

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 723 965**

51 Int. Cl.:

<b>H04R 3/12</b>	(2006.01)
<b>H04N 5/64</b>	(2006.01)
<b>H04N 5/74</b>	(2006.01)
<b>H04N 7/06</b>	(2006.01)
<b>H04R 5/02</b>	(2006.01)
<b>G11B 20/10</b>	(2006.01)
<b>H04S 7/00</b>	(2006.01)
<b>H04S 3/00</b>	(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **25.03.2015 PCT/EP2015/056417**
- 87 Fecha y número de publicación internacional: **01.10.2015 WO15144766**
- 96 Fecha de presentación y número de la solicitud europea: **25.03.2015 E 15716439 (3)**
- 97 Fecha y número de publicación de la concesión europea: **20.02.2019 EP 3123741**

54 Título: **Aparato y procedimiento para el remapeo de objetos de audio relacionado con la pantalla**

30 Prioridad:

**26.03.2014 EP 14161819**  
**08.12.2014 EP 14196769**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**04.09.2019**

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN FORSCHUNG E.V. (100.0%)**  
**Hansastraße 27c**  
**80686 München, DE**

72 Inventor/es:

**FÜG, SIMONE;**  
**PLOGSTIES, JAN;**  
**DICK, SASCHA;**  
**HILPERT, JOHANNES;**  
**ROBILLIARD, JULIEN;**  
**KUNTZ, ACHIM y**  
**HÖLZER, ANDREAS**

74 Agente/Representante:

**SALVÀ FERRER, Joan**

ES 2 723 965 T3

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

## DESCRIPCIÓN

Aparato y procedimiento para el remapeo de objetos de audio relacionado con la pantalla

- 5 **[0001]** La presente invención se relaciona con el procesamiento de señales de audio, en particular, con un aparato y un procedimiento para el remapeo de objetos de audio y, más específicamente, con un aparato y un procedimiento para el remapeo de objetos de audio relacionado con la pantalla.
- 10 **[0002]** Con el creciente consumo de contenido de multimedia en la vida diaria, la demanda de soluciones multimedia sofisticadas aumenta constantemente. En este contexto, la integración del contenido visual y de audio desempeña un papel importante. Sería ventajoso un ajuste óptimo del contenido multimedia audiovisual a la configuración de reproducción audiovisual existente.
- 15 **[0003]** En el estado de la técnica, se conocen los objetos de audio. Los objetos de audio se pueden considerar, por ejemplo, como pistas de sonido con metadatos asociados. Los metadatos pueden describir, por ejemplo, las características de los datos de audio no procesados, por ejemplo, la posición de reproducción deseada o el nivel de volumen. Una ventaja del audio basado en objetos es que se puede reproducir un movimiento predefinido mediante un proceso de renderización especial del lado de la reproducción de la mejor manera posible para todas las disposiciones de altavoces para la reproducción.
- 20 **[0004]** Se pueden emplear metadatos geométricos para definir dónde se debe renderizar un objeto de audio, por ejemplo, los ángulos del azimut o elevación o las posiciones absolutas con respecto a un punto de referencia, por ejemplo, el oyente. Los metadatos son almacenados o transmitidos junto con las señales de audio de objetos.
- 25 **[0005]** En el contexto de MPEG-H, en la 105ª reunión de MPEG el grupo de audio revisó los requisitos y los plazos de diferentes normas de aplicación (MPEG = Moving Picture Experts Group, Grupo de expertos en imágenes en movimiento). Según esa revisión, sería esencial cumplir con ciertos puntos de tiempo y requisitos específicos para un sistema de transmisión de la próxima generación. Según eso, un sistema debería poder aceptar objetos de audio en la entrada del codificador. Además, el sistema debería admitir la señalización, transmisión y renderización
- 30 de objetos de audio y debería habilitar el control de los objetos por el usuario, por ejemplo, para la mejora de diálogos, pistas de lenguajes alternativos y lenguaje de descripción de audio.
- [0006]** En el estado de la técnica, se han dado a conocer diferentes conceptos. Según una primera técnica anterior, presentada en "*Method and apparatus for playback of a higher-order ambisonics audio signal*" (véase [1]),
- 35 se adapta la reproducción del audio orientado al campo del sonido espacial a sus objetos visibles asociados mediante la aplicación de un proceso de distorsión espacial (warping). En esa técnica anterior, el decodificador deforma el campo sonoro de tal manera que todos los objetos sonoros en la dirección de la pantalla se compriman o estiren según la relación de los tamaños de las pantallas pretendidas y de referencia. Se incluye la posibilidad de codificar y transmitir el tamaño de referencia (o el ángulo de visualización desde una posición de escucha de referencia) de la pantalla empleada en la producción del contenido en forma de metadatos junto con el contenido.
- 40 Por otro lado, se presume un tamaño de pantalla de referencia fijo al codificar y decodificar, y el decodificador conoce el tamaño real de la pantalla objetivo. En esta técnica anterior, el decodificador distorsiona el campo sonoro de tal manera que todos los objetos sonoros en la dirección de la pantalla se compriman o estiren según la relación del tamaño de la pantalla pretendida y el tamaño de la pantalla de referencia. Se utilizan las denominadas funciones
- 45 de warping o por lineal por segmentos de a dos segmentos ("two-segment piecewise linear"). El estiramiento se limita a las posiciones angulares de los elementos sonoros. En esa técnica anterior, en el caso de las pantallas centradas la definición de la función de distorsión o warping es similar a la definición de la función de mapeo correspondiente al remapeo relacionado con la pantalla. El primero y tercer segmentos de la función de mapeo lineal segmentario de tres segmentos se podrían definir como la función lineal segmentaria de dos segmentos. Sin
- 50 embargo, con esa técnica anterior, la aplicación se limita a HOA (señales HOA = ambisonics de orden superior) (orientado al campo sonoro) en el dominio espacial. Además, la función de distorsión sólo depende de la relación entre la pantalla de referencia y la pantalla de reproducción, y no se presenta definición alguna respecto de las pantallas no centradas.
- 55 **[0007]** En otra técnica anterior, "*Vorrichtung und Verfahren zum Bestimmen einer Wiedergabeposition*" (véase [2]), se describe un procedimiento para adaptar la posición de una fuente de sonido a la reproducción de vídeo. La posición de reproducción de la fuente de sonido se determina de forma individual por cada objeto sonoro dependiendo de la dirección y la distancia hasta el punto de referencia y de los parámetros de la cámara. Esa técnica anterior también describe una pantalla con un tamaño de referencia fijo presunto. Se lleva a cabo el escalado
- 60 lineal de todos los parámetros de posición (en coordenadas cartesianas) para adaptar la escena a una pantalla de reproducción que es más pequeña o más grande que la pantalla de referencia. Sin embargo, según esa técnica anterior, la incorporación de parámetros físicos de la cámara y de proyección es compleja, y no siempre se dispone de esos parámetros. Además, el procedimiento de esa técnica anterior funciona con coordenadas cartesianas (x,y,z), por lo que no sólo la posición sino también la distancia de un objeto cambia con el escalado de la escena.
- 65 Además, esta técnica anterior no es aplicable a la adaptación de la posición del objeto con respecto a los cambios

de tamaño relativo de la pantalla (ángulo de apertura, ángulo de visualización) en las coordenadas angulares.

**[0008]** En una técnica anterior adicional, "*Verfahren zur Audiocodierung*" (véase [3]), se describe un procedimiento que incluye una transmisión del ángulo de visualización horizontal y vertical en curso (variable en el tiempo) en el flujo de datos (ángulo de visualización de referencia, en relación con la posición del oyente en la escena original). Del lado de la reproducción, se analiza el tamaño y la posición de la reproducción y se optimiza la reproducción de los objetos de audio en forma individual para equiparlos a la pantalla de referencia. En otra técnica anterior, "*Acoustical Zooming Based on a parametric Sound Field Representation*" (véase [4]), se describe un procedimiento que proporciona la renderización de audio que sigue el movimiento de la escena visual ("zoom acústico"). El proceso de zoom acústico (acercamiento-alejamiento) se define como desplazamiento de la posición de grabación virtual. El modelo de escena para el algoritmo de zooming coloca todas las fuentes de sonido en un círculo con un radio fijo, aunque arbitrario. Sin embargo, el procedimiento de esa técnica anterior actúa en el dominio de los parámetros DirAC, se cambian la distancia y los ángulos (dirección de llegada), la función de mapeo no es lineal y depende de un factor/parámetro de zoom y no se admiten las pantallas no centradas. En otro documento de la técnica anterior US 2012183162 se describe el procedimiento para la reproducción de sonido localizado para audio inmersivo, comprendiendo el procedimiento: recibir una señal de audio que contiene un primer objeto de audio; acceder a los metadatos asociados con el primer objeto de audio para determinar una primera ubicación perceptual y una segunda ubicación perceptual, los metadatos que incluyen un tamaño de imagen de audio y una posición de imagen de audio; presentar el primer objeto de audio con al menos un primer transductor y un segundo transductor; utilizando la presentación factores de peso para colocar el primer objeto de audio en la primera ubicación perceptual; y mover la presentación del primer objeto de audio a la segunda ubicación perceptual mediante el ajuste de los factores de peso para el segundo transductor y un tercer transductor. El objetivo de la presente invención es proporcionar conceptos mejorados para la integración de contenido multimedia audiovisual empleando configuraciones de reproducción multimedia existentes. El objetivo de la presente invención se resuelve mediante un aparato según la reivindicación 1, mediante un dispositivo decodificador según la reivindicación 13, mediante un procedimiento según la reivindicación 14 y mediante un programa informático según la reivindicación 15.

**[0009]** La invención se define en las reivindicaciones independientes 1, 13 y 14. Las formas de realización preferidas se definen en las reivindicaciones dependientes correspondientes.

**[0010]** A continuación se describen las formas de realización de la presente invención con más detalle con referencia a las figuras, en las cuales:

fig. 1 es un aparato para generar señales de altavoces según una forma de realización,  
 fig. 2 ilustra un renderizador de objetos según una forma de realización,  
 fig. 3 ilustra un procesador de metadatos de objeto según una forma de realización,  
 fig. 4 ilustra el remapeo del azimut según las formas de realización,  
 fig. 5 ilustra el remapeo de la elevación según las formas de realización,  
 fig. 6 ilustra el remapeo del azimut según algunas formas de realización,  
 fig. 7 ilustra el remapeo de la elevación según otras formas de realización,  
 fig. 8 ilustra una vista general de un codificador de audio 3D,  
 fig. 9 ilustra una vista general de un decodificador de audio 3D según una forma de realización,  
 fig. 10 ilustra una estructura de un conversor de formato,  
 fig. 11 ilustra la renderización de audio basado en objetos según una forma de realización,  
 fig. 12 ilustra un pre-procesador de metadatos de objeto según una forma de realización,  
 fig. 13 ilustra el remapeo del azimut según una forma de realización,  
 fig. 14 ilustra el remapeo de los ángulos de elevación según una forma de realización,  
 fig. 15 ilustra el remapeo de los ángulos azimutales según una forma de realización,  
 fig. 16 ilustra el remapeo de la elevación según otras formas de realización, y  
 fig. 17 ilustra el remapeo de la elevación según otras formas de realización.

**[0011]** fig. 1 ilustra un aparato para el remapeo de objetos de audio según una forma de realización. El aparato comprende un procesador de metadatos de objeto 110 y un renderizador de objetos 120.

**[0012]** El renderizador de objetos 120 está configurado para recibir un objeto de audio.

**[0013]** El procesador de metadatos de objetos 110 está configurado para recibir metadatos, que comprenden una indicación de si el objeto de audio tiene relación con la pantalla, y que además comprende una primera posición del objeto de audio. Además, el procesador de metadatos de objetos 110 está configurado para calcular una segunda posición del objeto de audio dependiendo de la primera posición del objeto de audio y dependiendo del tamaño de una pantalla, si en los metadatos se indica que el objeto de audio tiene relación con la pantalla.

**[0014]** El renderizador de objetos 120 está configurado para generar las señales de altavoces dependiendo del objeto de audio y dependiendo de información de posición.

**[0015]** El procesador de metadatos de objetos 110 está configurado para alimentar la primera posición del objeto de audio como información sobre la posición al renderizador de objetos 120, si en los metadatos se indica que el objeto de audio no está relacionado con la pantalla.

5

**[0016]** Además, el procesador de metadatos de objetos 110 está configurado para alimentar la segunda posición del objeto de audio como información sobre la posición al renderizador de objetos 120, si en los metadatos se indica que el objeto de audio tiene relación con la pantalla.

10 **[0017]** Según una forma de realización, el procesador de metadatos de objetos 110 puede estar configurado, por ejemplo, para no calcular la segunda posición del objeto de audio, si en los metadatos se indica que el objeto de audio no está relacionado con la pantalla.

15 **[0018]** En una forma de realización, el renderizador de objetos 120 puede estar configurado, por ejemplo, para no determinar si información sobre la posición es la primera posición del objeto de audio o la segunda posición del objeto de audio.

20 **[0019]** Según una forma de realización, el renderizador de objetos 120 puede estar configurado, por ejemplo, para generar las señales de altavoces dependiendo además del número de los altavoces de un entorno de reproducción.

**[0020]** En una forma de realización, el renderizador de objetos 120 puede estar configurado, por ejemplo, para generar las señales de altavoces dependiendo además de una posición de altavoces de cada uno de los altavoces del entorno de reproducción.

25

**[0021]** Según una forma de realización, el procesador de metadatos de objetos 110 está configurado para calcular la segunda posición del objeto de audio dependiendo de la primera posición del objeto de audio y dependiendo del tamaño de la pantalla, si en los metadatos se indica que el objeto de audio tiene relación con la pantalla, donde la primera posición indica la primera posición en un espacio tridimensional, y donde la segunda posición indica la segunda posición en el espacio tridimensional.

30

**[0022]** En una forma de realización, el procesador de metadatos de objetos 110 puede estar configurado, por ejemplo, para calcular la segunda posición del objeto de audio dependiendo de la primera posición del objeto de audio y dependiendo del tamaño de la pantalla, si en los metadatos se indica que el objeto de audio tiene relación con la pantalla, donde la primera posición indica un primer azimut, una primera elevación y una primera distancia, y donde la segunda posición indica un segundo azimut, una segunda elevación y una segunda distancia.

35

**[0023]** Según una forma de realización, el procesador de metadatos de objetos 110 puede estar configurado, por ejemplo, para recibir los metadatos, que comprenden la indicación de si el objeto de audio tiene relación con la pantalla como una primera indicación, y que además comprende una segunda indicación, si el objeto de audio tiene relación con la pantalla, indicando dicha segunda indicación si el objeto de audio es un objeto en pantalla. El procesador de metadatos de objetos 110 puede estar configurado, por ejemplo, para calcular la segunda posición del objeto de audio dependiendo de la primera posición del objeto de audio y dependiendo del tamaño de la pantalla, de tal manera que la segunda posición asuma un primer valor en un área de pantalla de la pantalla, si la segunda indicación indica que el objeto de audio es un objeto en pantalla.

45

**[0024]** En una forma de realización, el procesador de metadatos de objetos 110 puede estar configurado, por ejemplo, para calcular la segunda posición del objeto de audio dependiendo de la primera posición del objeto de audio y dependiendo del tamaño de la pantalla, de tal manera que la segunda posición asuma un segundo valor, que está en el área de pantalla o no está en el área de pantalla, si la segunda indicación indica que el objeto de audio no es un objeto en pantalla.

50

**[0025]** Según una forma de realización, el procesador de metadatos de objetos 110 puede estar configurado, por ejemplo, para recibir los metadatos, que comprenden la indicación de si el objeto de audio tiene relación con la pantalla como primera indicación, y que además comprende una segunda indicación, donde si el objeto de audio tiene relación con la pantalla, dicha segunda indicación indica si el objeto de audio es un objeto en pantalla. El procesador de metadatos de objetos 110 puede estar configurado, por ejemplo, para calcular la segunda posición del objeto de audio dependiendo de la primera posición del objeto de audio, dependiendo del tamaño de la pantalla, y dependiendo de una primera curva de mapeo como curva de mapeo, si la segunda indicación indica que el objeto de audio es un objeto en pantalla, donde la primera curva de mapeo define el mapeo de las posiciones originales de los objetos en un primer intervalo de valores con las posiciones remapeadas de los objetos en un segundo intervalo de valores. Además, el procesador de metadatos de objetos 110 puede estar configurado, por ejemplo, para calcular la segunda posición del objeto de audio dependiendo de la primera posición del objeto de audio, dependiendo del tamaño de la pantalla, y dependiendo de una segunda curva de mapeo como curva de mapeo si la segunda indicación indica que el objeto de audio no es un objeto en pantalla, donde la segunda curva de mapeo define el

60

65

mapeo de las posiciones originales de los objetos en el primer intervalo de valores con las posiciones remapeadas de los objetos en un tercer intervalo de valores, y donde dicho segundo intervalo de valores está comprendido en el tercer intervalo de valores, y donde dicho segundo intervalo de valores es menor que dicho tercer intervalo de valores.

