

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 435 669**

51 Int. Cl.:

H04N 13/02 (2006.01)

H04N 13/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **17.07.2009 E 11164109 (8)**

97 Fecha y número de publicación de la concesión europea: **11.09.2013 EP 2362671**

54 Título: **Gestión de subtítulos en visualización en 3D**

30 Prioridad:

25.07.2008 EP 08161152

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

20.12.2013

73 Titular/es:

**KONINKLIJKE PHILIPS N.V. (100.0%)
High Tech Campus 5
5656 AE Eindhoven, NL**

72 Inventor/es:

**NEWTON, PHILIP, S.;
BOLIO, DENNIS, D., R., J.;
SCALORI, FRANCESCO;
VAN DER HEIJDEN, GERARDUS, W., T.;
VAN DOVEREN, HENRICUS, F., P., M.;
DE HAAN, WIEBE y
MOLL, HENDRIK, F.**

74 Agente/Representante:

ZUAZO ARALUZE, Alexander

ES 2 435 669 T3

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

DESCRIPCIÓN

Gestión de subtítulos en visualización en 3D.

5

Campo de la invención

Esta invención se refiere a la creación y a la presentación de una señal de imagen tridimensional (3D). En una realización, la invención proporciona un posicionamiento óptimo automático de subtítulos en una visualización en 3D que reduce la fatiga del espectador.

10

Antecedentes de la invención

Actualmente, existe un renovado interés en la televisión en 3D, esto se relaciona con un avance reciente en la tecnología de visualización que permite una buena reproducción de vídeo en 3D para múltiples espectadores. Una de éstas es la pantalla lenticular 3D autoestereoscópica, aunque también hay otros tipos de pantallas, tales como las pantallas basadas en barreras autoestereoscópicas y las pantallas estéreo multiplexadas en el tiempo basadas en tecnología de retroproyección. Normalmente estos tipos de pantallas usan uno de dos formatos de vídeo básicos como entrada para crear una impresión en 3D al espectador. Las pantallas basadas en estéreo usan un entrelazado secuencial en el tiempo y gafas para visualizar dos vistas separadas, una para cada ojo, y por tanto esperan un vídeo estéreo como entrada. Ejemplos de éstas son las pantallas estéreo basadas en retroproyección multiplexadas en el tiempo y este tipo de sistema también se usa en el cine 3D. La alternativa principal a esto son las pantallas autoestereoscópicas de múltiples vistas, que no requieren gafas y a menudo usan un formato de entrada conocido como imagen + profundidad como entrada para generar una impresión en 3D. Puede encontrarse más información sobre tecnología de visualización en 3D en el capítulo 13 de "3D video communication - Algorithms, concepts and real time systems in human centered communication" por Oliver Shreer *et al.* (Wiley 2005).

15

20

25

El formato de vídeo estéreo es sencillo puesto que proporciona dos imágenes, una para cada ojo. Normalmente estas dos imágenes se entrelazan, o bien espacialmente o bien de una manera secuencial en el tiempo y luego se alimentan a la pantalla. El formato alternativo, denominado imagen + profundidad, es diferente porque es una combinación de una imagen en 2D con una denominada "profundidad", o mapa de disparidad. Esto es normalmente una imagen en escala de grises, donde el valor de la escala de grises de un píxel indica la cantidad de disparidad (o profundidad en caso de un mapa de profundidad) para el correspondiente píxel en la imagen en 2D asociada. La pantalla, cuando presenta la imagen en 3D usa la disparidad o mapa de profundidad para calcular las vistas adicionales tomando la imagen en 2D como entrada. Esto puede realizarse de diversas maneras, en la manera más simple se trata de desplazar píxeles a la izquierda o derecha dependiendo del valor de disparidad asociado a esos píxeles. El artículo titulado "Depth image based rendering, compression and transmission for a new approach on 3D TV" por Christoph Fen proporciona una perspectiva general de la tecnología.

30

35

Un problema con las pantallas 3D autoestereoscópicas y estéreo (basadas en secuencias en el tiempo) es lo que se denomina desajuste de acomodación-convergencia. Este es un problema en el que los ojos del espectador convergen en la ubicación virtual de los objetos que se visualizan, mientras que al mismo tiempo los ojos se acomodan (para ver la imagen nítida) en la superficie de la propia pantalla. Este desajuste puede provocar dolores de cabeza y otros síntomas asociados con mareos. Además cualquier diferencia geométrica (especialmente cualquier paralaje vertical), así como eléctrica (brillo, contraste, etc.) entre las vistas para los ojos izquierdo y derecho puede provocar adicionalmente fatiga visual. Sin embargo, si la cantidad de disparidad se mantiene pequeña, es decir menor que un grado, entonces estos problemas son menos graves y los usuarios pueden visualizar el contenido sin ningún problema significativo. Véase "Two factors in visual fatigue caused by stereoscopic HDTV images", Sumio Yano *et al.* Displays 2004 páginas 141 a 150 Elsevier, para más detalle.

45

50

Puede producirse un problema similar a los problemas anteriores cuando el dispositivo de presentación está visualizando texto tal como subtítulos o subtítulo codificado. Si el texto no está situado de manera apropiada en la pantalla, lo que depende del tipo y ajustes de la pantalla, el texto puede aparecer borroso, por ejemplo debido a una diafonía entre las vistas de los ojos izquierdo y derecho, y el espectador puede sentirse cansado. También se da el caso de que la visión borrosa puede afectar a la legibilidad del texto. Según E. Legge (véase "Psychophysics of Reading: I. Normal Vision", Gordon E. Legge *et al.* Vision Research, vol. 25, n.º 2, páginas 239 a 252, 1985) la lectura se ve afectada si el ancho de banda del texto está por debajo de dos ciclos por carácter. La visión borrosa es un problema con las pantallas autoestereoscópicas, puesto que normalmente se sacrifica la resolución para generar múltiples vistas, y para pantallas estéreo en general existe un problema con la separación subóptima entre las dos vistas, que puede añadirse a la visión borrosa de la imagen. Además según Yano (mencionado anteriormente), el movimiento de profundidad aumenta la fatiga visual.

55

60

Otro problema que se prevé, es que los espectadores pueden ajustar la cantidad de disparidad y posición relativa del plano de profundidad en una televisión en 3D (a través de, por ejemplo, algunos botones en su mando a distancia). Estos ajustes significan que el texto puede volverse borroso, cuando se aleja de una posición neutra de profundidad, o aumenta en "profundidad" de manera que provoca fatiga visual.

65

La publicación de solicitud de patente estadounidense US 2005/0140676 da a conocer un método para visualizar datos de texto de múltiples niveles en un mapa tridimensional. En el sistema descrito en esta publicación, se visualiza un mapa tridimensional en una pantalla, y se visualizan datos de texto con diferentes niveles de densidad según distancias desde un punto de vista del mapa tridimensional visualizado a nodos donde se visualizarán los datos de texto, mejorando así la legibilidad de los datos de texto. Además, es posible visualizar los datos de texto ajustando localmente la densidad de los datos de texto en la pantalla. El mapa tridimensional se visualiza en la pantalla de un panel de visualización convirtiendo los datos de mapa con coordenadas bidimensionales en aquéllos con coordenadas tridimensionales por medio de un método de proyección de perspectiva. Los datos de texto que van a visualizarse junto con el mapa tridimensional se convierten en aquéllos en un sistema de coordenadas tridimensional con un origen definido por el punto de vista del mapa tridimensional. Los datos de texto convertidos se proyectan en un plano bidimensional para su conversión en aquéllos con coordenadas de pantalla. Luego, se clasifican las distancias desde el punto de vista del mapa tridimensional visualizado a los nodos donde se visualizarán los datos de texto. Las distancias clasificadas se determinan para los datos de texto convertidos con las coordenadas de pantalla. Los datos de texto de niveles correspondientes a las distancias determinadas se visualizan en la pantalla del panel de visualización en la que se visualiza el mapa tridimensional.

Aunque en el contexto de la representación de un mapa tridimensional en un dispositivo de visualización bidimensional, la gestión de los datos de texto según esta publicación se dispone y ajusta a escala de una manera significativa para el usuario, no aborda ninguno de los problemas identificados anteriormente, en relación con la visualización de texto en un dispositivo de visualización tridimensional.

