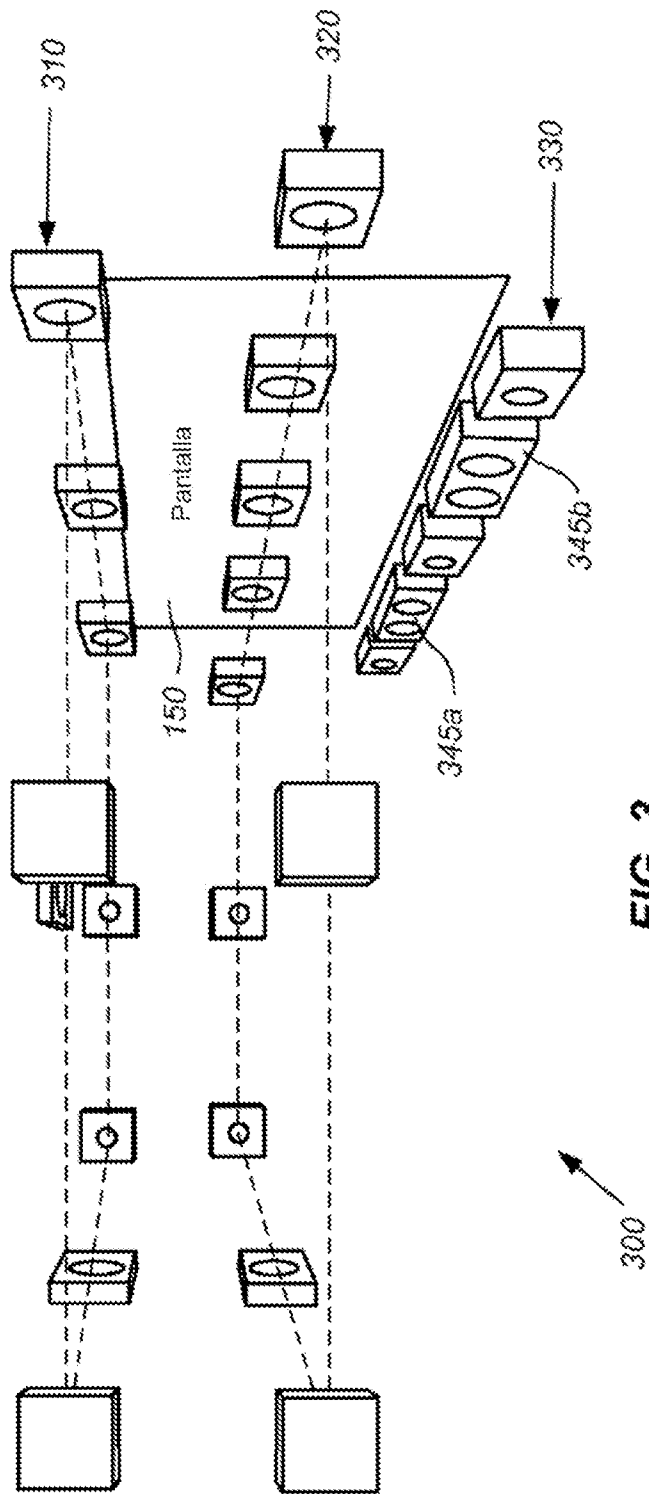


FIG. 3

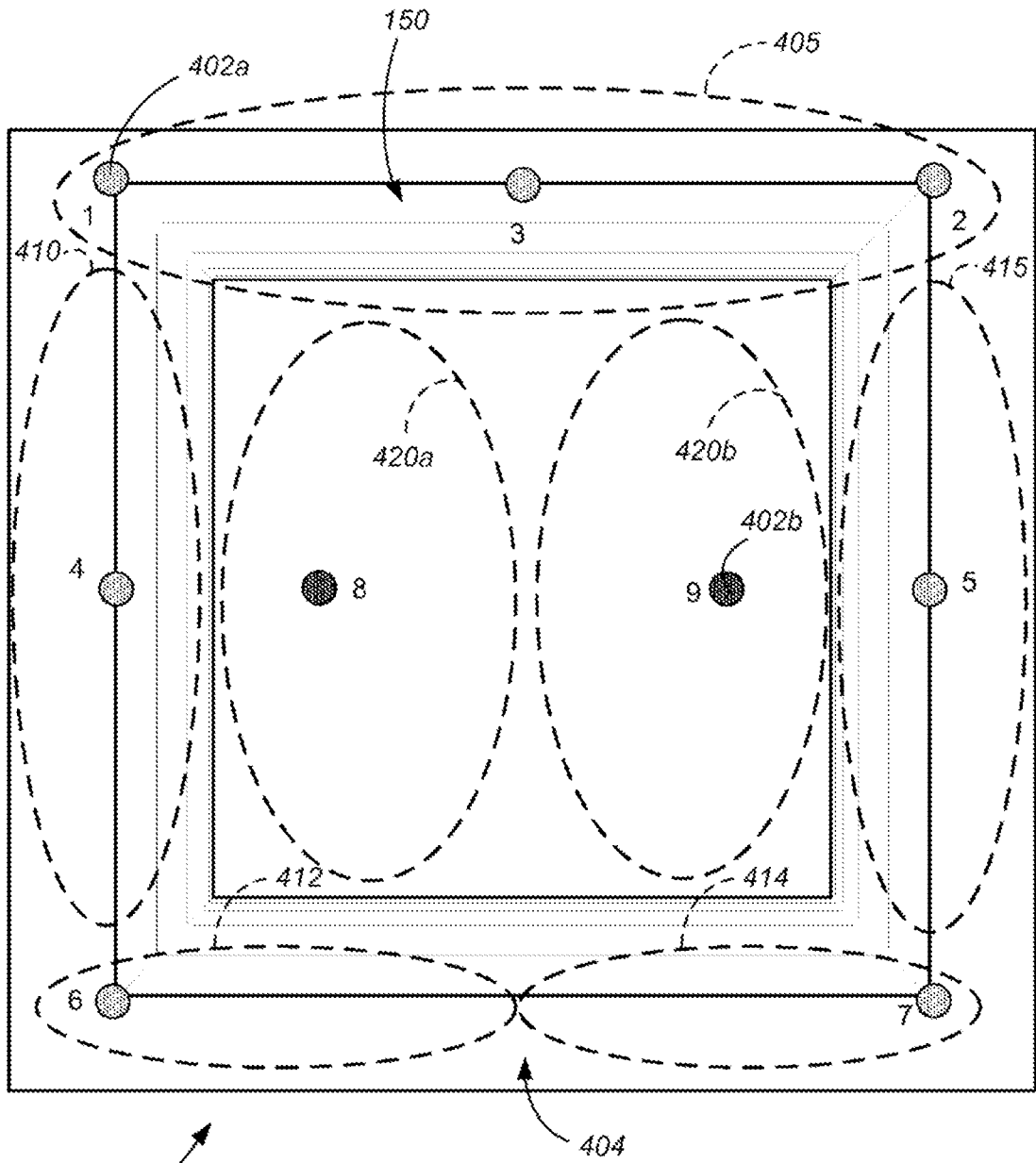


FIG. 4A

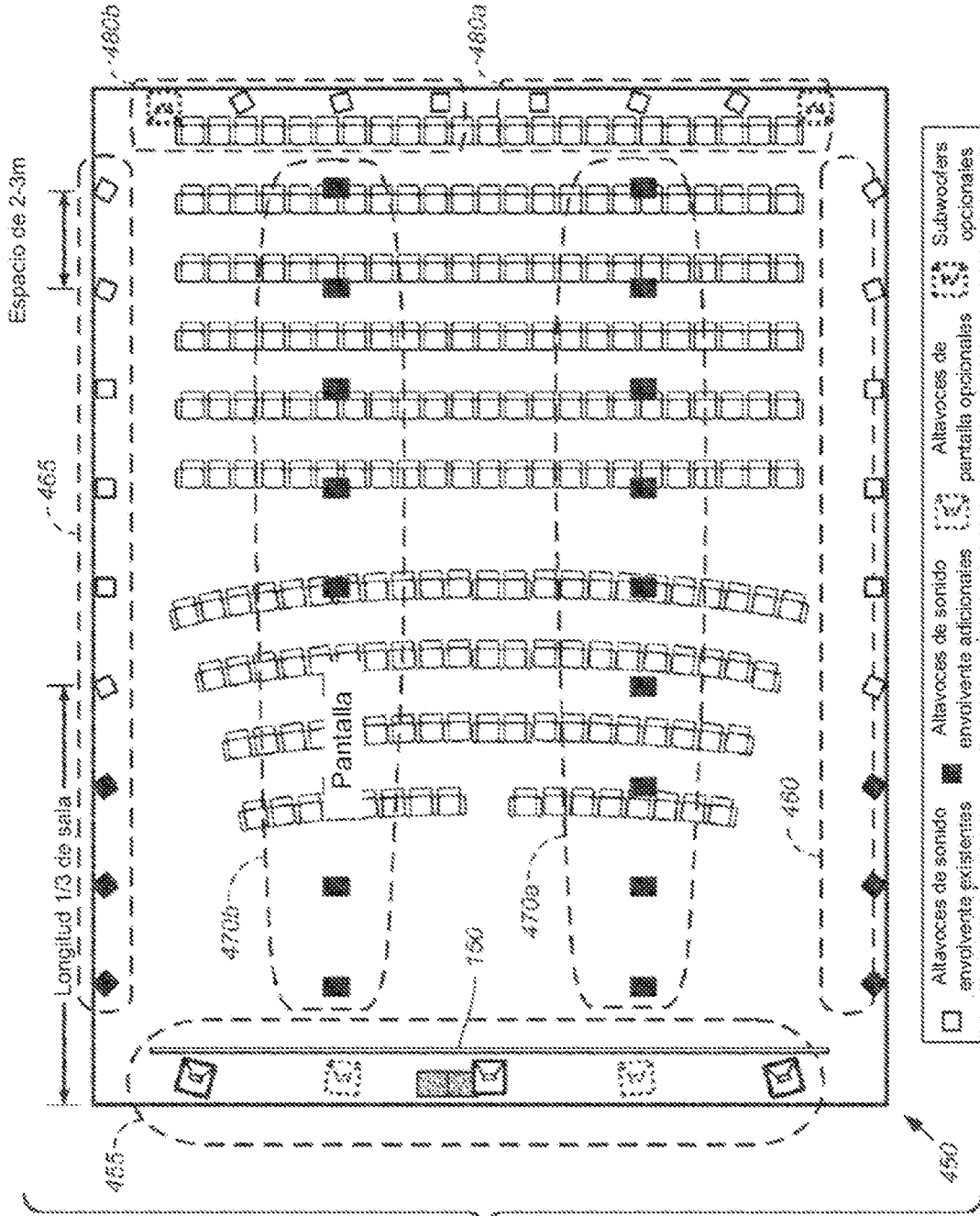
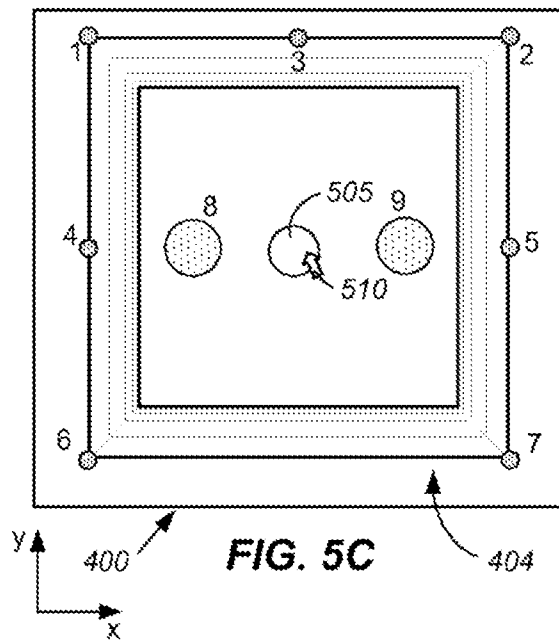