5 **[0026]** En una forma de realización, cada uno del primer intervalo de valores y el segundo intervalo de valores y el tercer intervalo de valores puede ser, por ejemplo, un intervalo de valores de ángulos azimutales, o cada uno del primer intervalo de valores y el segundo intervalo de valores y el tercer intervalo de valores puede ser, por ejemplo, un intervalo de valores de ángulos de elevación.

10 **[0027]** A continuación se describen formas de realización específicas de la presente invención y las características opcionales de una pluralidad de formas de realización de la presente invención.

15 **[0028]** Podría haber objetos de audio (señal de audio asociada a una posición en el espacio 3D, por ejemplo, el azimut, la elevación y la distancia dados) que no están destinados a una posición fija, sino cuya posición debería cambiar con el tamaño de una pantalla en la configuración de reproducción.

20 **[0029]** Si se señala un objeto como relacionado con la pantalla (por ejemplo, por medio de una bandera en los metadatos), su posición se remapea/recalcula con respecto al tamaño de la pantalla según una regla específica.

25 **[0030]** La fig. 2 ilustra un renderizador de objetos según una forma de realización.

**[0031]** Como introducción, se señala lo siguiente:  
 En los formatos de audio basado en objetos se almacenan o transmiten metadatos junto con las señales de objeto.  
 Los objetos de audio se renderizan del lado de la reproducción empleando los metadatos e información acerca del entorno de reproducción. Esa información es por ejemplo el número de altavoces o el tamaño de la pantalla.

Tabla 1: Ejemplo de metadatos:

	ObjectID
OAM dinámicos	Azimut Elevación Ganancia Distancia
Interactividad	AllowOnOff AllowPositionInteractivity AllowGainInteractivity DefaultOnOff DefaultGain InteractivityMinGain InteractivityMaxGain InteractivityMinAzOffset InteractivityMaxAzOffset InteractivityMinEIOffset InteractivityMaxEIOffset InteractivityMinDist InteractivityMaxDist
Reproducción	IsSpeakerRelatedGroup SpeakerConfig3D  AzimuthScreenRelated  ElevationScreenRelated  ClosestSpeakerPlayout
Contenido	ContentKind ContentLanguage
Grupo	GroupID GroupDescription GroupNumMembers

	GroupMembers Priority
Conmutación Grupo	SwitchGroupID SwitchGroupDescription SwitchGroupDefault SwitchGroupNumMembers SwitchGroupMembers
Escena de audio	NumGroupsTotal IsMainScene NumGroupsPresent NumSwitchGroups

**[0032]** En el caso de los objetos se pueden utilizar metadatos geométricos para definir de qué manera se deben renderizar, por ejemplo los ángulos de azimut o elevación o las posiciones absolutas con respecto a un punto de referencia, por ejemplo el oyente. El renderizador calcula las señales de altavoces sobre la base de los datos geométricos y los altavoces disponibles y su posición.

**[0033]** Las formas de realización según la presente invención surgen de lo anterior de la siguiente manera.

**[0034]** Para controlar la renderización relacionada con la pantalla, un campo de metadatos adicional controla la manera de interpretar los metadatos geométricos: Si se ajusta el campo a OFF (apagado), el renderizador interpreta por los metadatos geométricos que debe computar las señales de altavoces.

**[0035]** Si se ajusta el campo a ON (encendido) los metadatos geométricos son mapeados de los datos nominales con otros valores. El remapeo se realiza sobre los metadatos geométricos, de tal manera que el renderizador que sigue el procesador de metadatos de objetos desconoce el pre-procesamiento de los metadatos de objeto y opera sin cambios. En las siguientes tablas se presentan ejemplos de dichos campos de metadatos.

Tabla 2: Ejemplo de metadatos para controlar la renderización relacionada con la pantalla y su significado:

<b>AzimuthScreenRelated</b>	Se ajusta el azimut al tamaño de la pantalla
<b>ElevationScreenRelated</b>	Se ajusta la elevación al tamaño de la pantalla
<b>isScreenRelatedObject</b>	Se remapea el azimut y la elevación para renderizar objetos en relación con la pantalla
<b>isOnScreenObject</b>	Se relaciona la señal de objeto con un objeto colocado en pantalla

**[0036]** Además, se podría enviar el tamaño nominal de pantalla o el tamaño de la pantalla utilizado durante la producción del contenido de audio como información de metadatos.

<b>NominalScreenSize</b>	tamaño de pantalla utilizado durante la producción del contenido de audio
--------------------------	---

**[0037]** La siguiente tabla presenta un ejemplo de cómo se deben codificar esos metadatos de manera eficiente.

Tabla 3 — Sintaxis de ObjectMetadataConfig() según una forma de realización:

Sintaxis	N.º de bits	Mnemónica
ObjectMetadataConfig() { ...  <b>hasScreenRelatedObjects;</b> si( hasScreenRelatedObjects ) { para ( o = 1; o <= num_objects; o++ ) {	<b>1</b>	<b>bslbf</b>

```

        isScreenRelativeObject[o];      1      bs1bf
    si( !isScreenRelativeObject ){
        isOnScreenObject[o];          1      bs1bf
    }
}
}

```

**hasOnScreenObjects** Esta bandera especifica si hay presencia de objetos relacionados con la pantalla.

5 **isScreenRelatedObject** Esta bandera define si la posición de un objeto está relacionada con la pantalla (la posición debe ser renderizada de manera diferente, de tal manera que se remapee su posición, aunque de todos modos puede contener todos los valores angulares válidos.

10 **isOnScreenObject** Esta bandera define que el correspondiente objeto está “onscreen” (en pantalla). Los objetos donde esta bandera es igual a 1 deben ser renderizados de modo diferente, de tal manera que su posición sólo pueda asumir valores en el área de pantalla. De conformidad con una alternativa, no se utiliza la bandera, sino que se define un ángulo de la pantalla de referencia. Si **isScreenRelativeObject=1** entonces todos los ángulos se consideran con respecto a este ángulo de referencia. Podría haber otros casos de uso cuando es necesario saber que el objeto de audio está en pantalla.

15 **[0038]** Cabe señalar, con respecto a **isScreenRelativeObject**, que según una forma de realización, hay dos posibilidades: el remapeo de la posición, aunque de todos modos puede asumir todos los valores (relacionados con la pantalla) y el remapeo de tal modo que sólo pueda contener valores que estén en área de pantalla (en pantalla).

20 **[0039]** El remapeo se realiza en un procesador de metadatos de objeto que tiene en cuenta el tamaño de pantalla local y ejecuta el mapeo de los metadatos geométricos.

**[0040]** La fig. 3 ilustra un procesador de metadatos de objeto según una forma de realización.

25 **[0041]** En cuando a la modificación de los metadatos geométricos relacionados con la pantalla, se dice lo siguiente.

**[0042]** Dependiendo de la información *isScreenRelativeObject* e *isOnScreenObject* hay dos posibilidades de señalización de los elementos de audio relacionados con la pantalla:

- 30 a) Elementos de audio relativos a la pantalla
- b) Elementos de audio en pantalla

**[0043]** En ambos casos, los datos de posición de los elementos de audio son remapeados por el procesador de metadatos de objetos. Se aplica una curva que mapea los ángulos azimutales y de elevación originales de la posición con un azimut remapeado y un ángulo de elevación remapeado

35 **[0044]** La referencia es el tamaño de pantalla nominal en los metadatos o un tamaño de pantalla por defecto presunto.

40 **[0045]** Por ejemplo, se puede utilizar un ángulo de visualización definido en ITU-R REC-BT.2022 (Condiciones de visualización generales para la evaluación subjetiva de la calidad de imágenes de televisión SDTV y HDTV en pantallas de panel plano).

45 **[0046]** La diferencia entre los dos tipos de relación con la pantalla es la definición de la curva de remapeo.

**[0047]** En el caso a) el azimut remapeado puede asumir valores de entre -180° y 180° y la elevación remapeada puede asumir valores de entre -90° y 90°. La curva se define de tal manera que los valores de azimut entre un azimut del borde izquierdo por defecto y un azimut del borde derecho por defecto se mapeen (compriman o expandan) con el intervalo entre el borde izquierdo dado de la pantalla y el borde derecho dado de la pantalla (y de manera consiguiente en el caso de la elevación). Los demás valores de azimut y elevación se comprimen o expanden de manera correspondiente, de tal manera que se cubra el intervalo total de los valores.

**[0048]** La fig. 4 ilustra el remapeo del azimut según las formas de realización.

55 **[0049]** En el caso b) el azimut y la elevación remapeados sólo pueden asumir valores que describen posiciones en el área de pantalla (Azimut(borde izquierdo de la pantalla) Azimut(remapeo) Azimut(borde derecho de la pantalla) y Elevación(borde inferior de la pantalla) Elevación(remapeo) Elevación(borde superior de la

pantalla)).

**[0050]** Hay diferentes posibilidades de tratar los valores fuera de estos intervalos. Podrían ser mapeados con los bordes de la pantalla de tal manera que todos los objetos entre  $-180^\circ$  de azimut y el borde izquierdo de la pantalla terminen en el borde izquierdo de la pantalla y todos los objetos entre el borde derecho de la pantalla y  $180^\circ$  de azimut terminen a la derecha de la pantalla. Otra posibilidad consiste en mapear los valores del hemisferio posterior con el hemisferio anterior. En el hemisferio izquierdo se mapean entonces las posiciones entre  $-180^\circ + \text{Azimut}(\text{borde izquierdo de la pantalla})$  y  $\text{Azimut}(\text{borde izquierdo de la pantalla})$  con el borde izquierdo de la pantalla. Los valores entre  $-180^\circ$  y  $-180^\circ + \text{Azimut}(\text{borde izquierdo de la pantalla})$  se mapean con los valores entre  $0^\circ$  y  $\text{Azimut}(\text{borde izquierdo de la pantalla})$ . De la misma manera se trata el hemisferio derecho y los ángulos de elevación.

**[0051]** La fig. 5 ilustra el remapeo de la elevación según las formas de realización.

**[0052]** Los puntos  $-x1$  y  $+x2$  (que podrían ser diferentes o iguales a  $+x1$ ) de la curva en que el gradiente cambia pueden ser fijados en valores por defecto (tamaño de pantalla asumido por defecto + posición) o pueden estar presentes en los metadatos (por ejemplo, por el productor, que podría poner ahí entonces el tamaño de pantalla de la producción).

**[0053]** También existen funciones de mapeo posibles que no consisten en segmentos lineales sino que, por el contrario, son curvos.

**[0054]** Otros metadatos podrían controlar la forma de remapeo, por ejemplo, definiendo desviaciones o coeficientes no lineales para dar cuenta del comportamiento de paneo o la resolución de la audición.

**[0055]** También se podría señalar la manera de ejecutar el mapeo, por ejemplo, "proyectando" todos los objetos destinados para la parte posterior de la pantalla.

**[0056]** Tales procedimientos de mapeo alternativos están enumerados en las siguientes figuras.

**[0057]** En ellas, la fig. 6 ilustra el remapeo del azimut según las formas de realización.

**[0058]** La fig. 7 ilustra el remapeo de la elevación según las formas de realización.

**[0059]** En lo que respecta al comportamiento de un tamaño de pantalla desconocido: si no se da el tamaño de pantalla de reproducción,

- se asume entonces un tamaño de pantalla por defecto, o bien

- no se aplica mapeo alguno, incluso si un objeto está marcado como relacionado con la pantalla o en pantalla.

**[0060]** Volviendo a la fig. 4, en otra forma de realización, en el caso b) el azimut y la elevación remapeados sólo pueden asumir valores que describen posiciones en el área de pantalla ( $\text{Azimut}(\text{borde izquierdo de la pantalla}) \leq \text{Azimut}(\text{remapeo}) \leq \text{Azimut}(\text{borde derecho de la pantalla})$  y  $\text{Elevación}(\text{borde inferior de la pantalla}) \leq \text{Elevación}(\text{remapeada}) \leq \text{Elevación}(\text{borde superior de la pantalla})$ ). Hay diferentes posibilidades para tratar los valores fuera de estos intervalos: en algunas formas de realización, podrían ser mapeados con los bordes de la pantalla de tal manera que todos los objetos entre  $+180^\circ$  de azimut y el borde izquierdo de la pantalla terminen en el borde izquierdo de la pantalla y todos los objetos entre el borde derecho de la pantalla y  $-180^\circ$  de azimut terminen en el borde derecho de la pantalla. Otra posibilidad consiste en mapear los valores del hemisferio posterior con el hemisferio anterior.

**[0061]** En el hemisferio izquierdo entonces, se mapean las posiciones entre  $+180^\circ - \text{Azimut}(\text{borde izquierdo de la pantalla})$  y  $\text{Azimut}(\text{borde izquierdo de la pantalla})$  con el borde izquierdo de la pantalla. Los valores entre  $+180^\circ$  y  $+180^\circ - \text{Azimut}(\text{borde izquierdo de la pantalla})$  se mapean con los valores entre  $0^\circ$  y  $\text{Azimut}(\text{borde izquierdo de la pantalla})$ . El hemisferio derecho y los ángulos de elevación son tratados de la misma manera.

**[0062]** La fig. 16 ilustra una figura similar a la fig. 5. En las formas de realización ilustradas por la fig. 16, en ambos diagramas, se expone un intervalo de valores en el eje de las abscisas de  $-90^\circ$  a  $+90^\circ$  y un intervalo de valores en el eje de las ordenadas de  $-90^\circ$  a  $+90^\circ$ .

**[0063]** La fig. 17 ilustra una figura similar a la fig. 7. En las formas de realización ilustrada por la fig. 17, en ambos diagramas, se expone un intervalo de valores en el eje de las abscisas de  $-90^\circ$  a  $+90^\circ$  y un intervalo de valores en el eje de las ordenadas de  $-90^\circ$  a  $+90^\circ$ .

**[0064]** En lo sucesivo, se describen otras formas de realización de la invención y características opcionales de otras formas de realización con referencia a la fig. 8 – fig. 15.

**[0065]** Según algunas formas de realización, el remapeo de elementos relacionados con la pantalla sólo se puede procesar, por ejemplo, si el flujo de bits contiene elementos relacionados con la pantalla (bandera isScreenRelativeObject == 1 con respecto al menos a un elemento de audio) que estén acompañados por datos OAM (datos OAM = metadatos de objetos asociados) y si se señala el decodificador el tamaño de pantalla local a través de la interfaz LocalScreenSize().

**[0066]** Los datos geométricos de posición (los datos OAM antes de que haya tenido lugar alguna modificación de la posición) pueden ser, por ejemplo, mapeados con un rango diferente de valores mediante la definición y uso de una función de mapeo. El remapeo puede cambiar, por ejemplo, los datos geométricos de posición como paso de procesamiento previo a la renderización por lo que el renderizador no tiene conocimiento del remapeo y opera sin cambios.

**[0067]** Para el remapeo se puede tomar en cuenta, por ejemplo, el tamaño de pantalla de una pantalla de referencia nominal (utilizada en el proceso de mezcla y monitoreo) y/o la información de tamaño de pantalla en el recinto de la reproducción.

**[0068]** Si no se da ningún tamaño de pantalla nominal de referencia, se pueden utilizar, por ejemplo, los valores de referencia por defecto, por ejemplo presumiendo una pantalla de 4k y una distancia de visualización óptima.