El documento JP 2004-274125 describe una generación de señal en 3D. Un multiplexor multiplexa dos componentes de imagen y datos de texto, que se superpondrán en el extremo de recepción. La señal de imagen en 3D tiene una única secuencia de datos de texto (componente D). Se incluye un correspondiente parámetro de profundidad (componente E) en la señal para permitir el posicionamiento de la secuencia de texto en el extremo de recepción.

El documento WO2008/044191 describe la creación de datos de gráficos en 3D. Se describe una secuencia de datos multiplexados que comprende vídeo y gráficos. La secuencia de datos comprende dos secuencias diferentes de datos gráficos (secuencia de gráficos de presentación y secuencia de gráficos interactivos), que se presentan en un decodificador en respectivos planos de gráficos separados. En la secuencia para cada objeto gráfico (segmento de definición de objeto que contiene gráficos en 2D) hay una correspondiente estructura de datos que contiene la información de mapa de profundidad.

El documento US2008/043095 A1 da a conocer una disposición de cámara de campo luminoso en la que la disparidad máxima entre dos vistas adyacentes se calcula en el lado de codificador y se transmite al receptor usando mensajes de SEI, sin embargo esta información no se usa para la decodificación de vídeo.

Sumario de la invención

Por tanto, un objeto de la invención es mejorar la técnica conocida.

Según un primer aspecto de la presente invención, se proporciona un método según la reivindicación 1, un método según la reivindicación 3, un dispositivo según la reivindicación 5, un dispositivo según la reivindicación 6, una señal de imagen tridimensional según la reivindicación 8, un disco para un sistema de reproducción tridimensional según la reivindicación 10 y un producto de programa informático en un medio legible por ordenador para crear una señal de imagen tridimensional según la reivindicación 11. Ventajosamente, la ubicación Z para subtítulos basados tanto en texto como en gráficos de presentación es la misma y sólo requiere que se almacene una vez por secuencia (es decir, por idioma de subtítulo). Realizaciones prácticas, por ejemplo el disco BD, para almacenar esta ubicación Z compartida de subtítulos se definen en las reivindicaciones dependientes.

Con la invención es posible mejorar la legibilidad de texto tal como subtítulos en una pantalla 3D. La legibilidad mejorada se basa en el hecho de que se envían parámetros adicionales a la pantalla 3D de manera que la pantalla puede tratar la parte de la imagen que contiene los subtítulos de manera diferente del resto de la imagen. Entonces la pantalla puede garantizar que los subtítulos se sitúen automáticamente de la mejor manera con respecto a profundidad, nitidez y legibilidad general del texto. La aplicación de la invención puede aplicarse a sistemas tales como un reproductor de disco Blu-Ray y un reproductor de DVD o HD-DVD que esté conectado a una pantalla 3D, y que muestra contenido 3D y subtítulos asociados. Resultará evidente para el experto que la señal de imagen puede obtenerse adicionalmente de manera parcial o completa a través de una red digital, tal como Internet, o una intranet.

El texto, especialmente subtítulos, debe visualizarse dentro de un alcance limitado de profundidad desde la pantalla y no deben cambiar en profundidad durante toda la presentación. Si la profundidad del texto debe permanecer constante, entonces esto también representa un problema para el posicionamiento, puesto que la profundidad del vídeo puede variar y por tanto podría ocultar partes del texto durante ciertas escenas. En resumen, deben tenerse en cuenta los siguientes factores cuando se visualizan subtítulos o subtítulo codificado en una pantalla estereoscópica 3D, la cantidad de disparidad debe ser menor que un grado, el ancho de banda del texto debe

permanecer por encima de dos ciclos por carácter, el texto debe permanecer a una profundidad constante relativa a la pantalla, y el texto no debe ocultarse por objetos de vídeo.

Para cumplir con las limitaciones de ancho de banda, el reproductor debe garantizar que la resolución del texto sea lo suficientemente alta, que el efecto fantasma se mantenga mínimo y que la velocidad del texto en movimiento no sea demasiado alta. Para cumplir con estos factores el dispositivo de reproducción debe enviar los subtítulos o subtítulos codificados en una resolución que sea suficiente para la visualización, y debe ajustar la profundidad de manera que se minimice el efecto fantasma. Esto significa normalmente que la profundidad del texto debe mantenerse neutra (lo más cercana posible a la pantalla). Sin embargo, esto puede provocar un problema puesto que partes del vídeo pueden ocultar partes del texto puesto que la profundidad del vídeo cambia dinámicamente. Esto se resuelve ajustando dinámicamente la profundidad del texto para garantizar que permanezca por delante. Sin embargo, esto significa que el texto variaría en profundidad, lo que según Yano puede provocar fatiga visual. Estos problemas pueden superarse mediante la transmisión de la información acerca de la ubicación del texto, y el ajuste de parámetros de 3D de la imagen en 3D en la ubicación del texto.

Aunque preferiblemente la profundidad del texto debe ser fija para periodos de tiempo más largos, pueden permitirse variaciones por ejemplo con el fin de conseguir efectos en 3D particulares.

Breve descripción de los dibujos

A continuación se describirán realizaciones de la presente invención, sólo a modo de ejemplo, con referencia a los dibujos adjuntos, en los que se usan números de referencia similares para indicar elementos o funciones similares:

la figura 1 es un diagrama esquemático que ilustra la creación de la imagen en 3D,

la figura 2 es un diagrama esquemático de un sistema de reproducción en 3D,

la figura 3 es un diagrama esquemático que muestra el flujo de datos y contenido en el sistema de reproducción en 3D,

la figura 4 es un diagrama esquemático de un dispositivo de reproducción del sistema de reproducción en 3D,

la figura 5 es un diagrama esquemático que muestra el uso de planos en la creación de una salida,

la figura 6 es un diagrama esquemático de una señal de imagen en 3D,

la figura 7 es un diagrama esquemático de una señal de imagen en 3D mejorada,

la figura 8 es un diagrama esquemático de una señal de imagen en 3D mejorada adicional,

la figura 9 es un diagrama esquemático, similar a la figura 4, de una segunda realización de un dispositivo de reproducción, y

la figura 10 es un diagrama esquemático de texto colocado en una imagen en 3D.

Descripción detallada

En la figura 1 se ilustra de manera esquemática la creación de una imagen tridimensional. Esencialmente, si el dispositivo de visualización de presentación final es autoestereoscópico, o funciona con pares estéreo (lo que requiere que el usuario lleve gafas especiales), el proceso es el mismo. Se proporcionan una primera componente 10 de imagen y una segunda componente 12 para crear una imagen 14 tridimensional. La segunda componente 12 en combinación con la primera componente 10 de imagen se procesan conjuntamente para obtener la salida 14 final. En todos los sistemas, la primera componente 10 de imagen es un fotograma de imagen bidimensional convencional de cualquier norma adecuada. En el sistema de par estéreo, la segunda componente 12 también es una imagen, y en los sistemas autoestereoscópicos, la segunda componente 12 es un mapa de profundidad, o un mapa de disparidad.

Es importante entender que la salida 14 final no es necesariamente un único fotograma. Por ejemplo, en el sistema de par estéreo (en el que la imagen 10 es para el ojo izquierdo, y la imagen 12 es para el ojo derecho), entonces estas dos componentes 10 y 12 pueden mostrarse de manera secuencial. En los sistemas autoestereoscópicos no secuenciales en el tiempo, se usan la imagen 10 y el mapa 12 de profundidad para crear múltiples vistas similares desde la imagen 10, usándose el mapa de profundidad para generar los cambios en la imagen 10 necesarios para realizar las vistas individuales. Entonces estas vistas se fusionan en un único fotograma 14. Por ejemplo, el dispositivo de visualización de presentación final puede controlarse para mostrar cuatro vistas individuales, todas generadas desde la misma imagen 10. Estas vistas se unen entonces en bandas, teniendo cada una una resolución que es un cuarto de la resolución del dispositivo de visualización.