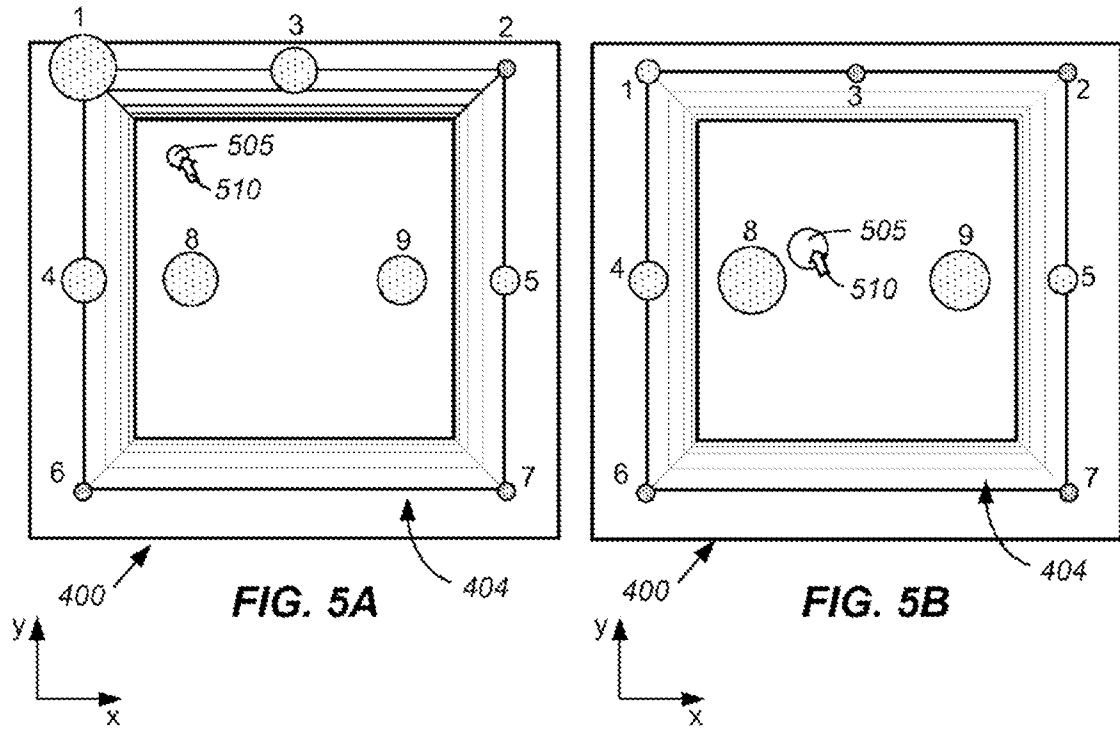
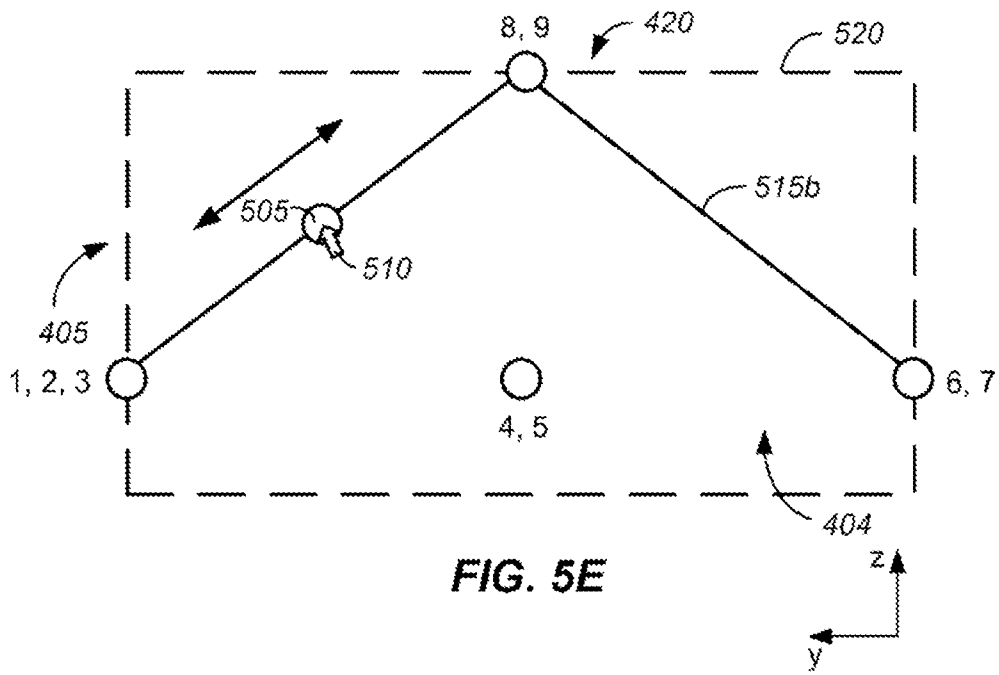
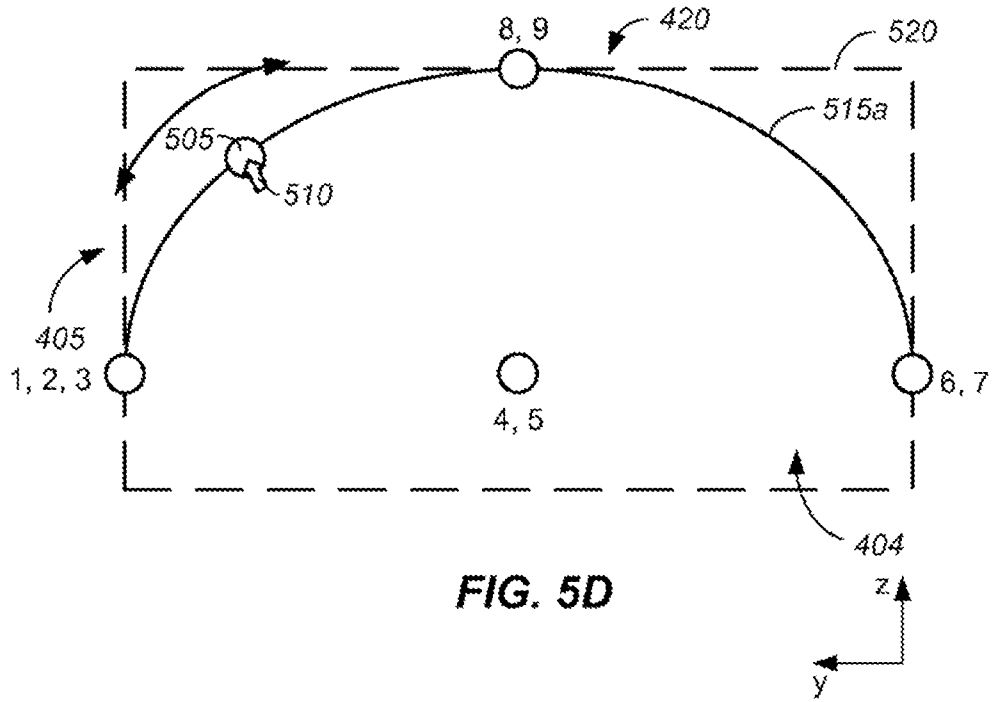
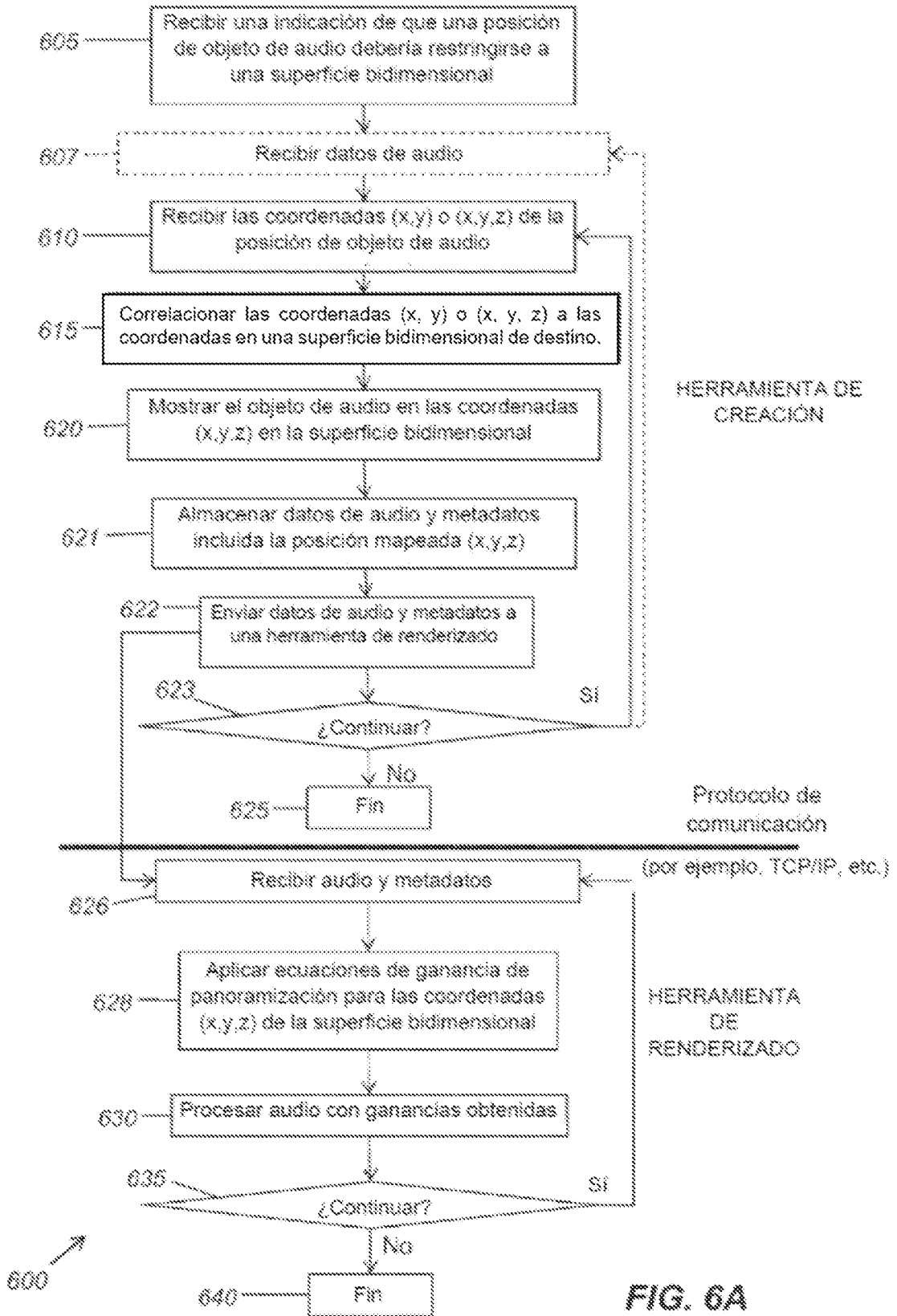