**[0069]** En caso de no darse ningún tamaño de pantalla local, entonces por ejemplo no se aplica el remapeo.

**[0070]** Se pueden definir dos funciones de mapeo lineal, por ejemplo, para el remapeo de los valores de elevación y azimut:  
 Los bordes de la pantalla del tamaño de pantalla nominal pueden estar dados, por ejemplo, por:

$$\varphi_{izq.}^{nominal}, \varphi_{der.}^{nominal}, \theta_{superior}^{nominal}, \theta_{inferior}^{nominal}$$

**[0071]** Los bordes de la pantalla de reproducción se pueden abreviar, por ejemplo, de la siguiente manera:

$$\varphi_{izq.}^{repro}, \varphi_{der.}^{repro}, \theta_{sup.}^{repro}, \theta_{inf.}^{repro}$$

**[0072]** El remapeo de los datos de posición por azimut y elevación se pueden definir, por ejemplo, por las siguientes funciones de mapeo lineal:

$$\varphi' = \begin{cases} \frac{\varphi_{der.}^{repro} + 180^\circ}{\varphi_{der.}^{nominal} + 180^\circ} \cdot (\varphi + 180^\circ) - 180^\circ & \text{para } -180^\circ \leq \varphi < \varphi_{der.}^{nominal} \\ \frac{\varphi_{izq.}^{repro} - \varphi_{der.}^{repro}}{\varphi_{izq.}^{nominal} - \varphi_{der.}^{nominal}} \cdot (\varphi - \varphi_{der.}^{nominal}) + \varphi_{der.}^{repro} & \text{para } \varphi_{der.}^{nominal} \leq \varphi < \varphi_{izq.}^{nominal} \\ \frac{180^\circ - \varphi_{left}^{repro}}{180^\circ - \varphi_{izq.}^{nominal}} \cdot (\varphi - \varphi_{izq.}^{nominal}) + \varphi_{izq.}^{repro} & \text{para } \varphi_{izq.}^{nominal} \leq \varphi < 180^\circ \end{cases}$$

$$\theta' = \begin{cases} \frac{\theta_{inf}^{repro} + 90^\circ}{\theta_{inf}^{nominal} + 90^\circ} \cdot (\theta + 90^\circ) - 90^\circ & \text{para } -90^\circ \leq \theta < \theta_{inf}^{nominal} \\ \frac{\theta_{sup}^{repro} - \theta_{inf}^{repro}}{\theta_{sup}^{nominal} - \theta_{inf}^{nominal}} \cdot (\theta - \theta_{inf}^{nominal}) + \theta_{inf}^{repro} & \text{para } \theta_{inf}^{nominal} \leq \theta < \theta_{sup}^{nominal} \\ \frac{90^\circ - \theta_{sup}^{repro}}{90^\circ - \theta_{sup}^{nominal}} \cdot (\theta - \theta_{sup}^{nominal}) + \theta_{sup}^{repro} & \text{para } \theta_{sup}^{nominal} \leq \theta < 90^\circ \end{cases}$$

- [0073]** La fig. 13 ilustra una función de remapeo de los datos de posición según una forma de realización. En particular, en la fig. 13, se ilustra una función de mapeo para el mapeo del azimut. En la fig. 13, la curva se define de tal manera que los valores de azimut entre el azimut de borde izquierdo nominal de referencia y el azimut de borde derecho nominal de referencia se mapean (comprimen o expanden) con el intervalo entre el borde izquierdo de la pantalla local dado y el borde derecho de la pantalla local dado. Los demás valores de azimut se comprimen o expanden de modo correspondiente, de tal manera que se cubra la totalidad del intervalo de valores.
- 5 **[0074]** El azimut remapeado puede asumir valores, por ejemplo, de entre -180° y 180° y la elevación remapeada puede asumir valores, por ejemplo, de entre -90° y 90°.
- [0075]** Según una forma de realización, por ejemplo, si la bandera isScreenRelativeObject es ajustada a cero, no se aplica el remapeo de los elementos relacionados con la pantalla para el correspondiente elemento y el renderizador utiliza directamente los datos geométricos de posición (datos OAM más el cambio de posición por la interactividad del usuario) para computar las señales de reproducción.
- 15 **[0076]** Según algunas formas de realización, las posiciones de todos los elementos relacionados con la pantalla pueden ser, por ejemplo, remapeadas según el tamaño de la pantalla de reproducción como adaptación al recinto de la reproducción. Por ejemplo, si no se proporciona información sobre el tamaño de la pantalla de reproducción o no existe ningún elemento relacionado con la pantalla, no se aplica el remapeo.
- 20 **[0077]** El remapeo se puede definir, por ejemplo, por funciones de mapeo lineal que tienen en cuenta la información sobre el tamaño de la pantalla de reproducción en el recinto de la reproducción e información de tamaño de la pantalla de una pantalla de referencia, por ejemplo, usada en el proceso de mezcla y monitoreo.
- 25 **[0078]** En la fig. 13 se ilustra una función de mapeo del azimut según una forma de realización. En dicha fig. 13, se ilustra una función de mapeo de los ángulos azimutales. Como en la fig. 13, se puede definir de tal manera, por ejemplo, que se mapeen los valores de azimut entre el borde izquierdo y el borde derecho de la pantalla de referencia (se compriman o expandan) con el intervalo entre el borde izquierdo y el borde derecho de la pantalla de reproducción. Los demás valores de azimut se comprimen o expanden de tal manera que se cubra el intervalo total de valores.
- 30 **[0079]** De manera correspondiente, se puede definir una función de mapeo de la elevación, por ejemplo, (véase la fig. 14). El procesamiento relacionado con la pantalla puede tomar en cuenta asimismo, por ejemplo, un área de zooming (acercamiento o alejamiento) en el contenido de vídeo de alta resolución. El procesamiento relacionado con la pantalla puede definirse sólo, por ejemplo, con respecto a los elementos que van acompañados por datos dinámicos de posición y están marcados como relacionados con la pantalla.
- 35 **[0080]** A continuación se presenta una reseña general de sistema de un sistema de códec de audio 3D. Las formas de realización de la presente invención pueden ser utilizadas en ese tipo de sistema de códec de audio 3D. El sistema de códec de audio 3D se puede basar, por ejemplo, en un Códec de MPEG-D USAC para la codificación de señales de canales y objetos.
- 40 **[0081]** Según las formas de realización, para aumentar la eficiencia para la codificación de una gran cantidad de objetos, se ha adaptado la tecnología de MPEG SAOC (SAOC = Spatial Audio Object Coding, Codificación de Objetos de Audio Espacial). Por ejemplo, según algunas formas de realización, tres tipos de renderizadores pueden ejecutar, por ejemplo, las tareas de renderización de objetos a canales, renderización de canales a auriculares o renderización de canales a una configuración de altavoces diferente.
- 45 **[0082]** Cuando se transmiten explícitamente señales de objeto o se codifican paramétricamente empleando SAOC, la información de Metadatos de Objeto correspondiente se comprime y multiplexa en el flujo de bits de audio
- 50

3D.

**[0083]** La fig. 8 y la fig. 9 muestran los diferentes bloques algorítmicos del sistema de Audio 3D. En particular, la fig. 8 ilustra una vista general de un codificador de audio 3D. La fig. 9 ilustra una vista general de un decodificador de audio 3D según una forma de realización.

**[0084]** Se describen ahora las posibles formas de realización de los módulos de la fig. 8 y la fig. 9.

**[0085]** En la fig. 8, se ilustra un prerenderizador 810 (al que también se hace referencia como mezclador). En la configuración de la fig. 8, el prerenderizador 810 (mezclador) es opcional. El prerenderizador 810 también se puede utilizar de manera opcional para convertir una escena de entrada de Canal + Objeto en una escena de canal antes de la codificación. Funcionalmente, el prerenderizador 810 del lado del codificador puede estar relacionado, por ejemplo, con la funcionalidad del renderizador/mezclador de objetos 920 del lado del decodificador, que se describe a continuación. La prerenderización de objetos garantiza una entropía de la señal determinística en la entrada del codificador que es básicamente independiente del número de señales de objeto activas de forma simultánea. Con la prerenderización de objetos, no es necesaria la transmisión de metadatos de objeto. Se renderizan las señales de objeto a la Disposición de Canales que el codificador está configurado para usar. Las ponderaciones de los objetos por cada canal se obtienen de los metadatos de objeto asociados (OAM).

**[0086]** El códec núcleo correspondiente a las señales de canales de los altavoces, las señales de objeto discretas, las señales de mezcla descendente de objetos y las señales prerenderizadas se basa en la tecnología MPEG-D USAC (Códec de Núcleo de USAC). El codificador de USAC 820 (por ejemplo, ilustrado en la fig. 8) trata la codificación de la multitud de señales mediante la creación de información de mapeo de canales y objetos basada en la información geométrica y semántica del canal de entrada y la asignación de objetos. Esta información de mapeo describe, cómo se mapean los canales de entrada y objetos con los Elementos de Canales de USAC (CPEs, SCEs, LFEs) y cómo se transmite la información correspondiente al decodificador.

**[0087]** Todas las cargas útiles adicionales como datos o metadatos de objeto de SAOC han sido pasadas por elementos de extensión y se pueden considerar, por ejemplo, en el control de la tasa del codificador de USAC.

**[0088]** La codificación de objetos es posible de diferentes maneras, dependiendo de los requisitos de tasa/distorsión y los requerimientos de interactividad del renderizador. Las siguientes variantes de codificación de objetos son posibles:

- 35 - Objetos prerenderizados: las señales de objeto son prerenderizadas y mezcladas con las señales de canales 22.2 antes de la codificación. La cadena de codificación subsiguiente ve señales de canales 22.2.
- Formas de onda de objetos discretos: los objetos son suministrados como formas de onda monofónicas al codificador de USAC 820. El codificador de USAC 820 utiliza elementos de canales únicos SCEs para transmitir los objetos además de las señales de canal. Los objetos decodificados son renderizados y mezclados del lado del receptor. Se transmite la información comprimida de metadatos de objeto al receptor/renderizador de forma paralela.
- 40 - Formas de onda paramétricas de objetos: las propiedades de los objetos y su relación entre sí se describen por medio de parámetros de SAOC. La mezcla descendente de las señales de objeto es codificada con USAC por el codificador de USAC 820. La información paramétrica se transmite de forma paralela. Se elige el número de canales de mezcla descendente según el número de objetos y de la tasa de datos general. Se transmite la información comprimida de metadatos de objeto al renderizador SAOC.

**[0089]** Del lado del decodificador, un codificador USAC 910 lleva a cabo la decodificación según USAC.

**[0090]** Además, según las formas de realización, se presenta un dispositivo decodificador, véase la fig. 9. El dispositivo decodificador comprende un decodificador USAC 910 para decodificar un flujo de bits a fin de obtener uno o más canales de entrada de audio, para obtener uno o más objetos de audio, para obtener metadatos de objeto comprimidos y para obtener uno o más canales de transporte para SAOC.

**[0091]** Además, el dispositivo decodificador comprende un decodificador SAOC 915 para decodificar dichos uno o más canales de transporte para SAOC para obtener un primer grupo de uno o más objetos de audio renderizados.

**[0092]** Además, el dispositivo decodificador comprende un aparato 917 según las formas de realización anteriormente descritas con respecto a las fig. 1 a 7 o como se describe más adelante con respecto a las fig. 11 a 15. El aparato 917 comprende un decodificador de metadatos de objetos 918 que es, por ejemplo, el procesador de metadatos de objetos 110 del aparato de la fig. 1, y que está implementado para decodificar los metadatos de objeto comprimidos para obtener metadatos no comprimidos.

**[0093]** Además, el aparato 917 según las formas de realización anteriormente descritas comprende un renderizador de objetos 920 que es, por ejemplo, el renderizador de objetos 120 del aparato de la fig. 1, para

renderizar dichos uno o más objetos de audio de entrada dependiendo de los metadatos no comprimidos para obtener un segundo grupo de uno o más objetos de audio renderizados.

**[0094]** Por añadidura, el dispositivo decodificador comprende un conversor de formato 922 para convertir 5 dichos uno o más canales de entrada de audio para obtener uno o más canales convertidos.

**[0095]** Además, el dispositivo decodificador comprende un mezclador 930 para mezclar dichos uno o más objetos de audio del primer grupo de uno o más objetos de audio renderizados, dichos uno o más objetos de audio del segundo grupo de uno o más objetos de audio renderizados y dichos uno o más canales convertidos para 10 obtener uno o más canales de audio decodificados.

**[0096]** En la fig. 9 se ilustra una forma de realización específica de un dispositivo decodificador. El codificador SAOC 815 (el codificador SAOC 815 es opcional, véase la fig. 8) y el decodificador SAOC 915 (véase la fig. 9) para las señales de objeto se basan en la tecnología de MPEG SAOC. El sistema tiene la capacidad de recrear, modificar 15 y renderizar un número de objetos de audio basándose en un número menor de canales transmitidos y datos paramétricos adicionales (OLDs, IOCs, DMGs) (OLD = object level difference (diferencia de nivel de los objetos), IOC = inter object correlation (correlación entre los objetos), DMG = downmix gain (ganancia de mezcla descendente)). Los datos paramétricos adicionales exhiben una tasa de datos significativamente más baja que la necesaria para transmitir todos los objetos de forma individual, lo que confiere mucha eficiencia a la codificación. 20

**[0097]** El codificador SAOC 815 toma como entrada señales de objeto/canales como formas de onda monofónicas y emite la información paramétrica (que está incluida en el flujo de bits de audio 3D) y los canales de transporte para SAOC (que se codifican empleando elementos de canales únicos y se transmiten).

25 **[0098]** El decodificador SAOC 915 reconstruye las señales de objeto/canales procedentes de los canales de transporte para SAOC decodificados e información paramétrica y genera la escena de audio de salida basándose en la disposición de la reproducción, en la información descomprimida de metadatos de objeto y opcionalmente en la información de interacción con el usuario.

30 **[0099]** En lo que respecta al códec de metadatos de objeto, por cada objeto, los metadatos asociados que especifican la posición geométrica y la expansión del objeto en el espacio 3D son codificados de manera eficiente mediante la cuantificación de las propiedades del objeto en tiempo y espacio, por ejemplo, por medio del codificador de metadatos 818 de la fig. 8. Los metadatos de objeto comprimidos cOAM (cOAM = metadatos de objeto de audio comprimidos) son transmitidos al receptor en forma de información lateral. En el receptor, los cOAM son 35 decodificados por el decodificador de metadatos 918.

**[0100]** Por ejemplo, en la fig. 9, el decodificador de metadatos 918 puede implementar por ejemplo, un procesador de metadatos de objeto según una de las formas de realización anteriormente descritas.

40 **[0101]** Un renderizador de objetos, por ejemplo, el renderizador de objetos 920 de la fig. 9, utiliza los metadatos de objeto comprimidos para generar formas de onda de objetos según el formato de reproducción dado. Se renderiza cada objeto a ciertos canales de salida según sus metadatos. La salida de este bloque es el resultado de la suma de los resultados parciales.

45 **[0102]** Por ejemplo, en la fig. 9, el renderizador de objetos 920 puede ser implementado, por ejemplo, según una de las formas de realización anteriormente descritas.

**[0103]** En la fig. 9, el decodificador de metadatos 918 puede ser implementado, por ejemplo, en forma procesador de metadatos de objeto descrito según una de las formas de realización anteriormente descritas o las 50 formas de realización descritas a continuación, en referencia a las fig. 1 a 7, y las fig. 11 a 15, y el renderizador de objetos 920 puede ser implementado, por ejemplo, como el renderizador de objetos descrito según una de las formas de realización anteriormente descritas o las formas de realización descritas a continuación, en referencia a las fig. 1 a 7, y de la fig. 11 a la fig. 15. El decodificador de metadatos 918 y el renderizador de objetos 920 pueden implementar, en conjunto, por ejemplo, un aparato 917 para generar señales de altavoces como el descrito 55 anteriormente o según se describe más adelante en referencia a las fig. 1 a 7, y de la fig. 11 a la fig. 15.

**[0104]** Si se decodifica tanto contenido basado en canales como objetos discretos/paramétricos, se mezclan las formas de onda basadas en canales y las formas de onda de objetos renderizados antes de emitir las formas de onda resultantes, por ejemplo, mediante el mezclador 930 de la fig. 9 (o antes de alimentarlas a un módulo post- 60 procesador como el renderizador binaural o el módulo renderizador de altavoces).

**[0105]** Un módulo renderizador binaural 940, puede producir, por ejemplo, una mezcla descendente binaural del material de audio multicanal, de tal manera que cada canal de entrada esté representado por una fuente de sonido virtual. El procesamiento se lleva a cabo trama por trama en el dominio QMF. La binauralización se puede 65 basar, por ejemplo, en respuestas al impulso medidas del recinto binaural.

