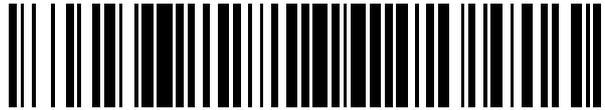


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 714 905**

51 Int. Cl.:

H04S 7/00 (2006.01)

H03G 3/00 (2006.01)

H03G 3/30 (2006.01)

G10L 19/008 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **27.02.2015 PCT/US2015/018045**

87 Fecha y número de publicación internacional: **03.09.2015 WO15131063**

96 Fecha de presentación y número de la solicitud europea: **27.02.2015 E 15754660 (7)**

97 Fecha y número de publicación de la concesión europea: **02.01.2019 EP 3111677**

54 Título: **Gestión de la sonoridad de audio basada en objetos**

30 Prioridad:

27.02.2014 US 201461945734 P
26.02.2015 US 201514632997

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
30.05.2019

73 Titular/es:

DTS, INC. (100.0%)
5220 Las Virgenes Road
Calabasas, CA 91302, US

72 Inventor/es:

MALAK, FADI;
KATSIANOS, THEMIS y
JOT, JEAN-MARC

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

ES 2 714 905 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

Gestión de la sonoridad de audio basada en objetos

5 Antecedentes

Campo

10 La presente descripción se refiere en general a la gestión de la sonoridad, y más particularmente, a la gestión de la sonoridad de audio basada en objetos.

Antecedentes

15 La sonoridad es un problema recurrente principalmente en la entrega de tipo de flujo de entretenimiento lineal (en tiempo real). El entretenimiento lineal puede incluir programación lineal de transmisión, video a pedido (VOD) y emisión en continuo en tiempo real (OTT). En el pasado, varias organizaciones internacionales de estándares, formadas por una variedad mundial de ingenieros y expertos en audio, han definido métodos para medir con precisión la sonoridad percibida de una mezcla de audio de transmisión. Aunque inicialmente este trabajo fue realizado por organizaciones de estándares, eventualmente se involucraron agencias reguladoras de los gobiernos nacionales. Estas agencias reguladoras emitieron regulaciones para implementar un conjunto estandarizado de especificaciones técnicas, estableciendo su uso y recomendando las mejores prácticas. Sin embargo, este trabajo se ha realizado solo en el mundo del audio basado en canales del estéreo y, más recientemente, en el sonido envolvente de 5.1 canales.

20 El documento WO 2014/025752 A1 describe métodos para procesar señales de audio basadas en objetos de acuerdo con los preámbulos de las reivindicaciones 1 y 13.

25 MAX NEUENDORF ET AL: "Contribución a Audio MPEG-H 3D Versión 1", 106. ENCUESTRO DE MPEG; 28-10-2013 - 1-11-2013; GINEBRA; (GRUPO DE EXPERTOS DE IMAGEN EN MOVIMIENTO O ISO/IEC JTC1/SC29/WG11), n.º m31360, 23 de octubre de 2013 (2013-10-23), XP030059813 analiza la normalización de la sonoridad que escala una señal de entrada de audio, en donde un factor de escala se deriva de la diferencia de un nivel de referencia del programa y un nivel objetivo del decodificador.

Sumario

35 La invención proporciona un método de señales de audio basado en objetos de procesamiento para la reproducción a través de un sistema de reproducción con las características de la reivindicación 1 y un método de procesamiento de señales de audio objeto de base para al menos uno de difusión, de distribución de ficheros, o emisión en continuo, con las características de la reivindicación 13. Las realizaciones de la invención se definen en las reivindicaciones dependientes.

40 Por consiguiente, en un aspecto de la divulgación, se proporcionan un método y un aparato para procesar señales de audio basadas en objetos para su reproducción a través de un sistema de reproducción. El aparato recibe una pluralidad de señales de audio basadas en objetos. Cada señal de audio basada en objetos de las señales de audio basadas en objetos incluye datos de forma de onda de audio y metadatos de objeto asociados con los datos de forma de onda de audio. Los metadatos del objeto incluyen al menos uno de un parámetro de sonoridad o un parámetro de potencia asociado con los datos de la forma de onda de audio. El aparato determina una métrica de sonoridad basada en las señales de audio basadas en objetos recibidas y en el al menos uno de los parámetros de sonoridad o el parámetro de potencia para cada señal de audio basada en objetos de las señales de audio basadas en objetos recibidas. El aparato procesa las señales de audio basadas en objetos recibidas en un conjunto de señales de salida basadas en la métrica de sonoridad determinada.

45 En un aspecto de la divulgación, un método y un aparato para procesar señales de audio basadas en objetos para al menos uno de difusión, están dentro de distribución de ficheros, o emisión en continuo. El aparato recibe una pluralidad de señales de audio basadas en objetos. Cada señal de audio basada en objetos de las señales de audio basadas en objetos incluye datos de forma de onda de audio y metadatos de objeto asociados con los datos de forma de onda de audio. Los metadatos del objeto incluyen al menos uno de un parámetro de sonoridad o un parámetro de potencia asociado con los datos de la forma de onda de audio. El aparato determina una métrica de sonoridad basada en las señales de audio basadas en objetos recibidas y en el al menos uno de los parámetros de sonoridad o el parámetro de potencia para cada señal de audio basada en objetos de las señales de audio basadas en objetos recibidas. El aparato transmite las señales de audio basadas en objetos recibidas en función de la métrica de sonoridad determinada.

Breve descripción de los dibujos

65 Con referencia ahora a los dibujos en los que números de referencia similares representan partes correspondientes en todo:

La figura 1 ilustra una construcción a modo de ejemplo de un flujo de bits de audio multidimensional para su uso por realizaciones del sistema y método de gestión de sonoridad de audio basado en objetos (OBA) de normalización espacial.

5 La figura 2 es un diagrama de bloques que ilustra una descripción general de una implementación a modo de ejemplo del proceso de creación de flujo de bits de transmisión de audio multidimensional (MDA).

La figura 3 es un diagrama de bloques que ilustra una visión general de la generación de metadatos de sonoridad global sumada.

10 La figura 4 es un diagrama de bloques que ilustra los usos contemplados de la métrica de sonoridad de normalización espacial básica y la métrica de sonoridad de normalización espacial mejorada calculada por realizaciones del sistema y el método de gestión de sonoridad del OBA de normalización espacial.

La figura 5 es un diagrama de bloques que ilustra una visión general del cálculo de la métrica de sonoridad de normalización espacial mostrada en la figura 4.

15 La figura 6 es un diagrama de bloques que ilustra una descripción general de la etapa de creación/codificación de contenido.

La figura 7 es un diagrama que ilustra la orientación y la geometría de la cabeza de un oyente en relación con un objeto y se utiliza durante una etapa de monitorización en asociación con el sistema y el método de gestión de la sonoridad del OBA de normalización espacial.

20 La figura 8 es un diagrama de flujo que ilustra el funcionamiento general de las tres etapas de las realizaciones del sistema y el método de gestión de la sonoridad del OBA de normalización espacial.

La figura 9 es un diagrama de flujo de un método de procesamiento de señales de audio basadas en objetos para su reproducción a través de un sistema de reproducción.

La figura 10 es un diagrama de flujo de un método de procesamiento de señales de audio basadas en objetos para al menos uno de transmisión, entrega de archivos o emisión en continuo.

25 La figura 11 es un diagrama de flujo de datos conceptual que ilustra el flujo de datos entre diferentes módulos/medios/componentes en un aparato a modo de ejemplo.

Descripción detallada

30 La descripción detallada expuesta a continuación en conexión con los dibujos adjuntos pretende ser una descripción de varias configuraciones y no pretende representar las únicas configuraciones en las que los conceptos descritos en el presente documento pueden ponerse en práctica. La descripción detallada incluye detalles específicos con el propósito de proporcionar una comprensión completa de varios conceptos. Los expertos en la técnica entenderán, sin embargo, que se pueden practicar diversas realizaciones sin estos detalles específicos. En algunos casos, las estructuras y componentes bien conocidos se muestran en forma de diagrama de bloques para evitar oscurecer tales conceptos. Los aparatos y métodos se describirán en la siguiente descripción detallada y se pueden ilustrar en los dibujos adjuntos mediante varios bloques, módulos, componentes, circuitos, etapas, procesos, algoritmos, elementos, etc.

40 En la siguiente descripción de realizaciones de un sistema de gestión de la normalización de sonoridad OBA espacial y el método, se hace referencia a los dibujos adjuntos. Estos dibujos muestran a modo de ilustración ejemplos específicos de cómo se pueden practicar las formas de realización del sistema y el método de gestión de la sonoridad del OBA de normalización espacial. Se entiende que se pueden utilizar otras realizaciones y se pueden hacer cambios estructurales sin apartarse del alcance del objeto reivindicado.

45 I. Introducción

50 Con el advenimiento del OBA, nuevas oportunidades y desafíos han emergido. Un tema clave es cómo medir y administrar la sonoridad cuando puede existir un número arbitrario de objetos de audio en una mezcla. La capacidad de medir y administrar la sonoridad es especialmente importante cuando se introduce un control interactivo en el hogar, lo que permite que un consumidor agregue o suelte objetos de audio. Si bien la flexibilidad del OBA tiene muchas ventajas, el OBA presenta desafíos porque los métodos existentes de gestión y control de la sonoridad basados en el canal son inaceptables.

55 La figura 1 ilustra una construcción a modo de ejemplo de un flujo de bits de programa MDA 100 para uso por realizaciones del sistema y el método de gestión de sonoridad del OBA de normalización espacial. La figura 2 es un diagrama de bloques 200 que ilustra una visión general de una implementación a modo de ejemplo del proceso de creación de flujo de bits de difusión MDA. Un flujo de bits 100 del programa MDA puede incluir metadatos específicos del programa 102 y una pluralidad de objetos de audio 104. Un objeto de audio 104 es una o más formas de onda de audio con metadatos específicos de objetos dinámicos o estáticos que describen ciertas características de esas formas de onda. Estas características pueden incluir la ubicación posicional en un espacio tridimensional (3D) en un punto determinado en el tiempo, los valores de sonoridad medidos, la naturaleza del objeto (como un instrumento, efecto, música, fondo o diálogo), el lenguaje del diálogo, cómo mostrar el objeto y los metadatos en forma de instrucciones sobre cómo procesar, transformar o reproducir el objeto. En el OBA puro, los objetos de audio 104 no se asignan a un canal específico. De hecho, puede que no se sepa cuántos canales contiene la configuración de reproducción. En otras palabras, el objeto de audio 104 está destinado a procesarse de manera unitaria independiente de cualquier

configuración de reproducción fija o predefinida particular de los altavoces de transformación. En estas situaciones, el proceso de transformación se realiza más tarde para convertir y mezclar en canales de reproducción (como se define en la configuración de reproducción).

5 En general, el volumen se define como un atributo de la sensación auditiva en términos de las cuales los sonidos pueden ser ordenados en una escala que se extiende desde silencio a alto. La sonoridad es una medida subjetiva que se ve afectada por el nivel de presión del sonido (SPL), la frecuencia, el ancho de banda, la duración y la proximidad. Además, el servicio de radiodifusión (BS) 1770 (UIT BS.1770) de la Unión Internacional de Telecomunicaciones (UIT) es un estándar de difusión para definir y calcular el volumen y la Unión Europea de Radiodifusión (EBU) R-128 define
10 cómo los emisores pueden medir y normalizar el audio.

Actualmente existen ambos ejemplos abierto y de propiedad del OBA. El sistema y el método de gestión de la sonoridad del OBA de normalización espacial que se describen en este documento utilizan el OBA con un amplio conjunto de metadatos que incluye un conjunto completo para los valores de la sonoridad. Un flujo de bits del OBA
15 abierto tiene una arquitectura abierta para que los metadatos sean legibles y accesibles en cualquier punto dado de la existencia del flujo de bits. A modo de ejemplo y no de limitación, el MDA es un formato tan abierto que incluye una representación de flujo de bits y una carga útil del OBA. El MDA es una plataforma de audio inmersiva en audio basada en objetos completamente abierta que permite a cualquier proveedor de contenido mezclar audio basado en objetos o cualquier combinación de audio basado en objetos y audio basado en canales. Por ejemplo, el contenido se puede
20 mezclar utilizando doce altavoces y el MDA asignará el contenido a cualquier configuración de reproducción, como 5.1 o estéreo. En este documento, se hará referencia al MDA como un ejemplo de aplicabilidad a las realizaciones del sistema y el método de gestión de la sonoridad del OBA de normalización espacial. Sin embargo, otros tipos de formatos de flujo de bits (por ejemplo, DTS: X) también pueden ser aplicables a las formas de realización del sistema y el método de gestión de sonoridad del OBA de normalización espacial. Si bien el MDA puede admitir objetos, canales
25 y audio basado en escenas (ambisonics de orden superior (HOA)), se debe tener en cuenta que, en este documento, el MDA se refiere principalmente a una carga útil del OBA.

II. Descripción operativa y del sistema

30 Es deseable definir nuevas técnicas o actualizar las técnicas existentes para hacer frente a la sonoridad como el mundo la producción de audio migra de audio basado en canal al OBA. En la actualidad, sin embargo, no hay un método conocido o acordado para medir el volumen del audio de un objeto en el espacio 3D. En un futuro próximo, los expertos de la industria mundial, como los de la EBU, indudablemente explorarán nuevos métodos para lidiar con la gestión de la sonoridad cuando tengan una carga de audio inmersiva, como Objetos, Canales + Objetos o HOA.

35 No sólo es deseable actualizar los métodos de medida de la sonoridad de aplicar a OBA, pero también es deseable diseñar técnicas que pueden definir y determinar una medición concreta y significativa de la sonoridad sin el conocimiento de la configuración de la representación de destino. Es incluso mejor si la técnica puede calcular una medición sin tener que transformar los objetos. Las realizaciones del sistema y el método de gestión de la sonoridad
40 del OBA de la normalización espacial cumplen estos objetivos.

Sonoridad global sumada

45 Se puede definir un parámetro de metadatos, llamado un valor de sonoridad sumada global, dentro de los metadatos específicos del programa de la corriente de bits del programa MDA. El valor de sonoridad sumada global puede representar el programa del OBA total o el valor de sonoridad combinada y medida. El único método conocido para lograr un valor que entendería la industria del audio hoy en día sería enviar la carga útil del objeto de audio en la secuencia a través de un procesamiento forzado al bucle de canal. En este caso, los objetos utilizan el renderizador de referencia MDA (como la panoramización de la amplitud de la base del vector (VBAP), consulte la figura 2) para
50 transformar a la configuración del diseño de altavoces 5.1 definida ITU. Esto esencialmente convierte todos los objetos en un canal 5.1. Luego, estos objetos se incorporan a un proceso de medición de sonoridad compatible con A85 de EBU R-128 o Comité de Sistemas de Televisión Avanzada (ATSC) A85. El valor medido (medido en sonoridad, ponderado en K, relativo a escala completa (LKFS) o unidades de sonoridad relativas a escala completa (LUFS)) se registra de nuevo en el flujo de bits MDA a nivel de flujo de bits, y no a nivel de objeto individual (consulte figura 1)
55 como el valor de la sonoridad sumada global del programa (por ejemplo, ITU5.1-23LUFS). Esto también se puede aplicar al estéreo.