FIG. 4B







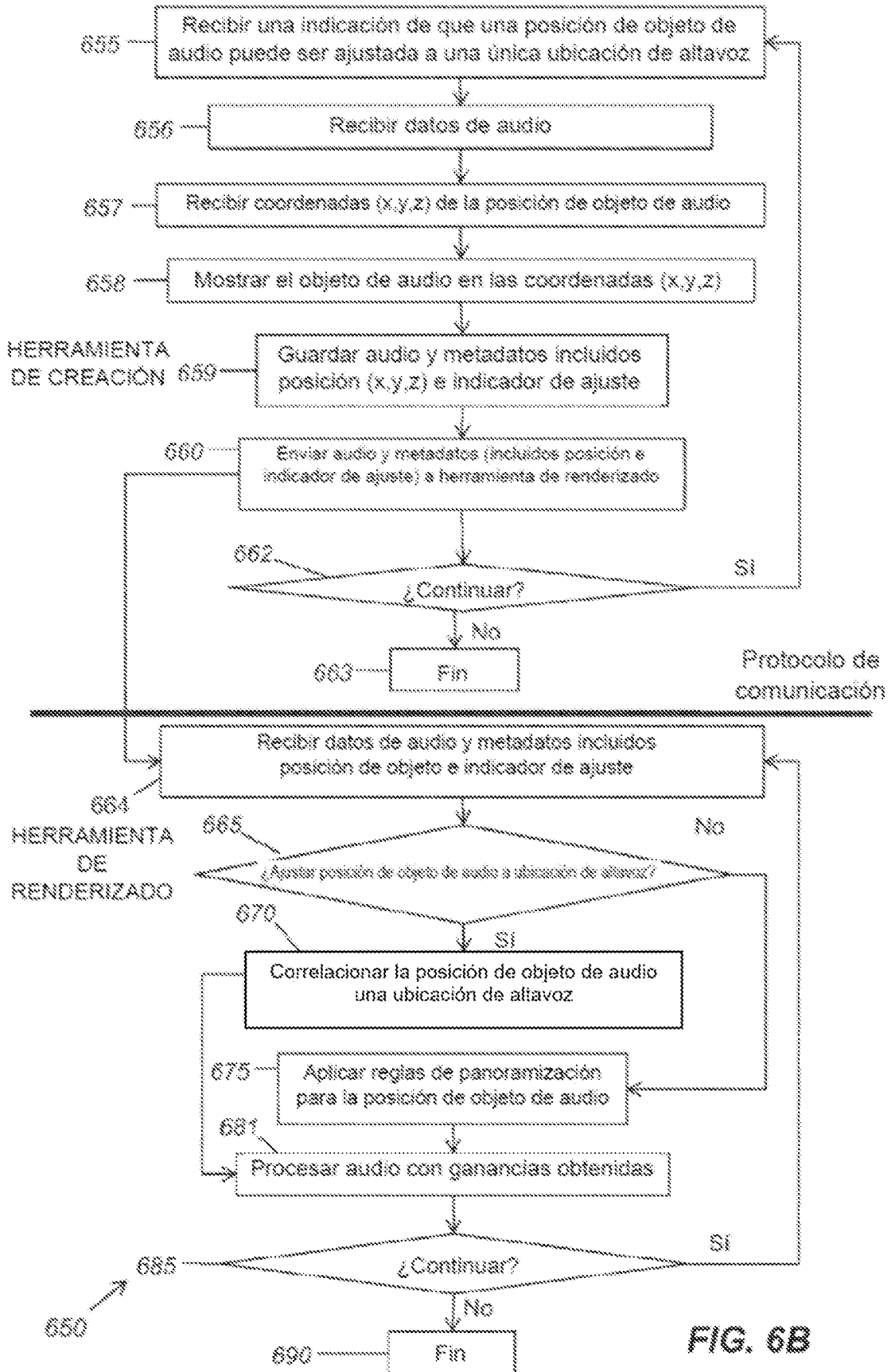


FIG. 6B

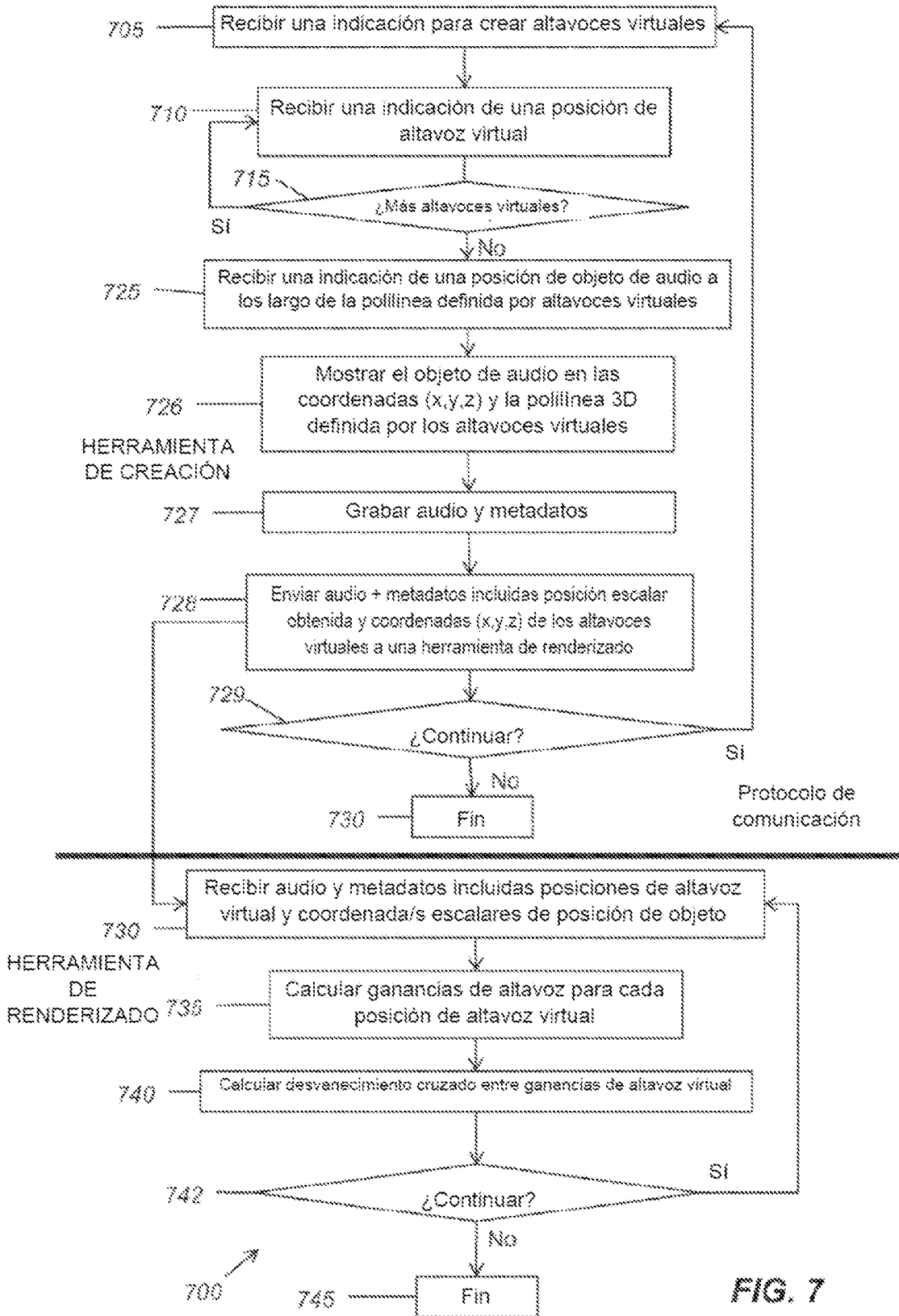
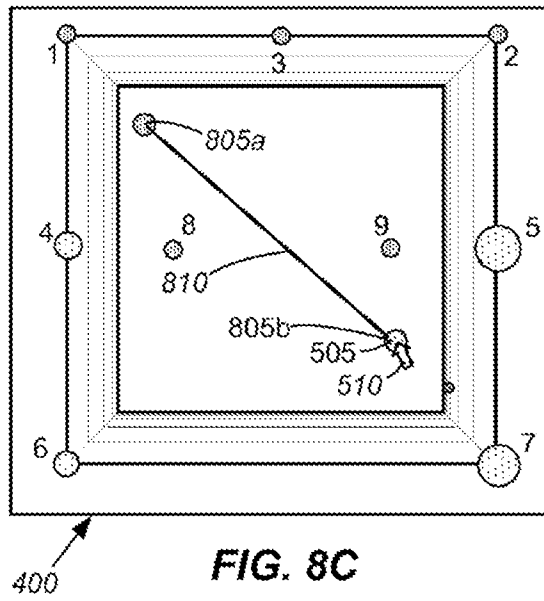
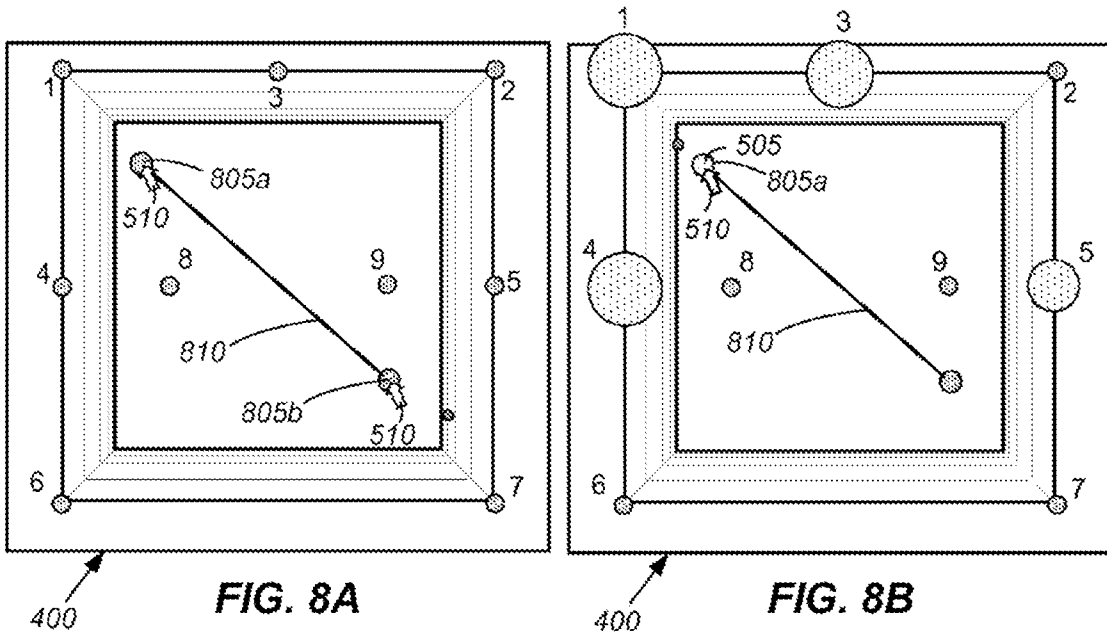