La primera componente 10 de imagen puede comprender, por ejemplo, una disposición bidimensional de elementos de píxeles/una imagen que comprende, por ejemplo, información RGB o YUV, que representa una vista de una escena, tal como sería el caso para un contenido de vista estéreo, contenido de múltiples vistas o contenido de imagen+profundidad. Tal como se mencionó anteriormente, la segunda componente 12 puede ser una disposición bidimensional de elementos de píxeles/una imagen. En el caso de contenido de vista estéreo o contenido de múltiples vistas esto puede ser una vista adicional de la misma escena, mientras que en el caso de contenido de imagen+profundidad esto podría ser una denominada disparidad/mapa de profundidad que comprende información de disparidad o profundidad. La salida 14 de imagen puede comprender una señal de imagen de vídeo, por ejemplo usando múltiples fotogramas que representan una o más escenas en el tiempo.

La figura 2 ilustra un sistema de reproducción en 3D, en el que está previsto un disco 16 que lleva las componentes 10 y 12 primera y segunda para una película en 3D, por ejemplo. El disco 16 podría ser un DVD de Blu-Ray convencional, por ejemplo. El disco 16 se reproduce por un dispositivo 18 de reproducción, tal como un reproductor de DVD Blu-Ray. Se usa una interfaz 20 de vídeo para comunicar las componentes 10 y 12 a una televisión 22 en 3D. La televisión 22 incluye una fase 24 de presentación, que procesa las componentes 10 y 12 en tiempo real para generar la salida 14 en 3D (según la figura 1 comentada anteriormente). El usuario 26 está viendo el dispositivo 22 de visualización de presentación, y puede suministrar ajustes 28 de usuario al dispositivo 22, a través de una interfaz de usuario adecuada y un dispositivo de mando a distancia (no mostrado).

En tales sistemas, se soporta el uso de subtítulos (o ciertamente cualquier texto tal como menús) por el dispositivo 18 de reproducción. Por ejemplo, es convencional que el disco 16 incluya subtítulos tales como subtítulos de idioma extranjero dentro de las secciones de datos de la información almacenada en el disco 16. Éstos pueden seleccionarse por el usuario a través de un menú en pantalla, antes de que comience la reproducción del contenido de la película real en el disco 16. Entonces estos subtítulos se visualizan por el dispositivo 22 de visualización, superponiendo el contenido en 3D que se presenta. Tal como se comentó anteriormente, en la revisión de la técnica anterior, en muchas situaciones, la visualización de los subtítulos en un sistema de visualización en 3D puede provocar tensión y cansancio en los ojos del espectador. El sistema de la presente invención proporciona un método para mostrar los subtítulos (o ciertamente cualquier texto) de tal manera que se mejoran los problemas de la técnica anterior.

La solución que se proporciona es permitir que la pantalla 22 identifique el texto (por ejemplo subtítulos) en la secuencia de vídeo entrante y así que trate el texto de una manera diferente del resto de la señal de vídeo. La pantalla 22 podría garantizar, por ejemplo, que el texto permanece delante del vídeo, o comprimir y ampliar la profundidad de la señal de vídeo para garantizar que el texto no sobresale (tal como se describe en la publicación de solicitud de patente internacional WO 2008/038205). Adicionalmente, la pantalla 22 puede proporcionar una sección de la pantalla que está diseñada específicamente para mostrar texto, o bien siendo conmutable, teniendo una configuración de lente diferente (en el caso de una pantalla lenticular), o bien soportando menos vistas por ejemplo. Puesto que la legibilidad de texto en una pantalla autoestereoscópica en 3D es un problema que es casi inherente a la tecnología de pantalla autoestereoscópica, existen expectativas de mejoras adicionales no previstas actualmente en la visualización que podrían beneficiarse de poder identificar los subtítulos en la señal de vídeo entrante.

Una manera de enviar el texto para subtítulos por separado a la pantalla es transmitiendo el texto como información de subtítulo codificado que se incluye en la línea 21 del vídeo analógico para NTSC, que se usa en Europa (PAL) para Ceefax o Teletext. Sin embargo, la información de subtítulo codificado no se soporta actualmente cuando se transmite vídeo de alta definición a través de HDMI. Actualmente para resolver esto, los dispositivos de reproducción tales como reproductores de discos de DVD o Blu-Ray decodifican la información de subtítulo codificado ubicada en la secuencia de MPEG y la disponen sobre el vídeo antes de la transmisión a la pantalla. Entonces para usar esta solución sería necesario extender la especificación de HDMI para incluir la transmisión de información de subtítulo codificado. Resultará evidente para un experto que lo anterior también puede aplicarse a otras interfaces de visualización digital.

Otra solución es proporcionar un enlace de comunicación bidireccional entre el dispositivo 18 de reproducción y la pantalla 22, de manera que cuando el usuario fija los ajustes de profundidad en la pantalla 22, se informa al dispositivo 18 de reproducción de este cambio. Como resultado el dispositivo 18 de reproducción puede ajustar el posicionamiento de los subtítulos. La realización preferida es permitir que el dispositivo 18 de reproducción gestione el posicionamiento y la disposición de los subtítulos sobre el vídeo e indique a la pantalla 22 que los subtítulos están presentes y dónde se ubican los subtítulos. Entonces la pantalla 22 puede garantizar que el valor de "profundidad" asociado con los subtítulos se mapea con una profundidad que es la más óptima para esa pantalla 22 particular y los ajustes 28 de usuario asociados. Esto tiene la ventaja añadida de que la pantalla 22 no requiere un decodificador de subtítulos o de subtítulo codificado en la fase de presentación.

En una realización, la mejora del sistema se consigue porque el dispositivo 18 de reproducción, tal como se muestra en la figura 2, ajusta el posicionamiento de los subtítulos, de manera que el texto está delante del vídeo, mientras se mantiene la cantidad de disparidad por debajo de un grado. Preferiblemente, el dispositivo 18 de reproducción inserta en la secuencia de salida información de metadatos que indica al dispositivo 22 de visualización si y dónde

están ubicados los subtítulos con respecto a la posición de x, y y z (“profundidad”) en el volumen de proyección de la pantalla. En este caso, la fase 24 de presentación en el dispositivo 22 de visualización ajusta el posicionamiento de los subtítulos en el volumen de proyección de la pantalla, dependiendo de la información de metadatos mencionada anteriormente y de los ajustes 28 preferidos por el usuario (con respecto a la cantidad de disparidad y posición relativa), mientras se mantiene el posicionamiento de manera que el efecto fantasma se mantiene mínimo y la cantidad de disparidad permanece por debajo de un grado. Adicionalmente, la pantalla 22, si está equipada para ello, sitúa los subtítulos en una parte especial de la superficie de visualización que puede conmutarse entre 2D y 3D, o que soporta menos efecto fantasma y una resolución óptica mayor (por ejemplo soportando menos vistas, o una cantidad limitada de disparidad).

En una realización, la componente 30 de texto es sustancialmente plana, y/o cuando un alcance de profundidad está asociado con el texto 30 entonces ese alcance de profundidad se limita por un umbral, de manera que la disparidad entre respectivas vistas se limita a un alcance específico, que puede ser un alcance predeterminado, tal como por ejemplo de uno a dos píxeles. En una realización preferida en la que la componente 30 de texto es sustancialmente plana, la componente de texto es una componente de texto basada en texto en lugar de una componente de texto basada en mapa de bits; se permite así una representación particularmente compacta de la componente 30 de texto.

El flujo de datos y contenido en el sistema de la figura 2 se ilustra en la figura 3. El disco 16 contiene la primera componente 10 de imagen y la segunda componente 12, más componente 30 de texto, que son subtítulos. Se recibe una instrucción 32 de usuario por el dispositivo 18 de reproducción, que indica que el usuario desea que se visualicen los subtítulos 30 junto con la película en 3D que va a ver. El dispositivo 18 de reproducción proporciona las componentes 10 y 12, con los subtítulos 30 y una componente 34 de datos que comprende información de ubicación que describe la ubicación de la componente 30 de texto dentro de la imagen 14 en 3D final, al presentador 24. El dispositivo de reproducción incluye un receptor para recibir las diversas componentes desde sus respectivas fuentes y un multiplexor para combinar estos cuatro elementos 10, 12, 30 y 34 en una señal 36 de imagen tridimensional, que se recibe por el presentador 24.