La figura 3 es un diagrama de bloques 300 que ilustra una visión general de la generación de metadatos de sonoridad global sumada. En 302, los metadatos OBA se generan en una etapa de generación de metadatos. Dichos metadatos
60 pueden generarse para MDA o DTS: X, por ejemplo. Posteriormente, en 304, se puede realizar un transformado o pretransformado en señales de objetos de audio para determinar una potencia o un volumen promedio de cada una de las señales de objetos de audio. En 306, se pueden realizar mediciones/monitorización de sonoridad para determinar un valor de sonoridad sumada global. Dichos monitoreos/mediciones pueden cumplir con la EBU R-128 o la Ley de Mitigación de Sonoridad de Anuncios Comerciales (CALM). Una vez que se calcula el valor de sonoridad sumada global, en 308, el valor de sonoridad sumada global calculado se puede comparar con un nivel de sonoridad
65 objetivo especificado en las normas internacionales de regulación, como -24 LKFS en la Ley CALM, o -23 LUFS en

EBU R -128. El nivel de sonoridad objetivo se puede llevar en los metadatos específicos del programa dentro del flujo de bits MDA. Sobre la base de la comparación, en 310, se puede calcular un desplazamiento y, en 312, ese desplazamiento se puede almacenar en los metadatos específicos del programa dentro del flujo de bits MDA como el desplazamiento de sonoridad sumada global. El desplazamiento de sonoridad sumada global se puede aplicar posteriormente en sentido descendente cuando el audio finalmente se transforma para la reproducción del consumidor.

Medición de la sonoridad espacial

Las realizaciones del sistema y el método de gestión de la sonoridad del OBA de la normalización espacial definen y determinan una medición de la sonoridad concreta y significativa sin el conocimiento de la configuración de representación de destino (por ejemplo, número de altavoces o configuración de los altavoces). Esto es importante ya que este será el caso del contenido OBA destinado a la entrega y reproducción del consumidor. Además, las realizaciones del sistema y el método computan esta medida de sonoridad sin tener que transformar los objetos.

La figura 4 es un diagrama de bloques 400 que ilustra los usos contemplados de la métrica de sonoridad de normalización espacial básica y la métrica de sonoridad de normalización espacial mejorada calculada por realizaciones del sistema y método de gestión de sonoridad del OBA de normalización espacial. La figura 5 es un diagrama de bloques 500 que ilustra una visión general del cálculo de la métrica de sonoridad de normalización espacial mostrada en la figura 4. El sistema y el método de gestión de la sonoridad del OBA de normalización espacial definen dos tipos de métricas de la sonoridad de la normalización espacial. Se puede calcular una métrica de sonoridad de normalización espacial básica conociendo la mezcla final (es decir, todos los objetos de audio utilizados para determinar la sonoridad) y el entorno de transformado objetivo. Esto alivia la necesidad de una lista fija de configuraciones de reproducción y se puede hacer en el último punto posible de la cadena en la casa del consumidor. La métrica de sonoridad de la normalización espacial mejorada se puede calcular basándose en información adicional, como la ubicación y la orientación del oyente. La métrica de sonoridad de normalización espacial se puede determinar en cualquier punto de la cadena de producción/entrega/reproducción de difusión.

Ambas técnicas y valores tienen por objeto resolver dos problemas identificados. El primer problema es proporcionar a las emisoras que se encuentran bajo estrictas regulaciones nacionales para mantener el control de la sonoridad y garantizar una experiencia de sonoridad constante para sus consumidores con el contenido que envían al hogar. El segundo problema es abordar la necesidad de encontrar nuevas técnicas para manejar la entrega de OBA. En estas situaciones, la aplicación y los requisitos para tales sistemas (como los nuevos estándares internacionales de transmisión) requieren flexibilidad para adaptarse a cualquier ubicación/orientación del oyente en el hogar con respecto a las ubicaciones de los objetos de audio.

Las realizaciones del sistema y el método de gestión de la sonoridad del OBA de la normalización espacial permiten que la cadena de extremo a extremo sea más inteligente y más adaptable. Además, el sistema y el método desplazan gran parte del trabajo pesado a la cabecera o arquitectura basada en la nube. Algunos cálculos algorítmicos se mantienen en el lado del consumidor para adaptarse a cualquier entorno dado y cualquier combinación arbitraria, o mezcla modificada (modificada por la intervención de la interactividad del consumidor a través de enlaces tardíos o feeds de red doméstica). El sistema y el método también tienen en cuenta el entorno de reproducción específico del consumidor.

En situaciones donde se utiliza la métrica de normalización de sonoridad espacial básica, los organismos de radiodifusión pueden emplear muchos componentes del sistema de inspección al azar, para validar, o para corregir la sonoridad de los canales de audio. Algunos son sistemas basados en archivos y otros utilizan equipos en tiempo real. Las responsabilidades de mantener el cumplimiento (o no tocar cierto contenido certificado) se transfieren de una parte de la cadena de distribución a la siguiente a través de los contratos de derechos de contenido. En última instancia, las leyes están diseñadas para responsabilizar a ciertas compañías individuales. En los Estados Unidos, son los programadores y los difusores quienes originan el contenido. Las herramientas son necesarias para verificar, comprobar o ajustar la sonoridad, mientras que tienen un impacto mínimo o nulo en la calidad de audio (como el rango dinámico) del contenido de audio originalmente creado. El sistema y el método a modo de ejemplo son una forma no invasiva de "aprovechar" el flujo de bits en cualquier punto dado y obtener un cálculo de la sonoridad del programa del OBA sin tener que transformar o conocer la configuración del usuario final.

En situaciones donde se utiliza la métrica de normalización de sonoridad espacial mejorada, se conoce la exacta posición/orientación del oyente. En este entorno, en la etapa final de la cadena de entrega (en el hogar del consumidor, vea la figura 4), el sistema tiene conocimiento de dónde se van a representar los objetos en la sala, en relación con el oyente. Esto mejora la precisión del sistema y el método y sus cálculos.

Si se dispusiera de proximidad en la creación de la mezcla original, un procesador avanzado que puede aprovechar esos efectos también se utiliza. La proximidad puede ser utilizada por las realizaciones del sistema y el método para mediciones y compensaciones aún más precisas. El sistema y el método también pueden utilizar cualquier cambio en la posición del oyente en relación con la posición del objeto en el espacio de escucha 3D. Este cambio se da a conocer al sistema y al método a través de un sistema de retroalimentación ambiental. Si hay información adicional disponible para las realizaciones del sistema y el método, puede calcular un nivel de sonoridad "percibida" de todos los objetos renderizados en el espacio 3D en relación con el punto de vista del oyente.

Es importante tener en cuenta y entender que las técnicas métricas de sonoridad de normalización espacial básica y las técnicas métricas de sonoridad de normalización espacial mejorada son, ante todo, un nuevo proceso de medición. Las realizaciones del sistema y el método pueden aplicarse al flujo de bits del OBA en cualquier punto de la cadena, ya sea basándose en archivos o en tiempo real. Ambos tipos de métricas de sonoridad de normalización espacial se pueden calcular en cualquier punto de la etapa de creación y distribución, y también se pueden insertar como metadatos en el flujo de bits de OBA una vez que se calculan y crean. En MDA, la métrica de sonoridad de la normalización espacial puede reemplazar el valor de la sonoridad sumada global utilizando el método de transformado a UIT 5.1.

La métrica de sonoridad normalización espacial mejorada puede calcularse en el punto final en la cadena de entrega, en el entorno de reproducción de los consumidores. Este es el punto donde el sistema conoce más información sobre el oyente y el conjunto final de objetos en la mezcla, después de que haya tenido lugar la interactividad. La métrica de sonoridad de normalización espacial mejorada se puede diseñar en un sistema en tiempo real en un decodificador de audio inmersivo para el consumidor. Además, en última instancia, se puede utilizar para corregir cualquier inconsistencia de sonoridad no deseada. El sistema y el método pueden ser autorizados e implementados en productos de transmisión profesional diseñados para medir, monitorizar o corregir problemas de sonoridad en el flujo de trabajo de la cabecera de la emisión, sin afectar la esencia del audio o la intención artística (solo a través de la corrección de metadatos).

En términos generales, las realizaciones del sistema de gestión de sonoridad OBA y el método contienen técnicas para medir la sonoridad cuando el audio está en la forma de objetos de audio en el espacio 3D. Las realizaciones del sistema y el método utilizan metadatos de objetos de audio para crear un nuevo valor medido y una nueva unidad de referencia de la carga útil del OBA. Hoy en día existen problemas de sonoridad en el mundo, principalmente en la entrega de entretenimiento lineal de tipo transmisión, como la programación lineal de transmisión, VOD y emisión en continuo OTT. Las técnicas existentes simplemente administran la sonoridad para los escenarios de estéreo tradicional y 5.1 canales.

El sistema de gestión de sonoridad del OBA y métodos permite a los creadores y distribuidores de contenidos para medir y hacer referencia a un nuevo valor para la sonoridad de carga útil sin necesidad de generar los objetos de audio. Además, la configuración de reproducción no necesita ser conocida.

Las técnicas actuales carecen de la capacidad para medir su sonido envolvente, como el OBA. Las realizaciones del sistema y el método utilizan información espacial combinada con la información de sonoridad/potencia del objeto individual para crear un nuevo valor de referencia, sin necesidad de representar la carga útil. El sistema y el método a modo de ejemplo utilizan un flujo de bits que transporta información sobre la posición de los objetos en el espacio 3D y el volumen individual medido durante la creación de un objeto. A modo de ejemplo y no de limitación, el flujo de bits puede ser un flujo de bits MDA o una especificación abierta de OBA alternativa. El sistema y el método a modo de ejemplo también incluyen una técnica de normalización espacial que calcula una métrica de sonoridad de normalización espacial, dada la información que necesita (y sin procesamiento).

En general, la normalización de métrica de sonoridad espacial puede ser una de dos tipos. En algunas realizaciones, se utiliza una métrica de la sonoridad de la normalización espacial básica que se puede calcular en cualquier momento asumiendo una posición/orientación particular del oyente con respecto a la ubicación de los objetos de audio. Esto es cierto tanto en un entorno basado en archivos como en un entorno en tiempo real. En otras realizaciones, se utiliza una métrica de sonoridad de normalización espacial mejorada. En estas realizaciones se conoce la posición/orientación del oyente. En comparación con la métrica de sonoridad de normalización espacial básica, la métrica de sonoridad de normalización espacial mejorada puede crear un valor de sonoridad percibida más preciso y personalizado. Esta métrica de sonoridad mejorada se puede usar para administrar cualquier inconsistencia de sonoridad durante un programa, o de un programa a otro, o de un programa a otro. La métrica de sonoridad de la normalización espacial mejorada también puede tener en cuenta otra diversa información relevante para la sonoridad. A modo de ejemplo y no de limitación, esta información relevante para el volumen incluye información de proximidad. Además, en algunas realizaciones, la métrica de sonoridad mejorada puede explicar cualquier interactividad del lado del consumidor y donde el contenido se remezcla (agregando un objeto a la mezcla, eliminando un objeto de la mezcla o cambiando la ubicación de un objeto en la mezcla (por ejemplo, moviendo un objeto a una ubicación diferente o desplazando un objeto en la mezcla), todo lo cual cambia la sonoridad completa de la mezcla).

5 La métrica espacial básica de normalización de sonoridad tomará forma como código de equipo de difusión profesional que tiene licencia en productos profesionales que llevan a cabo la creación de audio, el procesamiento y equipos de codificación/decodificación/transcodificación. Puede integrarse en herramientas independientes (cajas de hardware o software), en herramientas de terceros, en codificadores, o como parte de equipos de procesamiento basados en servidor o en la nube que condicionan y normalizan el audio.

10 La métrica de sonoridad de normalización espacial mejorada se puede utilizar como parte de las soluciones de consumo con licencia integrados (codec suite o post-procesamiento). Estas herramientas forman parte de la difusión y la preparación OTT para soluciones de entrega y reproducción en la actualidad. Las implementaciones del lado del cliente de la métrica de sonoridad de la normalización espacial mejorada incluyen decodificadores y reproductores implementados en aplicaciones multipantalla, como PC, tabletas, teléfonos inteligentes móviles, televisores y decodificadores. Además, estos dispositivos ni siquiera necesitan altavoces, ya que la reproducción por auriculares también es aplicable.

15 **III. Detalles operativos y del sistema**

El sistema y método de gestión de sonoridad del OBA de normalización espacial describen la medida de la sonoridad y la gestión en el OBA. Los metadatos asociados con cada objeto de audio pueden ser, por ejemplo, la ubicación del objeto en el espacio 3D, un factor de escala de amplitud de la forma de onda que se aplicará a la forma de onda al representar el objeto, datos sobre la relación de correlación de los objetos involucrados o temporales, información sobre el objeto, como cuándo comienza y cuándo finaliza. En la siguiente discusión, el sistema y el método se analizarán en el contexto de tres etapas: a) etapa de creación (o codificación) de contenido, b) etapa de monitorización intermedio, y c) etapa de consumo de contenido.

25 **Etapa de codificación de metadatos**

La figura 6 es un diagrama de bloques 600 que ilustra una visión general de la etapa de creación/codificación de contenido. Durante la etapa de creación, los metadatos asociados con cada objeto de audio (602) (como la potencia o el volumen de cada objeto de audio) se miden (604) realizando algún tipo de "pre-transformación" del flujo de bits. En esta etapa se pueden realizar mediciones cortas, intermedias o basadas en archivos. La medida de potencia z_i para una señal de entrada y_i se define como sigue durante un período de tiempo T :

$$z_i = 1/T \int_0^T y_i^2 dt, \quad (1)$$

35 donde $i \in I$, e I es el conjunto de señales de objetos de audio. La información de medición de potencia puede almacenarse como metadatos (606) con las señales del objeto de audio. La información de medición de potencia puede almacenarse alternativamente como información de sonoridad. Se puede almacenar información adicional como metadatos específicos del objeto, incluidos la ganancia/amplitud, la ubicación del objeto de audio y la ubicación/orientación de un oyente. El procesamiento de la información de metadatos de medición de potencia/volumen para objetos de audio se describe *infra* en relación con la fase de monitorización.

40 **Etapa de monitorización**

Si I es el conjunto de canales de entrada para las mediciones de potencia de canales (no objetos de audio), la asignación de la potencia medida en un valor de sonoridad puede ser hecha por medio de una transformación logarítmica:

$$\text{Sonoridad, } L_K = -0,691 + 10 \log_{10} \sum_i^N G_i z_i \text{ LKFS}, \quad (2)$$

50 donde G_i es un coeficiente de ponderación para la i -ésima señal de audio-objeto.