FIG. 7



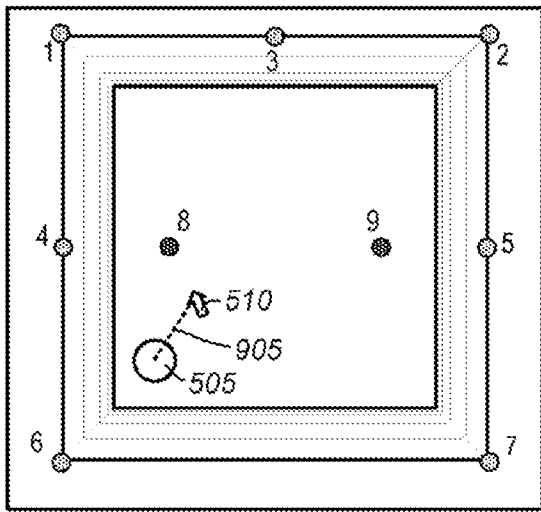


FIG. 9A

400

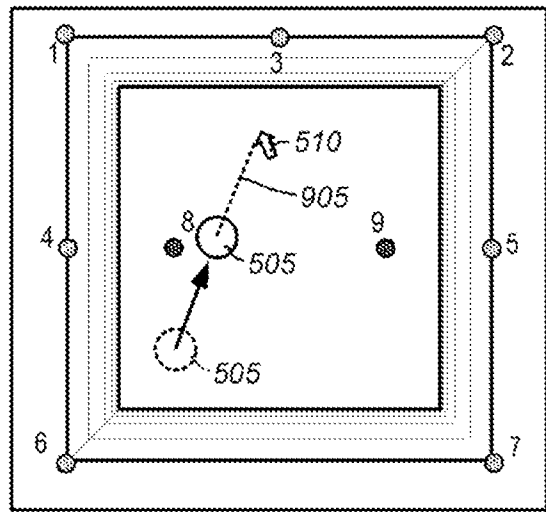


FIG. 9B

400

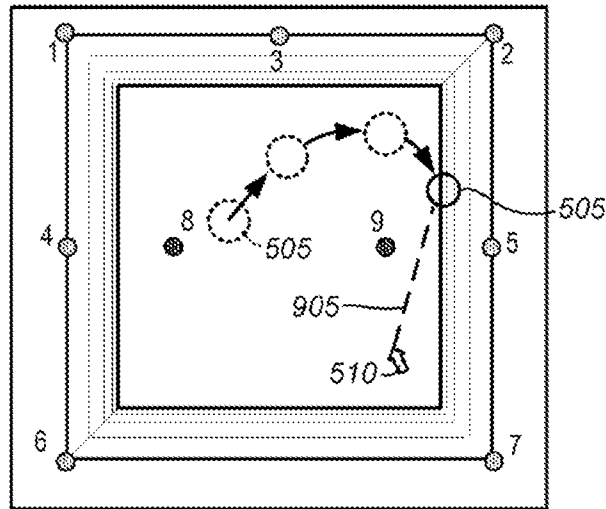


FIG. 9C

400

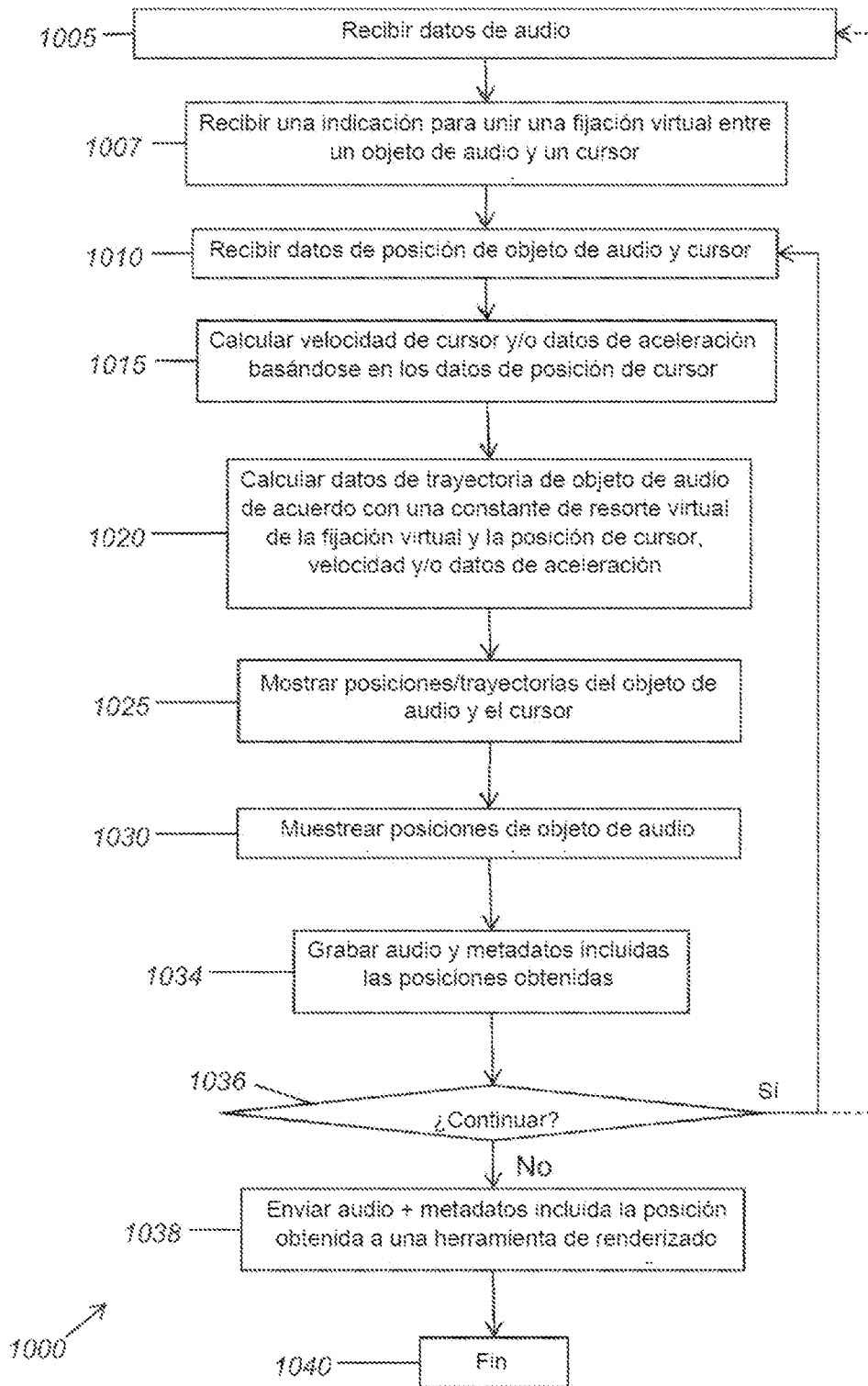


FIG. 10A

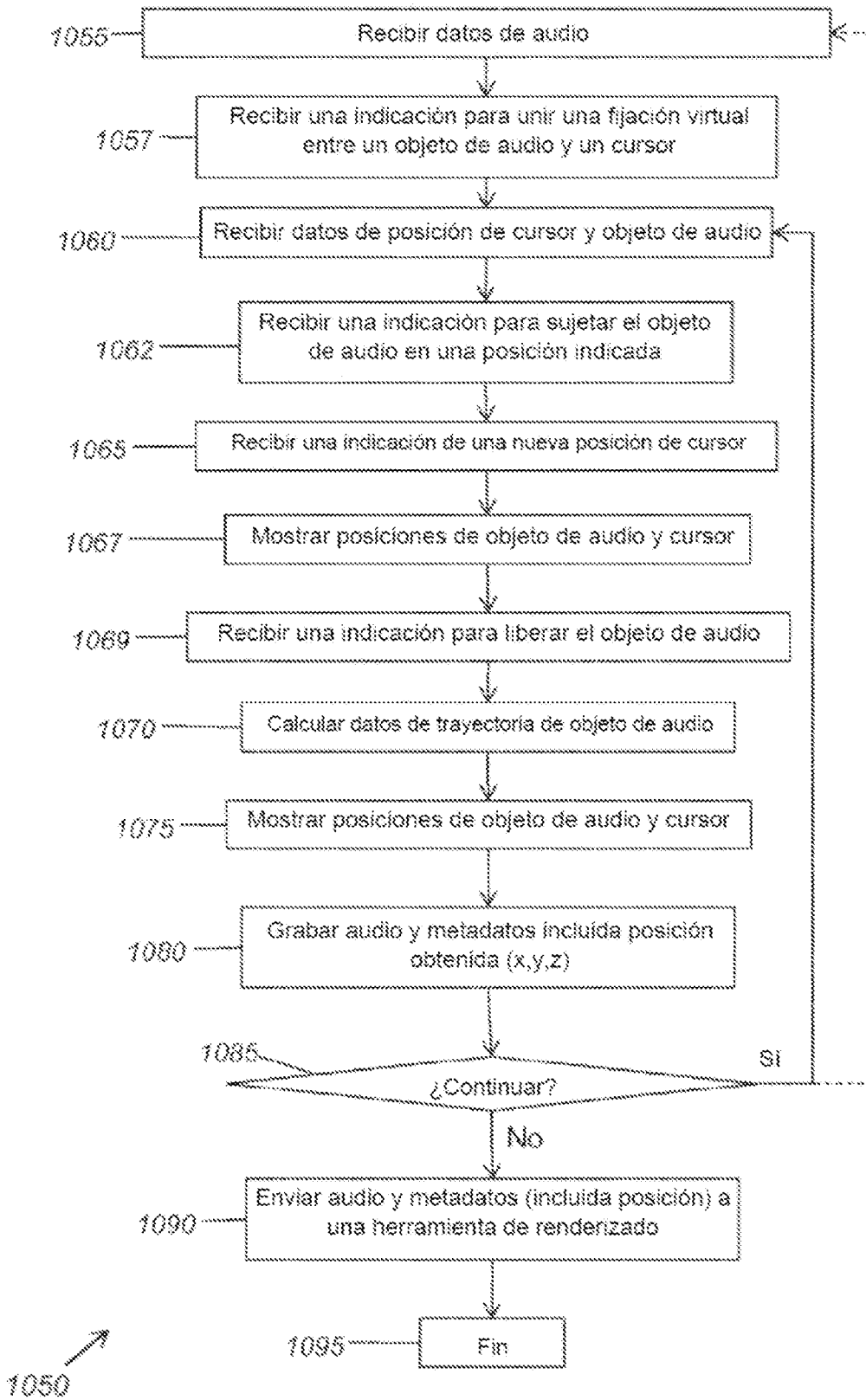
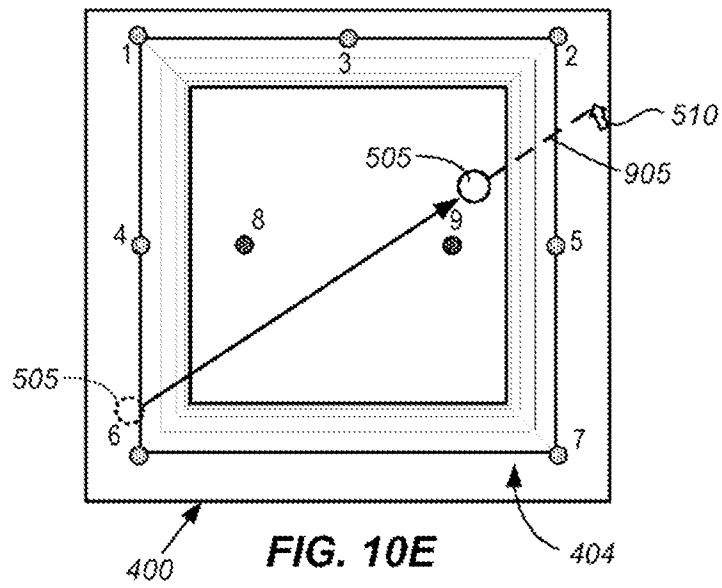
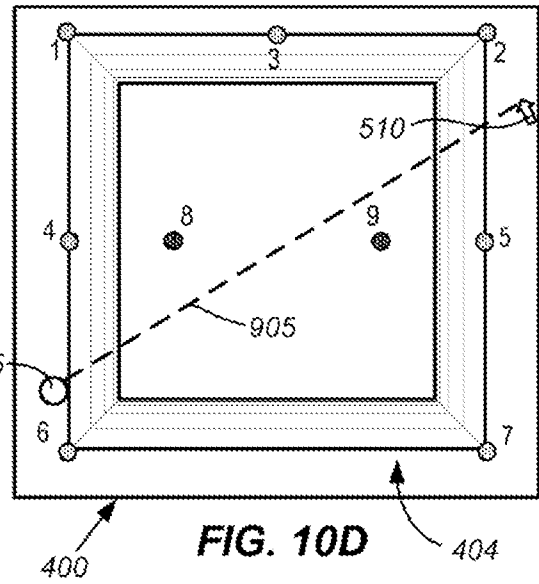
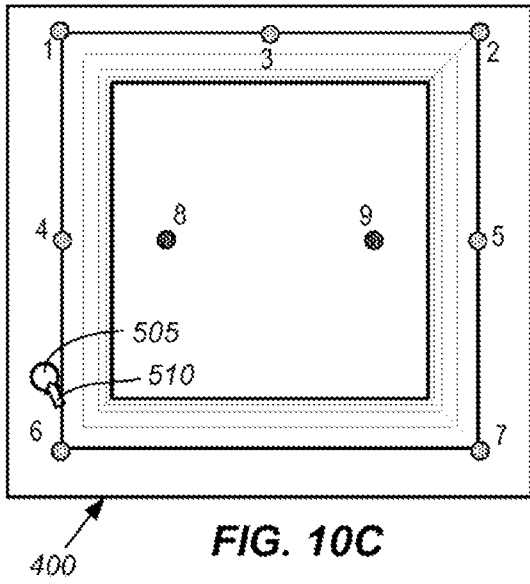