- [0106]** Un renderizador de altavoces 922 puede realizar la conversión, por ejemplo, entre la configuración de canales transmitida y el formato de reproducción deseado. Por consiguiente, a continuación se le denomina convertor de formato 922. El convertor de formato 922 realiza conversiones para reducir los números de canales de salida, por ejemplo, genera mezclas descendentes. El sistema genera automáticamente matrices optimizadas de mezcla descendente correspondientes a la combinación dada de formatos de entrada y salida y aplica estas matrices en un proceso de mezcla descendente. El convertor de formato 922 admite las configuraciones normales de altavoces, así como configuraciones aleatorias con posiciones de altavoces que no son las estándar.
- 10 **[0107]** La fig. 10 ilustra una estructura de un convertor de formato. La fig. 10 ilustra un configurador de mezcla descendente 1010 y un procesador de mezcla descendente para procesar la mezcla descendente en el dominio QMF (dominio QMF = quadrature mirror filter domain (dominio de filtros espejo en cuadratura)).
- 15 **[0108]** Según algunas formas de realización, el renderizador de objetos 920 puede estar configurado para realizar el remapeo de objetos de audio relacionado con la pantalla descrito con respecto a una de la pluralidad de formas de realización anteriormente descritas en referencia a la fig. 1 – fig. 7, o descritas con respecto a una de la pluralidad de formas de realización que se describen a continuación en referencia a la fig. 11 – fig. 15.
- 20 **[0109]** A continuación se describen otras formas de realización y conceptos de las formas de realización de la presente invención.
- [0110]** Según algunas formas de realización, el control de los objetos por el usuario puede emplear, por ejemplo, metadatos descriptivos, por ejemplo, información acerca de la existencia de un objeto dentro del flujo de bits y las propiedades de alto nivel de los objetos y puede emplear, por ejemplo, metadatos restrictivos, por ejemplo, información sobre cómo la interacción es posible o es habilitada por el creador de contenidos.
- 25 **[0111]** Según algunas formas de realización, la señalización, transmisión y renderización de objetos de audio pueden emplear, por ejemplo, metadatos de posición, metadatos estructurales, por ejemplo, el agrupamiento y jerarquía de los objetos, la capacidad de renderizar a un canal específico y para señalar el contenido del canal en forma de objetos y medios para adaptar la escena de los objetos al tamaño de la pantalla.
- 30 **[0112]** Las formas de realización incluyen nuevos campos de metadatos además de la posición geométrica ya definida y el nivel del objeto en el espacio 3D.
- 35 **[0113]** Si se reproduce una escena de audio basado en objetos en diferentes configuraciones de reproducción, según algunas formas de realización, las posiciones de las fuentes de sonido renderizadas pueden ser, por ejemplo, escaladas automáticamente a la dimensión de la reproducción. En caso de presentarse contenido audiovisual, la renderización normal de los objetos de audio a la reproducción puede llevar, por ejemplo, a una violación de la coherencia audiovisual posicional como las ubicaciones de las fuentes de sonido y la posición de lo que originó visualmente el sonido, por ejemplo, ya puede no ser coherente.
- 40 **[0114]** Para evitar este efecto, se puede emplear, por ejemplo, una posibilidad que consiste en señalar que los objetos de audio no están destinados a una posición fija en el espacio 3D, sino cuya posición debe cambiar con el tamaño de una pantalla en la configuración de reproducción. Según algunas formas de realización, un tratamiento especial de estos objetos de audio y una definición correspondiente a un algoritmo de escalado de escenas pueden dar lugar, por ejemplo, a una experiencia más inmersiva ya que la reproducción se puede optimizar, por ejemplo, respecto de las características locales del entorno de reproducción.
- 45 **[0115]** En algunas formas de realización, un renderizador o un módulo de preprocesamiento puede tomar en cuenta, por ejemplo, el tamaño de la pantalla local en el recinto de la reproducción y de esa manera puede, por ejemplo, preservar la relación entre audio y vídeo en el contexto de una película o un juego. En esas formas de realización, la escena de audio puede ser escalada automáticamente entonces, por ejemplo, según la configuración de reproducción, de tal manera que las posiciones de los elementos visuales y la posición de una fuente de sonido correspondiente coincidan. Se puede mantener, por ejemplo, la coherencia audiovisual posicional para pantallas con variación de tamaño.
- 50 **[0116]** Por ejemplo, según las formas de realización, el diálogo y la voz pueden ser percibidos, por ejemplo, desde la dirección de alguien que habla en la pantalla independientemente del tamaño de la pantalla de reproducción. Esto es posible entonces para fuentes fijas, así como en el caso de fuentes en movimiento en que las trayectorias del sonido y el movimiento de los elementos visuales tienen que corresponder.
- 60 **[0117]** Para controlar la renderización relacionada con la pantalla, se introduce un campo de metadatos adicional que permite marcar los objetos como relacionados con la pantalla. Si el objeto está marcado como relacionado con la pantalla, sus metadatos de posición geométrica se remapean con otros valores antes de la renderización. Por ejemplo, la fig. 13 ilustra un ejemplo de función de (re)mapeo correspondiente a los ángulos
- 65

azimutales.

**[0118]** Entre otras cosas, algunas formas de realización pueden lograr, por ejemplo, una sencilla función de mapeo que se define por actuar en el dominio angular (azimut, elevación).

5

**[0119]** Además, algunas formas de realización pueden lograr, por ejemplo, que la distancia de los objetos no cambie; no se lleva a cabo ningún alejamiento ni acercamiento ("zooming") ni movimiento virtual hacia la pantalla o alejándose de la pantalla, sino un escalado sólo de la posición de los objetos.

10 **[0120]** Además, algunas formas de realización pueden manipular, por ejemplo, pantallas de reproducción no centradas ( $\left| \varphi_{izq.}^{repro} \right| \neq \left| \varphi_{der.}^{repro} \right|$  y/o  $\left| \theta_{sup.}^{repro} \right| \neq \left| \theta_{inf.}^{repro} \right|$ ) ya que la función de mapeo no sólo se basa en la relación de pantalla, sino que toma en cuenta el azimut y elevación de los bordes de la pantalla

15 **[0121]** Además, algunas formas de realización pueden definir, por ejemplo, funciones de mapeo especiales para los objetos en pantalla. Según algunas formas de realización, las funciones de mapeo correspondientes al azimut y elevación pueden ser independientes, por ejemplo, por lo que se las puede elegir para remapear sólo los ángulos azimutales o de elevación.

**[0122]** A continuación se proporcionan otras formas de realización.

20

**[0123]** La fig. 11 ilustra la renderización del audio basado en objetos según una forma de realización. Los objetos de audio pueden ser renderizados, por ejemplo, del lado de la reproducción empleando los metadatos e información acerca del entorno de reproducción. Dicha información consiste, por ejemplo, en el número de altavoces o el tamaño de la pantalla. El renderizador 1100 puede calcular, por ejemplo, las señales de altavoces sobre la base

25

**[0124]** Se describe ahora un (pre)procesador de metadatos de objetos 1210 según una forma de realización con referencia a la fig. 12.

30 **[0125]** En la fig. 12, el procesador de metadatos de objetos 1210 está configurado para llevar a cabo el remapeo que tiene en cuenta el tamaño de pantalla local y realiza el mapeo de los metadatos geométricos.

**[0126]** Los datos de posición de los objetos relacionados con la pantalla son remapeados por el procesador de metadatos de objetos 1210. Se puede aplicar una curva, por ejemplo, que mapea los ángulos azimutales y de elevación originales de la posición con un azimut remapeado y un ángulo de elevación remapeado.

35

**[0127]** Se puede emplear el tamaño de pantalla de una pantalla nominal de referencia, por ejemplo, para el proceso de mezclado y monitoreo, y se puede tomar en cuenta la información de tamaño de la pantalla local en el recinto de la reproducción, por ejemplo, para el remapeo.

40

**[0128]** Se puede transmitir el tamaño de pantalla de referencia, al que se hace referencia como tamaño de pantalla de producción, por ejemplo, en los metadatos.

45 **[0129]** En algunas formas de realización, si no se da el tamaño de pantalla nominal, se puede presumir, por ejemplo, un tamaño de pantalla por defecto.

**[0130]** Por ejemplo, se puede utilizar un ángulo de visualización definido en ITU-R REC-BT.2022 (véase: General viewing conditions for subjective assessment of quality of SDTV y HDTV television pictures on flat panel displays).

50

**[0131]** En algunas formas de realización, se pueden definir, por ejemplo, dos funciones de mapeo lineal para el remapeo de los valores de elevación y azimut.

55 **[0132]** A continuación se describe la modificación de los metadatos geométricos relacionados con la pantalla según algunas formas de realización en referencia a la fig. 13 – fig. 15.

**[0133]** El azimut remapeado puede asumir valores de entre  $-180^\circ$  y  $180^\circ$  y la elevación remapeada puede asumir valores de entre  $-90^\circ$  y  $90^\circ$ . La curva de mapeo se define, en general, de tal manera que se mapeen (compriman o expandan) los valores de azimut entre un azimut de borde izquierdo por defecto y un azimut de borde derecho por defecto con el intervalo entre el borde izquierdo dado de la pantalla y el borde derecho dado de la pantalla (y, en consecuencia, en el caso de la elevación). Los demás valores de azimut y elevación se comprimen o expanden de modo correspondiente, de tal manera que se cubra la totalidad del intervalo de valores.

60

**[0134]** Como ya se ha descrito anteriormente, los bordes de pantalla del tamaño de pantalla nominal pueden

estar dados, por ejemplo, por:

$$\varphi_{izq.}^{nominal}, \varphi_{der.}^{nominal}, \theta_{sup.}^{nominal}, \theta_{inf.}^{nominal}$$

**[0135]** Los bordes de la pantalla de reproducción se pueden abreviar, por ejemplo, de la siguiente manera:

5

$$\varphi_{izq.}^{repro}, \varphi_{der.}^{repro}, \theta_{sup.}^{repro}, \theta_{inf.}^{repro}$$

**[0136]** El remapeo de los datos de posición de azimut y elevación se puede definir, por ejemplo, por las siguientes funciones de mapeo lineal:

10

$$\varphi' = \begin{cases} \frac{\varphi_{der}^{repro} + 180^\circ}{\varphi_{der}^{nominal} + 180^\circ} \cdot (\varphi + 180^\circ) - 180^\circ & \text{para } -180^\circ \leq \varphi < \varphi_{der}^{nominal} \\ \frac{\varphi_{izq.}^{repro} - \varphi_{der}^{repro}}{\varphi_{izq.}^{nominal} - \varphi_{der}^{nominal}} \cdot (\varphi - \varphi_{der}^{nominal}) + \varphi_{der}^{repro} & \text{para } \varphi_{der}^{nominal} \leq \varphi < \varphi_{izq.}^{nominal} \\ \frac{180^\circ - \varphi_{izq.}^{repro}}{180^\circ - \varphi_{izq.}^{nominal}} \cdot (\varphi - \varphi_{izq.}^{nominal}) + \varphi_{izq.}^{repro} & \text{para } \varphi_{izq.}^{nominal} \leq \varphi < 180^\circ \end{cases}$$

$$\theta' = \begin{cases} \frac{\theta_{inf.}^{repro} + 90^\circ}{\theta_{inf.}^{nominal} + 90^\circ} \cdot (\theta + 90^\circ) - 90^\circ & \text{para } -90^\circ \leq \theta < \theta_{inf.}^{nominal} \\ \frac{\theta_{sup.}^{repro} - \theta_{inf.}^{repro}}{\theta_{sup.}^{nominal} - \theta_{inf.}^{nominal}} \cdot (\theta - \theta_{inf.}^{nominal}) + \theta_{inf.}^{repro} & \text{para } \theta_{inf.}^{nominal} \leq \theta < \theta_{sup.}^{nominal} \\ \frac{90^\circ - \theta_{sup.}^{repro}}{90^\circ - \theta_{sup.}^{nominal}} \cdot (\theta - \theta_{sup.}^{nominal}) + \theta_{sup.}^{repro} & \text{para } \theta_{sup.}^{nominal} \leq \theta < 90^\circ \end{cases}$$

**[0137]** La función de mapeo correspondiente al azimut está ilustrada en la fig. 13 y la función de mapeo correspondiente a la elevación está ilustrada en la fig. 14.

**[0138]** Los puntos  $\varphi_{izq.}^{nominal}, \varphi_{der.}^{nominal}, \theta_{sup.}^{nominal}, \theta_{inf.}^{nominal}$  de las curvas en que los cambios de gradiente se pueden establecer como valores por defecto (tamaño de pantalla normal presunto por defecto y posición de pantalla normal presunta por defecto), o pueden estar presentes en los metadatos (por ejemplo, introducidos por el productor, que en ese caso podría poner el tamaño de pantalla de producción/monitoreo ahí).

**[0139]** Con respecto a la definición de metadatos de objeto para el remapeo relacionado con la pantalla, para controlar la renderización relacionada con la pantalla, se define una bandera de metadatos adicional denominada "isScreenRelativeObject". Esta bandera puede definir, por ejemplo, si un objeto de audio debe ser procesado/renderizado en relación con el tamaño de la pantalla de reproducción local.

**[0140]** Si hay elementos relacionados con la pantalla presentes en la escena de audio, se ofrece la posibilidad, en ese caso, de incluir la información de tamaño de la pantalla de una pantalla nominal de referencia que ha sido utilizada para mezclar y monitorear (tamaño de pantalla utilizado durante la producción del contenido de audio).

30



relacionado con la pantalla.

**[0142]** Algunas de las formas de realización admiten variaciones posibles.

5 **[0143]** En algunas formas de realización, se emplean funciones de mapeo no lineales. Estas funciones de mapeo posibles no consisten en segmentos lineales, sino que por el contrario, son curvos. En algunas formas de realización, hay metadatos adicionales que controlan la manera de remapear, por ejemplo, definiendo desviaciones o coeficientes no lineales para dar cuenta del comportamiento de paneo o la resolución de la audición.

10 **[0144]** Algunas formas de realización efectúan un procesamiento independiente del azimut y la elevación. El azimut y la elevación podrían ser marcados y procesados independientemente en relación con la pantalla. La Tabla 5 ilustra la sintaxis de ObjectMetadataConfig() según esa forma de realización.

Tabla 5: Sintaxis de ObjectMetadataConfig() según una forma de realización:

Sintaxis	N.º de bits	Mnemónica
ObjectMetadataConfig() {		
...		
<b>hasScreenRelatedObjects;</b>	<b>1</b>	<b>bslbf</b>
si( hasScreenRelatedObjects ) {		
...		
para ( o = 0; o <= num_objects-1; o++ ) {		
<b>AzimuthScreenRelated[o];</b>	<b>1</b>	<b>bslbf</b>
<b>ElevationScreenRelated[o];</b>	<b>1</b>	<b>bslbf</b>
}		
}		
}		

15

**[0145]** Algunas formas de realización emplean una definición de objetos en pantalla. Se puede distinguir entre objetos relacionados con la pantalla y objetos en pantalla. Entonces una sintaxis posible podría ser la siguiente de la tabla 6:

20

Tabla 6 — Sintaxis de ObjectMetadataConfig() según una forma de realización:

Sintaxis	N.º de bits	Mnemónica
ObjectMetadataConfig() {		
...		
<b>hasScreenRelatedObjects;</b>	<b>1</b>	<b>bslbf</b>
si( hasScreenRelatedObjects ) {		
...		
para ( o = 0; o <= num_objects-1; o++ ) {		
<b>isScreenRelativeObject[o];</b>	<b>1</b>	<b>bslbf</b>
si( !isScreenRelativeObject ){		
<b>isOnScreenObject[o];</b>	<b>1</b>	<b>bslbf</b>
}		
}		
}		
}		

**hasOnScreenObjects**  
**isScreenRelatedObject**

Esta bandera especifica si hay presencia de objetos relacionados con la pantalla. Esta bandera define si la posición de un objeto se relaciona con la pantalla (la posición debe ser renderizada de manera diferente, de tal manera que su posición se remapee, aunque de todas maneras contenga todos los valores angulares válidos).

25

**isOnScreenObject**

Esta bandera define si el objeto correspondiente es "onscreen" (en pantalla). Los objetos respecto de los cuales esta bandera es igual a 1 deben ser renderizados de manera diferente, de tal manera que su posición sólo pueda asumir valores en el área de pantalla.

30

**[0146]** En el caso de los objetos en pantalla, el azimut y la elevación remapeados sólo pueden asumir valores que describan posiciones en el área de pantalla ( $\varphi_{izq.}^{repro} \leq \varphi \leq \varphi_{der.}^{repro}$  y  $\theta_{sup.}^{repro} \geq \theta \geq \theta_{inf.}^{repro}$ ).

35

**[0147]** Según algunas formas de realización, hay diferentes posibilidades para tratar los valores fuera de estos intervalos. Se podrían mapear con los bordes de la pantalla. En el hemisferio izquierdo entonces, las

posiciones entre  $180^\circ$  y  $180^\circ - \varphi_{izq.}^{nominal}$  se mapean con el borde izquierdo de la pantalla  $\varphi_{izq.}^{nominal}$ . El hemisferio derecho y los ángulos de elevación se tratan de la misma manera (función de mapeo sin guiones 1510 de la fig. 15).

**[0148]** Otra posibilidad ofrecida por algunas de las formas de realización consiste en mapear los valores del hemisferio posterior con el hemisferio anterior. Los valores entre  $180^\circ$  y  $180^\circ - \varphi_{izq.}^{nominal}$  se mapean con los valores entre  $0^\circ$  y  $\varphi_{izq.}^{repro}$ . El hemisferio derecho y los ángulos de elevación se tratan de la misma manera (función de mapeo con guiones 1520 de la fig. 15).

**[0149]** La fig. 15 ilustra el remapeo de los ángulos azimutales (objetos en pantalla) según estas formas de realización.

**[0150]** La elección del comportamiento deseado podría ser señalizada mediante metadatos adicionales (por ejemplo, una bandera para "proyectar" todos los objetos en pantalla destinados a la parte posterior ( $[180^\circ$  y  $180^\circ - \varphi_{izq.}^{nominal}$ ] y  $[-180^\circ$  y  $-180^\circ - \varphi_{der.}^{nominal}$ ] sobre la pantalla).