La información 34 de subtítulo puede enviarse al presentador 24 separada de los datos 10 y 12 de imagen, es decir, quizás no en una zona de imagen activa sino en una cabecera, o en islas de datos, o en una parte del fotograma que no contiene datos de imagen. Por ejemplo puede ser que se envíe una secuencia de vídeo al doble de la velocidad de imagen normal, conteniendo un fotograma los datos 10 y 12 de imagen y el otro fotograma información de profundidad (que opcionalmente también incluye la no oclusión) y también contiene un área especial para los subtítulos 30 y la información 34 de posición.

Entonces el presentador 24 está dispuesto para presentar la imagen 14 tridimensional desde la primera componente 10 de imagen y la segunda componente 12, la presentación (incluyendo la presentación de la componente 30 de texto en la imagen 14 tridimensional), incluyendo la presentación de la componente 30 de texto ajustar uno o más parámetros tridimensionales de la imagen 14 tridimensional en la ubicación de la componente 30 de texto presentada. La propia componente 30 de texto puede permanecer totalmente inalterada por los ajustes realizados por el presentador 24. El factor importante es que el presentador ajustará los parámetros de las partes de la imagen 14 que se visualizan en la ubicación del texto 30. Esto puede comprender reducir la profundidad percibida de la imagen en 3D en esta área de la imagen 14, o puede comprender reducir el número de vistas (en un dispositivo de visualización autoestereoscópico) o una combinación de dos o más ajustes. También es posible que el presentador 24 incluya no sólo el cambio de la profundidad del área en la que van a visualizarse los subtítulos 30, sino que también incluya un desplazamiento, con el fin de mover todo el alcance de profundidad hacia delante o hacia atrás. Este desplazamiento puede controlarse por separado para los subtítulos 30 desde el resto de los datos 14 de imagen.

En la realización que usa una pantalla autoestereoscópica, la segunda componente 12 comprende un mapa de profundidad, y el método para presentar la imagen 14 en 3D puede comprender además un ajuste a escala de la segunda componente 12 antes de presentar la imagen 14 tridimensional, con el fin de permitir la colocación de la componente 30 de texto en la ubicación tal como se especifica por la componente 34 de datos. En este caso el presentador puede recibir además una componente de recomendación que comprende un ajuste a escala recomendado de la segunda componente 12, de manera que la información de ubicación puede permanecer igual para múltiples imágenes consecutivas. Como resultado del uso de esta componente de recomendación, resulta posible maximizar la profundidad de la escena cuando se desactivan los subtítulos/gráficos y cuando se activan los subtítulos/gráficos, para usar una componente de ajuste a escala predeterminada adaptada a la señal de imagen, para presentar el contenido que incluye subtítulos/gráficos.

Un ejemplo de un dispositivo 18 de reproducción se muestra en la figura 4, que es un decodificador 18 de disco Blu-Ray que muestra la decodificación y el uso de planos de presentación. Una unidad 38 ROM recibe el disco 16 y lee el contenido que se demodula y decodifica por el componente 40. Un conjunto de memorias 42 intermedias paralelas almacenan diferentes componentes de la señal decodificada, y éstas pasan por un conjunto de decodificadores 44 paralelos para proporcionar la salida que va a visualizarse por la pantalla 22, que es un vídeo 46 convencional y el contenido 48 de superposición (que son subtítulos y menús, etc.).

En el sistema de Blu-Ray, los subtítulos en disco pueden o bien estar basados en mapa de bits o bien estar basados en texto y fuentes relacionadas y adicionalmente el reproductor 18 proporciona soporte de información de subtítulo codificado. Técnicamente, ambos sistemas son algo similares, aunque los subtítulos codificados proporcionan a menudo más información y se proporcionan específicamente para personas con discapacidad auditiva. Los subtítulos tanto basados en texto como basados en mapa de bits y subtítulos codificados se decodifican y presentan en uno de los planos de presentación en el disco Blu-Ray. La figura 4 muestra el decodificador 18 y los planos, mostrando un ejemplo de texto combinado con vídeo. El texto se presenta siempre en el plano de gráficos de presentación (PG); este plano se dispone en el dispositivo 18 de reproducción sobre el vídeo (planos) y se presenta como una presentación combinada en la salida. La figura 5 muestra un ejemplo de una presentación 50 combinada. Esta salida 50 se forma a partir de un plano 52 de película principal, un plano 54 de presentación y un plano 56 interactivo. La salida 50 tal como se muestra en la figura 5 se envía a la pantalla 22 a través de la interfaz 20 de vídeo (tal como se muestra en la figura 2). En la realización preferida, el dispositivo 18 de reproducción puede enviar información adicional a la pantalla 22 que contiene los datos 34 en 3D, además de la salida 50. Esto permite que la pantalla 22 en 3D muestre una representación en 3D de la salida de vídeo, texto y gráficos combinada del dispositivo 18 de reproducción.

La figura 6 muestra un ejemplo de una salida combinada de este tipo para su envío a la pantalla 22 3D. El sistema en 3D que se usa es un ejemplo de una salida de imagen y profundidad. La primera componente 10 de imagen y la segunda componente 12 (que es un mapa de profundidad) se combinarán para realizar la imagen en 3D por el dispositivo 22 de visualización. La componente 30 de texto se incluye dentro de la primera componente 10, y ciertamente todo el contenido (las tres componentes 10, 12 y 30) puede formarse como un único fotograma de alta definición.

La figura 7 muestra la adición de una cabecera que indica la ubicación de los subtítulos 30. La cabecera comprende una componente 34 de datos que comprende información 58 de ubicación que describe la ubicación de la componente de texto dentro de la imagen tridimensional, que se muestra en este caso como coordenadas x e y, aunque también estará presente una componente z en virtud del contenido del mapa 12 de profundidad. La cabecera permite que el dispositivo de presentación final realice ajustes a la salida de 3D para tener en cuenta la presencia de los subtítulos 30.

Los parámetros adicionales que se transmiten, pueden estar contenidos en la cabecera por encima de los datos de imagen tal como se describe en "3D interface Specifications- white paper", Philips 3D solutions, <http://www.business-sites.philips.com/shared/assets/global/Downloadablefile/Philips-3D-Interface-White-Paper-13725.pdf>, en línea con la publicación de solicitud de patente internacional WO 2006/137000A1, o por ejemplo en la isla de datos de la norma de HDMI. Estos parámetros consisten en una indicación de si los subtítulos están presentes y su ubicación en el vídeo de entrada tal como se muestra en la figura 7. Entonces el dispositivo 22 de visualización puede garantizar que la disparidad de la parte de la imagen ubicada en esta posición no vaya más allá de un grado y permanezca constante, incluso cuando el usuario, a través de una instrucción de entrada, aumenta la cantidad de disparidad mostrada por el dispositivo 22 de visualización.

La calidad del resultado final puede mejorarse, si la salida del reproductor 18 también contiene información de oclusión o de fondo. Esto se muestra en la figura 8. Para evitar artefactos, los subtítulos 30 deben estar ubicados fuera de la parte más activa de la imagen, por ejemplo en la parte superior o parte inferior. La figura 8 muestra una imagen, profundidad y formato de salida de fondo. Las componentes 60 y 62 de fondo se proporcionan para la primera componente 10 de imagen y la segunda componente 12, respectivamente.

La ubicación de los subtítulos 30 dentro de la imagen 14 final puede incluir una referencia a la ubicación Z de los subtítulos. Por ejemplo, es posible extender el modelo de decodificador de subtítulos de texto de disco Blu-Ray e información de composición asociada (específicamente el estilo de diálogo establecido, sección 9.15.4.2.2) con una posición Z de manera que el autor del contenido original puede indicar dónde deben posicionarse los subtítulos en el espacio de proyección en dirección Z de la pantalla 3D. En la norma de disco Blu-Ray se define un decodificador de subtítulos de texto y secuencia relacionada. El decodificador de subtítulos de texto se define en la sección 8.9 y consiste en diversos elementos de procesamiento y memorias intermedias. La figura 9 muestra una imagen simplificada del modelo de decodificador, extendido para gestionar imágenes en 3D con una componente de texto asociada tal como un subtítulo 30.