FIG. 10B



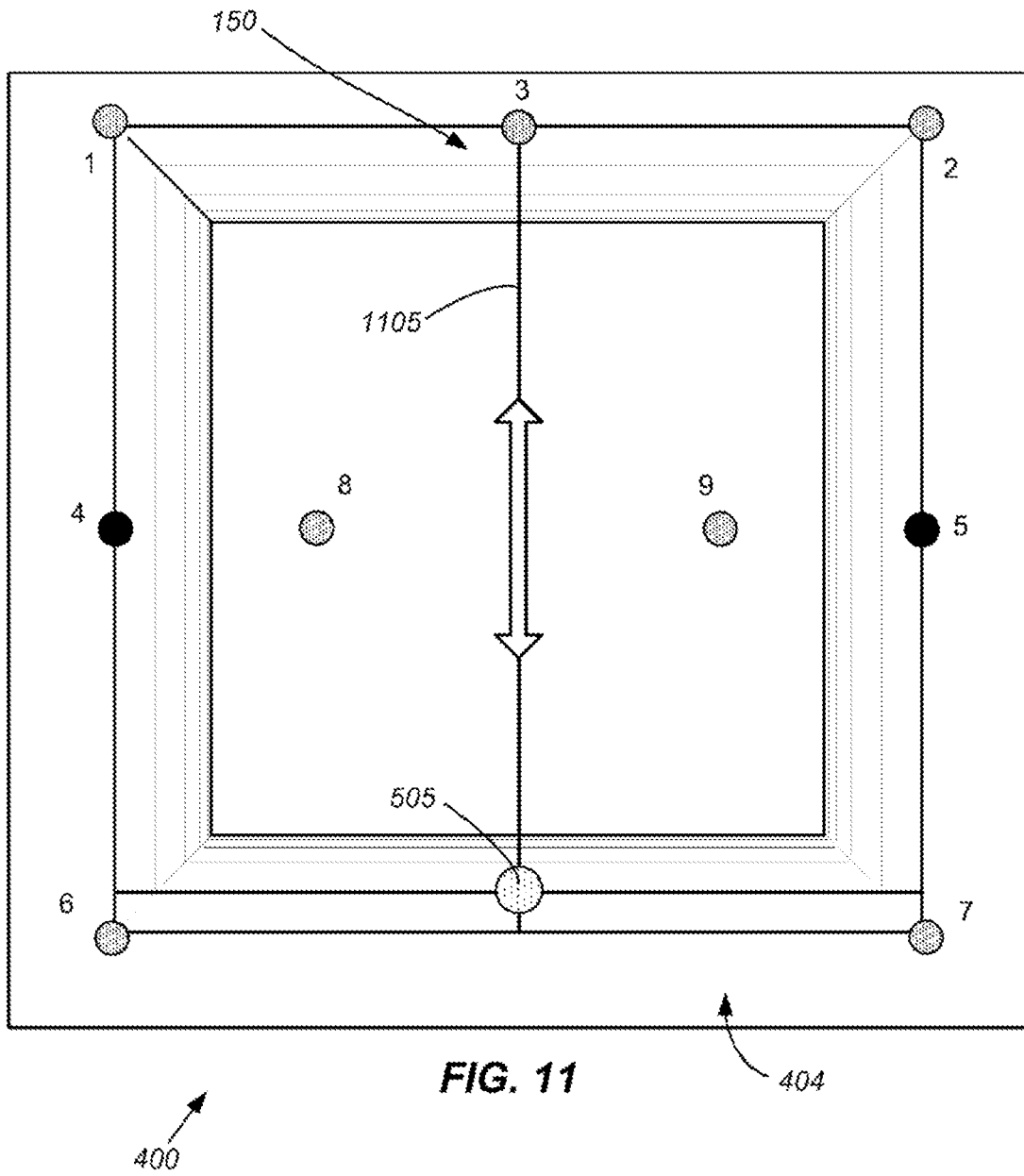


FIG. 11

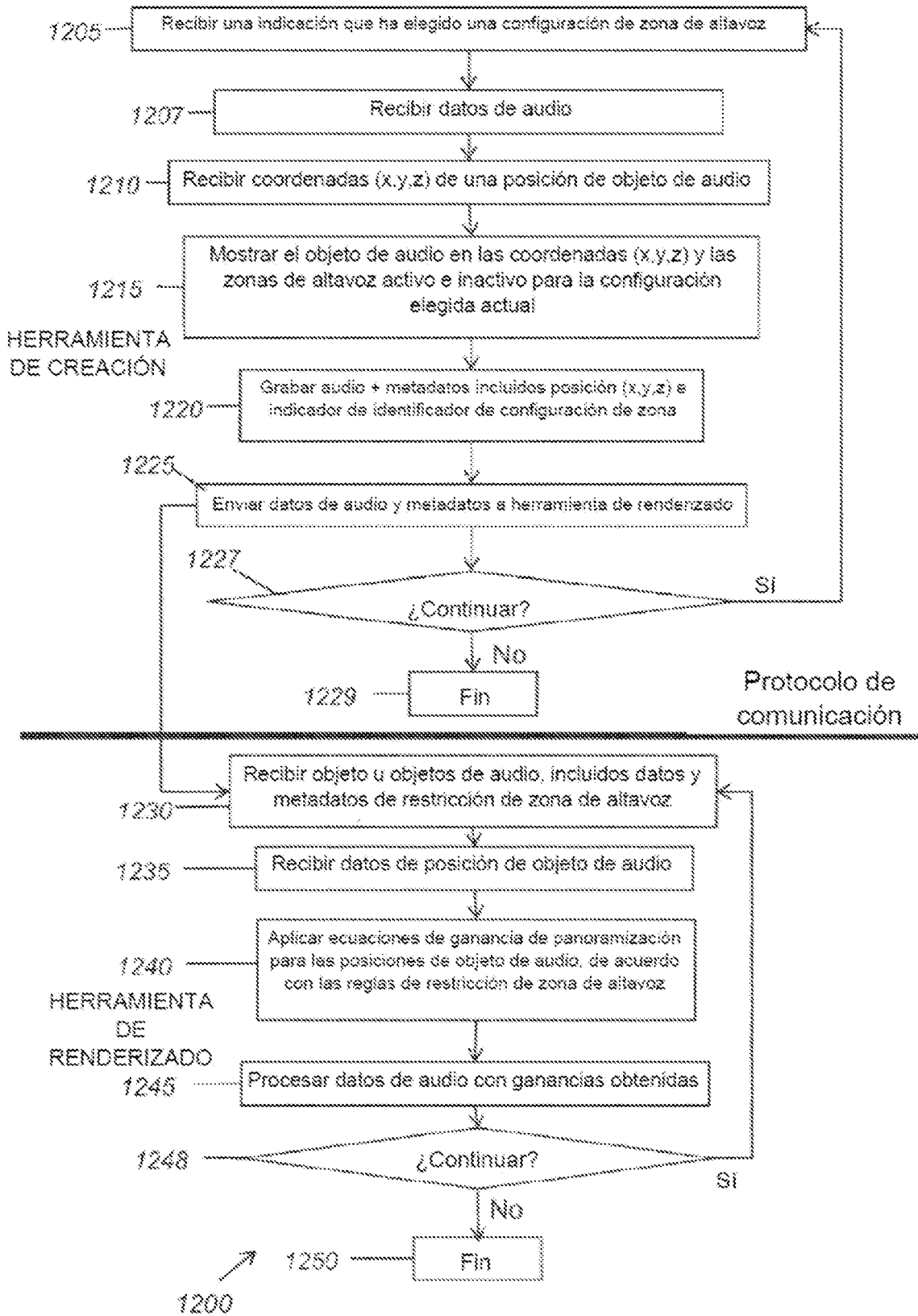
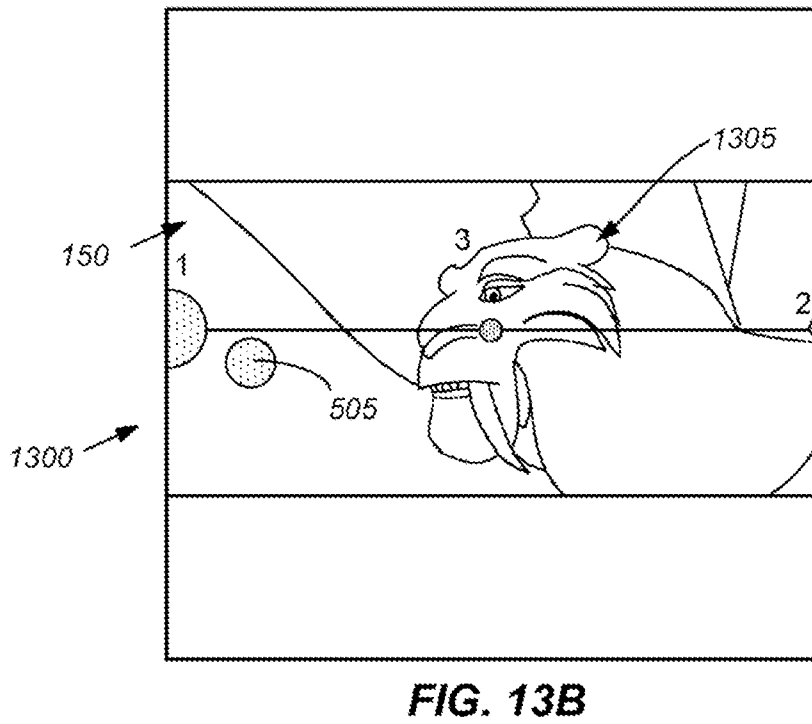
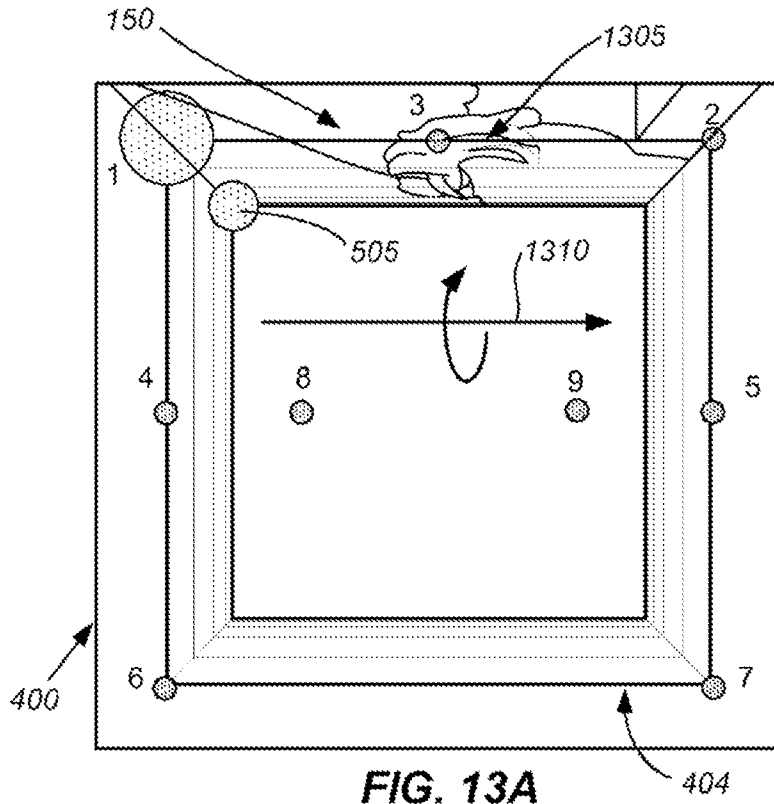