**[0151]** Aunque se han descrito algunos aspectos en el contexto de un aparato, es obvio que estos aspectos también representan una descripción del procedimiento correspondiente, en el cual un bloque o dispositivo corresponde a una etapa del procedimiento o a una característica de una etapa del procedimiento. De manera análoga, los aspectos descritos en el contexto de una etapa del procedimiento también representan una descripción de un bloque o elemento correspondiente o de una característica de un aparato correspondiente.

**[0152]** La señal descompuesta de la invención puede ser almacenada en un medio digital o se puede transmitir por un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión por cables tal como Internet.

**[0153]** Dependiendo de ciertos requisitos de implementación, las formas de realización de la invención pueden ser implementadas en hardware o en software. La implementación se puede realizar empleando un medio de almacenamiento digital, por ejemplo un disco flexible, un DVD, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene almacenadas en la misma señales de control legibles electrónicamente, que cooperan (o tienen capacidad para cooperar) con un sistema informático programable de tal manera que se ejecute el procedimiento respectivo.

**[0154]** Algunas formas de realización según la invención comprenden un soporte de datos que comprende señales de control legibles electrónicamente, con capacidad para cooperar con un sistema informático programable de tal manera que se ejecute uno de los procedimientos descritos en esta invención.

**[0155]** En general, las formas de realización de la presente invención pueden ser implementadas en forma de producto de programa informático con un código de programa, donde el código de programa cumple la función de ejecutar uno de los procedimientos al ejecutarse el programa informático en un ordenador. El código de programa puede ser almacenado, por ejemplo, en un soporte legible por una máquina.

**[0156]** Otras formas de realización comprenden el programa informático para ejecutar uno de los procedimientos descritos en esta invención, almacenado en un soporte legible por una máquina.

**[0157]** En otras palabras, una forma de realización del procedimiento de la invención consiste, por lo tanto, en un programa informático que consta de un código de programa para realizar uno de los procedimientos descritos en esta invención al ejecutarse el programa informático en un ordenador.

**[0158]** Otra forma de realización de los procedimientos de la invención consiste, por lo tanto, en un soporte de datos (o medio de almacenamiento digital, o medio legible por ordenador) que comprende, grabado en el mismo, el programa informático para ejecutar uno de los procedimientos descritos en esta invención.

**[0159]** Otra forma de realización del procedimiento de la invención es, por lo tanto, un flujo de datos o una secuencia de señales que representa el programa informático para ejecutar uno de los procedimientos descritos en esta invención. El flujo de datos o la secuencia de señales pueden estar configurados, por ejemplo, para ser transferidos a través de una conexión de comunicación de datos, por ejemplo, por Internet.

**[0160]** Otra forma de realización comprende un medio de procesamiento, por ejemplo un ordenador, o un dispositivo lógico programable, configurado o adaptado para ejecutar uno de los procedimientos descritos en esta invención.

**[0161]** Otra forma de realización comprende un ordenador en el que se ha instalado el programa informático para ejecutar uno de los procedimientos descritos en esta invención.

**[0162]** En algunas formas de realización, se puede utilizar un dispositivo lógico programable (por ejemplo una matriz de puertas programables en el campo) para ejecutar algunas o todas las funcionalidades de los procedimientos descritos en esta invención. En algunas formas de realización, una matriz de puertas programables en el campo puede cooperar con un microprocesador para ejecutar uno de los procedimientos descritos en esta invención. Por lo general, los procedimientos son ejecutados preferentemente por cualquier aparato de hardware.

**[0163]** Las formas de realización anteriormente descritas son meramente ilustrativas de los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y detalles descritos en esta invención serán evidentes para las personas expertas en la técnica. Por lo tanto, sólo es intención limitarse al alcance de las siguientes reivindicaciones de patente y no a los detalles específicos presentados a modo de descripción y explicación de las formas de realización presentadas en esta invención.

15

**[0164]**

Bibliografía

20 [1] "Method and apparatus for playback of a higher-order ambisonics audio signal", Solicitud de patente número EP20120305271

[2] "Vorrichtung und Verfahren zum Bestimmen einer Wiedergabeposition", Solicitud de patente número WO2004073352A1

[3] "Verfahren zur Audiocodierung", Solicitud de patente número EP20020024643

25 [4] "Acoustical Zooming Based on a Parametric Sound Field Representation"  
<http://www.aes.org/tmpFiles/elib/20140814/15417.pdf>

**REIVINDICACIONES**

1. Un aparato para generar señales de altavoces, que comprende:
- 5 un procesador de metadatos de objeto (110), y  
un renderizador de objetos (120),  
en el que el renderizador de objetos (120) está configurado para recibir un objeto de audio,  
en el que el procesador de metadatos de objetos (110) está configurado para recibir metadatos, que comprenden  
una indicación de si el objeto de audio tiene relación con la pantalla, y que además comprenden una primera  
10 posición del objeto de audio,  
en el que el procesador de metadatos de objetos (110) está configurado para calcular una segunda posición del  
objeto de audio dependiendo de la primera posición del objeto de audio y dependiendo del tamaño de la pantalla, si  
en los metadatos se indica que el objeto de audio tiene relación con la pantalla,  
en el que el renderizador de objetos (120) está configurado para generar las señales de altavoces dependiendo del  
15 objeto de audio y dependiendo de información de posición,  
en el que el procesador de metadatos de objetos (110) está configurado para alimentar la primera posición del objeto  
de audio como información sobre la posición al renderizador de objetos (120), si en los metadatos se indica que el  
objeto de audio no está relacionado con el tamaño de la pantalla, y  
en el que el procesador de metadatos de objetos (110) está configurado para alimentar la segunda posición del  
20 objeto de audio como información sobre la posición al renderizador de objetos (120), si en los metadatos se indica  
que el objeto de audio tiene relación con el tamaño de la pantalla,  
en el que la primera posición indica un primer azimut, una primera elevación y una primera distancia, en el que la  
segunda posición indica un segundo azimut, una segunda elevación y una segunda distancia,  
en el que el procesador de metadatos de objeto (110) está configurado para calcular la segunda posición del objeto  
25 de audio dependiendo de al menos una de una primera función de mapeo lineal y una segunda función de mapeo  
lineal, en el que la primera función de mapeo lineal se define para mapear un primer valor de azimut a un segundo  
valor de azimut, en el que la segunda función de mapeo lineal se define para mapear un primer valor de elevación a  
un segundo valor de elevación.
- 30 2. Un aparato según la reivindicación 1, en el que el procesador de metadatos de objetos (110) está  
configurado para no calcular la segunda posición del objeto de audio, si en los metadatos se indica que el objeto de  
audio no está relacionado con el tamaño de la pantalla.
3. Un aparato según la reivindicación 1 o 2, en el que el renderizador de objetos (120) está configurado  
35 para no determinar si información sobre la posición es la primera posición del objeto de audio o la segunda posición  
del objeto de audio.
4. Un aparato según una de las reivindicaciones anteriores, en el que el renderizador de objetos (120)  
está configurado para generar las señales de altavoces dependiendo además del número de los altavoces de un  
40 entorno de reproducción.
5. Un aparato según la reivindicación 4, en el que el renderizador de objetos (120) está configurado para  
generar las señales de altavoces dependiendo además de una posición de altavoces de cada uno de los altavoces  
del entorno de reproducción.
- 45 6. Un aparato según una de las reivindicaciones anteriores, en el que el procesador de metadatos de  
objetos (110) está configurado para calcular la segunda posición del objeto de audio dependiendo de la primera  
posición del objeto de audio y dependiendo del tamaño de la pantalla, si en los metadatos se indica que el objeto de  
audio tiene relación con el tamaño de la pantalla, en el que la primera posición indica la primera posición en un  
50 espacio tridimensional, y en el que la segunda posición indica la segunda posición en el espacio tridimensional.
7. Un aparato según la reivindicación 6, en el que el procesador de metadatos de objetos (110) está  
configurado para calcular la segunda posición del objeto de audio dependiendo de la primera posición del objeto de  
audio y dependiendo del tamaño de la pantalla, si en los metadatos se indica que el objeto de audio tiene relación  
55 con el tamaño de la pantalla, en el que la primera posición indica un primer azimut, una primera elevación y una  
primera distancia, y en el que la segunda posición indica un segundo azimut, una segunda elevación y una segunda  
distancia.
8. Un aparato según una de las reivindicaciones anteriores,  
60 en el que el procesador de metadatos de objetos (110) está configurado para recibir los metadatos que comprenden  
la indicación de si el objeto de audio tiene relación con el tamaño de la pantalla como una primera indicación, y que  
además comprenden una segunda indicación de si el objeto de audio tiene relación con el tamaño de la pantalla,  
indicando dicha segunda indicación si el objeto de audio es un objeto en pantalla y  
65

en el que el procesador de metadatos de objetos (110) está configurado para calcular la segunda posición del objeto de audio dependiendo de la primera posición del objeto de audio y dependiendo del tamaño de la pantalla, por lo que la segunda posición asume un primer valor en un área de pantalla de la pantalla, si la segunda indicación indica que el objeto de audio es un objeto en pantalla.

5

9. Un aparato según la reivindicación 8, en el que el procesador de metadatos de objetos (110) está configurado para calcular la segunda posición del objeto de audio dependiendo de la primera posición del objeto de audio y dependiendo del tamaño de la pantalla, por lo que la segunda posición asume un segundo valor, que está en el área de pantalla o no está en el área de pantalla, si la segunda indicación indica que el objeto de audio no es un objeto en pantalla.

10

10. Un aparato según una de las reivindicaciones 1 a 7,

en el que el procesador de metadatos de objetos (110) está configurado para recibir los metadatos que comprenden la indicación de si el objeto de audio tiene relación con el tamaño de la pantalla como una primera indicación, y que además comprende una segunda indicación de si el objeto de audio tiene relación con el tamaño de la pantalla, indicando dicha segunda indicación si el objeto de audio es un objeto en pantalla,

15

en el que el procesador de metadatos de objetos (110) está configurado para calcular la segunda posición del objeto de audio dependiendo de la primera posición del objeto de audio, dependiendo del tamaño de la pantalla, y dependiendo de una primera curva de mapeo como curva de mapeo, si la segunda indicación indica que el objeto de audio es un objeto en pantalla, en el que la primera curva de mapeo define un mapeo de las posiciones originales de los objetos en un primer intervalo de valores con las posiciones remapeadas de los objetos en un segundo intervalo de valores y

20

en el que el procesador de metadatos de objetos (110) está configurado para calcular la segunda posición del objeto de audio dependiendo de la primera posición del objeto de audio, dependiendo del tamaño de la pantalla, y dependiendo de una segunda curva de mapeo como curva de mapeo, si la segunda indicación indica que el objeto de audio no es un objeto en pantalla, en el que la segunda curva de mapeo define un mapeo de las posiciones originales de los objetos en el primer intervalo de valores con las posiciones remapeadas de los objetos en un tercer intervalo de valores, y en el que dicho segundo intervalo de valores está comprendido en el tercer intervalo de valores, y en el que dicho segundo intervalo de valores es menor que dicho tercer intervalo de valores.

25

30

11. Un aparato según la reivindicación 10,

en el que cada uno del primer intervalo de valores y el segundo intervalo de valores y el tercer intervalo de valores es un intervalo de valores de ángulos azimutales o en el que cada uno del primer intervalo de valores y el segundo intervalo de valores y el tercer intervalo de valores es un intervalo de valores de ángulos de elevación.

35

12. Un aparato según una de las reivindicaciones anteriores,

40

en el que  $\varphi_{izq.}^{nominal}$  indica una referencia al borde de pantalla del azimut izquierdo,

en el que  $\varphi_{der.}^{nominal}$  indica una referencia al borde de pantalla del azimut derecho,

en el que  $\theta_{sup.}^{nominal}$  indica una referencia al borde de pantalla de elevación superior,

en el que  $\theta_{inf.}^{nominal}$  indica una referencia al borde de pantalla de elevación inferior,

45

en el que  $\varphi_{izq.}^{repro}$  indica un borde azimutal izquierdo de la pantalla,

en el que  $\varphi_{der.}^{repro}$  indica un borde azimutal derecho de la pantalla,

en el que  $\theta_{sup.}^{repro}$  indica un borde de elevación superior de la pantalla,

en el que  $\theta_{inf.}^{repro}$  indica un borde de elevación inferior de la pantalla,

en el que  $\varphi$  indica el primer valor de azimut,

50

en el que  $\varphi'$  indica el segundo valor de azimut,

en el que  $\theta$  indica el primer valor de elevación,

en el que  $\theta'$  indica el segundo valor de elevación,

en el que el segundo valor de azimut  $\varphi'$  resulta de un primer mapeo del primer valor de azimut  $\varphi$  según la primera función de mapeo lineal según

55

$$\varphi' = \begin{cases} \frac{\varphi_{der}^{repro} + 180^\circ}{\varphi_{der}^{nominal} + 180^\circ} \cdot (\varphi + 180^\circ) - 180^\circ & \text{para } -180^\circ \leq \varphi < \varphi_{der}^{nominal} \\ \frac{\varphi_{der}^{repro} - \varphi_{der}^{repro}}{\varphi_{der}^{nominal} - \varphi_{der}^{nominal}} \cdot (\varphi - \varphi_{der}^{nominal}) + \varphi_{der}^{repro} & \text{para } \varphi_{der}^{nominal} \leq \varphi < \varphi_{der}^{nominal} \\ \frac{180^\circ - \varphi_{der}^{repro}}{180^\circ - \varphi_{der}^{nominal}} \cdot (\varphi - \varphi_{der}^{nominal}) + \varphi_{der}^{repro} & \text{para } \varphi_{der}^{nominal} \leq \varphi < 180^\circ \end{cases}$$

y

en el que el segundo valor de elevación  $\theta'$  resulta de un segundo mapeo del primer valor de elevación  $\theta$  según la segunda función de mapeo lineal según

$$\theta' = \begin{cases} \frac{\theta_{inf}^{repro} + 90^\circ}{\theta_{inf}^{nominal} + 90^\circ} \cdot (\theta + 90^\circ) - 90^\circ & \text{para } 0^\circ \leq \theta < \theta_{inf}^{nominal} \\ \frac{\theta_{sup}^{repro} - \theta_{inf}^{repro}}{\theta_{top}^{nominal} - \theta_{inf}^{nominal}} \cdot (\theta - \theta_{inf}^{nominal}) + \theta_{inf}^{repro} & \text{para } \theta_{inf}^{nominal} \leq \theta < \theta_{sup}^{nominal} \\ \frac{90^\circ - \theta_{sup}^{repro}}{90^\circ - \theta_{sup}^{nominal}} \cdot (\theta - \theta_{sup}^{nominal}) + \theta_{sup}^{repro} & \text{para } \theta_{sup}^{nominal} \leq \theta < 90^\circ \end{cases}$$

5

13. Un dispositivo decodificador que comprende:

un primer decodificador (910) para discurso unificado y decodificación de audio para decodificar un flujo de bits para obtener uno o más canales de entrada de audio, para obtener uno o más objetos de audio de entrada, para obtener metadatos de objeto comprimidos y para obtener uno o más canales de transporte,

10

un segundo decodificador (915) para la decodificación de objetos de audio espacial para decodificar dichos uno o más canales de transporte a fin de obtener un primer grupo de uno o más objetos de audio renderizados,

un aparato (917) según una de las reivindicaciones anteriores, en el que el aparato comprende

15

un tercer decodificador que es un decodificador de metadatos de objetos (918), que es el procesador de metadatos de objetos (110) del aparato según una de las reivindicaciones anteriores, y que es implementado para decodificar los metadatos de objeto comprimidos para obtener metadatos no comprimidos y

el renderizador de objetos (920; 120) del aparato según una de las reivindicaciones anteriores, para renderizar dichos uno o más objetos de audio de entrada dependiendo de los metadatos no comprimidos para obtener un

20

segundo grupo de uno o más objetos de audio renderizados,

un conversor de formato (922) para convertir dichos uno o más canales de entrada de audio para obtener uno o más canales convertidos y

un mezclador (930) para mezclar dichos uno o más objetos de audio del primer grupo de uno o más objetos de audio renderizados, dichos uno o más objetos de audio del segundo grupo de uno o más objetos de audio renderizados y dichos uno o más canales convertidos para obtener uno o más canales de audio decodificados.