La figura 9 muestra un modelo de decodificador de subtítulos de texto en 3D. Los segmentos de subtítulo de texto de izquierda a derecha entran en el decodificador 18 en el que se procesan y decodifican por el procesador 64 de secuencia de texto. Los datos de texto decodificados se disponen en la memoria 66 intermedia de diálogo, mientras que la información de composición de subtítulo de texto decodificado se dispone en la memoria 72 intermedia de composición. Un controlador 74 interpreta la información de composición y aplica esto al texto después de que se haya presentado por el presentador 68 de texto (usando un archivo 80 de fuente) y dispone el resultado en las memorias 70 intermedias de mapa de bits. En la etapa final, las imágenes de mapa de bits se componen por el reproductor 18 sobre los planos de gráficos apropiados. La entrada 78 de usuario también se recibe por el controlador 78, que puede afectar, por ejemplo, a los parámetros de profundidad de la pantalla.

Además de subtítulos basados en texto, el disco Blu-ray (BD) también soporta subtítulos basados en imágenes de mapa de bits, los denominados gráficos de presentación (PG). De la misma manera, se requiere que la posición Z de los subtítulos PG se especifique en el disco, preferiblemente cerca de la información de posición X,Y ya definida. Puesto que estos últimos se almacenan en la estructura `composition_object()`, es lógico extender esto con la posición Z de los subtítulos, por ejemplo usando los bits reservados que comienzan en un bit 26 de desplazamiento. El término componente de texto tal como se usa en toda la presente solicitud se usa para hacer referencia a subtítulos basados en texto y/o tales imágenes de mapa de bits.

El decodificador 18 se ha extendido con una memoria 76 intermedia de mapa de bits adicional que contiene la profundidad o mapa de disparidad que indica dónde debe estar ubicado el subtítulo 30 de texto en la dirección Z. La información de profundidad o disparidad puede estar contenida en el estilo de diálogo establecido tal como se define en la especificación de Blu-Ray. Un método para conseguir esto es extender el estilo de diálogo establecido con una `region_depth_position`. La `region_depth_position` es un campo de 8 bits con valores entre 0 y 255, pero en el futuro esto puede extenderse. El controlador 74 determina los valores de píxeles de mapa de disparidad o profundidad correctos basándose en el valor del campo `region_depth_position`. La traslación de este valor a un valor de color para el mapa de disparidad o profundidad depende de los metadatos en 3D que se llevan en el campo de metadatos en 3D en la lista de reproducción o que se llevan en la tabla de mapa de programa de MPEG-2 (ISO/IEC 13818-1). La sintaxis de esta información se define por MPEG en ISO/IEC 23002-3.

Entonces la posición Z de los subtítulos 30 se proporciona al presentador 24, que puede tener en cuenta la posición Z de los subtítulos 30, cuando se presenta la imagen 14 en 3D, para su visualización por el dispositivo 22 de visualización en 3D. La figura 10 muestra un ejemplo de un espacio en 3D con el posicionamiento de la profundidad a través del uso del campo `region_depth_position`. La imagen 14 incluye el subtítulo 30, que está colocado en el plano 82. El plano 82 indica la posición en el espacio a una profundidad cero, las otras líneas indican las tres direcciones x, y y z en el espacio en 3D.

Además de tener incluida una zona de profundidad para indicar la región en el espacio en la que deben situarse los subtítulos, también es posible ir incluso más allá añadiendo una profundidad separada para el propio texto 30, y especificar la región no como un plano sino como un cubo en el espacio. Esto puede realizarse, por ejemplo, extendiendo los campos de posición de región en el estilo de diálogo establecido con `region_depth_position` y `region_depth_length`. Lo mismo puede realizarse para las posiciones de casilla de texto que indican la ubicación exacta del texto dentro de la región.

El archivo 80 de fuente puede usarse para incluir un estilo de fuente en relieve, puesto que se ha mostrado que mejora la legibilidad del texto resultante en 3D. Los estilos de fuente se describen en la tabla 9-75 de la especificación de Blu-Ray, alternativamente se usa el campo de grosor de diseño de fuente por este motivo, que es la tabla 9-76 de la especificación de Blu-Ray. Ambas tablas tienen campos reservados al final que podrían usarse con este fin. El estilo de fuente en relieve sería el valor 0x08 en la tabla 9-75 y para el grosor de diseño de fuente en relieve sería un valor 0x04 en la tabla 9-76.

Ventajosamente, la ubicación Z para subtítulos tanto basados en texto como basados en PG es la misma y sólo es necesario que se almacene una vez por secuencia (es decir por idioma de subtítulo). Es evidente para el experto que hay muchas ubicaciones alternativas en el disco BD para almacenar esta ubicación Z compartida de subtítulos. Ejemplos de tales ubicaciones alternativas se describen a continuación. El documento "White paper Blu-ray Disc Format 2.B Audio Visual Application Format Specifications for BD- ROM March 2005" disponible a través de Internet http://www.blu-raydisc.com/Assets/Downloadablefile/2b_bdrom_audiovisualapplication_0305-12955-15269.pdf, contiene información de fondo adicional sobre el formato de BD y las estructuras en el formato mencionado a continuación.

La ubicación Z compartida de subtítulos podría almacenarse por ejemplo en una nueva tabla definida como datos de extensión de la lista de reproducción. La lista de reproducción en la especificación de BD es una lista que indica una secuencia de ítems de contenido audiovisual que conjuntamente forman la presentación de un título, tal como una película. Esta estructura de lista de reproducción proporciona un mecanismo para futuras extensiones a través de sus datos de extensión. La ubicación "Z" del plano de subtítulos (plano de PG) para las diversas secuencias de idioma puede incluirse aquí en una nueva tabla denominada tabla de metadatos de desplazamiento. La tabla 1 muestra los detalles de esta tabla.

Además en el caso en el que una lista de reproducción contiene ítems de reproducción que permiten la reproducción de datos de vídeo multiángulo estereoscópico entonces la ubicación "Z" para la superposición de gráficos tales como subtítulos puede ser diferente para cada clip de vídeo multiángulo estereoscópico. Por tanto los metadatos de desplazamiento también deben permitir ubicaciones "Z" diferentes para cada clip de vídeo multiángulo estereoscópico mencionado en el ítem de reproducción. En tal caso la tabla de metadatos de desplazamiento contiene ubicaciones "Z" diferentes para superponer subtítulos para cada clip de vídeo multiángulo estereoscópico. Entonces estas ubicaciones "Z" diferentes pueden indicarse con un identificador para cada clip de vídeo multiángulo que requiere una ubicación "Z" diferente para superponer subtítulos. La tabla 2 muestra un ejemplo de cómo la tabla 1 puede extenderse para soportar ubicaciones "Z" diferentes para clips de vídeo multiángulo estereoscópico

diferentes. Las referencia StreamID y AngleID en la tabla 1 y 2 sirven como identificadores únicos de respectivamente una secuencia básica en el disco (que contiene contenido de subtítulo) y un clip de vídeo estereoscópico (multi)ángulo.

5 En lugar de usar los datos de extensión también es posible definir una nueva estructura de tabla de lista de reproducción específicamente para reproducir secuencias en 3D e incluir los parámetros de ubicación "Z" en las mismas. Los problemas con la reproducción en reproductores de herencia pueden evitarse proporcionando una nueva tabla de índice (una lista de títulos que puede seleccionar el usuario) en el disco o usando los datos de extensión de la tabla de índice que enumera los títulos que pueden reproducirse sólo por un reproductor que soporta 10 3D.

Alternativamente si la información debe proporcionarse en una tabla existente en la lista de reproducción entonces puede incluirse en STN_Table_SS(). Ésta es una tabla que enumera las secuencias de vídeo y gráficos asociadas al ítem de reproducción. Por ítem de reproducción contiene un bucle con información de codificación de secuencia 15 (atributos) para cada secuencia de subtítulos de texto y mapa de bits. Se propone incluir la información de ubicación "Z" en el mismo bucle como información de atributo de secuencia.