Para las mediciones de sonoridad intermedias, un tipo cerrado, de solapamiento (que puede ser ajustado a 75 %) de transformación se puede utilizar:

$$55 \text{ Sonoridad controlada, } L_{KG} = -0,691 + 10 \log_{10} \sum_i G_i \left(\frac{1}{|J_g|} \sum_{j \in J_g} z_{ij} \right) \text{ LKFS} \quad (3),$$

donde J_g es un conjunto de índices de bloque donde la sonoridad del bloque de activación está por encima del umbral de activación que generalmente se toma a -70 dBFS y $|J_g|$ es una serie de elementos en J_g :

$$60 J_g = \{j: l_j > \Gamma_a\}, \Gamma_a = -70 \text{ LKFS} \quad (4).$$

Para las medidas objeto de forma de onda de sonoridad de un solo canal, las sumas en el colapso ecuaciones *supra* a 1 (es decir, $N = 1$) y el coeficiente de escala G_i se puede establecer en 1,0. En algunas realizaciones, para un grupo

de objetos de un solo canal que representan una señal de audio multicanal donde cada canal se asigna a una ubicación fija en el espacio (denominada "lecho" multicanal), los coeficientes de escala G_i en las ecuaciones *supra* se pueden asignar a los pesos por canal que se encuentran en las especificaciones BS.1770-x y se definen solo para configuraciones de reproducción de altavoces multicanal estándar, como la configuración "5.1". Las mediciones de potencia/volumen (612) se pueden realizar en los objetos de un solo canal (610) y se pueden almacenar como metadatos (614). Aquí se debe tener en cuenta que se puede aplicar una etapa de prefiltrado antes de tomar la medición de sonoridad. Esto puede incluir filtros para tener en cuenta los efectos acústicos de la cabeza y los filtros revisados de peso B de baja frecuencia. Una vez que se mide la información de potencia/volumen por objeto, el valor medido se almacena como metadatos que se pueden usar más adelante. Además, los metadatos que representan la correlación de objetos se pueden calcular (608) en esta etapa, lo que posteriormente puede facilitar los cálculos de la métrica de sonoridad de normalización espacial. La correlación de objetos representa el enmascaramiento auditivo que se produce cuando la percepción de un sonido se ve afectada por la presencia de otro sonido. En el dominio de la frecuencia, la máscara auditiva puede denominarse máscara simultánea, máscara de frecuencia o máscara espectral. En el dominio del tiempo, el enmascaramiento auditivo puede denominarse enmascaramiento temporal o enmascaramiento no simultáneo.

Durante la etapa de monitorización, la información de potencia/volumen de metadatos por objeto se recupera (véase la ecuación (1)) y se utiliza para calcular un valor global de la sonoridad. Las mediciones de potencia/sonoridades individuales recuperadas se combinan/suman para obtener una medición de sonoridad global. Además, las mediciones individuales se pueden escalar en función de la información de metadatos de peso por objeto antes de ser combinadas/sumadas. La medición de la sonoridad global combinada/sumada se define como la métrica de la sonoridad de la normalización espacial (también conocida como SNM). Debe entenderse que la métrica de la sonoridad de la normalización espacial se puede calcular en un formato de potencia o asignación de la sonoridad.

Las ecuaciones (2) y (3), descritas *supra*, describen las mediciones de sonoridad por canal o un grupo de canales. Las ecuaciones (2) y (3) ahora se pueden generalizar para generar mediciones métricas de sonoridad de normalización espacial por objeto de audio o un grupo de objetos de audio.

En particular, la determinación de la métrica espacial normalización de sonoridad (SNM) sobre un intervalo T se puede generalizar como sigue:

$$SNM = -0,691 + 10 \log_{10} \sum_i^N FJJT(r_i, r_L, \theta_i, \phi_i, a_i, g_i, c_i) z_i \quad (5),$$

donde $i \in I$, I es el conjunto de señales de objeto de audio, y N es el número de señales de objeto de audio en el conjunto I de señales de objeto de audio.

La determinación de la métrica de sonoridad normalización espacial cerrada (cerrada SNM) sobre un intervalo T se puede generalizar como sigue:

$$SNM \text{ puerta} = -0,691 + 10 \log_{10} \sum_i^N FJJT(r_i, r_L, \theta_i, \phi_i, a_i, g_i, c_i) \left(\frac{1}{|J_g|} \sum_{j_g} z_{ij} \right) \quad (6),$$

donde $i \in I$, I es el conjunto de señales de objeto de audio, y N es el número de señales de objeto de audio en el conjunto I de señales de objeto de audio.

Para el SNM cerrada, el intervalo T se divide en un conjunto de la superposición de los intervalos de bloque de puerta. Un bloque de activación es un conjunto de muestras de audio contiguas de duración T_g , que pueden tener una duración de 400 ms. La superposición de cada bloque de activación puede ser del 75 % de la duración del bloque de activación T_g . Con una superposición del 75 % y una duración de ventana de 400 ms, el SNM sincronizado se determina en función de los 300 ms de una señal de objeto de audio utilizada en un cálculo previo del SNM sincronizado. Como tal, con una superposición del 75 % y una duración de ventana de 400 ms, el SNM sincronizado se determina cada 100 ms.

Se puede observar que las ecuaciones generalizadas (5) y (6) se derivan de las ecuaciones (2) y (3) mediante la introducción de la función $FJJT$, que se define como sigue:

$$FJJT(r_i, r_L, \theta_i, \phi_i, a_i, g_i, c_i) = c_i a_i g_i \left(\frac{1}{|\bar{r}_i - \bar{r}_L|^2} \right) O(\theta_i, \phi_i) \quad (7),$$

donde c_i es el factor de correlación del objeto $i^{\text{ésimo}}$ y es una medida de correlación de un enmascaramiento de audio del objeto $i^{\text{ésimo}}$ por uno o más de los otros objetos $N-1$, a_i es el factor de escala de amplitud (factor de escala para obtener amplitud prevista de la señal de audio-objeto) del $i^{\text{ésimo}}$ objeto transportado en el flujo de metadatos,

$\frac{1}{|\bar{r}_i - \bar{r}_L|^2}$ es opcional y sigue la ley de la distancia relativa cuadrada inversa del objeto $i^{\text{ésimo}}$ y el oyente, y g_i es un factor de peso opcional dependiente de la frecuencia que explica la sensibilidad de la sonoridad de la audición humana en

función de la ubicación angular del sonido con respecto a la cabeza, y en relación con una dirección de aspecto asumida (que normalmente coincide con la posición del canal "centro-frontal"). Este factor de peso puede considerarse como una forma generalizada de la etapa de prefiltrado descrita en BS.1770-x, diseñada para tener en cuenta los efectos acústicos de la cabeza. El factor de correlación c_i puede estar relacionado inversamente con el enmascaramiento de audio. Por ejemplo, cuando no hay enmascaramiento de audio, el factor de correlación c_i puede ser 1, y cuando hay un enmascaramiento de audio del 100 %, el factor de correlación c_i puede ser 0.

$O(\theta_i, \varphi_i)$ es un factor de corrección adicional que está ligado al factor de peso g_i descrito *supra*. $O(\theta_i, \varphi_i)$ utiliza los vectores derecho y ascendente de productos cruzados $R = F \times A$ y $U = R \times F$, respectivamente, que describen la geometría de la cabeza del oyente en relación con el objeto $i^{\text{ésimo}}$ como se muestra en la figura 7. Como se muestra en la figura 7, la dirección de un *Vector Derecho* es a lo largo de la dirección de la oreja derecha del oyente que apunta hacia afuera de su cabeza. La dirección del *Vector Arriba* está por encima de la cabeza del oyente.

Cuando hay un oyente con relación al objeto de acimut (θ)/cambio de elevación (φ) descrito por las transformaciones angulares $\theta - \left(\theta_R - \frac{\pi}{2}\right)$ y $\varphi_{\text{orientación}} = \varphi - \varphi'$ que se describen en relación con el *vector derecho* y el *vector ascendente*, luego $O(\theta_i, \varphi_i)$ calcula un factor de escala de corrección para la sonoridad del objeto $i^{\text{ésimo}}$.

Además, una tabla de salidas discretas de la función O continua (θ_i, φ_i) pueden ser generados para diferentes pares de acimut y elevación y esa tabla se puede consultar cuando un factor de peso de orientación necesita ser calculado. Se debe tener en cuenta que $O(\theta_i, \varphi_i) = 1$ para la posición central delantera predeterminada (correspondiente a la ubicación típica de un altavoz central frontal). Tanto la distancia relativa como los factores de orientación se pueden establecer en 1 en esta etapa y se pueden recalcular opcionalmente en la etapa de dispositivo del consumidor que se describe a continuación.

Tenga en cuenta que la ubicación del objeto $i^{\text{ésimo}}$ es la ubicación prevista para la señal de audio-objeto (por ejemplo, efecto, diálogo, instrumento, música, fondo, etc.) con respecto al jefe de un oyente y es independiente de una configuración de altavoz de reproducción. Se puede hacer referencia al jefe de un oyente como un punto de recepción en el que todos los objetos cuando se procesan están destinados a ser recibidos. Por ejemplo, un efecto particular de una señal de audio-objeto i puede estar destinado a ubicarse en un acimut de θ_i y una elevación de φ_i en relación con un punto de recepción deseado. Para otro ejemplo, en el acimut de θ_i y una elevación de φ_i relativo en relación con el punto de recepción, tal efecto particular de la señal de audio-objeto i puede tener una distancia prevista $|r_i - r_L|$ desde el punto de recepción.

Debe ser en cuenta que esta es una operación de "monitorización de la sonoridad". En otras palabras, no hay necesariamente ninguna representación del OBA que tenga lugar en esta etapa. Es una ventaja esencial del sistema y el método de gestión de la sonoridad del OBA de la normalización espacial que permite medir la sonoridad del contenido de OBA sin necesidad de decodificación de la forma de onda de audio y cálculo de potencia o sonoridad. En cambio, el cálculo de la métrica de sonoridad de la normalización espacial se facilita y simplifica considerablemente gracias a la provisión de metadatos de sonoridad por objeto en el formato de contenido OBA y el método de cálculo descrito *anteriormente*.

Etapa de dispositivo de consumo (consumo)

En esta etapa final, la sonoridad resumida global, o la métrica de sonoridad de normalización espacial, se pueden recalcular (si es necesario) con el fin de dar cuenta de cualquiera de creación de objetos, delección de objeto, modificación de objeto (por ejemplo, atenuación), o localización del oyente o cambios en la orientación de la cabeza. Esta actualización puede incluir, pero no se limita a, actualizaciones relativas de distancia y orientación según la descripción correspondiente dada en la etapa de monitorización.

Metadatos de sonoridad sumada global

En algunas realizaciones un valor de metadatos de sonoridad sumado global se calcula durante la producción de la secuencia de audio (por ejemplo, MDA o DTS:X). Esto se realiza mediante el uso de un procesador de referencia (como VBAP) para procesar el flujo a una configuración de destino (como, por ejemplo, pero no limitado a ITU 5.1). Una vez que se haya procesado el flujo, se pueden usar las herramientas existentes de monitorización/medición de sonoridad que cumplen con R128/CALM para medir las propiedades de volumen del flujo transformado. Estas propiedades pueden incluir, pero no se limitan a, sonoridad instantánea, sonoridad a corto plazo, pico real y rango de sonoridad, y pueden medirse con o sin el uso de elementos de anclaje, como el diálogo.

Estas mediciones pueden ser ponderadas por las especificaciones del modelo de sonoridad como BS.1770-x, pero no se limitan a seguir estos pesos. Además, se pueden tomar mediciones de correlación para identificar la relación de correlación de las señales transformadas, y en función de los modelos de relación de correlación distintos a la suma de potencia de BS.1770-x, se pueden emplear para calcular la sonoridad general. Una vez que se calcula la sonoridad, la sonoridad calculada se compara con los estándares de regulación internacionales, como -24 LKFS en CALM Act, o

-23 LUFS en EBU R-128. Sobre la base de esta comparación, se calcula un desplazamiento y ese desplazamiento es el volumen global sumado que luego se guarda como valor de metadatos en la transmisión que se produce.

5 Hay que señalar aquí que la sonoridad sumada global es probable que sea sobrescrito cuando se emplea una mayor normalización de métrica de sonoridad espacial. Esto se debe a la información dinámica adicional utilizada en tales casos.

IV. Realizaciones alternativas y entorno operativo a modo de ejemplo

10 Muchas otras variaciones que los descritos en el presente documento serán evidentes a partir de este documento. Por ejemplo, dependiendo de la realización, ciertos actos, eventos o funciones de cualquiera de los métodos y algoritmos descritos en este documento se pueden realizar en una secuencia diferente, se pueden agregar, fusionar o omitir por completo (de manera que no todos los actos o los eventos descritos son necesarios para la práctica de los métodos y algoritmos). Además, en ciertas realizaciones, los actos o eventos se pueden realizar de manera concurrente, como a través del procesamiento de subprocesos múltiples, el procesamiento de interrupciones o procesadores múltiples o núcleos de procesadores o en otras arquitecturas paralelas, en lugar de en forma secuencial. Además, diferentes máquinas y sistemas informáticos pueden funcionar juntos para realizar diferentes tareas o procesos.

20 La figura 8 es un diagrama de flujo 800 que ilustra el funcionamiento general de las tres etapas de las realizaciones del sistema y el método de gestión de la sonoridad del OBA de la normalización espacial. En la etapa de generación de metadatos 802, se generan metadatos para los objetos de audio dentro de un flujo de bits del programa. Dichos metadatos pueden incluir información utilizada en las ecuaciones 1 y 5-7, como se explicó *anteriormente*. En el bloque 804, se realizan mediciones de potencia/volumen. Las mediciones de potencia/volumen pueden realizarse de acuerdo con la ecuación 1. Las propias mediciones de potencia/volumen pueden almacenarse en los metadatos de los objetos de audio. Posteriormente, en el bloque 806, se puede determinar una métrica de sonoridad de normalización espacial. La métrica de la sonoridad de la normalización espacial se puede determinar en función de los metadatos y las suposiciones acerca de la posición/ubicación de un oyente con respecto a los objetos de audio. Por ejemplo, se puede suponer que el oyente está en la posición 0,0,0 en el espacio 3D con respecto a los objetos de audio, que se encuentran alrededor del oyente en un radio/distancia particular. Posteriormente, en 808, los objetos de audio se transmiten (por ejemplo, se transmiten, se envían por entrega de archivos, se emiten) a un dispositivo de consumidor 810. En el dispositivo del consumidor, en 812, la métrica de sonoridad de normalización espacial se puede volver a determinar en función de los objetos de audio incluidos en la mezcla. Por ejemplo, un oyente puede actualizar 814 (por ejemplo, agregar, eliminar) objetos de audio de la mezcla o cambiar objetos de audio en la mezcla. En 812, una métrica de sonoridad de normalización espacial mejorada se puede determinar basándose en información adicional, como la posición 816 actualizada del oyente. Por ejemplo, un oyente puede cambiar su ubicación cuando escucha los objetos de audio de manera que el oyente ya no está en la posición 0,0,0 en el espacio 3D, o el oyente puede cambiar su orientación de la cabeza al recostarse antes que sentarse. La métrica de normalización espacial mejorada se puede determinar en función de la elevación actualizada, el acimut y la información de ubicación del oyente con respecto a las ubicaciones previstas de los objetos de audio.

40 La figura 9 es un diagrama de flujo 900 de un método para procesar señales de audio basadas en objetos para su reproducción a través de un sistema de reproducción. El método puede ser realizado por un procesador, como una unidad central de procesamiento (CPU). El procesador puede estar dentro del sistema de reproducción, o puede estar dentro de un sistema que procesa señales de audio basadas en objetos para su posterior reproducción a través de un sistema de reproducción. El procesador se conoce como un aparato en la figura 9. En 902, el aparato recibe una pluralidad de señales de audio basadas en objetos. Cada señal de audio basada en objetos de las señales de audio basadas en objetos incluye datos de forma de onda de audio y metadatos de objeto asociados con los datos de forma de onda de audio. Los metadatos del objeto incluyen al menos uno de un parámetro de sonoridad o un parámetro de potencia asociado con los datos de la forma de onda de audio. Por ejemplo, los metadatos del objeto pueden incluir parámetros de potencia z_i , como se explicó *supra*. Alternativamente, los metadatos del objeto pueden incluir parámetros de sonoridad (en LKFS o LUFS) que son una función de los parámetros de potencia z_i . En 904, el aparato determina una métrica de sonoridad basada en las señales de audio basadas en objetos recibidas y en el al menos uno del parámetro de sonoridad o el parámetro de potencia para cada señal de audio basada en objetos de las señales de audio recibidas basadas en objetos. Por ejemplo, el aparato puede determinar una métrica de sonoridad de normalización espacial basada en las ecuaciones 5, 6 y 7. En 910, el aparato procesa las señales de audio basadas en objetos recibidas en un conjunto de señales de salida basadas en la métrica de sonoridad determinada.