25

14. Un procedimiento para generar señales de altavoces, que comprende:

recibir un objeto de audio,

30

recibir metadatos que comprenden una indicación de si el objeto de audio tiene relación con el tamaño de la pantalla, y que además comprenden una primera posición del objeto de audio,

calcular una segunda posición del objeto de audio dependiendo de la primera posición del objeto de audio y dependiendo de un tamaño de una pantalla, si en los metadatos se indica que el objeto de audio tiene relación con el tamaño de la pantalla,

35

generar las señales de altavoces dependiendo del objeto de audio y dependiendo de información de posición,

en el que la información sobre la posición es la primera posición del objeto de audio, si en los metadatos se indica

- que el objeto de audio no está relacionado con el tamaño de la pantalla y  
en el que la información sobre la posición es la segunda posición del objeto de audio, si en los metadatos se indica  
que el objeto de audio tiene relación con el tamaño de la pantalla,  
en el que el cálculo de la segunda posición del objeto de audio se realiza en función de la primera posición del objeto  
5 de audio y en función del tamaño de la pantalla, si el objeto de audio se indica en los metadatos como dependiente  
del tamaño de la pantalla, en el que la primera posición indica un primer azimut, una primera elevación y una primera  
distancia,  
en el que la segunda posición indica un segundo azimut, una segunda elevación y una segunda distancia,  
en el que el cálculo de la segunda posición del objeto de audio se realiza dependiendo de al menos una de una  
10 primera función de mapeo lineal y una segunda función de mapeo lineal, en el que la primera función de mapeo  
lineal se define para mapear un primer valor de azimut a un segundo valor de azimut, en el que la segunda función  
de mapeo lineal se define para asignar un primer valor de elevación a un segundo valor de elevación.
15. Un programa informático para implementar el procedimiento de la reivindicación 14 al ejecutarse en un  
15 ordenador o un procesador de señales.

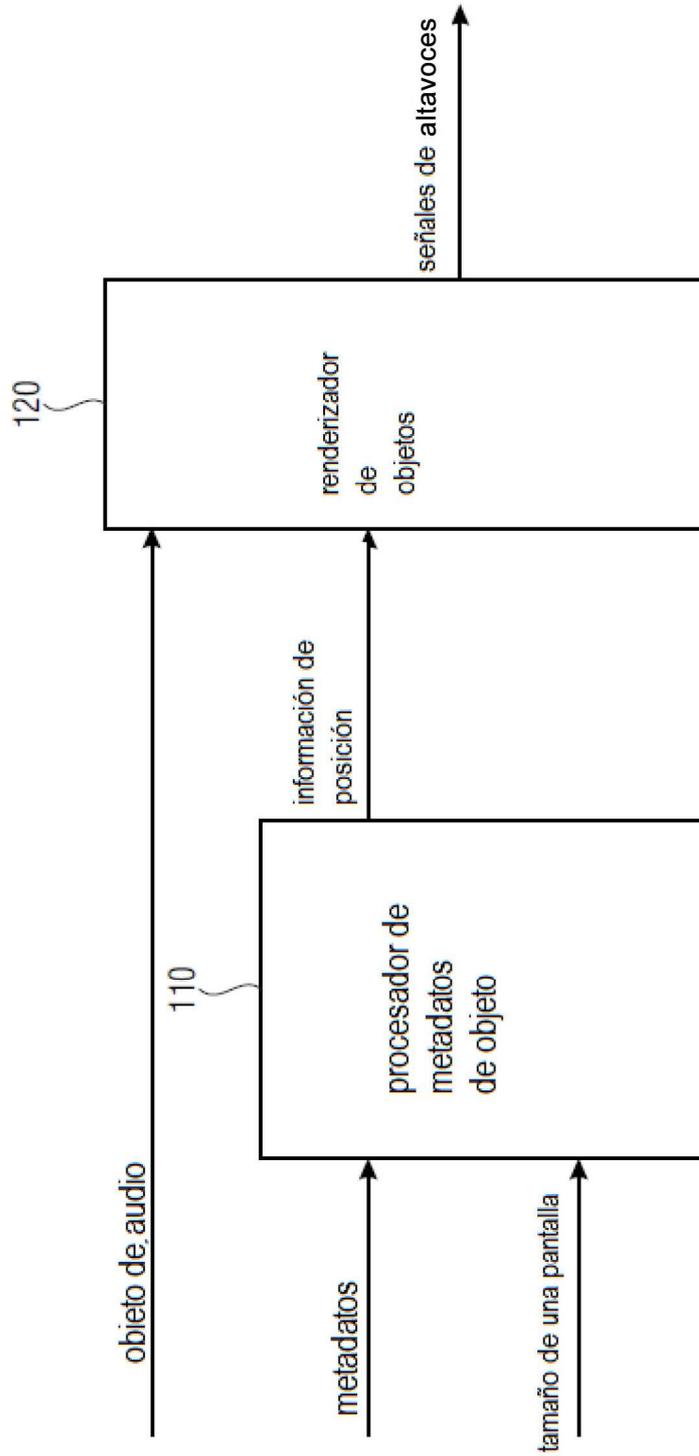
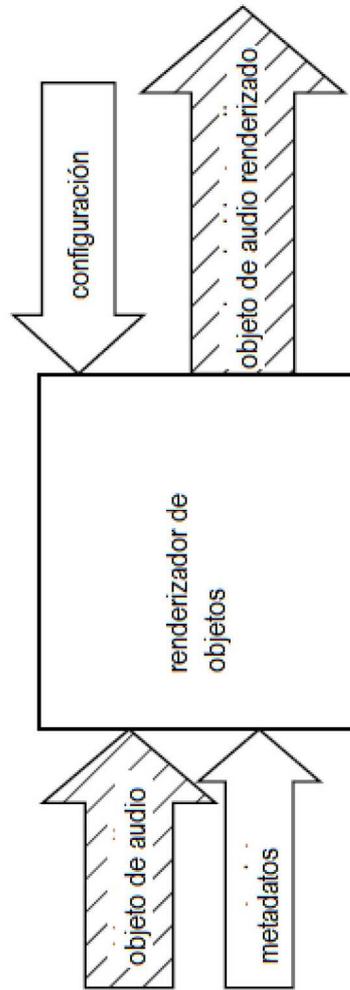


FIG. 1



**FIG. 2**

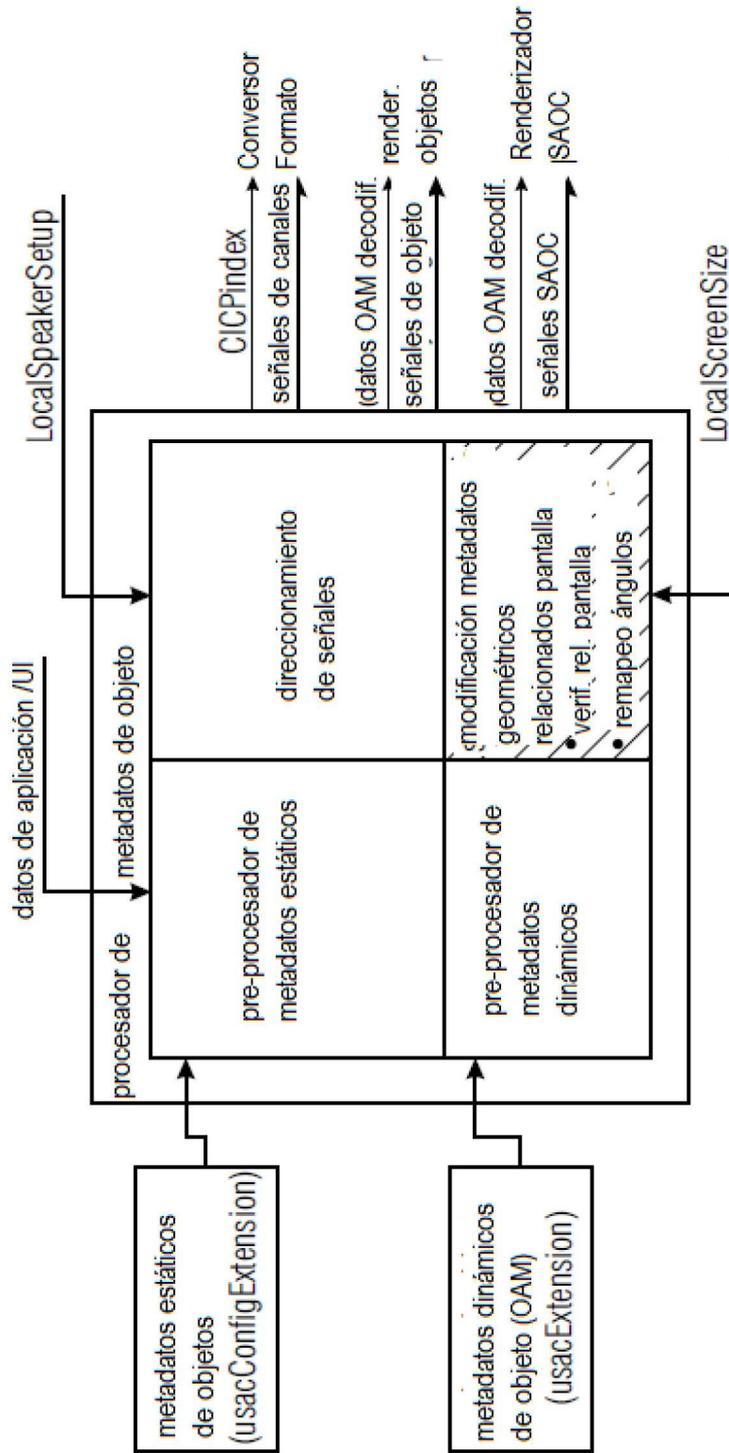
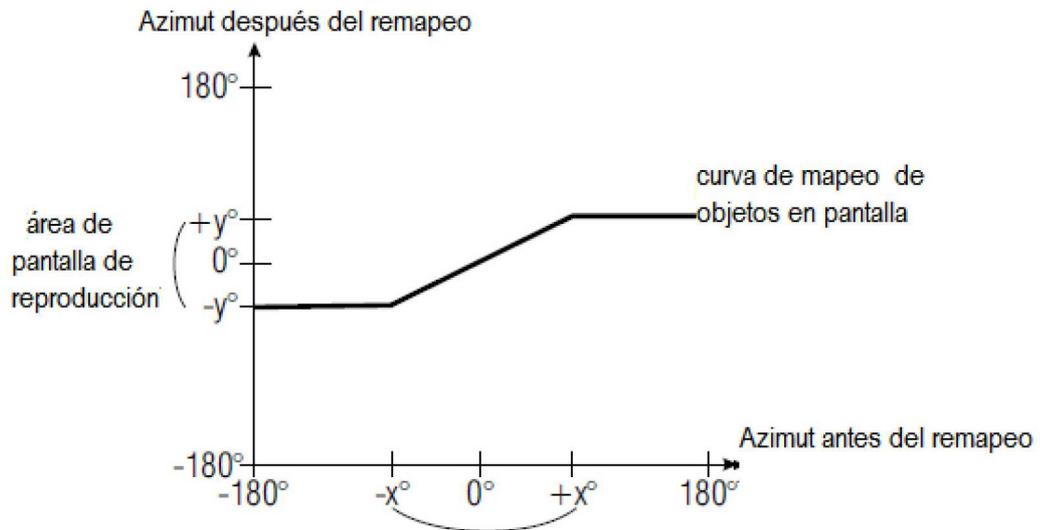
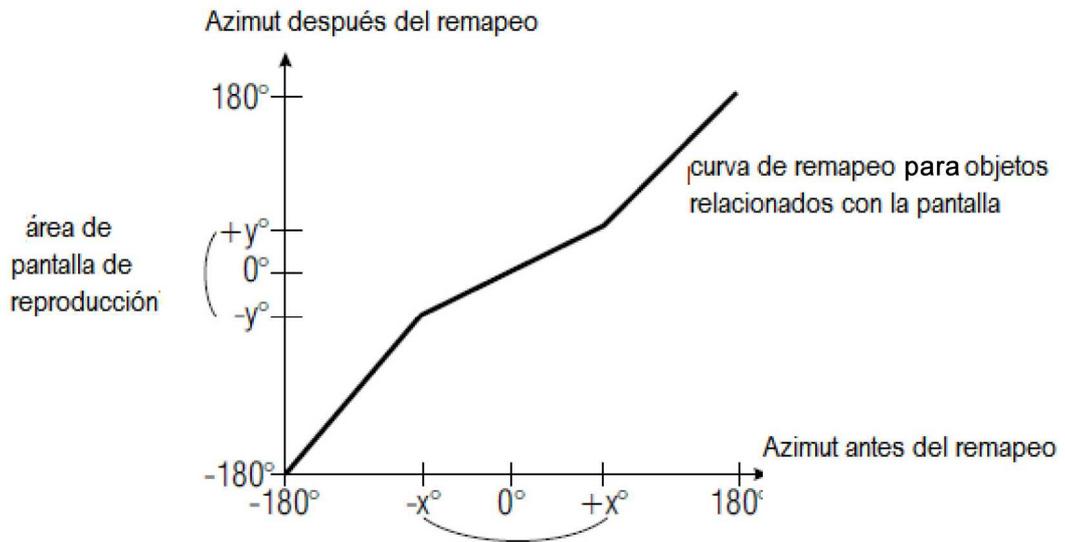
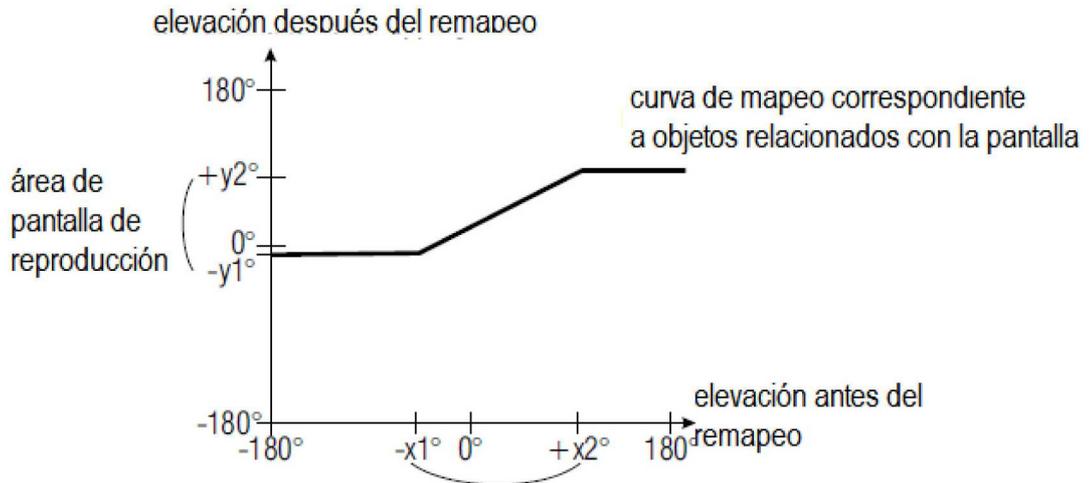
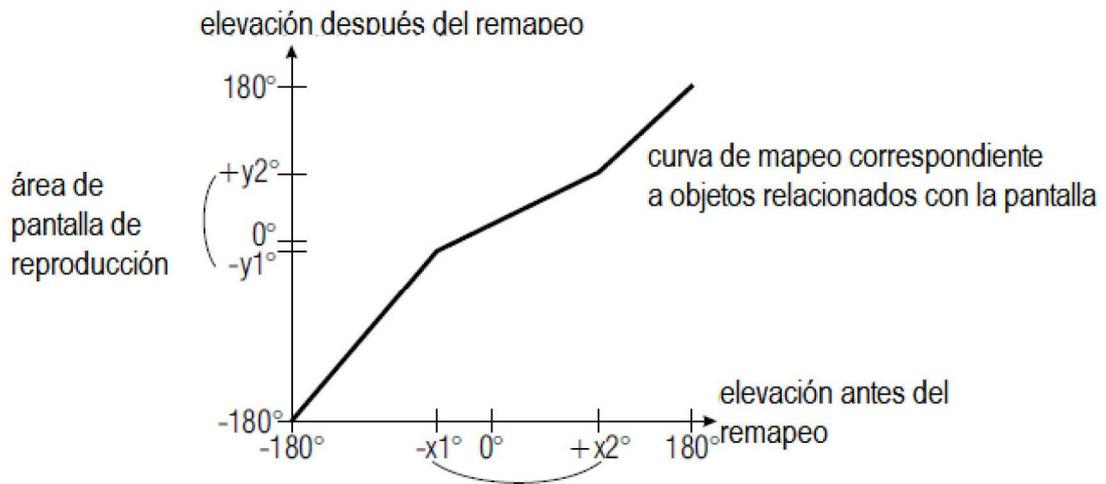