FIG. 12



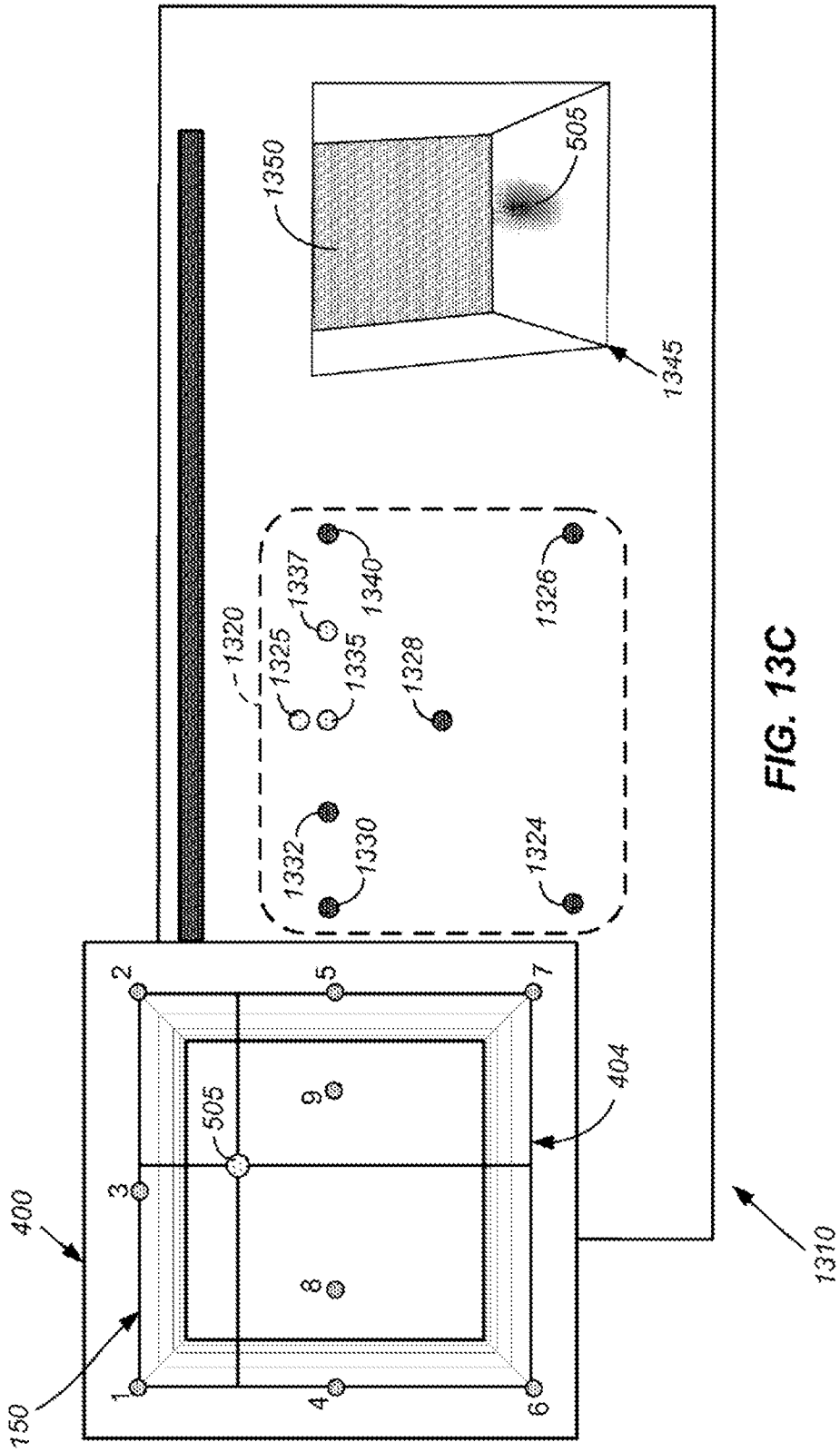
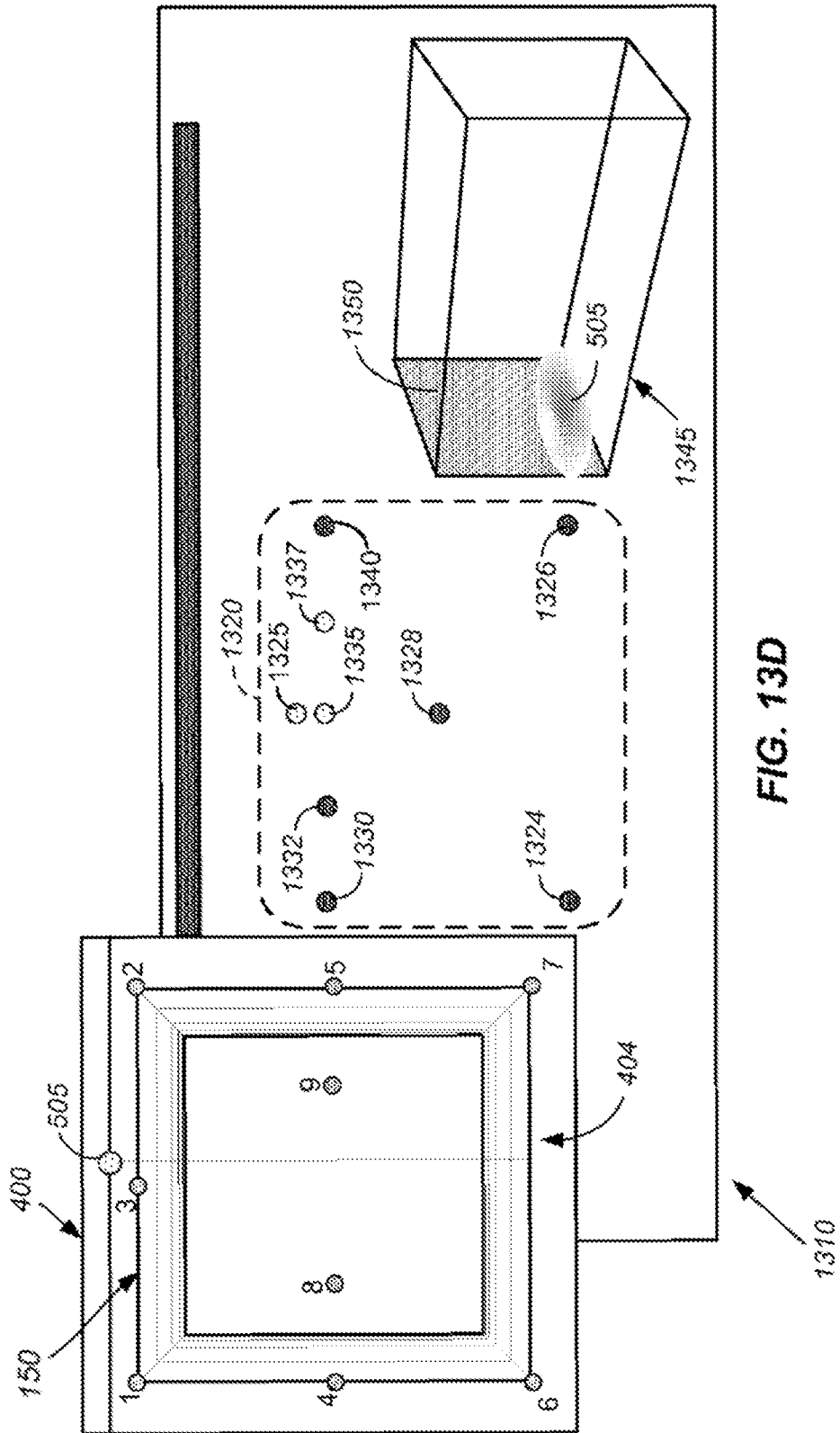
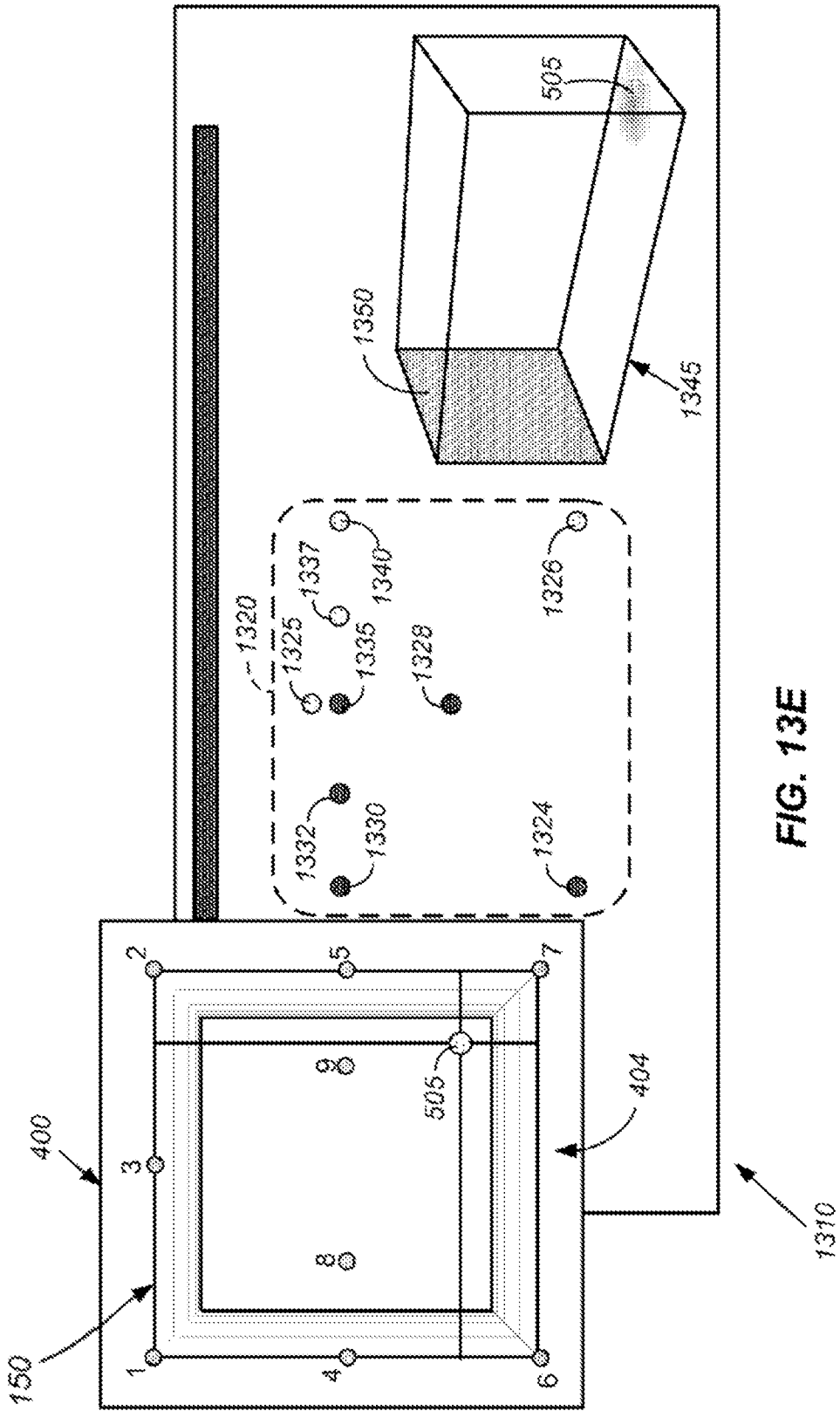
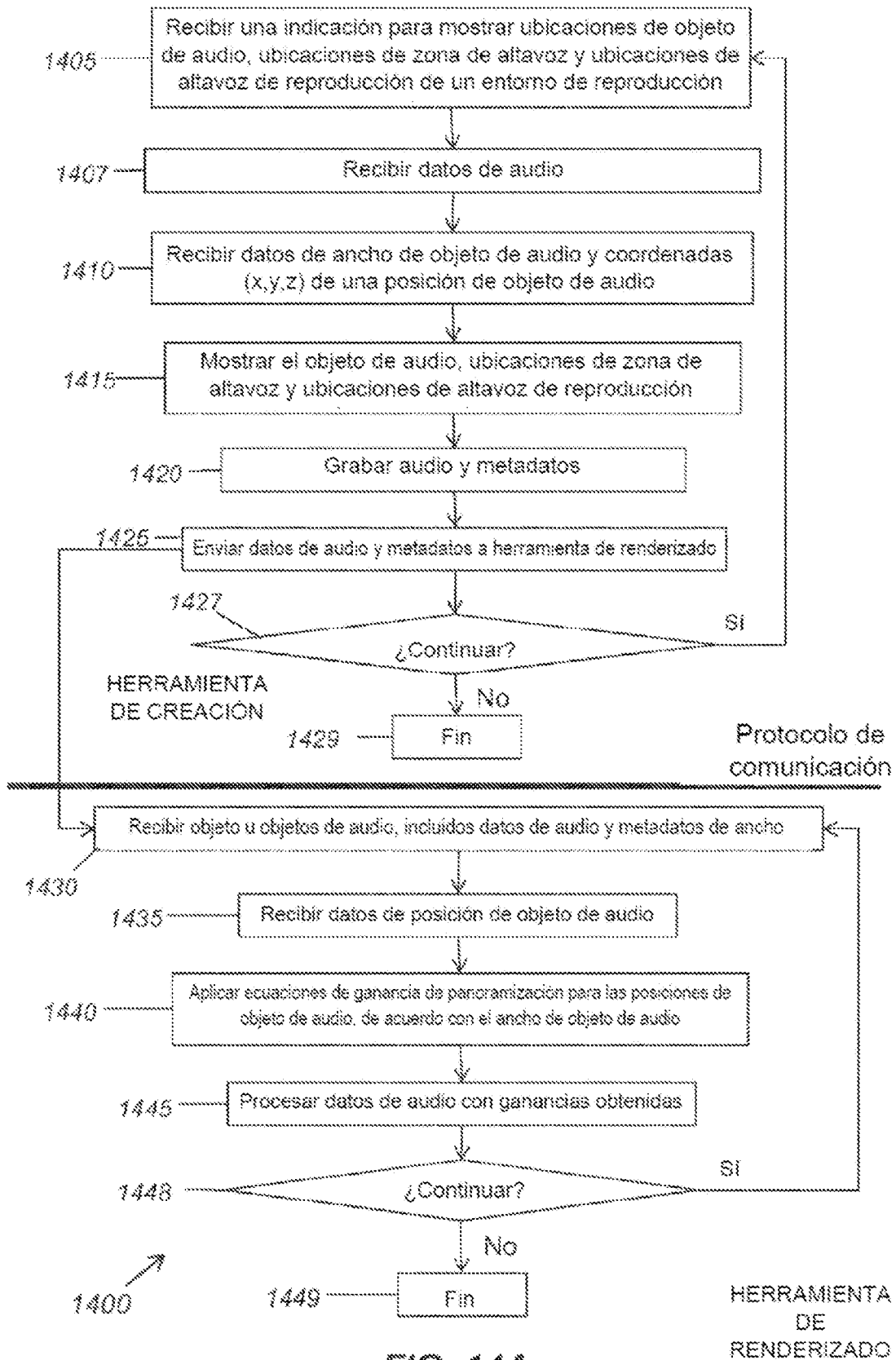