60 En una configuración, en 910, cuando la prestación de las señales de audio basadas en objetos recibidas, el aparato puede ajustar una amplitud a_i de al menos una de las señales de audio basadas en objetos recibidas basándose en la intensidad métrica determinada. Por ejemplo, en 906, el aparato puede determinar un desplazamiento de la métrica de sonoridad basándose en una comparación entre la métrica de sonoridad y una métrica de sonoridad objetivo. La métrica de sonoridad objetivo puede ser un nivel de sonoridad objetivo especificado en los estándares de regulación internacional, como -24 LKFS en la Ley CALM, o -23 LUFS en EBU R-128, como se explicó *supra*. La amplitud a_i de la al menos una de las señales de audio basadas en objetos recibidas puede ajustarse en función de la comparación entre la métrica de sonoridad y la métrica de sonoridad objetivo. Específicamente, en 908, el aparato puede determinar si la métrica de sonoridad es mayor que la métrica de sonoridad objetivo. En 908, cuando se determina que la métrica

de sonoridad es mayor que la métrica de sonoridad objetivo, en 910, el aparato puede ajustar/escalar la amplitud a_i de al menos una de las señales de audio basadas en objetos recibidas para reducir la sonoridad para acercarse a la métrica de sonoridad objetivo.

5 En una configuración, la métrica de sonoridad se determina en base más adelante, para cada señal de audio basada en objetos de las señales de audio basadas en objetos recibidas, datos de posición de la señal de audio basada en objetos con respecto a un punto de recepción. El punto de recepción supuesto, como 0,0,0 en el espacio 3D, o una ubicación real de un oyente dentro de un espacio 3D. Los datos de posición de la señal de audio basada en objetos es una ubicación prevista de la señal de audio basada en objetos con respecto al punto de recepción 0,0,0. Por ejemplo, una primera señal de audio basada en un objeto puede estar pensada para ubicarse a la derecha de un oyente, en una primera elevación y un primer acimut con respecto al punto de recepción, y una segunda señal de audio basada en un objeto puede estar destinada a ubicado a la izquierda de un oyente, en una segunda elevación y segundo acimut con respecto al punto de recepción.

15 En una configuración, los datos de posición incluyen al menos uno de un acimut entre una localización de la señal de audio basada en objetos y el punto de recepción, o una elevación entre la ubicación de la señal de audio basada en objetos y el punto de recepción. Los datos de posición pueden incluir además una distancia entre la ubicación de la señal de audio basada en objetos y el punto de recepción. En una configuración, se puede suponer que todas las señales de audio basadas en objetos tienen una distancia/radio fija desde el punto de recepción 0,0,0.

20 En una configuración, la métrica de sonoridad se determina en base más adelante, para cada señal de audio basada en objetos de las señales de audio basadas en objetos recibidas, al menos uno de un factor de correlación c_i de la señal de audio basada en objetos con respecto a una u otras señales de audio basadas en objetos más de las señales de audio recibidas basadas en objetos, un factor de escala de amplitud a_i de la señal de audio basada en objetos, o un factor de peso dependiente de la frecuencia g_i de la señal de audio basada en objetos para tener en cuenta para escuchar la sensibilidad de la sonoridad en función de una ubicación angular de la señal de audio basada en objetos con respecto a un punto de recepción.

30 En una configuración, la métrica de sonoridad se determina basándose en $\sum_i^N FJJT(r_i, r_L, \theta_i, \phi_i, a_i, g_i, c_i)z_i$, donde $i \in I$, I es el conjunto de señales de audio basadas en objetos, N es una serie de señales de audio basadas en objetos en las señales de audio recibidas basadas en objetos, z_i se dice al menos uno de los parámetros de sonoridad o potencia para una i ésima señal de audio basada en objetos, r_i es una ubicación asociada con la i ésima señal de audio basada en objetos, r_L es una ubicación asociada con un punto de recepción, θ_i es un acimut entre la ubicación de la i ésima señal de audio basada en objeto y el punto de recepción, ϕ_i es una elevación entre la ubicación de la i ésima señal de audio basada en el objeto y el punto de recepción, a_i es un factor de escala de amplitud de la i ésima señal de audio basada en el objeto, g_i es un factor de peso dependiente de la frecuencia de la señal de audio basada en el objeto de i ésima para dar cuenta de la sensibilidad de la audición de la audición como una función de una ubicación angular de la señal de audio basada en el objeto de i ésima con respecto al punto de recepción, y c_i es una correlación factor de de la i ésima señal de audio basada en objetos con respecto a uno o más otras señales de audio basadas en objetos de las señales de audio recibidas basadas en objetos.

45 En una configuración, las señales de audio basadas en objetos recibidas son especificadas por el usuario. Es decir, un usuario puede agregar, eliminar o cambiar las señales de audio basadas en objetos. Como tal, en 912, el aparato puede recibir información que indica un nuevo conjunto de señales de audio basadas en objetos. La información que indica el conjunto de señales de audio basadas en objetos se recibe en función de la entrada del usuario. Posteriormente, el aparato puede volver a determinar (904) la métrica de sonoridad en función del conjunto de señales de audio basadas en objetos y en función de al menos uno del parámetro de sonoridad o el parámetro de potencia para cada señal de audio basada en objetos del conjunto de señales de audio basadas en objetos.

50 Como se ha discutido *supra*, los metadatos de objeto de cada señal de audio basada en objetos incluyen al menos uno de los datos de posición de la señal de audio basada en objetos con respecto a un punto de recepción, o un factor de escala de amplitud de la señal de audio basada en objetos. Los datos de posición de la señal de audio basada en objetos con respecto al punto de recepción pueden incluir al menos una distancia entre una ubicación de la señal de audio basada en objetos y el punto de recepción, un acimut entre la ubicación de la señal de audio basada en objetos y el punto de recepción, o una elevación entre la ubicación de la señal de audio basada en objetos y el punto de recepción.

60 En una configuración, la métrica de sonoridad es un volumen promedio sobre T ms, y la métrica de sonoridad se determina cada T ms. En una configuración, la métrica de sonoridad es una sonoridad media sobre T ms, y la métrica de sonoridad se determina cada D ms donde $D < T$. Por ejemplo, la métrica de sonoridad puede ser una sonoridad promedio de más de 400 ms, y la métrica de sonoridad puede ser determinada cada 100 ms. En una configuración, la métrica de sonoridad se determina a lo largo de una pluralidad de períodos de ventana, cada uno de los cuales se superponen con una ventana anterior. Por ejemplo, cada uno de los períodos de ventana puede tener una duración de 400 ms y superponerse a otros períodos de ventana en 300 ms.

65

En una configuración, las señales de audio basadas en objetos recibidas incluyen N señales de audio basadas en objetos, las señales de audio basadas en objetos recibidas se prestan a través de altavoces que tienen n canales, y N está correlacionado con n . En particular, el número de señales de audio basadas en objetos N no está correlacionada completamente con el número de canales n .

5 Con referencia nuevamente a 910, en una configuración, después de ajustar/modificar la amplitud, si es necesario, de una o más de las señales de audio basadas en objetos recibidas, el aparato puede asignar señales de audio basadas en objetos a altavoces particulares (o canales) al generar el conjunto de señales de salida. Cada señal de audio basada en objetos se puede asignar a un conjunto de altavoces más cercano (por ejemplo, tres altavoces en un patrón de triángulo) en función de la ubicación de la señal de audio basada en objetos y en función de las ubicaciones de los altavoces más cercanas a la señal de audio basada en objetos. En otra configuración, el conjunto de señales de salida se proporciona a otro aparato que realiza la asignación a altavoces particulares (canales).

15 La figura 10 es un diagrama de flujo 1000 de un método de procesamiento de señales de audio basadas en objetos para al menos uno de transmisión, entrega de archivos o emisión en continuo. El método puede ser realizado por un procesador, como una CPU. El procesador se conoce como un aparato en la figura 10. En 1002, el aparato recibe una pluralidad de señales de audio basadas en objetos. Cada señal de audio basada en objetos de las señales de audio basadas en objetos incluye datos de forma de onda de audio y metadatos de objeto asociados con los datos de forma de onda de audio. Los metadatos del objeto incluyen al menos uno de un parámetro de sonoridad o un parámetro de potencia asociado con los datos de la forma de onda de audio. En 1004, el aparato determina una métrica de sonoridad basada en las señales de audio basadas en objetos recibidas y en el al menos uno del parámetro de sonoridad o el parámetro de potencia para cada señal de audio basada en objetos de las señales de audio recibidas basadas en objetos. En 1012, el aparato transmite las señales de audio basadas en objetos recibidas en función de la métrica de sonoridad determinada.

25 En una configuración, en 1006, el aparato puede determinar un desplazamiento de métrica de sonoridad sobre la base de una comparación entre la métrica de sonoridad y una métrica de sonoridad objetivo. En 1008, el aparato puede determinar si la métrica de sonoridad es mayor que la métrica de sonoridad objetivo. Si se determina que la métrica de sonoridad es mayor que la métrica de sonoridad objetivo, en 1010, el aparato puede ajustar una amplitud de al menos una de las señales de audio basadas en objetos recibidas en función de la métrica de sonoridad determinada. Posteriormente, en 1012, el aparato puede transmitir las señales de audio basadas en objetos recibidas con la amplitud ajustada de al menos una de las señales de audio basadas en objetos recibidas. Alternativamente, en 1010, el aparato puede modificar una amplitud de al menos una de las señales de audio basadas en objetos recibidas en metadatos específicos del programa de un flujo de bits de programa basado en la métrica de sonoridad determinada. Posteriormente, en 1012, el aparato puede transmitir las señales de audio basadas en objetos recibidas con la amplitud modificada en los metadatos específicos del programa.

40 La figura 11 es un diagrama de flujo de datos conceptual 1100 que ilustra el flujo de datos entre diferentes módulos/medios/componentes en un aparato a modo de ejemplo 1102. El aparato es para procesar señales de audio basadas en objetos para su reproducción a través de un sistema de reproducción o para transmisión a través de transmisión, entrega de archivos o emisión en continuo. El aparato 1102 incluye un módulo de recepción 1104 que está configurado para recibir una pluralidad de señales de audio basadas en objetos. Cada señal de audio basada en objetos de las señales de audio basadas en objetos incluye datos de forma de onda de audio y metadatos de objeto asociados con los datos de forma de onda de audio. Los metadatos del objeto incluyen al menos uno de un parámetro de sonoridad o un parámetro de potencia asociado con los datos de la forma de onda de audio. El aparato 1102 incluye además un módulo de determinación de la métrica de sonoridad 1106 que está configurado para determinar una métrica de sonoridad basada en las señales de audio basadas en objetos recibidas y en el al menos uno del parámetro de sonoridad o el parámetro de potencia para cada señal de audio basada en objetos de cada señal de audio basada en objetos recibida. El aparato 1102 incluye además un módulo de transformación/transmisión 1108 que puede configurarse para procesar las señales de audio basadas en objetos recibidas en un conjunto de señales de salida basadas en la métrica de sonoridad determinada. Alternativamente, el módulo de transformación/transmisión 1108 está configurado para transmitir las señales de audio basadas en objetos recibidas en función de la métrica de sonoridad determinada.

55 En una configuración, el módulo de representación/transmisión 1108 está configurado para ajustar una amplitud de al menos una de las señales de audio basadas en objetos recibidas basándose en la métrica de sonoridad determinada. En una configuración, el módulo de determinación de la métrica de sonoridad 1106 puede configurarse para determinar un desplazamiento de la métrica de sonoridad basándose en una comparación entre la métrica de sonoridad y una métrica de sonoridad objetivo. El módulo de transformación/transmisión 1108 puede configurarse para determinar la amplitud de al menos una de las señales de audio basadas en objetos recibidas basándose en la comparación entre la métrica de sonoridad y la métrica de sonoridad objetivo. En una configuración, el módulo de determinación de la métrica de sonoridad 1106 puede configurarse para determinar si la métrica de sonoridad es mayor que la métrica de sonoridad objetivo. El módulo de transformación/transmisión 1108 puede configurarse para ajustar la amplitud de al menos una de las señales de audio basadas en objetos recibidas cuando se determina que la métrica de sonoridad es mayor que la métrica de sonoridad objetivo.

En una configuración, el módulo de representación/transmisión 1108 puede ajustar una amplitud de al menos una de las señales de audio basadas en objetos recibidas basándose en la métrica de sonoridad determinada antes de transmitir las señales de audio basadas en objetos recibidas. En una configuración, las señales de audio basadas en objetos están asociadas con un flujo de bits del programa, y el módulo de transformación/transmisión 1108 está configurado para modificar una amplitud de al menos una de las señales de audio basadas en objetos recibidas en metadatos específicos del programa del flujo de bits del programa basado en la métrica de sonoridad determinada antes de transmitir las señales de audio basadas en objetos recibidas.

En una configuración, el módulo de determinación de la métrica de sonoridad 1106 puede estar configurado para determinar la métrica de sonoridad más basado en, para cada señal de audio basada en objetos de las señales de audio basadas en objetos recibidas, datos de posición de la señal de audio basada en objetos con respecto a un punto de recepción. Los datos de posición pueden incluir al menos uno de un acimut entre una ubicación de la señal de audio basada en objetos y el punto de recepción, o una elevación entre la ubicación de la señal de audio basada en objetos y el punto de recepción. Los datos de posición pueden incluir además una distancia entre la ubicación de la señal de audio basada en objetos y el punto de recepción. El módulo de determinación de la métrica de sonoridad 1106 puede determinar la métrica de la sonoridad basándose además en, para cada señal de audio basada en objetos de las señales de audio basadas en objetos recibidas, al menos uno de un factor de correlación de la señal de audio basada en objetos con respecto a uno o más otras señales de audio basadas en objetos de las señales de audio recibidas basadas en objetos, un factor de escala de amplitud de la señal de audio basada en objetos, o un factor de peso dependiente de la frecuencia de la señal de audio basada en objetos para tener en cuenta la sensibilidad de la audición como función de una ubicación angular de la señal de audio basada en objetos con respecto a un punto de recepción.

En una configuración, el módulo de determinación de la métrica de sonoridad 1106 puede determinar la métrica de sonoridad basado en $\sum_i^N FJJT(r_i, r_L, \theta_i, \phi_i, a_i, g_i, c_i)z_i$, donde $i \in I$, I es el conjunto de señales de audio basadas en objetos, N es una serie de señales de audio basadas en objetos en las señales de audio recibidas basadas en objetos, z_i se dice al menos uno de los parámetros de sonoridad o potencia para una i -ésima señal de audio basada en objetos, r_i es una ubicación asociada con la i -ésima señal de audio basada en objetos, r_L es una ubicación asociada con un punto de recepción está un acimut entre la ubicación de la i -ésimo objeto basada en la señal de audio y el punto de recepción, θ_i es una elevación entre la ubicación de la i -ésima señal de audio basada en el objeto y el punto de recepción, ϕ_i es un factor de escala de amplitud de la i -ésima señal de audio basada en el objeto, g_i es un factor de peso dependiente de la frecuencia de la señal de audio basada en el objeto de i -ésimo para dar cuenta de la sensibilidad de la audición de la audición como una función de una ubicación angular de la señal de audio basada en el objeto de i -ésimo con respecto al punto de recepción, y c_i es una correlación factor de la i -ésima señal de audio basada en objetos con respecto a uno o más otras señales de audio basadas en objetos de las señales de audio recibidas basadas en objetos.