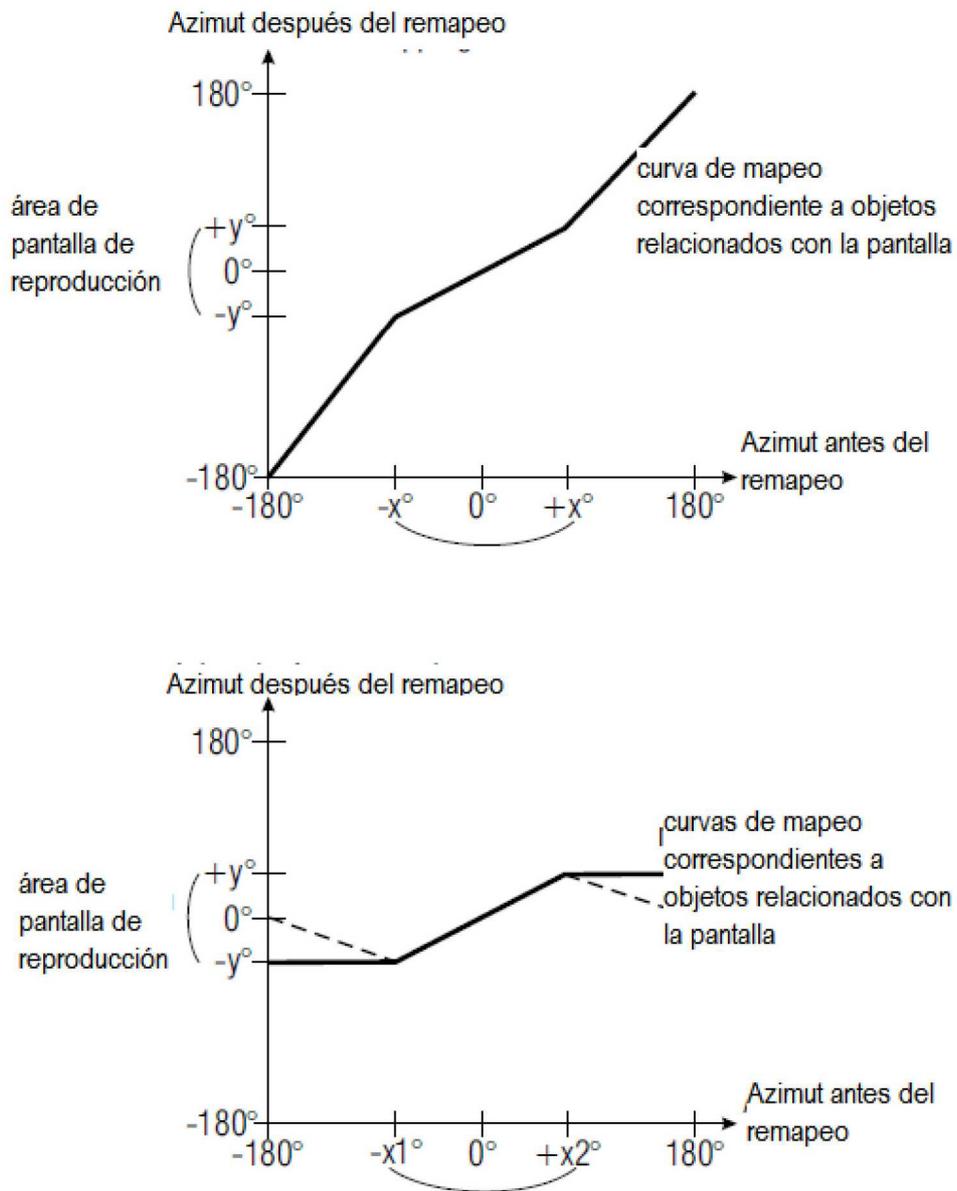
FIG. 3



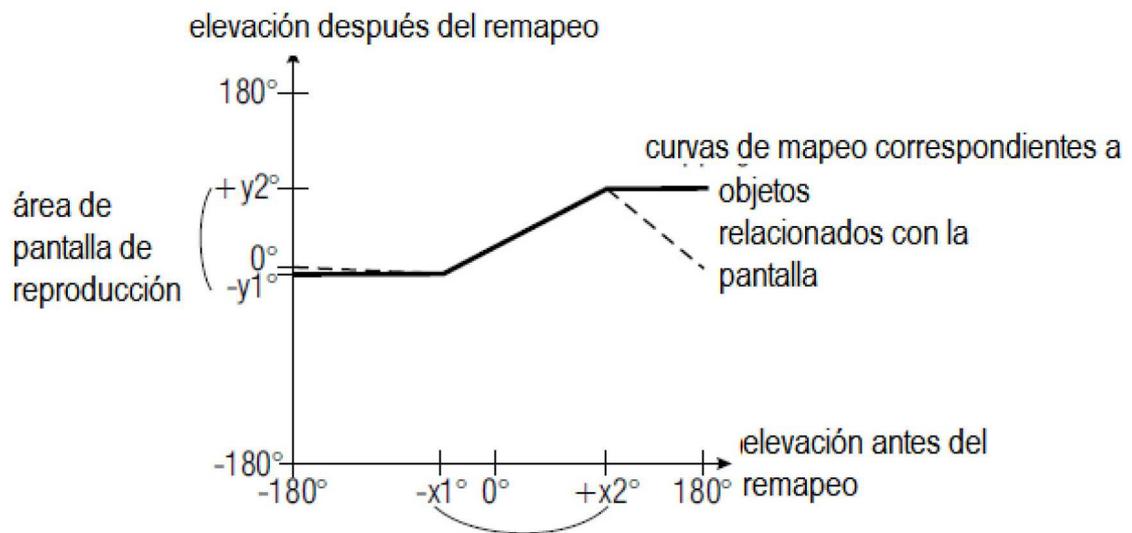
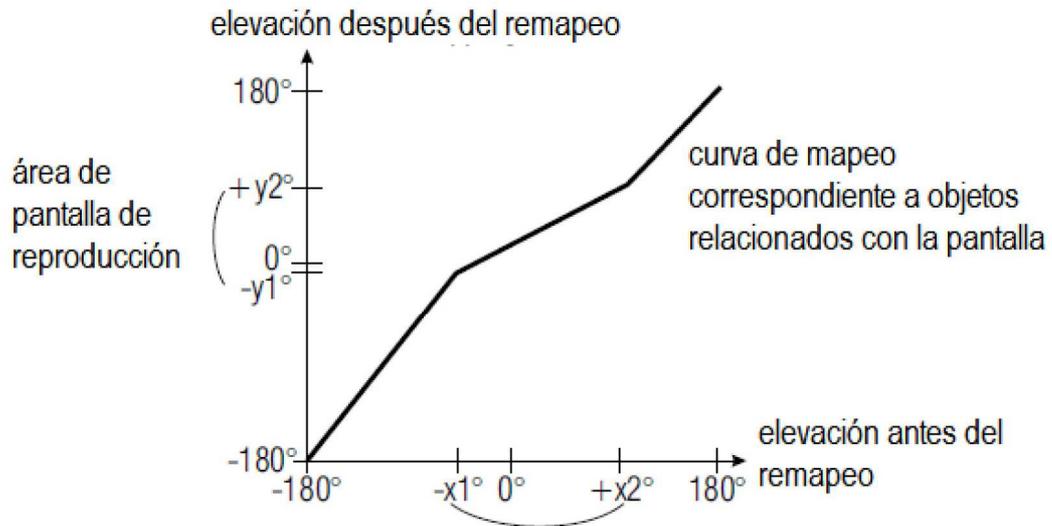
**FIG. 4**



**FIG. 5**



**FIG. 6**



**FIG. 7**

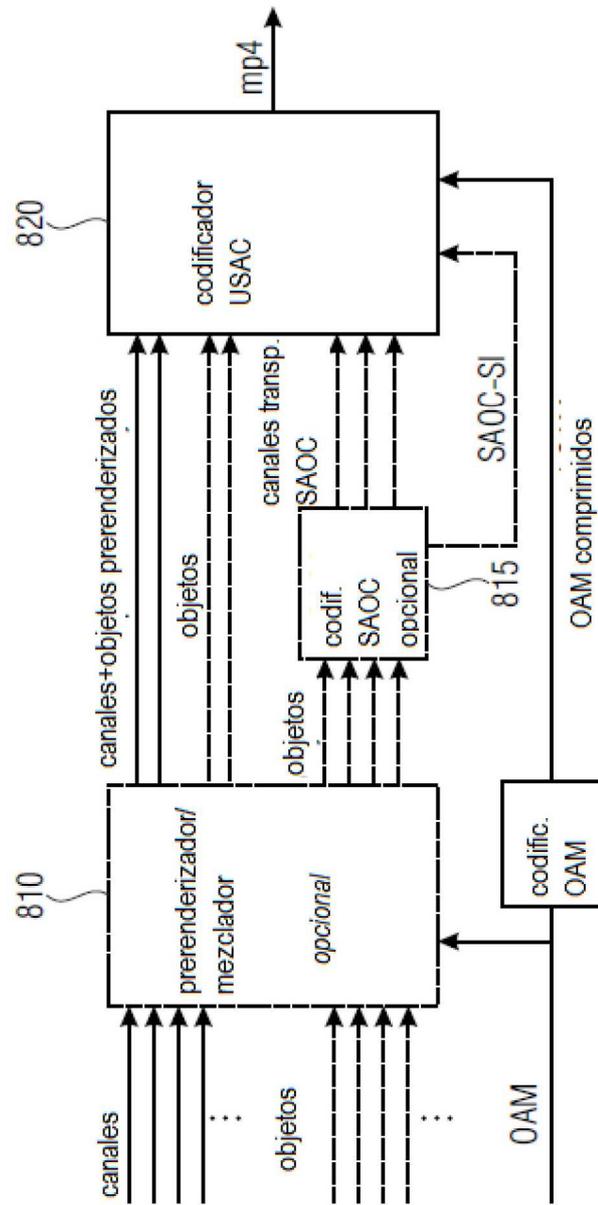


FIG. 8

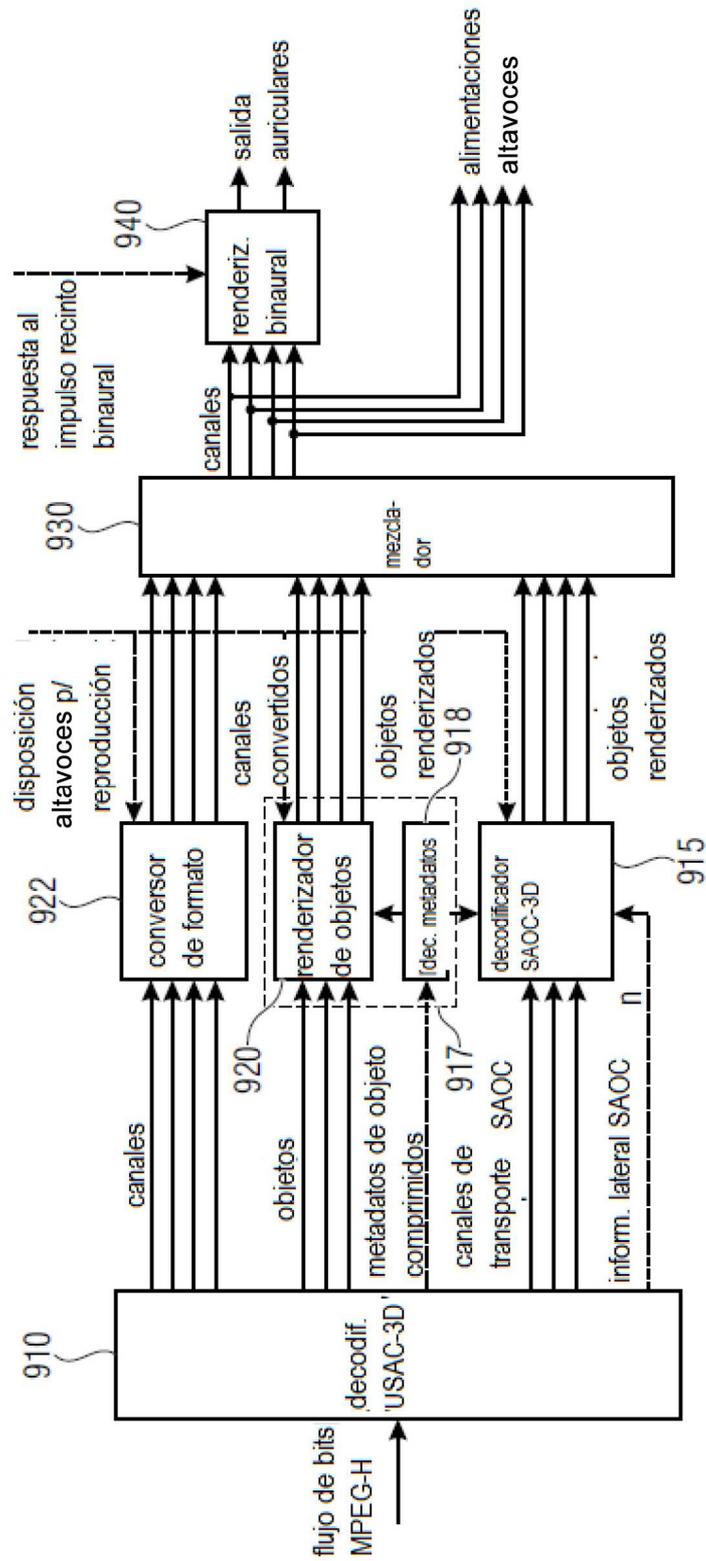
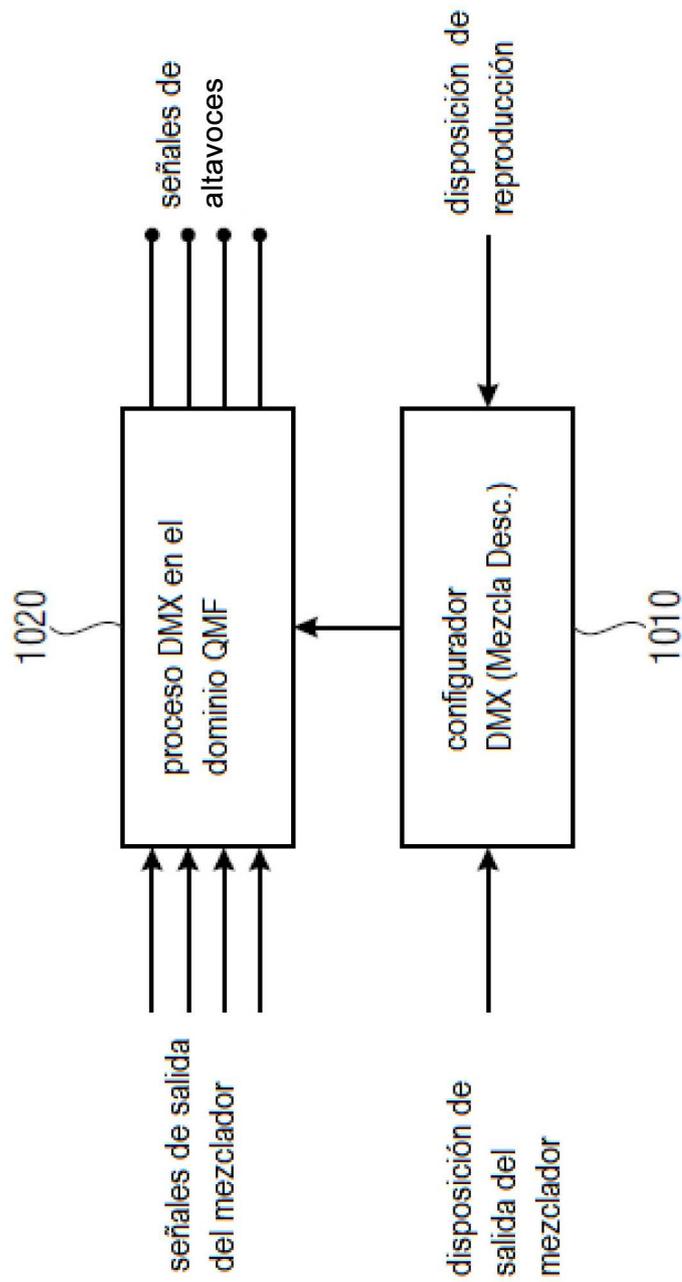
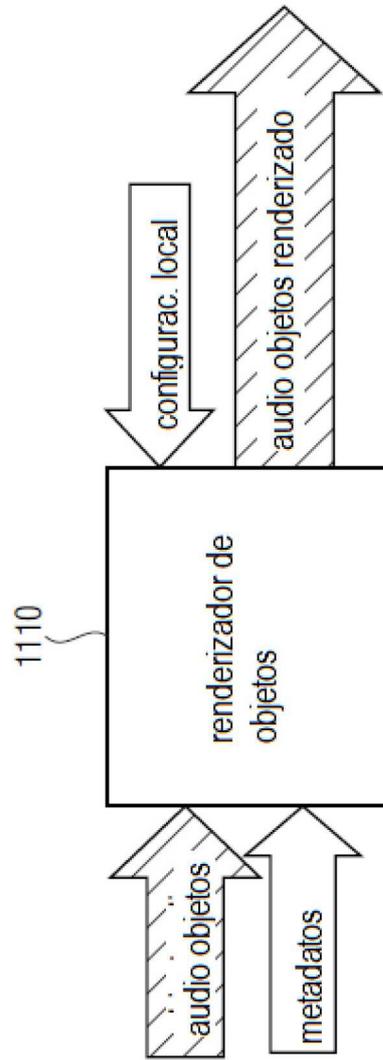