Un problema con el uso de la lista de reproducción para almacenar los parámetros "Z" para cada secuencia de subtítulos es la duplicación de datos. Múltiples listas de reproducción pueden hacer referencia a las mismas 20 secuencias de subtítulos. Esta limitación puede superarse incluyendo los metadatos de ubicación "Z" en el archivo de información de clip. El archivo de información de clip enumera metadatos relacionados con la secuencia de transporte de MPEG que contiene el contenido A/V y las secuencias de subtítulos. El archivo de información de clip puede extenderse con una tabla similar a la que se propone para los datos de extensión de lista de reproducción. Sin embargo, puesto que el archivo de información de clip está asociado a una secuencia de transporte de MPEG, 25 enumera las secuencias básicas de subtítulo por identificador de paquetes (PID). Por tanto, se propone que los metadatos de ubicación "Z" se proporcionen para cada listado de PID que indica el PID de una secuencia de subtítulos.

Alternativamente en lugar de definir una nueva tabla en los datos de extensión, la ubicación "Z" se almacena en la 30 tabla ProgramInfo en el archivo de información de clip. La tabla ProgramInfo proporciona una lista de secuencias básicas que forman conjuntamente una presentación de contenido A/V. Contiene información similar a la tabla PMT definida en los sistemas de MPEG ISO/IEC 13818-1. Se propone que para cada secuencia básica que contiene información de subtítulo se proporcionen metadatos de ubicación "Z". Los metadatos pueden estar incluidos en la propia ProgramInfo o en una subtabla de la tabla ProgramInfo, la tabla StreamCodingInfo. Proporciona detalles 35 sobre la codificación y códigos de idioma para la secuencia de subtítulos presente en la secuencia de transporte. La tabla StreamCodingInfo también tiene algunos campos reservados, se propone usar estos campos reservados para llevar los parámetros de posición "Z" asociados a la secuencia de subtítulo para un código de idioma particular.

Si se requiere cambiar la ubicación "Z" cada ciertos segundos entonces debe usarse la tabla CPI() en el archivo de 40 información de clip. La información de CPI es una tabla que enumera los puntos de entrada en las secuencias de vídeo para una reproducción no estándar. Los listados de los puntos de entrada podrían extenderse con un parámetro de ubicación "Z" que indica por punto de entrada la ubicación en "profundidad" de dónde superponer cualquier gráfico tal como subtítulos.

45 Alternativamente si la información de ubicación "Z" es precisa en cuanto al fotograma entonces debe estar incluida en los mensajes de SEI de la secuencia de vídeo dependiente. SEI (información de mejora complementaria, denominada mensajes de señalización) son paquetes incrustados en una secuencia básica de MPEG que llevan parámetros que pueden usarse por un decodificador para ayudar a decodificar el contenido. MPEG también permite 50 incrustar mensajes de SEI de datos de usuarios privados; éstos se ignorarán por los decodificadores convencionales pero pueden usarse por un decodificador modificado. Un mensaje de SEI de datos de usuario puede definirse para llevar la ubicación "Z" de cualquier superposición que deba superponerse sobre el vídeo tal como subtítulos. Para proporcionar la información "Z" por secuencia (idioma) se propone enumerar los valores "Z" para todas las secuencias de subtítulos previstas en el disco.

offset_metadata() – Sintaxis
offset_metadata() {
Longitud
para(i=0; i < number_of_playitem; i++) {
number_of Plane_offsets[i]
para (j=0; j < number_of Plane_offsets[i]; j++) {
PG Plane_offset [pts]
Otros campos no relevantes 1
Otros campos no relevantes n
is_PG_Offsets
si(is_PG_Offsets==1 _b) {

<pre> number_of_Streams[n] para(k=0; k < number_of_Streams[n] _id; k++) { streamIDList[k] </pre>
PG Plane "Z" metadatos de posición
(datos de profundidad o disparidad)
<pre> } } </pre>
<pre> } } } } </pre>

Tabla 1, ejemplo de una tabla para almacenar metadatos de desplazamiento.

5 Resultará evidente para el experto que las etapas de procesamiento tales como la presentación de imágenes en 3D, o etapas de multiplexación pueden implementarse en una amplia variedad de plataformas de procesamiento. Estas plataformas de procesamiento pueden variar desde circuitos dedicados y/o dispositivos semiconductores, dispositivos lógicos programables, procesadores de señal digital, o incluso procesadores de uso general. De la misma manera las implementaciones de software y hardware combinadas también pueden usarse de manera ventajosa.

10 Aunque la invención se ha ilustrado y descrito en detalle en los dibujos y la descripción anterior, tal ilustración y descripción deben considerarse ilustrativas o a modo de ejemplo y no restrictivas; la invención no se limita a las realizaciones dadas a conocer.

offset_metadata() – Sintaxis
offset_metadata() {
Longitud
para(i=0; i < number_of_playitem; i++) {
number_of Plane offsets[i]
para (j=0; j < number_of Plane offsets[i]; j++) {
PG Plane_offset [pts]
Otros campos no relevantes 1
Otros campos no relevantes n
is_PG_Offsets
si(is_PG_Offsets==1 _b) {
number_of_Z_values
para(k=0; k < number_of_Z_values
k++) {
number_of_streams
number_of_angles
Para (int m=0; m<number_of_streams; m++) {
StreamID;
}
Para (int o=0; o<number_of_angles; o++) {
angleID;
}
PG Plane "Z" metadatos de posición
(datos de profundidad o disparidad)
}
}
}
}
}
}

15 Tabla 2, ejemplo de una tabla para almacenar metadatos de desplazamiento extendidos con metadatos de ubicación "Z" de subtítulo diferentes por segmento de vídeo multiángulo estereoscópico.

20 Otras variaciones de las realizaciones dadas a conocer pueden entenderse y efectuarse por los expertos en la técnica al poner en práctica la invención reivindicada, a partir de un estudio de los dibujos, la descripción y las reivindicaciones adjuntas. En las reivindicaciones, la expresión "que comprende/comprendiendo" no excluye otros elementos o etapas, y el artículo indefinido "un" o "una" no excluye una pluralidad. Un único procesador u otra unidad puede cumplir con las funciones de varios elementos mencionados en las reivindicaciones. El mero hecho de que ciertas medidas se mencionen en reivindicaciones dependientes diferentes entre sí no indica que no pueda

5 usarse ventajosamente una combinación de estas medidas. Un programa informático puede almacenarse/distribuirse en un medio adecuado, tal como un medio de almacenamiento óptico o un medio de estado sólido suministrado junto con o como parte de otro hardware, pero también puede distribuirse de otras maneras, tal como a través de Internet u otros sistemas de telecomunicación por cable o inalámbricos. Ningún símbolo de referencia en las reivindicaciones debe interpretarse como que limita el alcance.