En una configuración, el módulo de recepción 1104 puede recibir la entrada específica del usuario que indica las señales de audio basadas en objetos recibidas. En una configuración, el módulo de recepción 1104 puede recibir información que indica un nuevo conjunto de señales de audio basadas en objetos. La información que indica el conjunto de señales de audio basadas en objetos se puede recibir según la entrada del usuario. En tal configuración, el módulo de determinación de la métrica de sonoridad 1106 puede volver a determinar la métrica de sonoridad basándose en el conjunto de señales de audio basadas en objetos y en la base de al menos uno de los parámetros de sonoridad o el parámetro de potencia para cada señal de audio basada en objetos del conjunto de señales de audio basadas en objetos.

En una configuración, los metadatos de objeto de cada señal de audio basada en objetos incluyen al menos uno de los datos de posición de la señal de audio basada en objetos con respecto a un punto de recepción, o un factor de escala de amplitud de la señal de audio basado en objetos. En una configuración, los datos de posición de la señal de audio basada en objetos con respecto al punto de recepción incluyen al menos uno de una distancia entre una ubicación de la señal de audio basada en objetos y el punto de recepción, un acimut entre la ubicación del objeto basada en la señal de audio y el punto de recepción, o una elevación entre la ubicación de la señal de audio basada en objetos y el punto de recepción.

En una configuración, la métrica de sonoridad es un volumen promedio sobre T ms, y la métrica de sonoridad se determina cada T ms. En una configuración, la métrica de sonoridad es una sonoridad promedio sobre T ms, y la métrica de sonoridad se determina cada D ms donde $D < T$. En una configuración, la medición de sonoridad se determina sobre una pluralidad de períodos de ventana, cada uno de los cuales se superpone una ventana anterior. En una configuración, cada uno de los períodos de ventana tiene una duración de 400 ms y se superpone a otros períodos de ventana 300 ms. En una configuración, las señales de audio basadas en objetos recibidas incluyen N señales de audio basadas en objetos, las señales de audio recibidas basadas en objetos se procesan a través de altavoces que tienen n canales y N no está correlacionada con n .

El aparato puede incluir módulos adicionales que realizan cada uno de los bloques del algoritmo en los diagramas de flujo mencionados anteriormente de las figuras. 9, 10. Como tal, cada bloque en los diagramas de flujo de las figuras 9, 10 puede ser realizado por un módulo y el aparato puede incluir uno o más de esos módulos. Los módulos pueden

ser uno o más componentes de hardware específicamente configurados para llevar a cabo los procesos/algoritmos establecidos, implementados por un procesador configurado para realizar los procesos/algoritmos establecidos, almacenados en un medio legible por ordenador para ser implementados por un procesador, o alguna combinación de ellos.

5 En una configuración, el aparato para procesar señales de audio basadas en objetos para la reproducción a través de un sistema de reproducción incluye medios para recibir una pluralidad de señales de audio basadas en objetos. Cada señal de audio basada en objetos de las señales de audio basadas en objetos incluye datos de forma de onda de audio y metadatos de objeto asociados con los datos de forma de onda de audio. Los metadatos del objeto incluyen
10 al menos uno de un parámetro de sonoridad o un parámetro de potencia asociado con los datos de la forma de onda de audio. El aparato incluye además medios para determinar una métrica de sonoridad basándose en las señales de audio basadas en objetos recibidas y en el al menos uno del parámetro de sonoridad o el parámetro de potencia para cada señal de audio basada en objetos de las señales de audio recibidas basadas en objetos. El aparato incluye además medios para representar las señales de audio basadas en objetos recibidas a un conjunto de señales de salida basadas en la métrica de sonoridad determinada.

En una configuración, los medios para hacer las señales de audio basadas en objetos recibidas están configurados para ajustar una amplitud de al menos una de las señales de audio basadas en objetos recibidas basándose la métrica de sonoridad determinada. En una configuración, el aparato incluye medios para determinar un desplazamiento de la métrica de sonoridad basándose en una comparación entre la métrica de sonoridad y una métrica de sonoridad objetivo. En tal configuración, la amplitud de la al menos una de las señales de audio basadas en objetos recibidas se ajusta en función de la comparación entre la métrica de sonoridad y la métrica de sonoridad objetivo. En una configuración, el aparato incluye medios para determinar si la métrica de sonoridad es mayor que la métrica de sonoridad objetivo. En tal configuración, la amplitud de la al menos una de las señales de audio basadas en objetos recibidas se ajusta cuando se determina que la métrica de sonoridad es mayor que la métrica de sonoridad objetivo.

En una configuración, un aparato para procesar señales de audio basadas en objetos para al menos uno de difusión, de distribución de ficheros, o emisión en continuo incluye medios para recibir una pluralidad de señales de audio basadas en objetos. Cada señal de audio basada en objetos de las señales de audio basadas en objetos incluye datos de forma de onda de audio y metadatos de objeto asociados con los datos de forma de onda de audio. Los metadatos del objeto incluyen al menos uno de un parámetro de sonoridad o un parámetro de potencia asociado con los datos de la forma de onda de audio. El aparato incluye además medios para determinar una métrica de sonoridad basándose en las señales de audio basadas en objetos recibidas y en el al menos uno del parámetro de sonoridad o el parámetro de potencia para cada señal de audio basada en objetos de las señales de audio recibidas basadas en objetos. El aparato incluye además medios para transmitir las señales de audio basadas en objetos recibidas basándose la métrica de sonoridad determinada.

Los diversos bloques, módulos, métodos y procesos de algoritmos lógicos ilustrativos y las secuencias descritas en conexión con las realizaciones descritas en este documento pueden implementarse como hardware electrónico, software informático, o combinaciones de ambos. Para ilustrar claramente esta capacidad de intercambio de hardware y software, varios componentes, bloques, módulos y acciones de proceso ilustrativos se han descrito *supra* en general en términos de su funcionalidad. Si dicha funcionalidad se implementa como hardware o software depende de la aplicación particular y las restricciones de diseño impuestas en el sistema en general. La funcionalidad descrita puede implementarse de diferentes maneras para cada aplicación en particular, pero tales decisiones de implementación no deben interpretarse como una desviación del alcance de este documento.

Los diversos bloques y módulos lógicos ilustrativos descritos en conexión con las realizaciones dadas a conocer en el presente documento pueden implementarse o realizarse por una máquina, tal como un procesador de propósito general, un procesador de señal digital (DSP), un circuito integrado de aplicación específica (ASIC), una matriz de compuerta programable de campo (FPGA) u otro dispositivo lógico programable, compuerta discreta o lógica de transistor, componentes de hardware discretos o cualquier combinación de los mismos diseñada para realizar las funciones descritas en este documento. Un procesador de propósito general puede ser un microprocesador, pero en la alternativa, el procesador puede ser un controlador, un microcontrolador o una máquina de estados, combinaciones de los mismos o similares. Un procesador también puede implementarse como una combinación de dispositivos informáticos, como una combinación de un DSP y un microprocesador, una pluralidad de microprocesadores, uno o más microprocesadores en conjunto con un núcleo de DSP, o cualquier otra configuración similar.

Las realizaciones del sistema de gestión de la sonoridad OBA de normalización espacial y el método descrito en el presente documento son operativos dentro de numerosos tipos de entornos o configuraciones de sistemas informáticos de propósito general o de propósito especial. En general, un entorno informático puede incluir cualquier tipo de sistema informático, incluido, entre otros, un sistema informático basado en uno o más microprocesadores, un ordenador central, un procesador de señales digitales, un dispositivo informático portátil, un organizador personal, un controlador de dispositivo, un motor informático dentro de un dispositivo, un teléfono móvil, un ordenador de escritorio, un ordenador móvil, una tableta, un teléfono inteligente y dispositivos con un ordenador integrado, por nombrar algunos.

Tales dispositivos informáticos pueden ser normalmente encontrados en los dispositivos que tienen al menos alguna capacidad informática mínima, incluyendo, pero no limitados a, ordenadores personales, ordenadores servidores, dispositivos informáticos de mano, ordenadores portátiles o móviles, dispositivos de comunicación tales como teléfonos y PDA, sistemas multiprocesador, sistemas basados en microprocesadores, decodificadores, dispositivos electrónicos de consumo programables, PC de red, miniordenadores, ordenadores centrales, reproductores de medios de audio o video, etc. En algunas realizaciones, los dispositivos informáticos incluirán uno o más procesadores. Cada procesador puede ser un microprocesador especializado, como un DSP, una palabra de instrucción muy larga (VLIW) u otro microcontrolador, o puede ser una CPU convencional que tenga uno o más núcleos de procesamiento, incluyendo una unidad de procesamiento de gráficos especializada (GPU). núcleos en una CPU de múltiples núcleos.

Las acciones del proceso de un método, proceso o algoritmo descrito en conexión con las realizaciones descritas en este documento pueden realizarse directamente en hardware, en un módulo de software ejecutado por un procesador, o en cualquier combinación de los dos. El módulo de software puede estar contenido en medios legibles por ordenador a los que se puede acceder mediante un dispositivo informático. Los medios legibles por ordenador incluyen medios volátiles y no volátiles que pueden ser removibles, no removibles o una combinación de ellos. Los medios legibles por ordenador se utilizan para almacenar información como instrucciones legibles por ordenador o ejecutables por ordenador, estructuras de datos, módulos de programas u otros datos. A modo de ejemplo, y no de limitación, los medios legibles por ordenador pueden comprender "medios de almacenamiento informático" y "medios de comunicación".

Los medios de almacenamiento informático incluyen, pero no se limitan a, equipo o máquina de medios o dispositivos de almacenamiento legible tal como dispositivos ópticos de almacenamiento, discos Blu-ray (BD), discos versátiles digitales (DVD), discos compactos (CD), disquetes, unidades de cinta, discos duros, unidades ópticas, dispositivos de memoria de estado sólido, memoria de acceso aleatorio (RAM), memoria ROM, memoria EPROM, memoria EEPROM, memoria flash u otra tecnología de memoria, casetes magnéticos, cintas magnéticas, almacenamiento en disco magnético, u otros dispositivos de almacenamiento magnético, o cualquier otro dispositivo que se pueda utilizar para almacenar la información deseada y al que se pueda acceder mediante uno o más dispositivos informáticos.

Un módulo de software puede residir en la memoria RAM, memoria flash, memoria ROM, memoria EPROM, memoria EEPROM, registros, disco duro, un disco extraíble, un CD-ROM o cualquier otra forma de no transitoria legible por ordenador medio de almacenamiento, medios o almacenamiento físico de ordenador conocido en la técnica. Un medio de almacenamiento a modo de ejemplo se puede acoplar al procesador de tal manera que el procesador pueda leer y escribir información en el medio de almacenamiento. En la alternativa, el medio de almacenamiento puede ser integral al procesador. El procesador y el medio de almacenamiento pueden residir en un ASIC. El ASIC puede residir en un terminal de usuario. Alternativamente, el procesador y el medio de almacenamiento pueden residir como componentes discretos en un terminal de usuario.

La frase "no transitorio" como se usa en este documento significa "duradero o de larga vida." La frase "medios legibles por ordenador no transitorios" incluye todos los medios legibles por ordenador, con la única excepción de una señal de propagación transitoria. Esto incluye, a modo de ejemplo y no como limitación, medios legibles por ordenador no transitorios, como la memoria de registro, la memoria caché del procesador y la RAM.

La retención de información tal como instrucciones legibles por ordenador o ejecutables por ordenador, estructuras de datos, módulos de programas, etc., también se puede lograr usando varios medios de comunicación para codificar una o más señales de datos modulados, ondas electromagnéticas (como ondas portadoras), u otros mecanismos de transporte o protocolos de comunicaciones, e incluye cualquier mecanismo de entrega de información por cable o inalámbrico. El término "señal de datos modulados" significa una señal que tiene una o más de sus características configuradas o modificadas de tal manera que codifican información en la señal. Por ejemplo, los medios de comunicación incluyen medios cableados, como una red cableada o una conexión cableada directa que transportan una o más señales de datos modulados, y medios inalámbricos, como medios acústicos, de radiofrecuencia (RF), infrarrojos, láser y otros medios inalámbricos para transmitir, recibir, o ambos, una o más señales de datos modulados u ondas electromagnéticas. Las combinaciones de cualquiera de los anteriores también deben incluirse dentro del alcance de los medios de comunicación.

Además, una o cualquier combinación de software, programas, productos de programa de ordenador que incorporan algunas o todas de las diversas realizaciones del sistema de gestión de sonoridad espacial OBA de normalización y el método descritos en este documento, o porciones de los mismos, pueden ser almacenados, recibidos, transmitido, o lea de cualquier combinación deseada de medios de almacenamiento o dispositivos de comunicación y medios de comunicación de ordenador o máquina en forma de instrucciones ejecutables por ordenador u otras estructuras de datos.

Las realizaciones del sistema de gestión de la sonoridad del OBA de normalización espacial y el método descrito en el presente documento pueden describirse adicionalmente en el contexto general de instrucciones ejecutables por ordenador, tales como módulos de programa, que se ejecutan mediante un dispositivo informático. Generalmente, los módulos de programa incluyen rutinas, programas, objetos, artefactos físicos, estructuras de datos, etc. que realizan tareas particulares o implementan tipos de datos abstractos particulares. Las realizaciones descritas en el presente

5 documento también pueden ponerse en práctica en entornos informáticos distribuidos donde las tareas son realizadas por uno o más dispositivos de procesamiento remoto, o dentro de una nube de uno o más dispositivos, que están vinculados a través de una o más redes de comunicaciones. En un entorno informático distribuido, los módulos de programa pueden ubicarse en medios de almacenamiento informáticos locales y remotos, incluidos los dispositivos de almacenamiento multimedia. Aún más, las instrucciones mencionadas anteriormente pueden implementarse, en parte o en su totalidad, como circuitos lógicos de hardware, que pueden incluir o no un procesador.

10 El lenguaje condicional utilizado en este documento, tal como, entre otros, "puede", "podría", "por ejemplo," y similares, a menos que se indique específicamente lo contrario, o entendido de otra manera dentro del contexto tal como se usa, generalmente pretende transmitir que ciertas realizaciones incluyen, mientras que otras realizaciones no incluyen ciertas características, elementos y/o estados. Por lo tanto, dicho lenguaje condicional generalmente no pretende implicar que las características, elementos y/o estados sean de alguna manera necesarios para una o más realizaciones o que una o más realizaciones necesariamente incluyan lógica para decidir, con o sin la aportación del autor o las preguntas, ya sea estas características, elementos y/o estados están incluidos o se realizarán en cualquier realización particular. Los términos "que comprende", "que incluye", "que tienen" y similares son sinónimos y se usan inclusive, de manera abierta, y no excluyen elementos, características, actos, operaciones, etc. adicionales. Además, el término "o" se usa en su sentido inclusivo (y no en su sentido exclusivo) de modo que cuando se usa, por ejemplo, para conectar una lista de elementos, el término "o" significa uno, algunos o todos los elementos de la lista.