FIG. 9



**FIG. 10**



**FIG. 11**

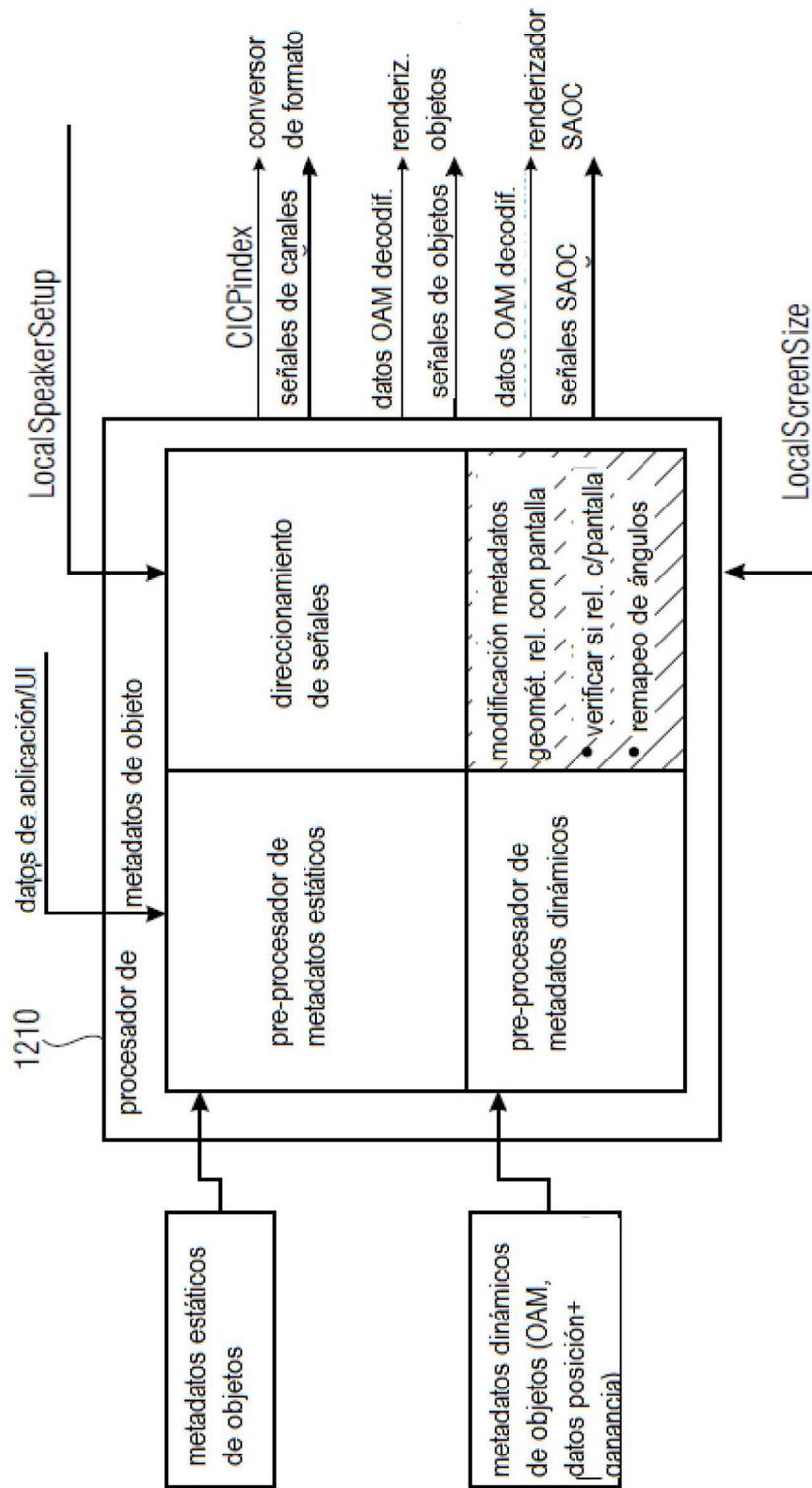
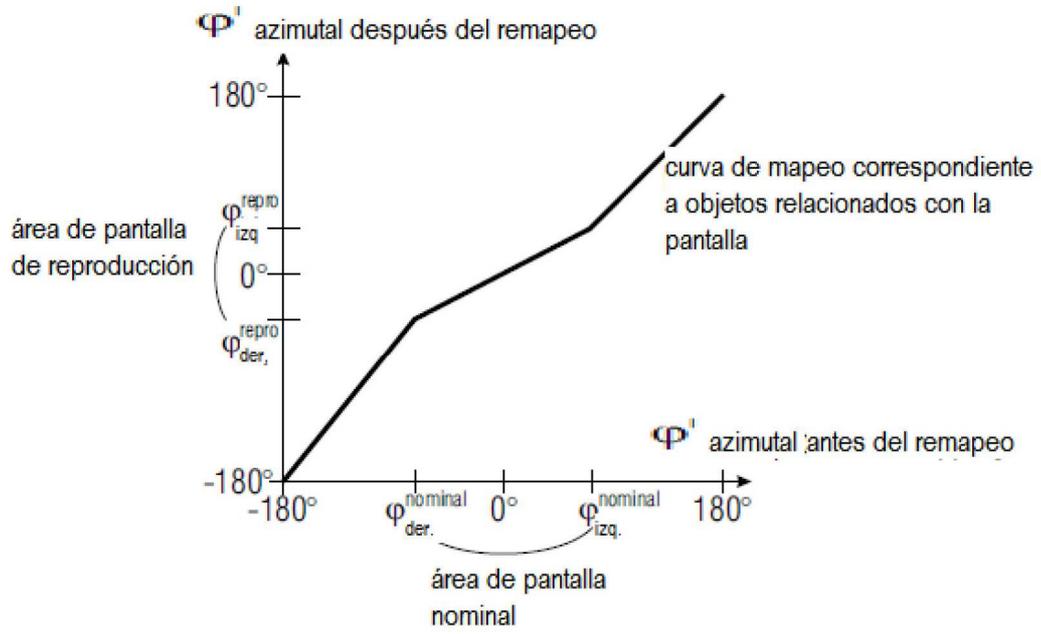
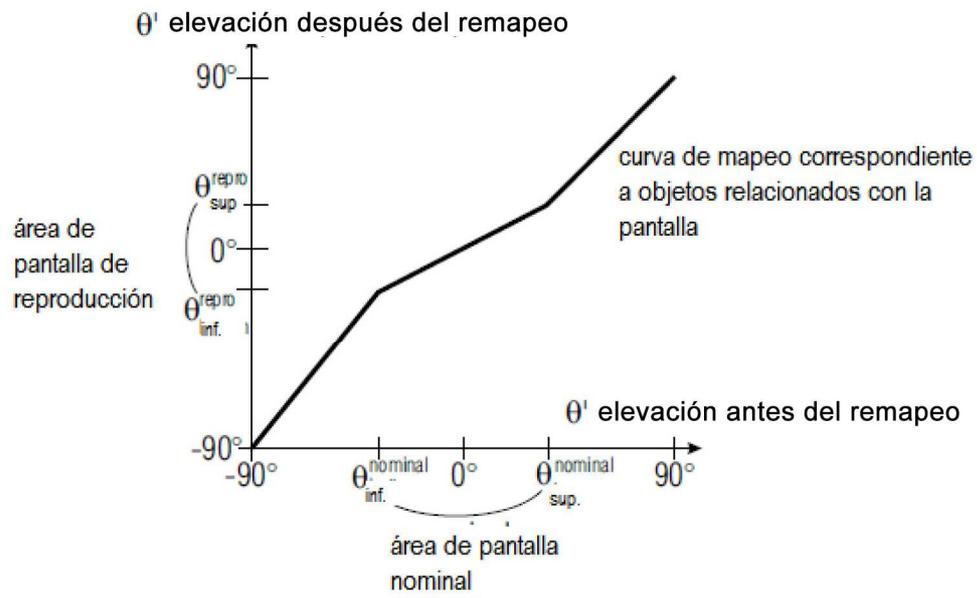


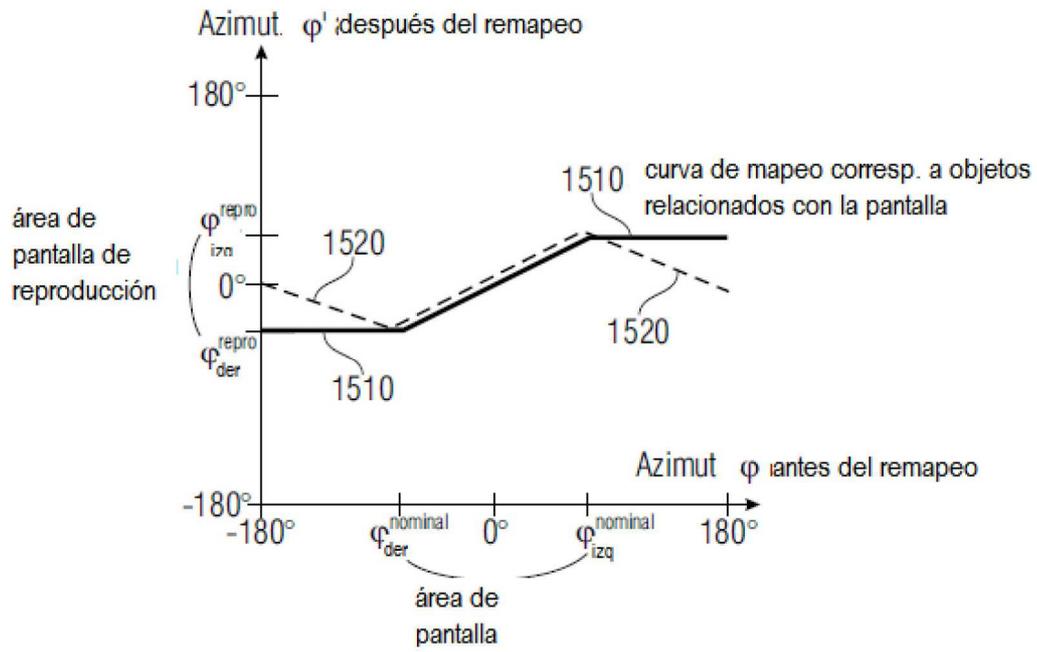
FIG. 12



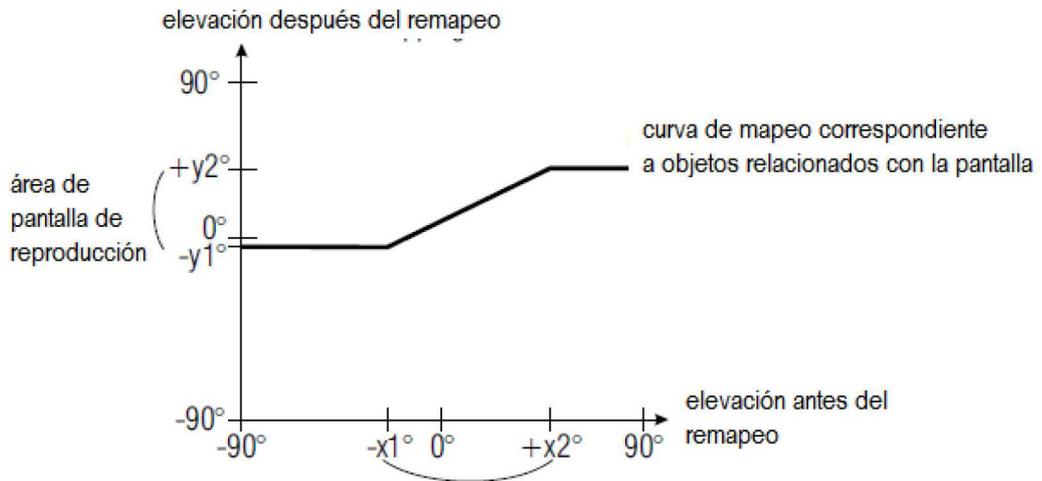
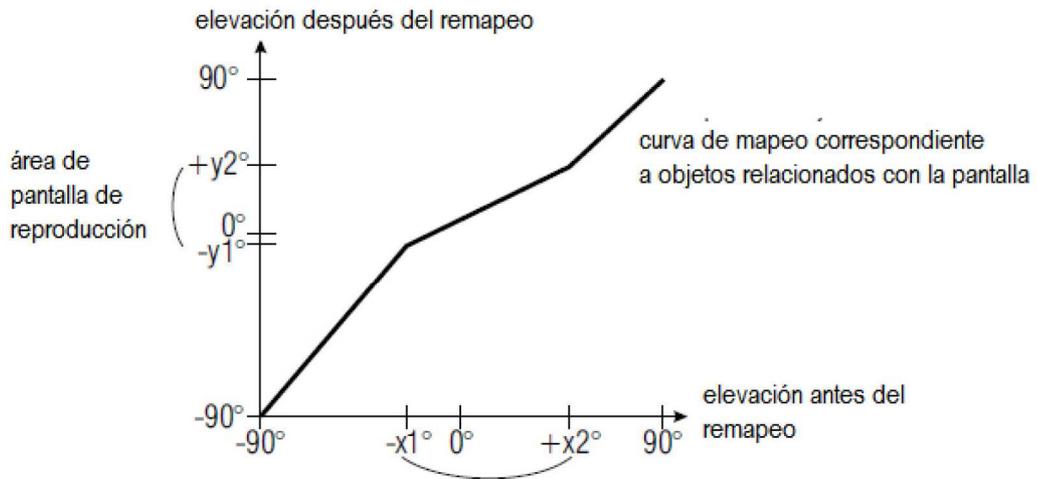
**FIG. 13**



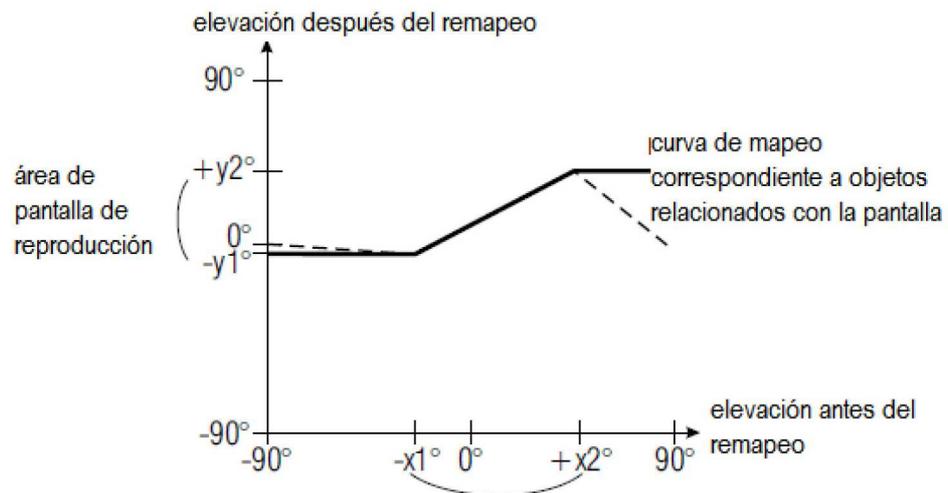
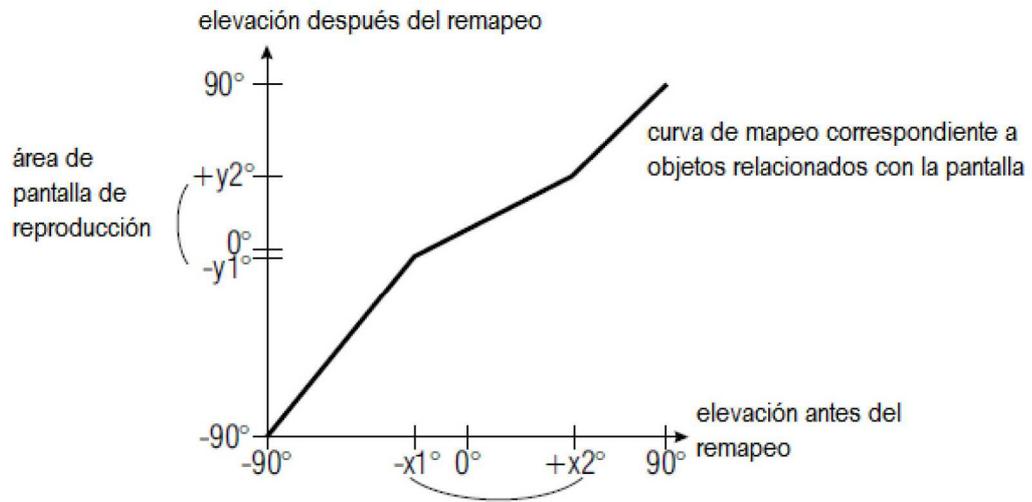
**FIG. 14**



**FIG. 15**



**FIG. 16**



**FIG. 17**