REIVINDICACIONES

1. Método para crear una señal de vídeo tridimensional que comprende:
- 5 - recibir una primera componente de vídeo que comprende primeras imágenes,
- recibir una segunda componente de vídeo que comprende segundas imágenes, representando las respectivas primeras imágenes y correspondientes segundas imágenes pares estéreo,
- 10 - recibir una primera componente de texto y una segunda componente de texto, comprendiendo la primera componente de texto subtítulos basados en texto y comprendiendo la segunda componente de texto imágenes de mapa de bits basadas en gráficos de presentación para su inclusión en el vídeo tridimensional,
- 15 - recibir una componente de ubicación Z compartida que comprende información de ubicación Z que describe la posición de profundidad dentro del vídeo tridimensional de los subtítulos basados en texto y las imágenes de mapa de bits basadas en gráficos de presentación tanto de la primera componente de texto como de la segunda componente de texto usando valores de profundidad o valores de disparidad, y
- 20 - crear una señal de vídeo tridimensional que comprende la primera componente de vídeo, la segunda componente de vídeo, ambos componentes de texto y la componente de ubicación Z compartida, y
- 25 en el que la información de ubicación Z es precisa en cuanto al fotograma y está comprendida en la señal de vídeo tridimensional en mensajes de señalización [SEI] que son paquetes incrustados en una secuencia básica de MPEG de contenido audiovisual, llevando los paquetes parámetros que pueden usarse por un decodificador para decodificar el contenido.
2. Método según la reivindicación 1, en el que los subtítulos basados en texto se prevén en secuencias de subtítulos y en el que la información de ubicación Z compartida comprende valores Z enumerados para todas las secuencias de subtítulos.
- 30
3. Método para presentar una señal de vídeo tridimensional que comprende:
- 35 - recibir una señal de vídeo tridimensional que comprende una primera componente de vídeo que comprende primeras imágenes, una segunda componente de vídeo que comprende segundas imágenes, representando las respectivas primeras imágenes y correspondientes segundas imágenes pares estéreo, una primera componente de texto y una segunda componente de texto, comprendiendo la primera componente de texto subtítulos basados en texto y comprendiendo la segunda componente de texto imágenes de mapa de bits basadas en gráficos de presentación para su inclusión en el vídeo tridimensional, y
- 40 - presentar la primera componente de vídeo y la segunda componente de vídeo para proporcionar un vídeo tridimensional, incluyendo la presentación presentar los subtítulos basados en texto o imágenes de mapa de bits basadas en gráficos de presentación en el vídeo tridimensional,
- 45 y en el que la señal de vídeo tridimensional comprende además una componente de ubicación Z compartida que comprende información de ubicación Z que describe la posición de profundidad dentro del vídeo tridimensional de los subtítulos basados en texto y las imágenes de mapa de bits basadas en gráficos de presentación tanto de la primera componente de texto como de la segunda componente de texto usando valores de profundidad o valores de disparidad, y
- 50 en el que la información de ubicación Z es precisa en cuanto al fotograma y está comprendida en la señal de vídeo tridimensional en mensajes de señalización [SEI] que son paquetes incrustados en una secuencia básica de MPEG de contenido audiovisual, llevando los paquetes parámetros que pueden usarse por un decodificador para decodificar el contenido, y
- 55 en el que la presentación de los subtítulos basados en texto o imágenes de mapa de bits basadas en gráficos de presentación incluye ajustar la posición de profundidad de los subtítulos basados en texto o imágenes de mapa de bits basadas en gráficos de presentación de una manera precisa en cuanto al fotograma basándose en la componente de ubicación Z compartida.
- 60
4. Método según la reivindicación 3, en el que los subtítulos basados en texto se prevén en secuencias de subtítulos y en el que la información de ubicación Z compartida comprende valores Z enumerados para todas las secuencias de subtítulos.
- 65
5. Dispositivo para crear una señal de vídeo tridimensional que comprende:

- un receptor dispuesto para recibir una primera componente de vídeo que comprende primeras imágenes, una segunda componente de vídeo que comprende segundas imágenes, representando las respectivas primeras imágenes y correspondientes segundas imágenes pares estéreo, una primera componente de texto y una segunda componente de texto, comprendiendo la primera componente de texto subtítulos basados en texto y comprendiendo la segunda componente de texto imágenes de mapa de bits basadas en gráficos de presentación para su inclusión en el vídeo tridimensional, y

en el que el receptor está dispuesto además para recibir una componente de ubicación Z compartida que comprende información de ubicación Z que describe la posición de profundidad dentro del vídeo tridimensional de los subtítulos basados en texto y las imágenes de mapa de bits basadas en gráficos de presentación tanto de la primera componente de texto como de la segunda componente de texto usando valores de profundidad o valores de disparidad, y

comprendiendo además el dispositivo:

- un multiplexor dispuesto para crear una señal de vídeo tridimensional que comprende la primera componente de vídeo, la segunda componente de vídeo, la componente de texto y la componente de ubicación Z compartida y

en el que la información de ubicación Z es precisa en cuanto al fotograma y está comprendida en la señal de vídeo tridimensional en mensajes de señalización [SEI] que son paquetes incrustados en una secuencia básica de MPEG de contenido audiovisual, llevando los paquetes parámetros que pueden usarse por un decodificador para decodificar el contenido.

6. Dispositivo para presentar una señal de vídeo tridimensional que comprende:

- un receptor (18) dispuesto para recibir una señal de vídeo tridimensional que comprende una primera componente de vídeo que comprende primeras imágenes, una segunda componente de vídeo que comprende segundas imágenes, representando las respectivas primeras imágenes y correspondientes segundas imágenes pares estéreo, una primera componente de texto y una segunda componente de texto, comprendiendo la primera componente de texto subtítulos basados en texto y comprendiendo la segunda componente de texto imágenes de mapa de bits basadas en gráficos de presentación para su inclusión en el vídeo tridimensional, y

- un presentador (24) dispuesto para presentar la primera componente de vídeo y la segunda componente de vídeo para proporcionar un vídeo tridimensional, incluyendo la presentación presentar los subtítulos basados en texto o imágenes de mapa de bits basadas en gráficos de presentación en el vídeo tridimensional, y

en el que la señal de imagen tridimensional comprende además una componente de ubicación Z compartida que comprende información de ubicación Z que describe la posición de profundidad dentro del vídeo tridimensional de los subtítulos basados en texto y las imágenes de mapa de bits basadas en gráficos de presentación tanto de la primera componente de texto como de la segunda componente de texto usando valores de profundidad o valores de disparidad y

en el que la información de ubicación Z es precisa en cuanto al fotograma y está comprendida en la señal de vídeo tridimensional en mensajes de señalización [SEI] que son paquetes incrustados en una secuencia básica de MPEG de contenido audiovisual, llevando los paquetes parámetros que pueden usarse por un decodificador para decodificar el contenido, y

en el que la presentación de los subtítulos basados en texto o imágenes de mapa de bits basadas en gráficos de presentación incluye ajustar la posición de profundidad de los subtítulos basados en texto o imágenes de mapa de bits basadas en gráficos de presentación de una manera precisa en cuanto al fotograma basándose en la componente de ubicación Z compartida.

7. Método según la reivindicación 6, en el que los subtítulos basados en texto se prevén en secuencias de subtítulos y en el que la información de ubicación Z compartida comprende valores Z enumerados para todas las secuencias de subtítulos.

8. Señal de imagen tridimensional que comprende una primera componente de vídeo que comprende primeras imágenes, una segunda componente de vídeo que comprende segundas imágenes, representando las respectivas primeras imágenes y correspondientes segundas imágenes pares estéreo, una primera componente de texto y una segunda componente de texto, comprendiendo la primera componente de texto subtítulos basados en texto y comprendiendo la segunda componente de texto imágenes de mapa de bits basadas en gráficos de presentación para su inclusión en el vídeo

tridimensional, y

5 en la que la señal de vídeo tridimensional comprende además una componente de ubicación Z compartida que comprende información de ubicación Z que describe la posición de profundidad dentro del vídeo tridimensional de los subtítulos basados en texto y las imágenes de mapa de bits basadas en gráficos de presentación tanto de la primera componente de texto como de la segunda componente de texto usando valores de profundidad o valores de disparidad y

10 en la que la información de ubicación Z es precisa en cuanto al fotograma y está comprendida en la señal de vídeo tridimensional en mensajes de señalización [SEI] que son paquetes incrustados en una secuencia básica de MPEG de contenido audiovisual, llevando los paquetes parámetros que pueden usarse por un decodificador para decodificar el contenido.

15 9. Señal de vídeo tridimensional según la reivindicación 8, en la que los subtítulos basados en texto se prevén en secuencias de subtítulos y en la que la información de ubicación Z compartida comprende valores Z enumerados para todas las secuencias de subtítulos.

20 10. Disco para un sistema de reproducción tridimensional, comprendiendo el disco (16) la señal de vídeo tridimensional según la reivindicación 8 ó 9.

11. Producto de programa informático en un medio legible por ordenador, comprendiendo el producto instrucciones para realizar las etapas de uno cualquiera de los métodos según una cualquiera de las reivindicaciones 1 a 4 cuando el programa informático se ejecuta en un ordenador.