15

REIVINDICACIONES

1. Un método para procesar señales de audio basadas en objetos para su reproducción a través de un sistema de reproducción, que comprende:

5 recibir una pluralidad de señales de audio basadas en objetos, incluyendo cada señal de audio basada en objetos de las señales de audio basadas en objetos datos de forma de onda de audio y metadatos de objeto asociados a los datos de forma de onda de audio, incluyendo los metadatos de objeto al menos uno de un parámetro de sonoridad o un parámetro de potencia asociados a los datos de forma de onda de audio;

10 determinar una métrica de sonoridad basada en las señales de audio basadas en objetos recibidas y basada en dicho al menos uno del parámetro de sonoridad o el parámetro de potencia para cada señal de audio basada en objetos de las señales de audio basadas en objetos recibidas; y
convertir las señales de audio basadas en objetos recibidas en un conjunto de señales de salida basadas en la métrica de sonoridad determinada;

15 **caracterizado por que**
la métrica de sonoridad se determina además en función de, para cada señal de audio basada en objetos de las señales de audio basadas en objetos recibidas, datos posicionales de la señal de audio basada en objetos con respecto a un punto de recepción, en donde el punto de recepción es un punto supuesto en un espacio 3D o una ubicación real de un oyente dentro de un espacio 3D.

20 2. El método de la reivindicación 1, en el que la conversión de las señales de audio basadas en objetos recibidas comprende ajustar una amplitud de al menos una de las señales de audio basadas en objetos recibidas basándose en la métrica de sonoridad determinada.

25 3. El método de la reivindicación 2, que comprende además determinar un desplazamiento de la métrica de sonoridad basándose en una comparación entre la métrica de sonoridad y una métrica de sonoridad objetivo, en donde la amplitud de la al menos una de las señales de audio basadas en objetos recibidas se ajusta basándose en la comparación entre la métrica de sonoridad y la métrica de sonoridad objetivo.

30 4. El método de la reivindicación 1, en el que los datos de posición comprenden al menos uno de un acimut entre una ubicación de la señal de audio basada en objetos y el punto de recepción, o una elevación entre la ubicación de la señal de audio basada en objetos y el punto de recepción.

35 5. El método de la reivindicación 1, en el que la métrica de sonoridad se determina además en función de, para cada señal de audio basada en objetos de las señales de audio basadas en objetos recibidas, al menos uno de un factor de correlación de la señal de audio basada en objetos con respecto a una o más señales de audio basadas en objetos de las señales de audio basadas en objetos recibidas, un factor de escala de amplitud de la señal de audio basada en objetos o un factor de peso dependiente de la frecuencia de la señal de audio basada en objetos para tener en cuenta la sensibilidad de la audición como función de una ubicación angular de la señal de audio basada en objetos con respecto a un punto de recepción.

40 6. El método de la reivindicación 1, en el que las señales de audio basadas en objetos recibidas son especificadas por el usuario.

45 7. El método de la reivindicación 6, que comprende además recibir información que indica un nuevo conjunto de señales de audio basadas en objetos, siendo recibida la información que indica el nuevo conjunto de señales de audio basadas en objetos que se recibe sobre la base de la entrada del usuario, comprendiendo el método además volver a determinar la métrica de sonoridad basada en el nuevo conjunto de señales de audio basadas en objetos y basándose en dicho al menos uno de los parámetros de sonoridad o el parámetro de potencia para cada señal de audio basada en objetos del conjunto de señales de audio basadas en objetos.

50 8. El método de la reivindicación 1, en el que los metadatos de objeto de cada señal de audio basada en objetos incluyen al menos uno de los datos de posición de la señal de audio basada en objetos con respecto a un punto de recepción, o un factor de escala de amplitud de la señal de audio basada en objetos.

55 9. El método de la reivindicación 8, en el que los datos posicionales de la señal de audio basada en objetos con respecto al punto de recepción comprenden al menos una distancia entre una ubicación de la señal de audio basada en objetos y el punto de recepción, un acimut entre la ubicación de la señal de audio basada en objetos y el punto de recepción, o una elevación entre la ubicación de la señal de audio basada en objetos y el punto de recepción.

60 10. El método de la reivindicación 1, en el que la métrica de sonoridad es una sonoridad media sobre T ms, y la métrica de sonoridad se determina cada T ms.

65 11. El método de la reivindicación 1, en el que la métrica de sonoridad se determina a lo largo de una pluralidad de periodos de ventana, cada uno de los cuales se solapa con una ventana anterior.

12. El método de la reivindicación 1, en el que las señales de audio basadas en objetos recibidas comprenden N señales de audio basadas en objetos, las señales de audio recibidas basadas en objetos se convierten a través de altavoces que comprenden n canales, en donde el número de señales de audio basadas en objetos N no está correlacionado con el número de canales n .

5 13. Un método para procesar señales de audio basadas en objetos para al menos una transmisión, entrega de archivos o emisión en continuo, que comprende:

10 recibir una pluralidad de señales de audio basadas en objetos, incluyendo cada señal de audio basada en objetos de las señales de audio basadas en objetos datos de forma de onda de audio y metadatos de objeto asociados a los datos de forma de onda de audio, incluyendo los metadatos de objeto al menos uno de un parámetro de sonoridad o un parámetro de potencia asociado a los datos de forma de onda de audio;

15 determinar una métrica de sonoridad basada en las señales de audio basadas en objetos recibidas y basada en dicho al menos uno del parámetro de sonoridad o el parámetro de potencia para cada señal de audio basada en objetos de las señales de audio basadas en objetos recibidas; y

transmitir las señales de audio basadas en objetos recibidas en función de la métrica de sonoridad determinada;
caracterizado por que

20 la métrica de sonoridad se determina además en función de, para cada señal de audio basada en objetos de las señales de audio basadas en objetos recibidas, datos posicionales de la señal de audio basada en objetos con respecto a un punto de recepción, en donde el punto de recepción es un punto supuesto en un espacio 3D o una ubicación real de un oyente dentro de un espacio 3D.

25 14. El método de la reivindicación 13, que comprende además ajustar una amplitud de al menos una de las señales de audio basadas en objetos recibidas basándose en la métrica de sonoridad determinada antes de transmitir las señales de audio basadas en objetos recibidas.

30 15. El método de la reivindicación 13, en el que las señales de audio basadas en objetos están asociadas a un flujo de bits, y el método comprende además la modificación de una amplitud de al menos una de las señales de audio basadas en objetos recibidas en metadatos específicos del programa del flujo de bits basándose en la métrica de sonoridad determinada antes de transmitir las señales de audio basadas en objetos recibidas.