FIG. 13C







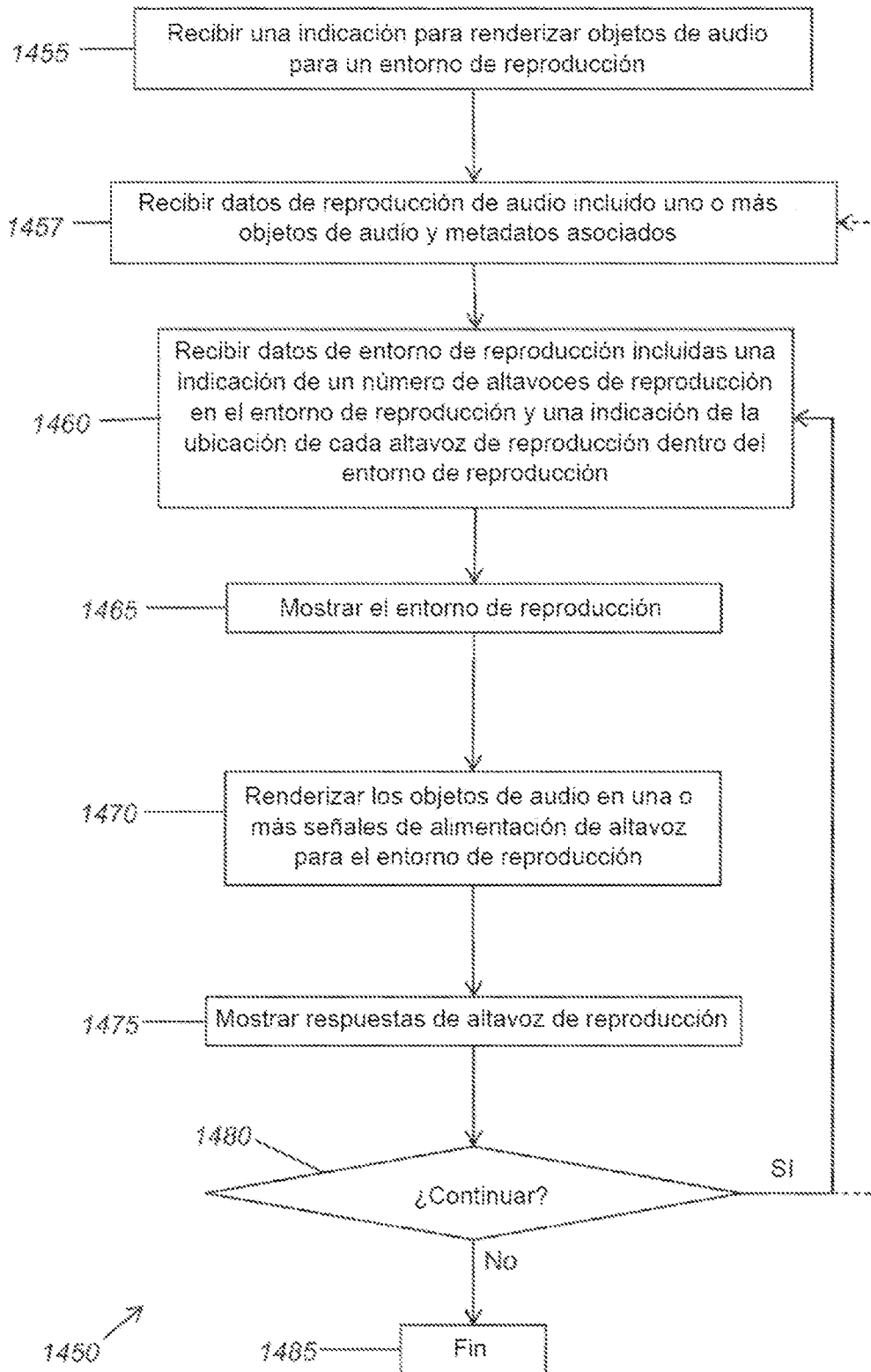


FIG. 14B

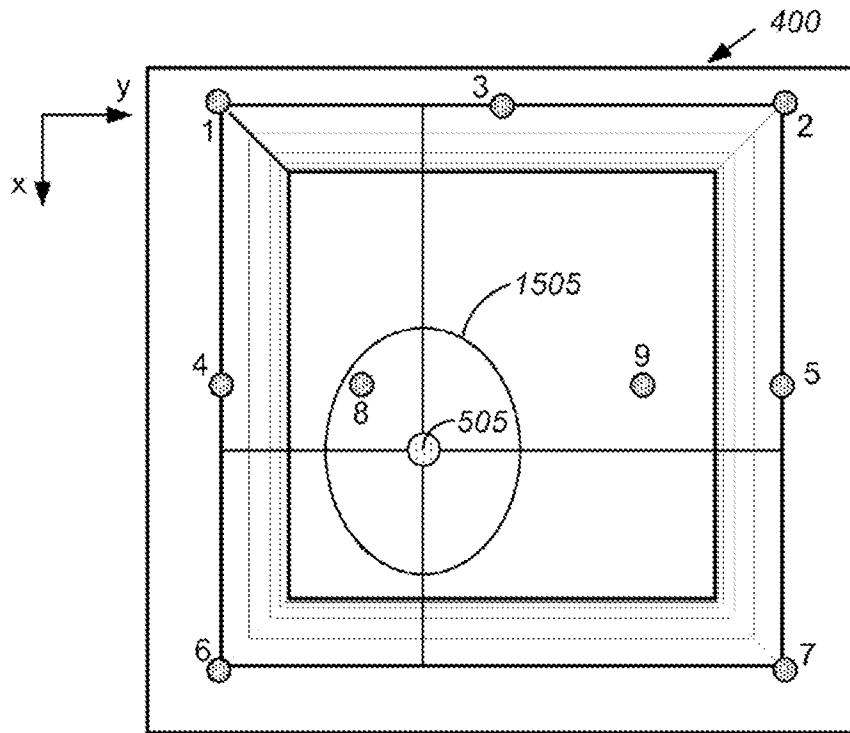


FIG. 15A

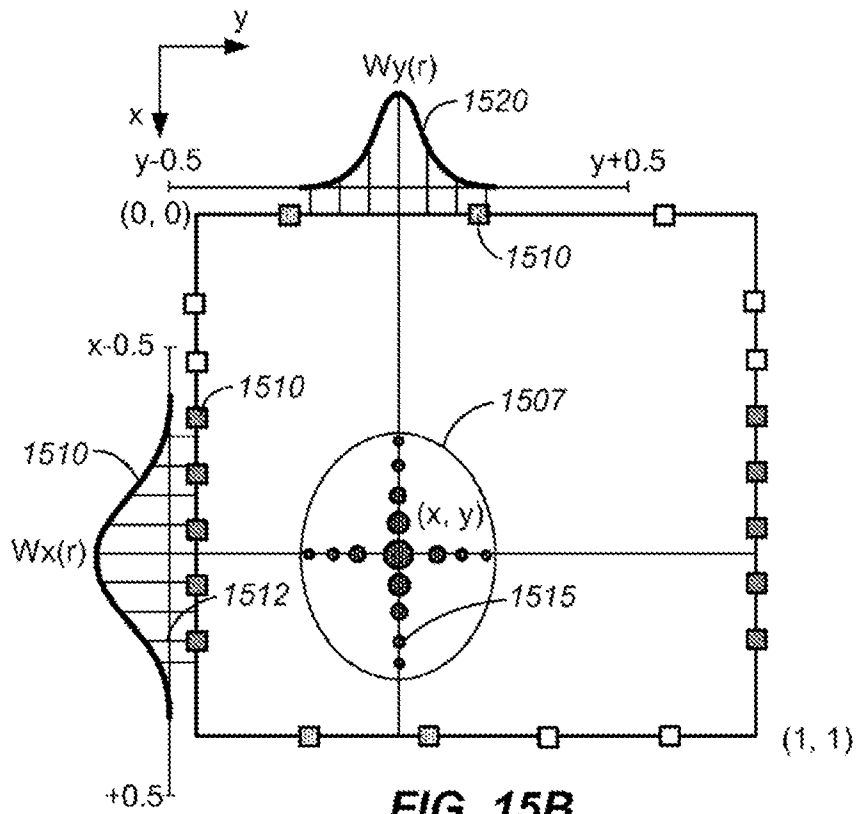


FIG. 15B

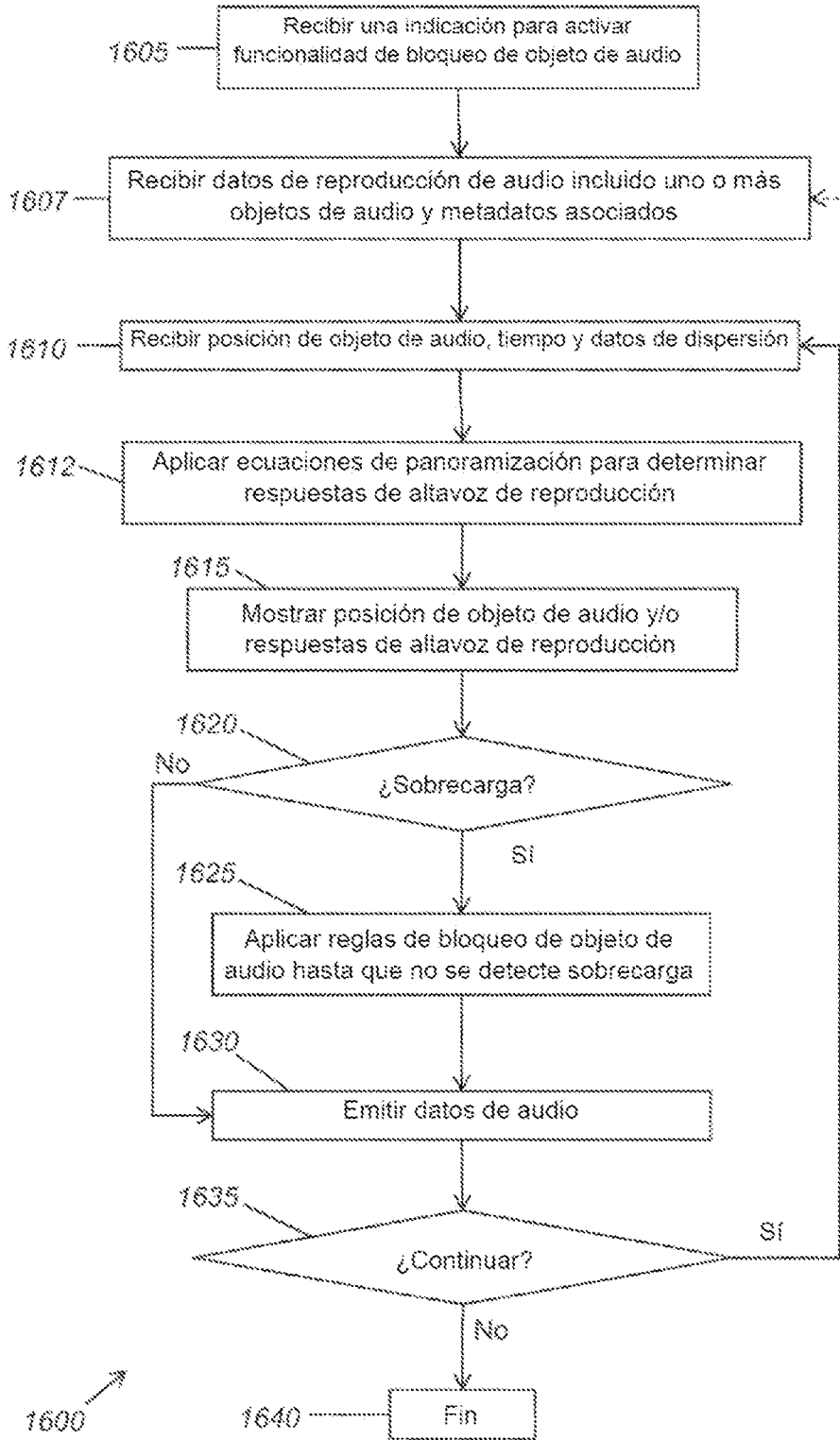
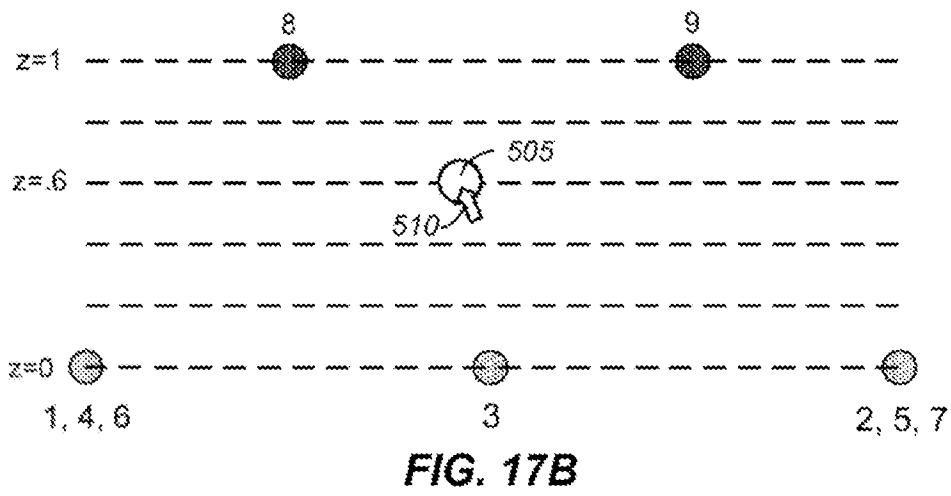
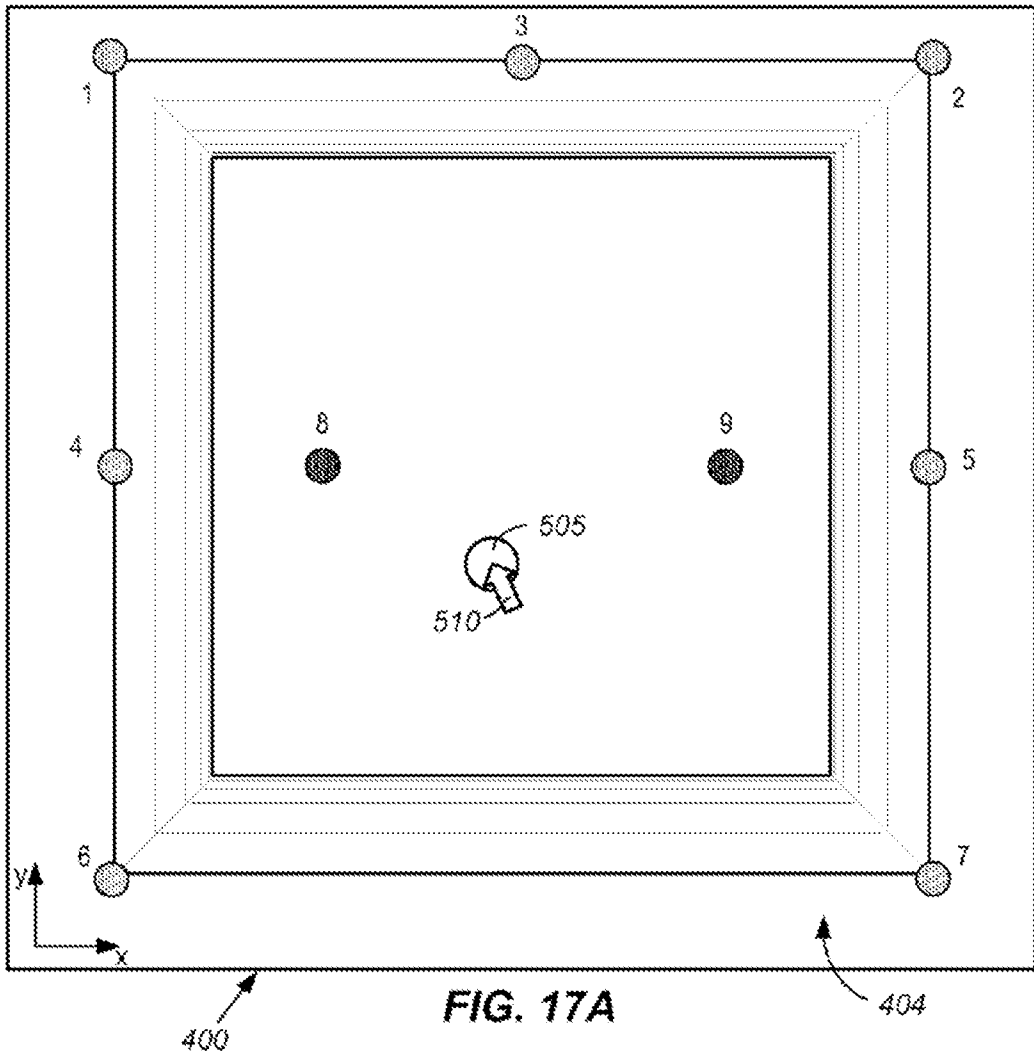


FIG. 16



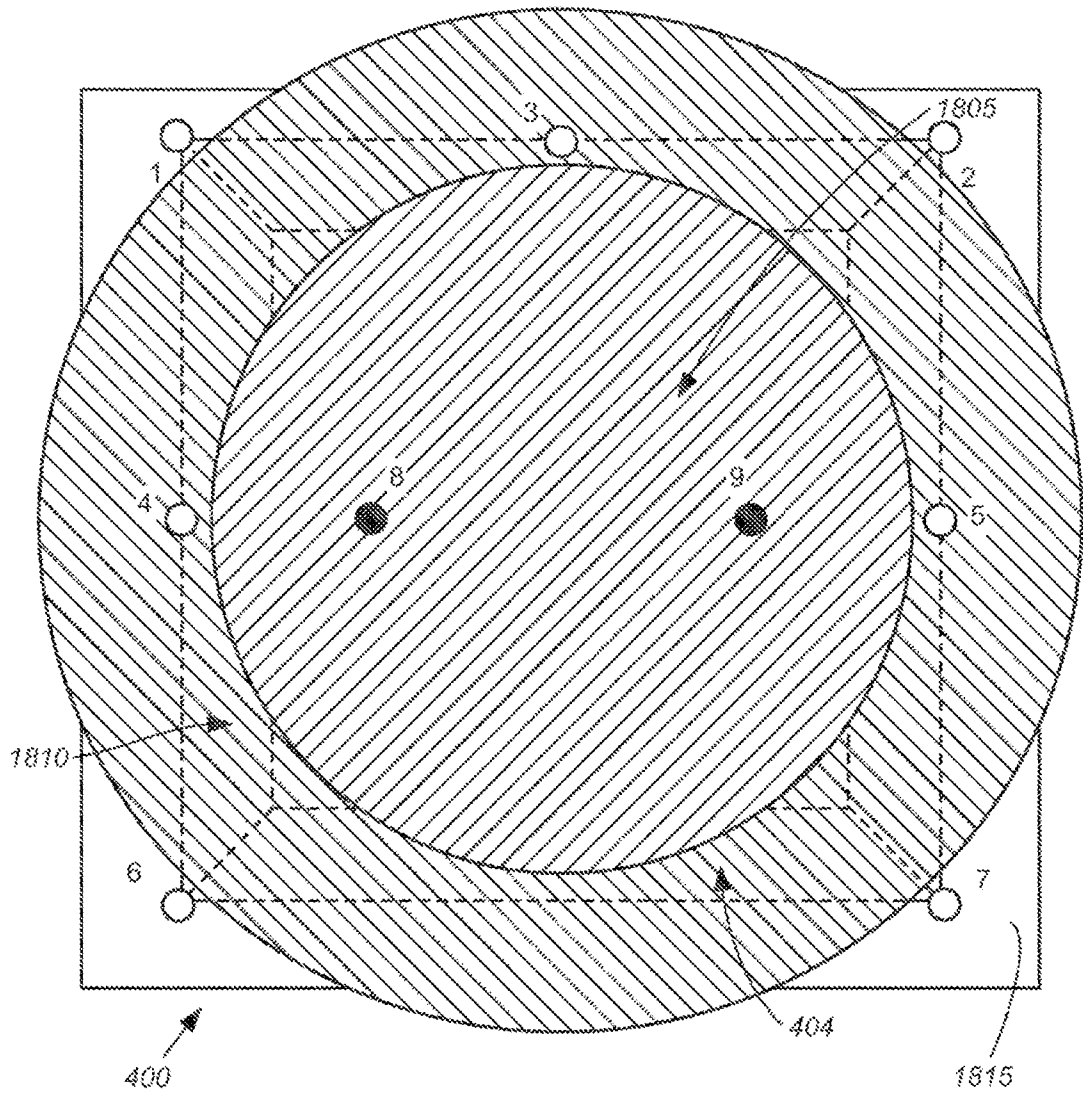


FIG. 18

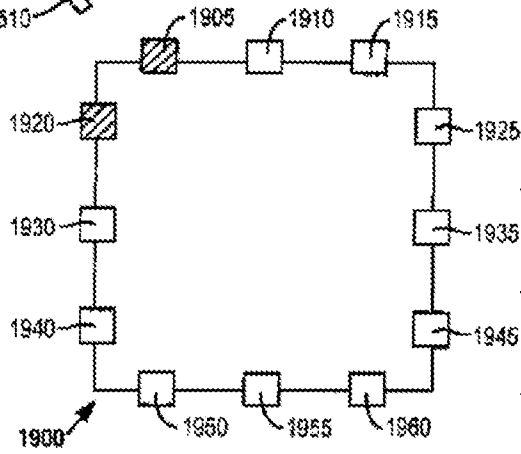
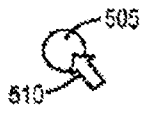