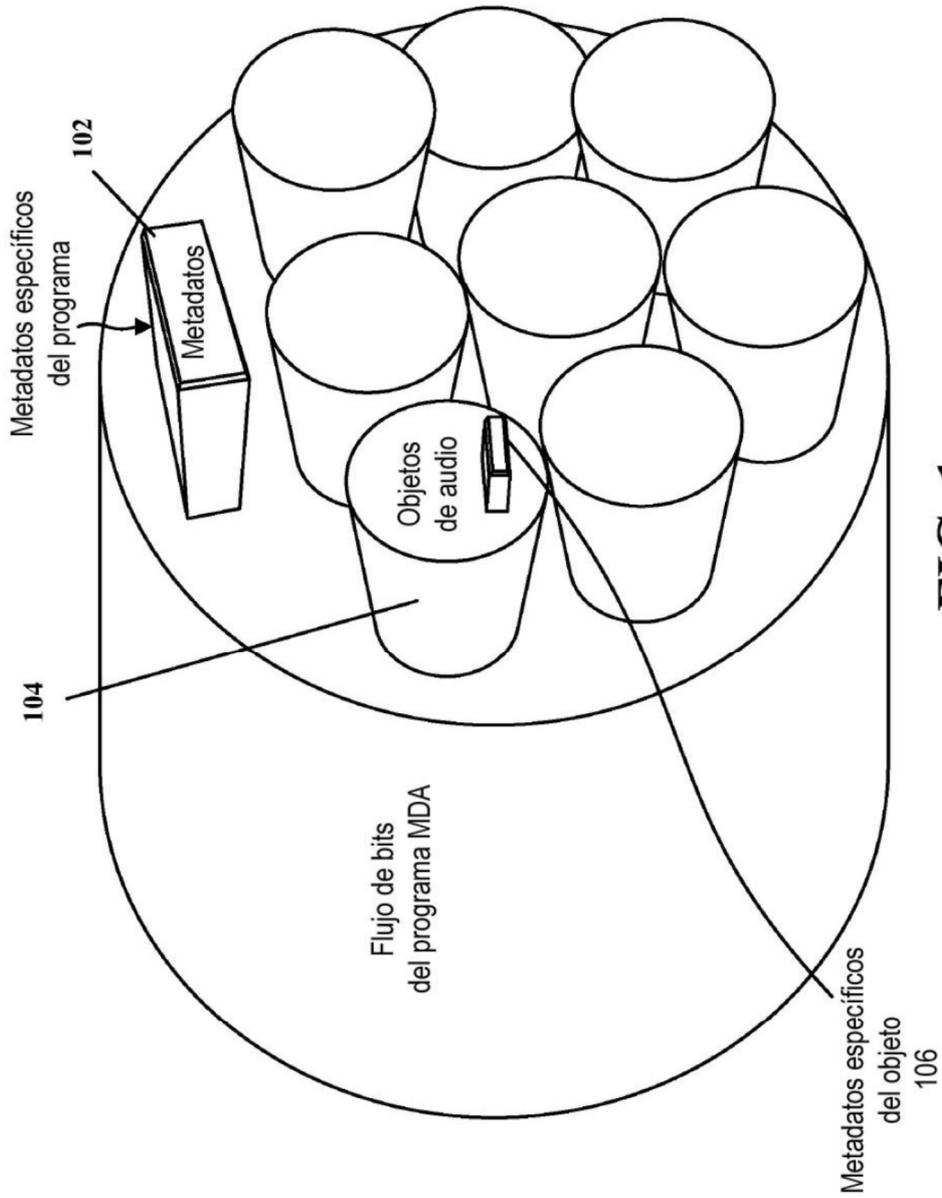


FIG. 1

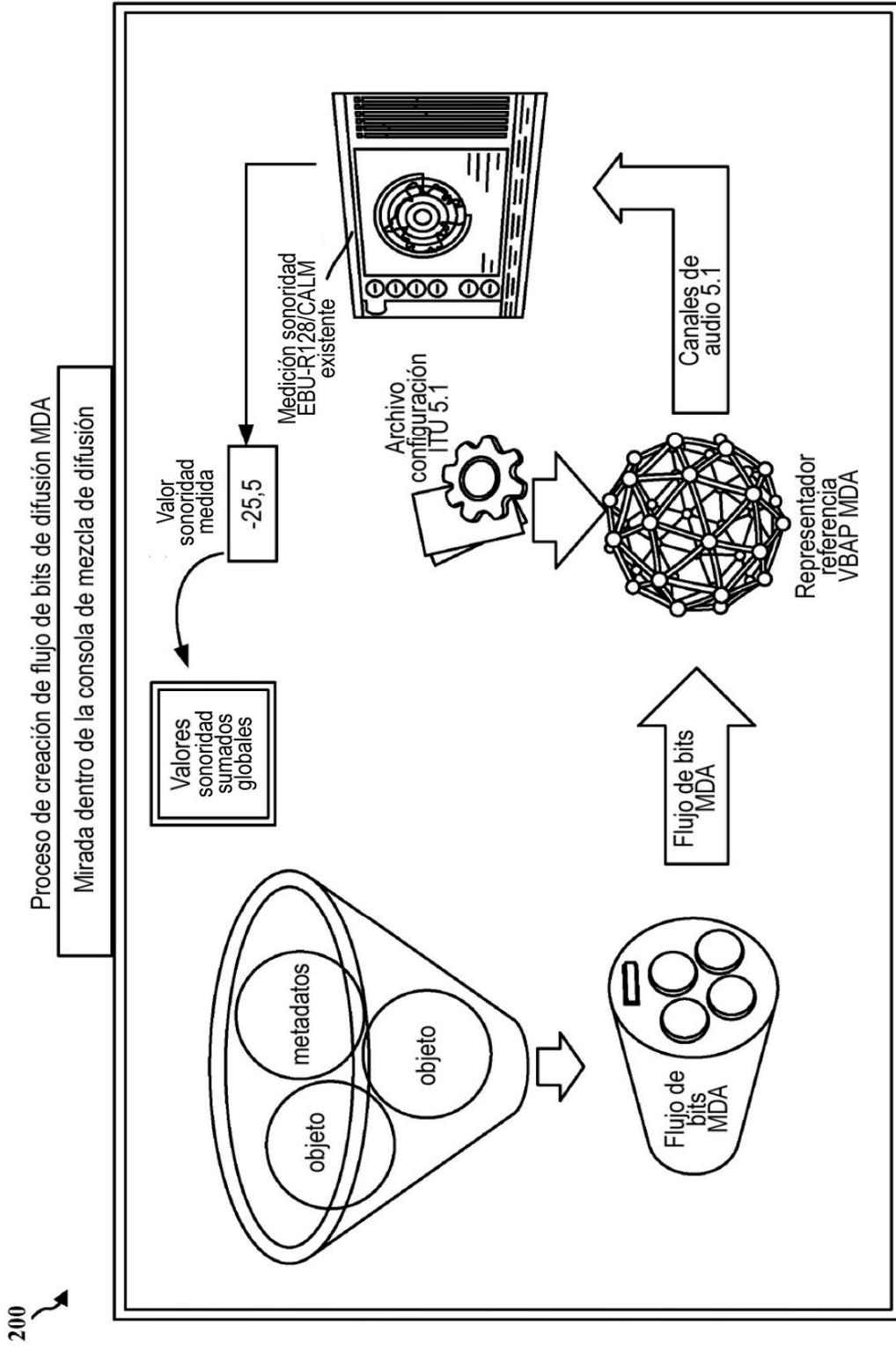


FIG. 2

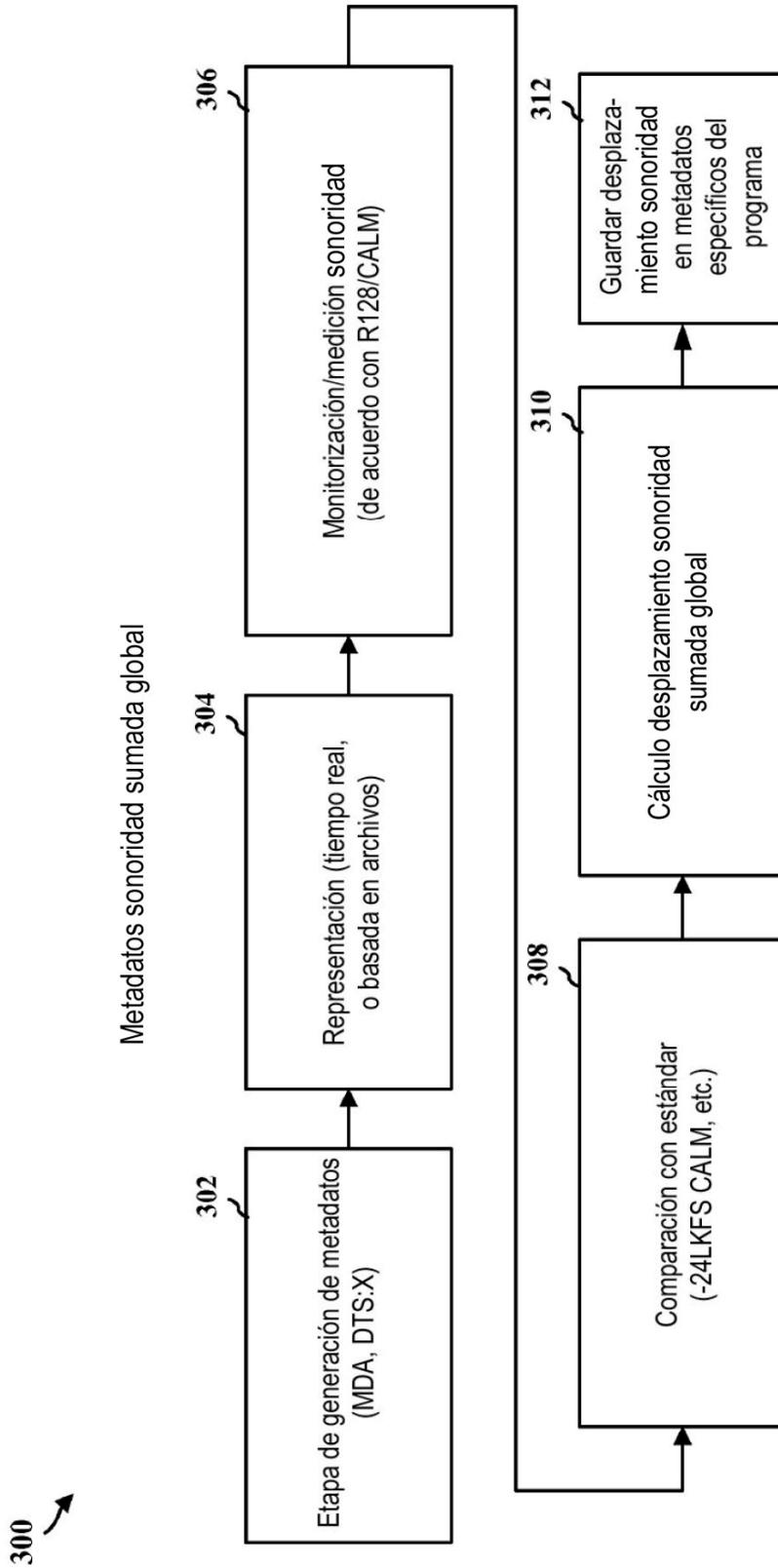


FIG. 3

400 ↗

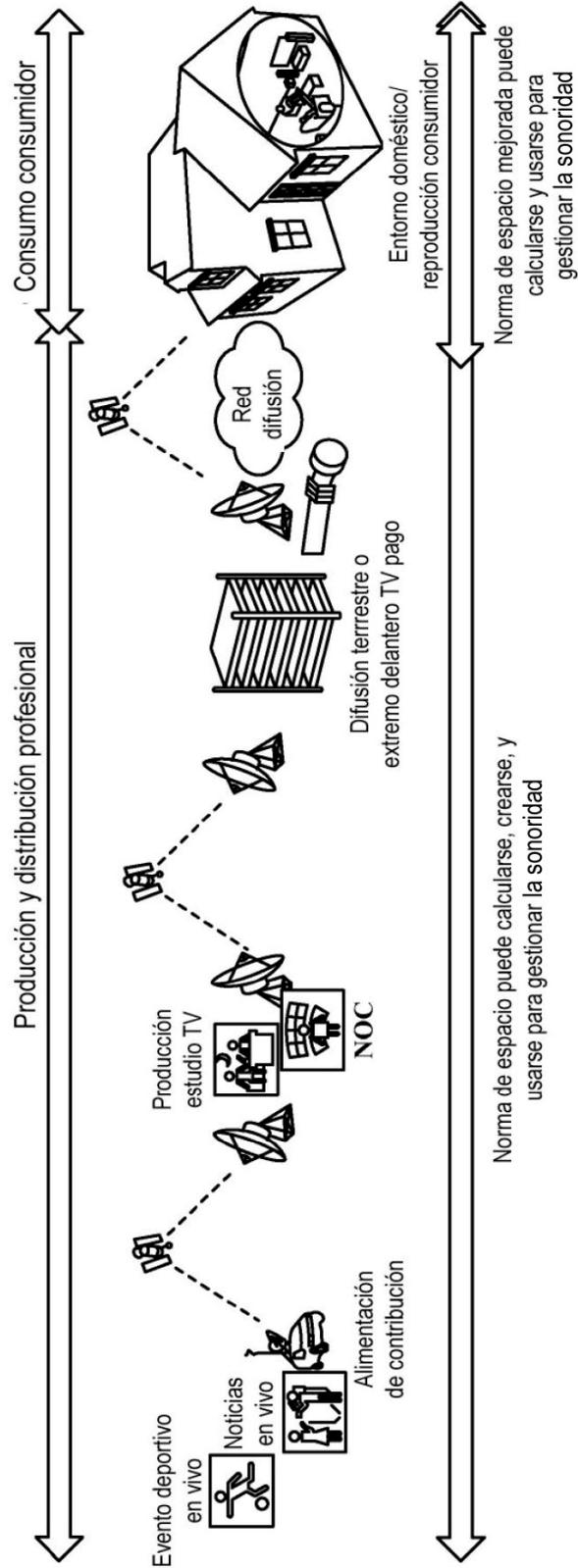


FIG. 4

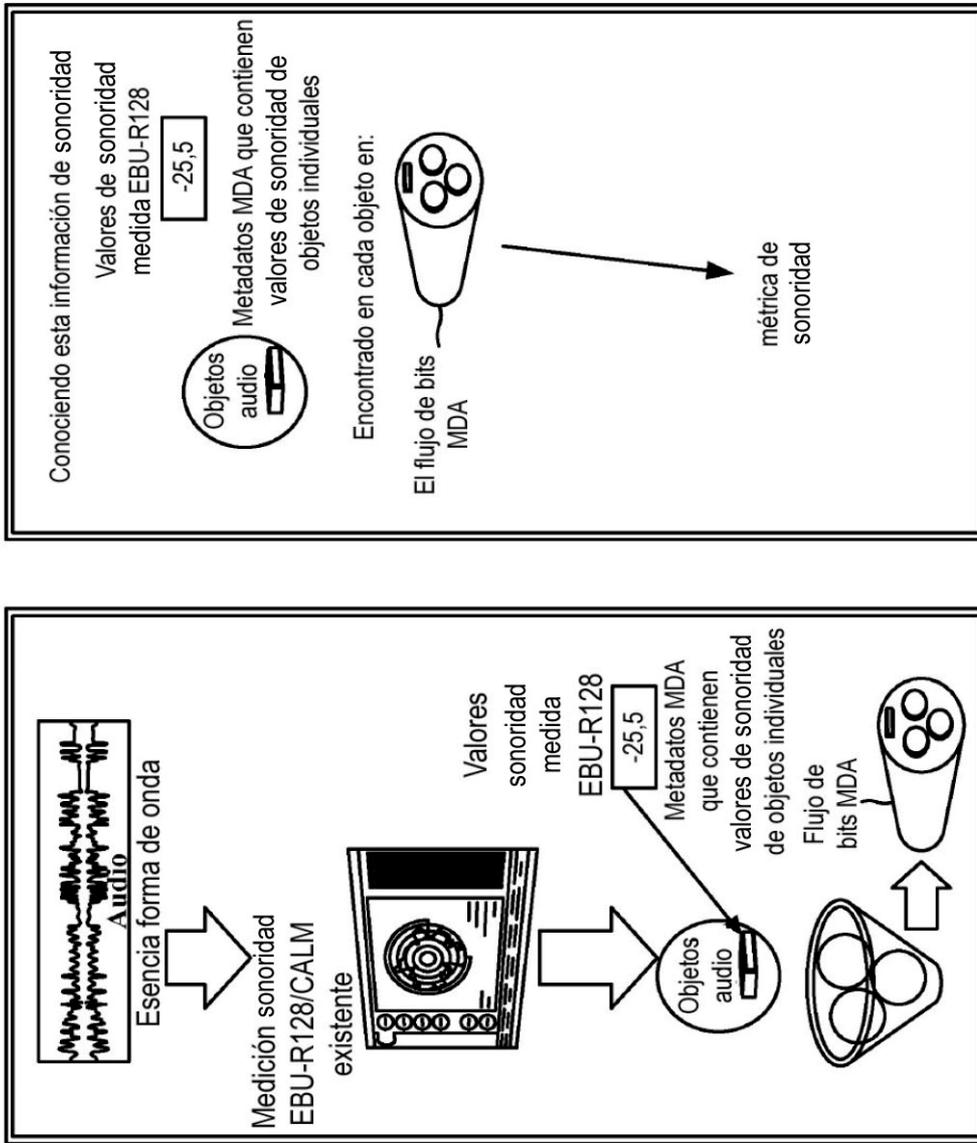


FIG. 5

500 ↗

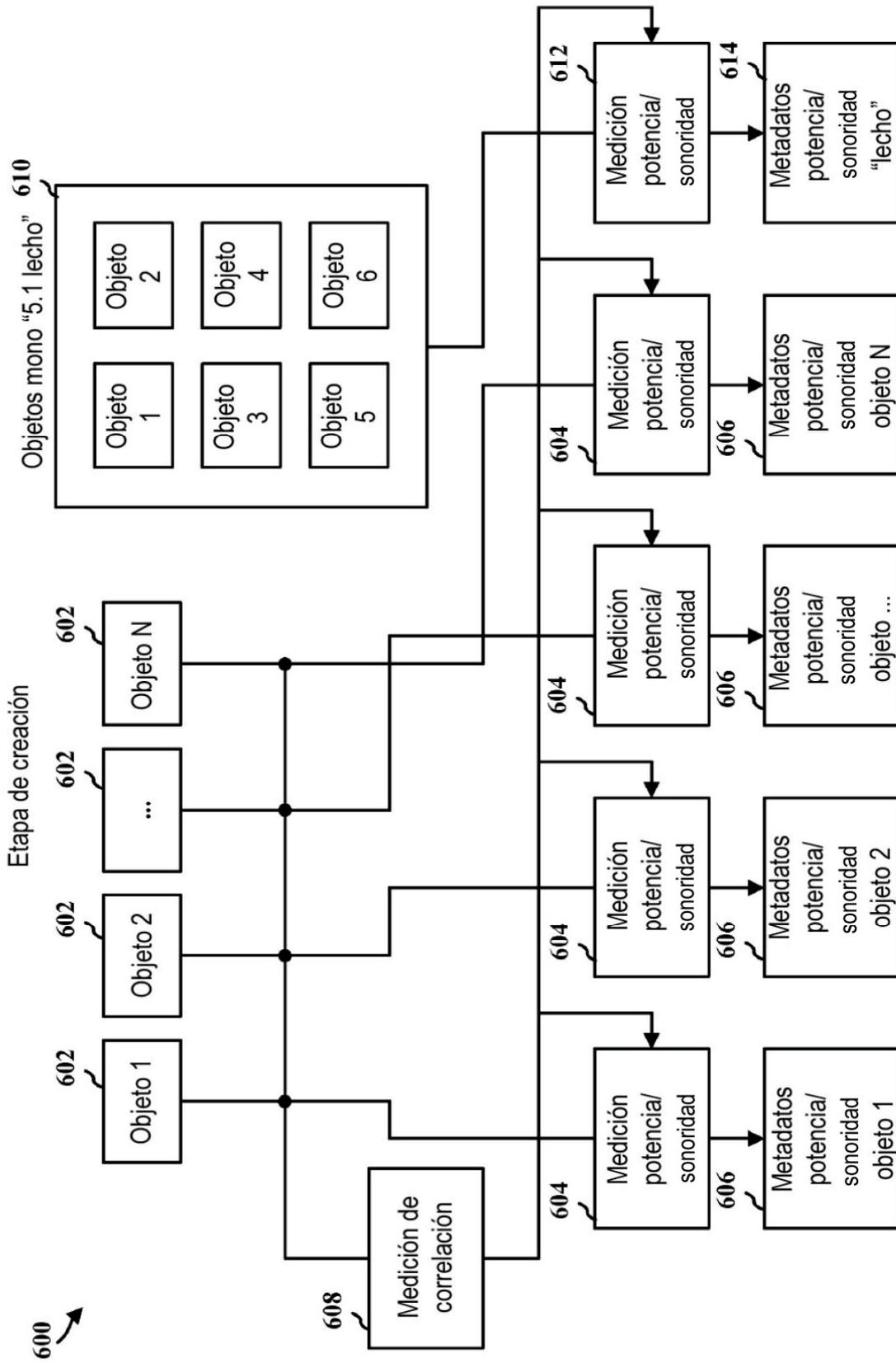
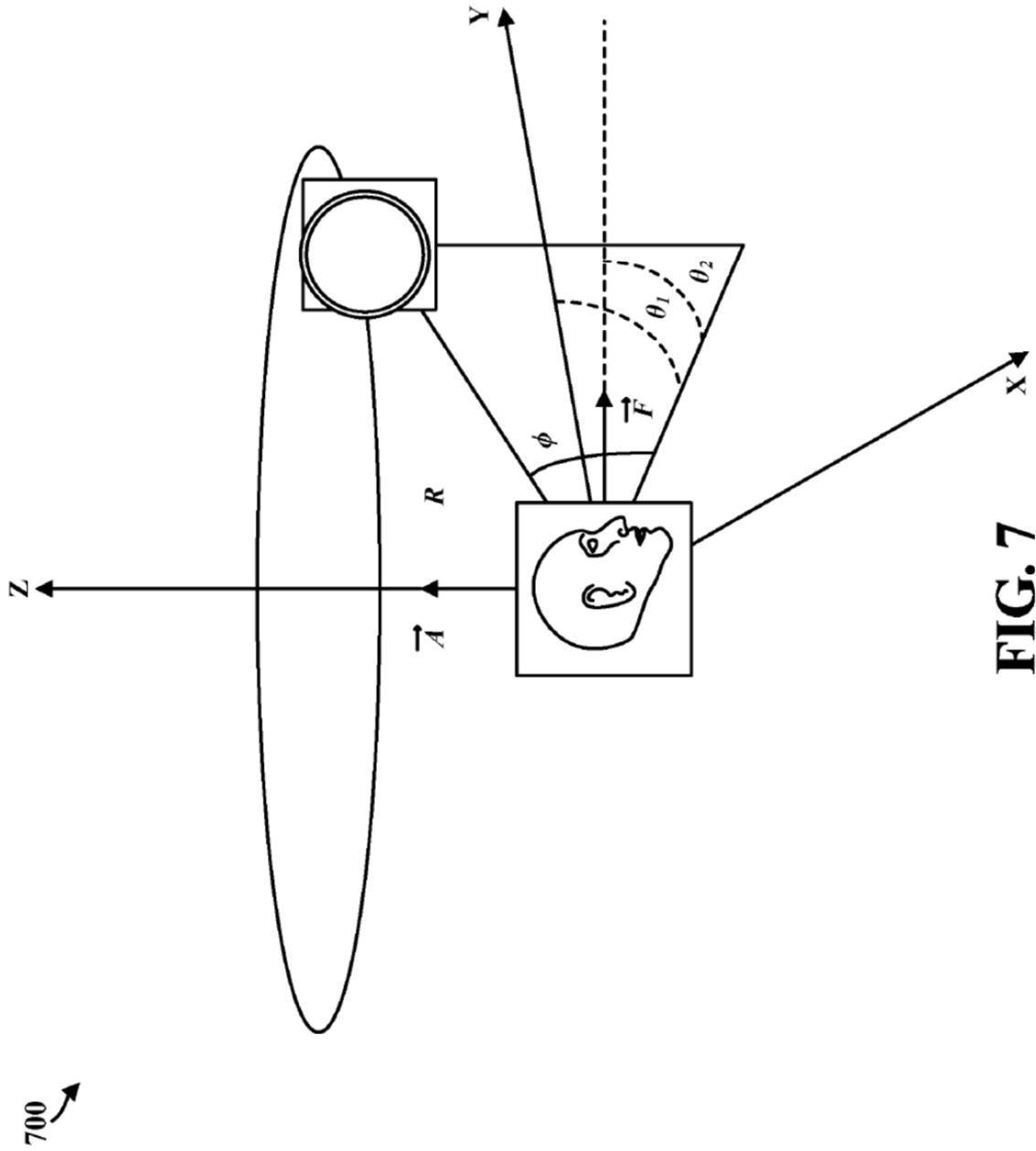


FIG. 6



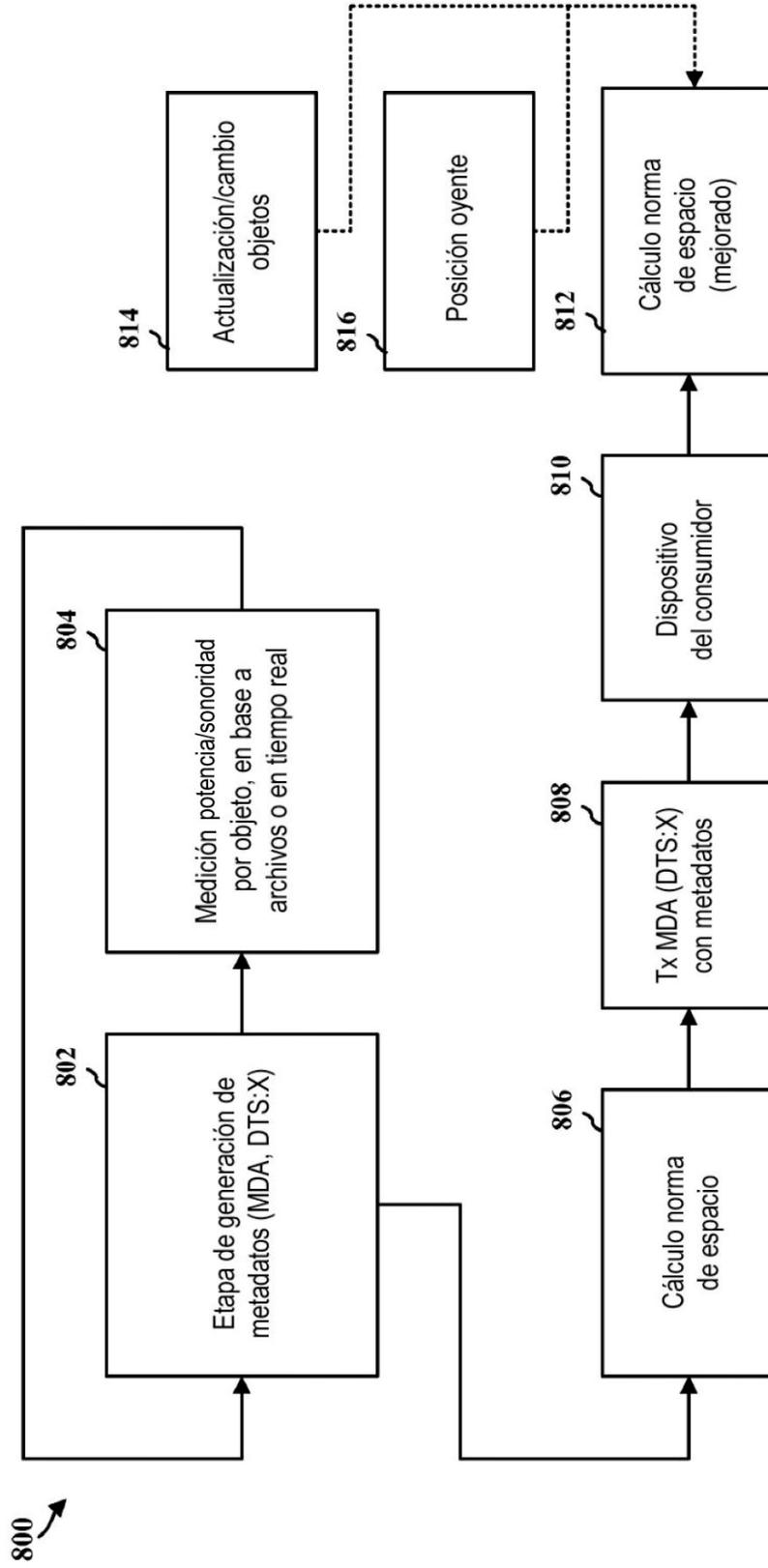


FIG. 8

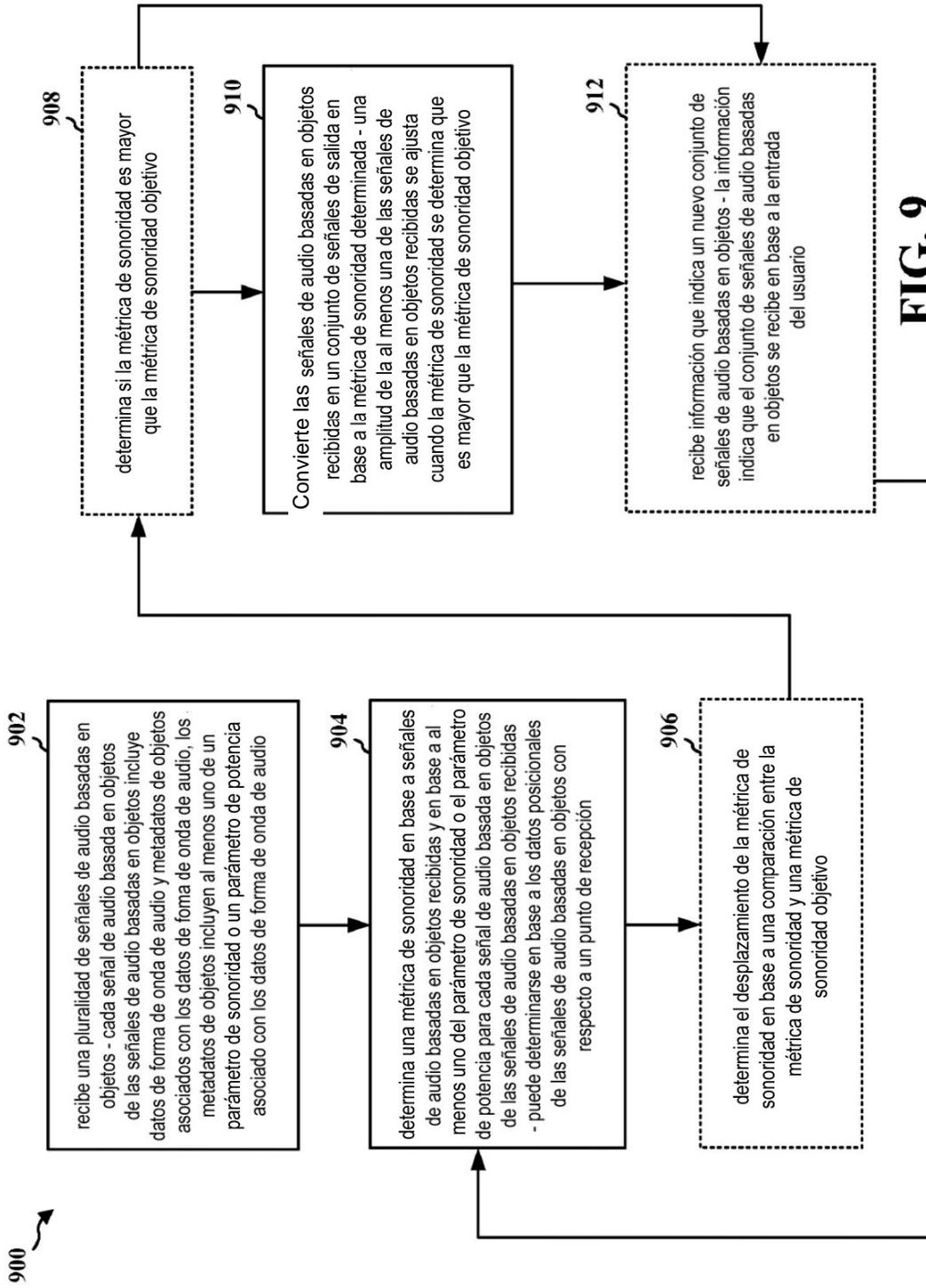


FIG. 9

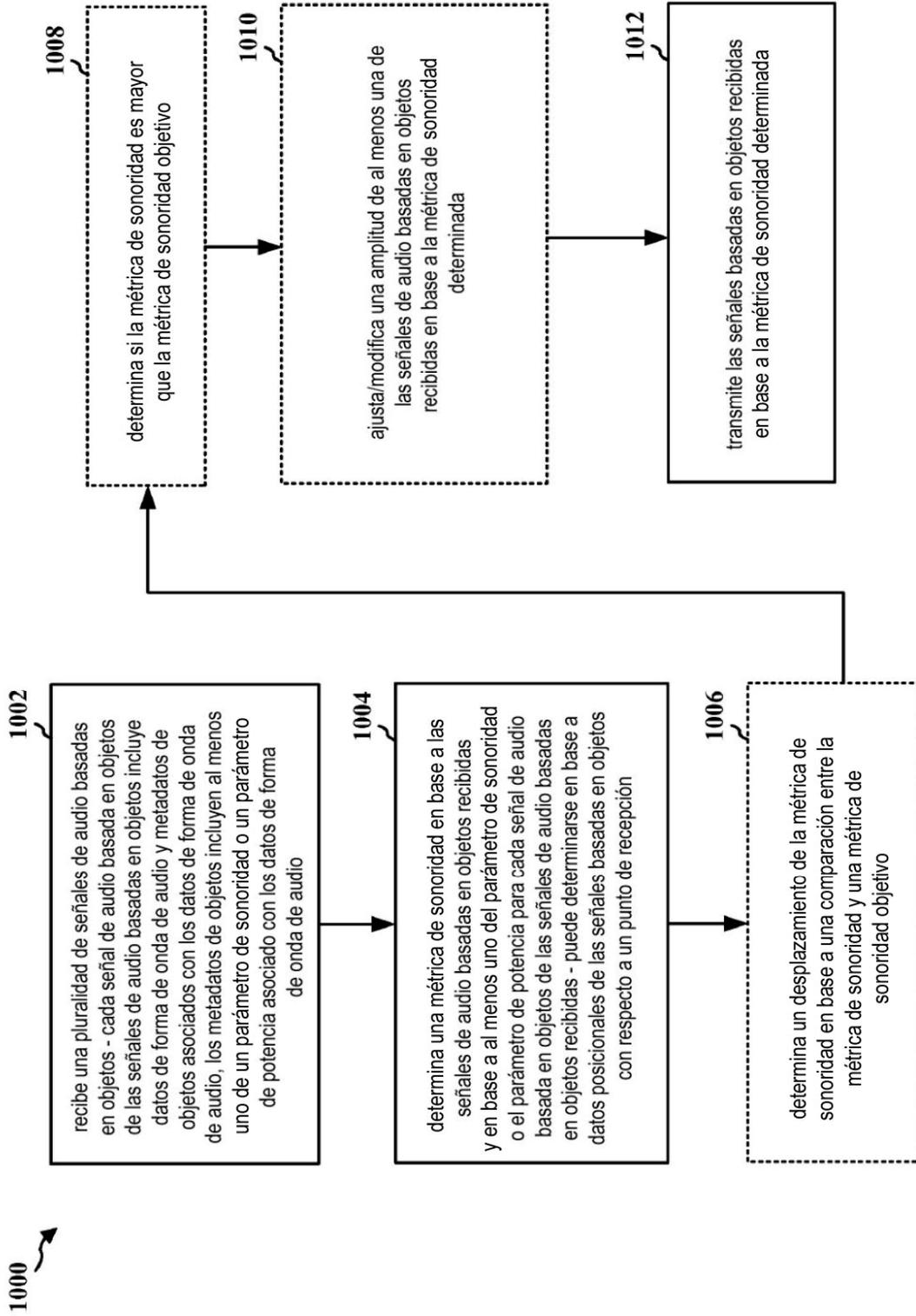


FIG. 10

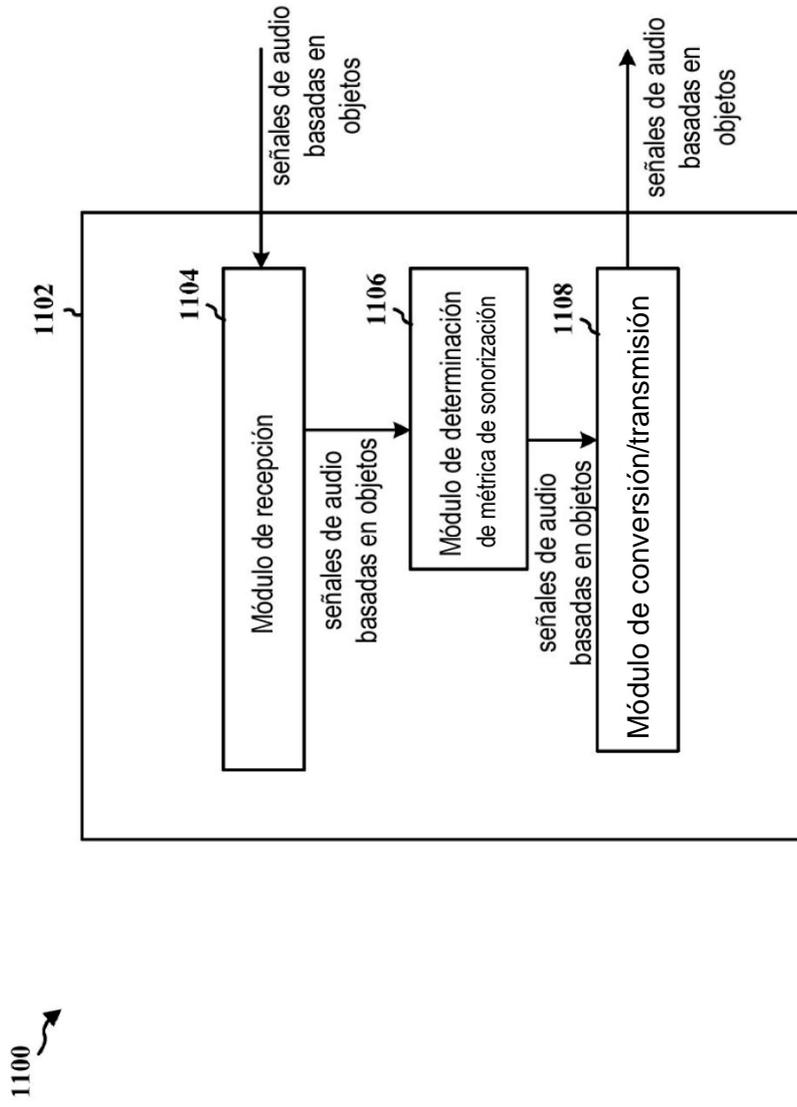


FIG. 11