Figura 19A

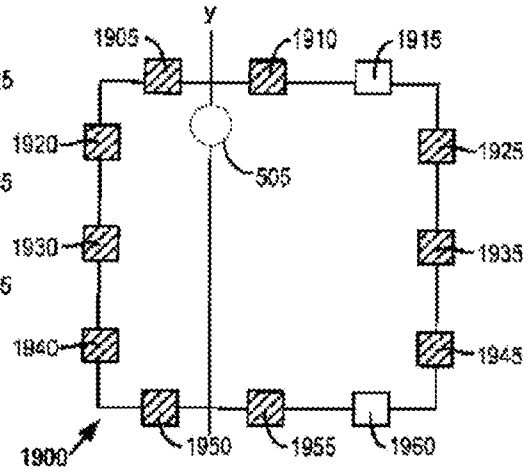


Figura 19B

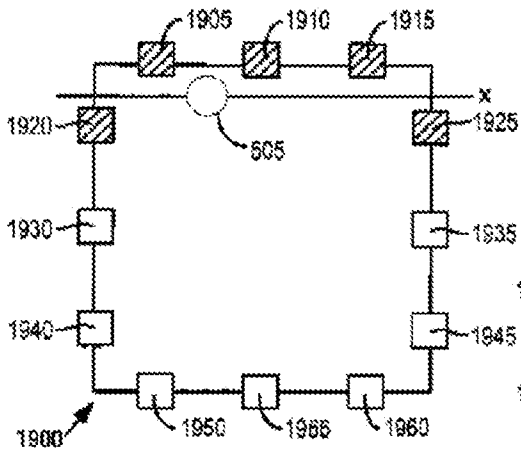


Figura 19C

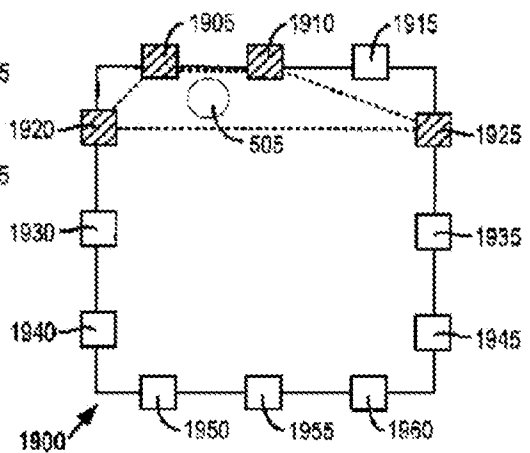


Figura 19D

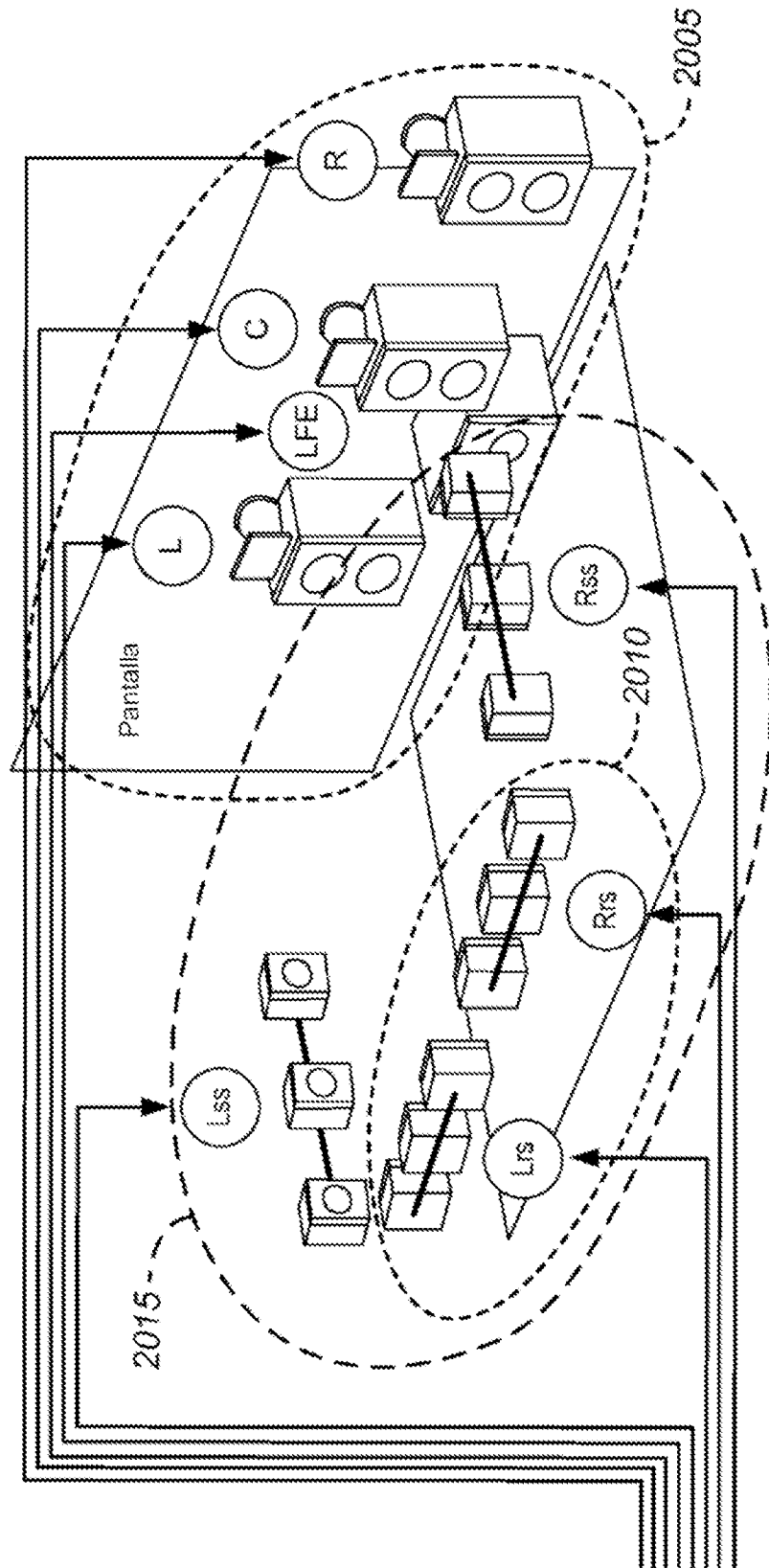


FIG. 20

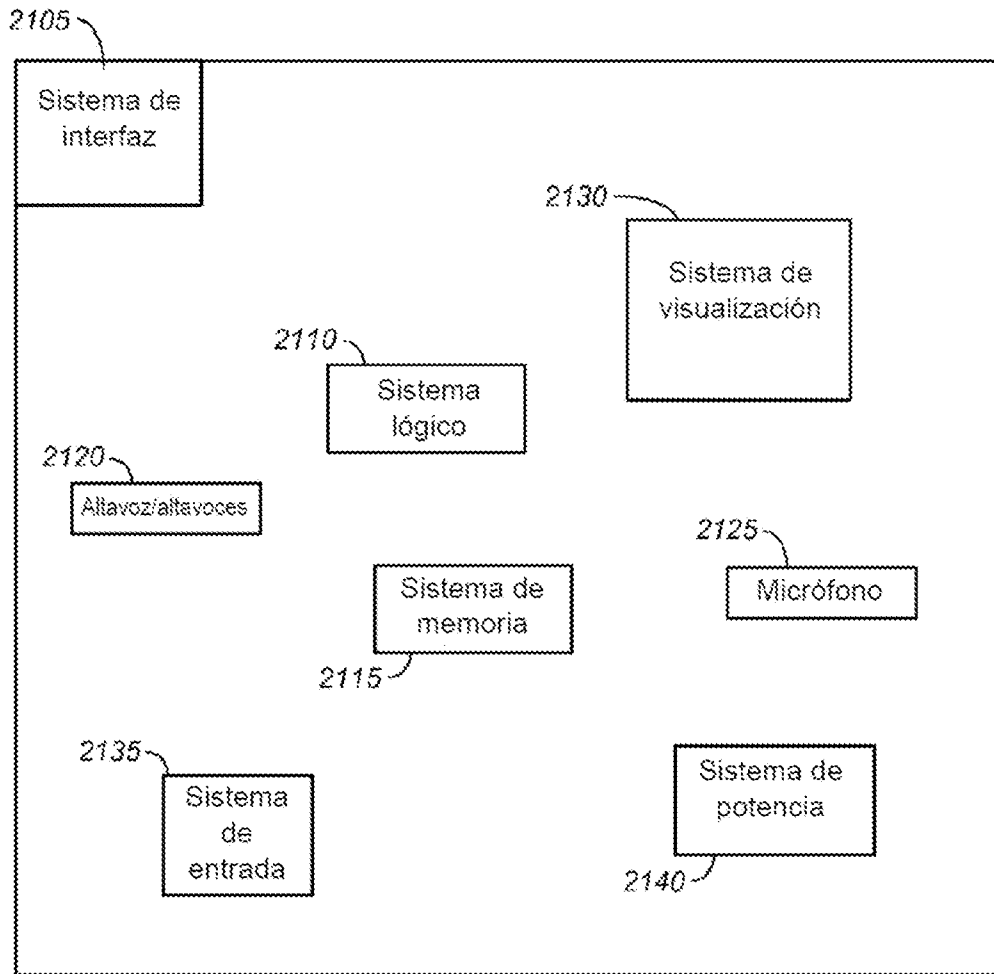


FIG. 21

2100 ↗

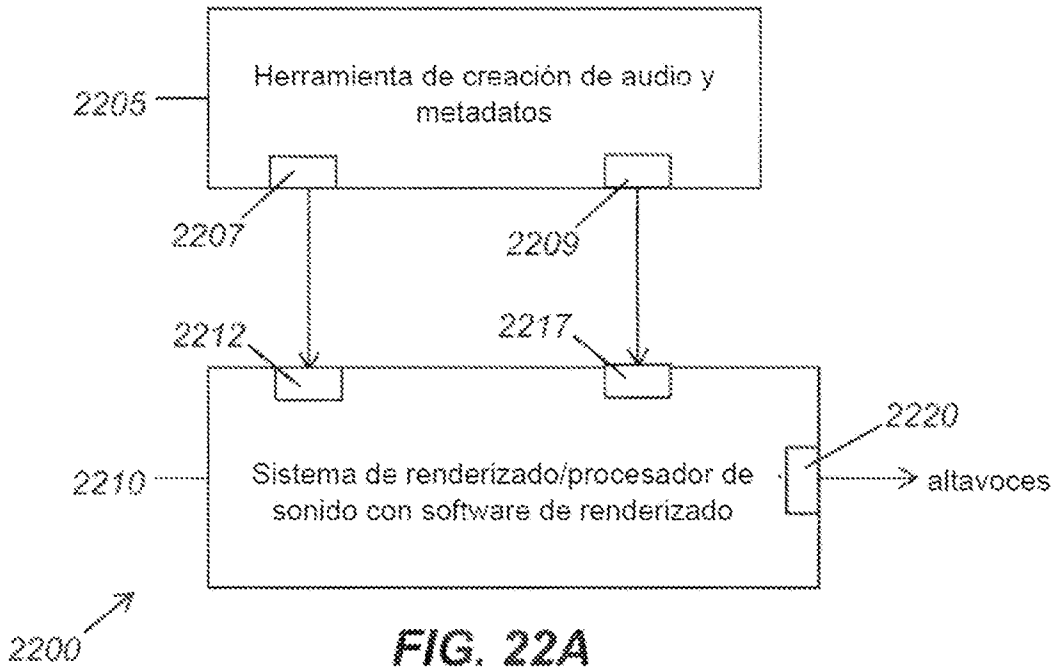


FIG. 22A

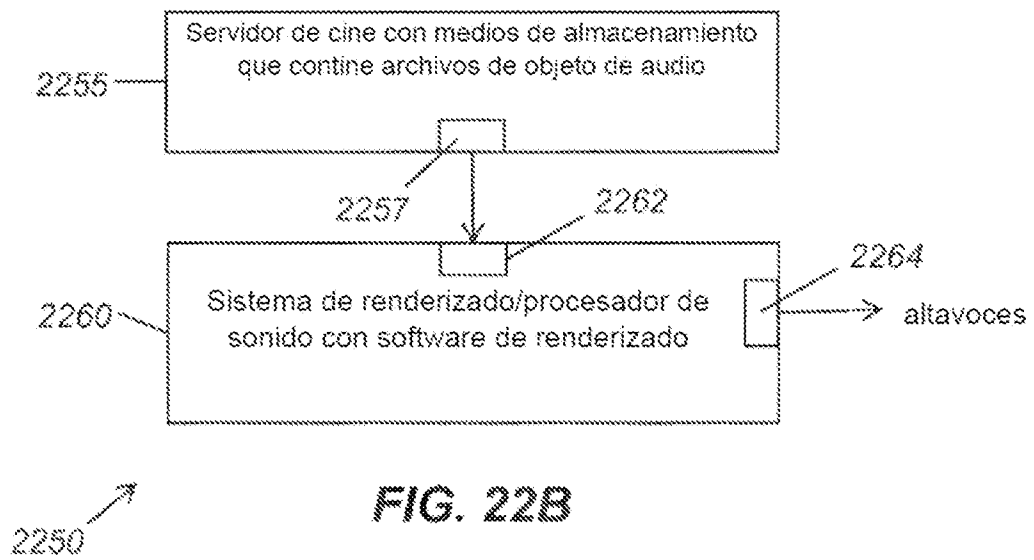


FIG. 22B