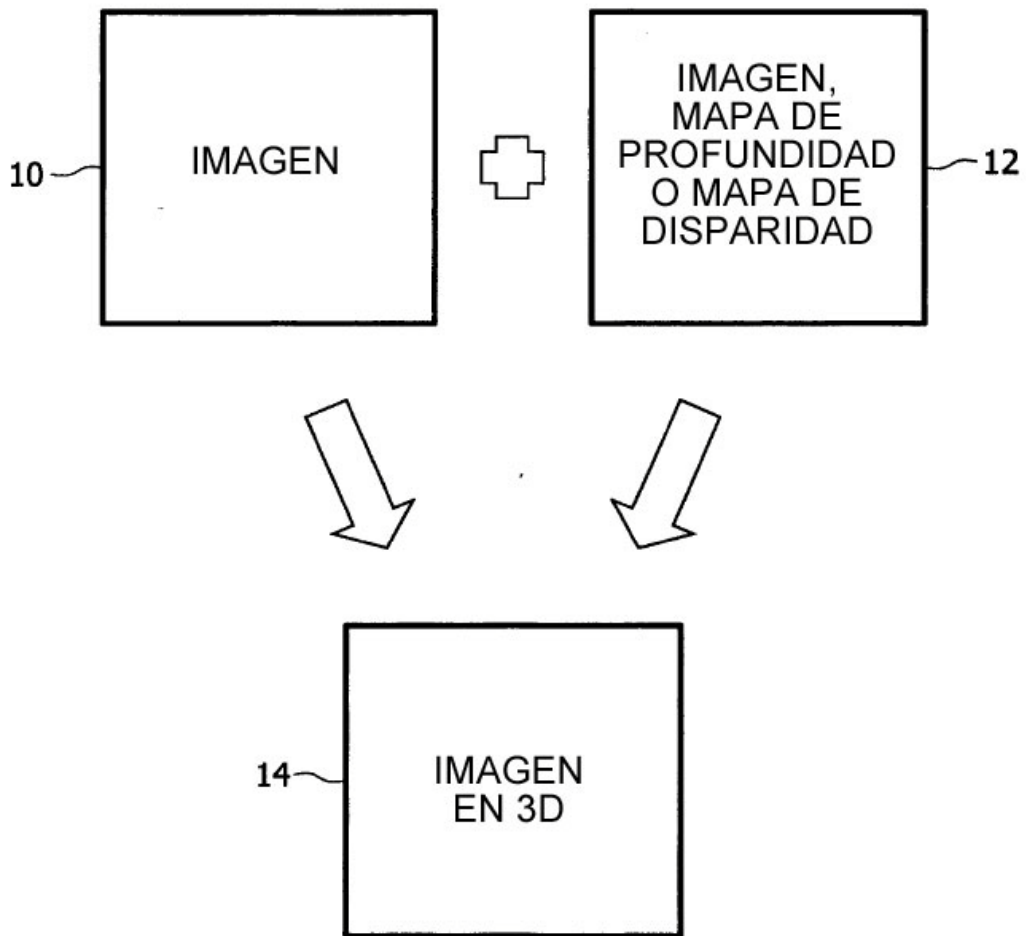


FIG. 1

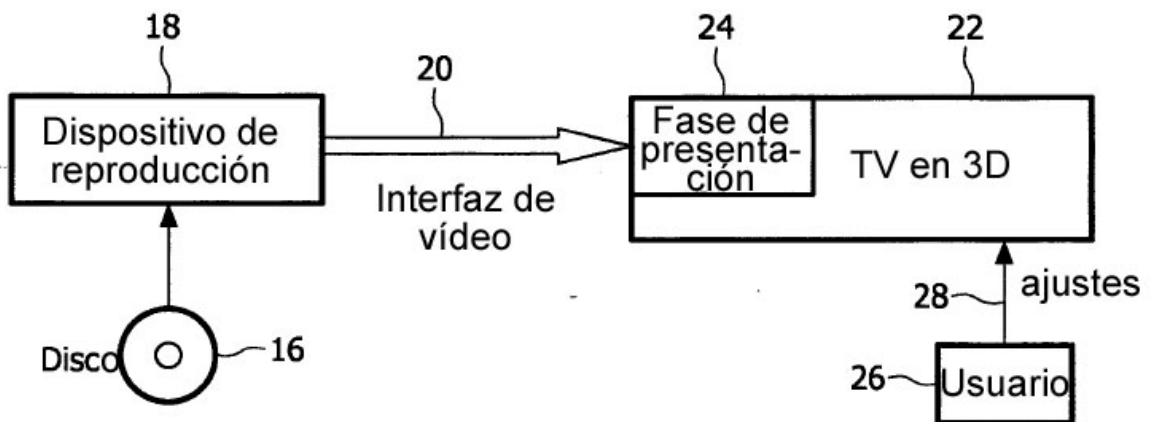


FIG. 2

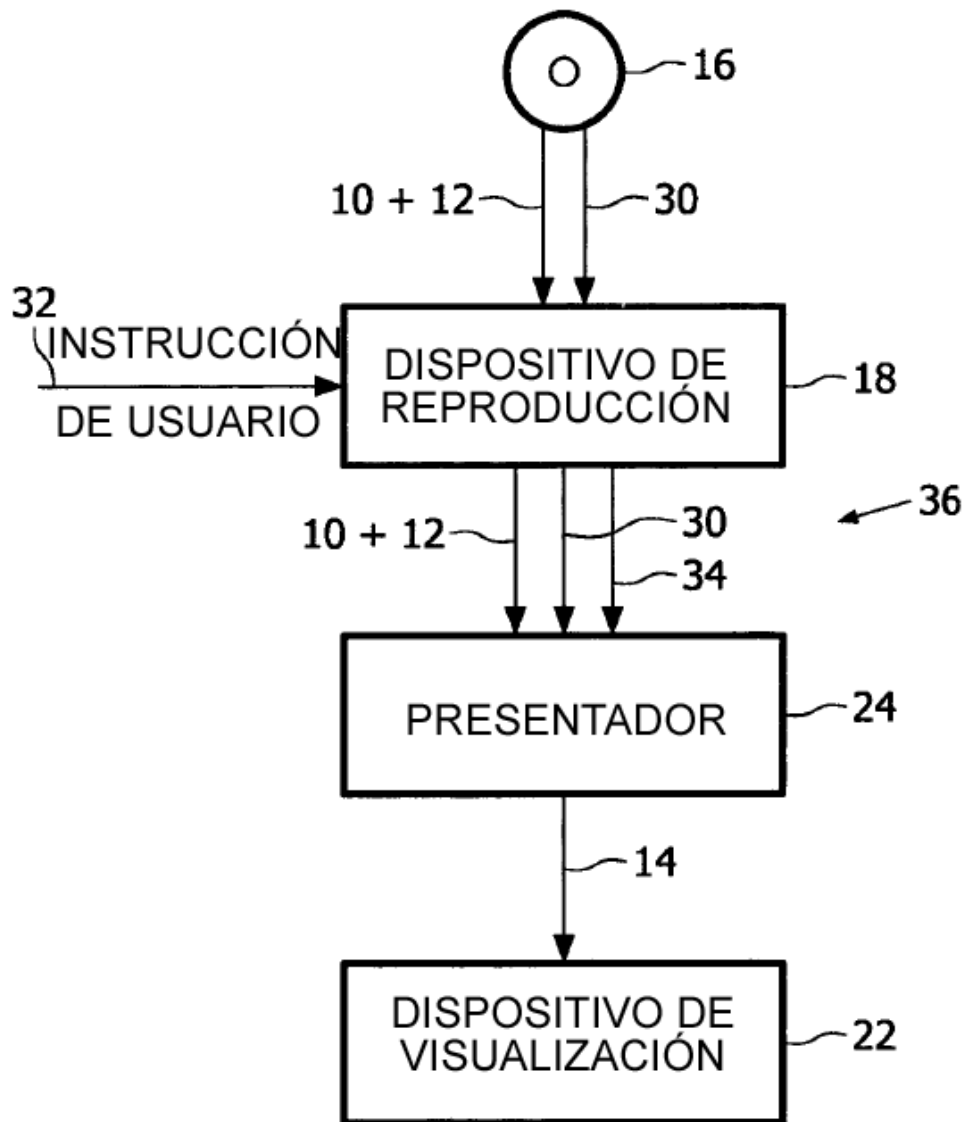


FIG. 3

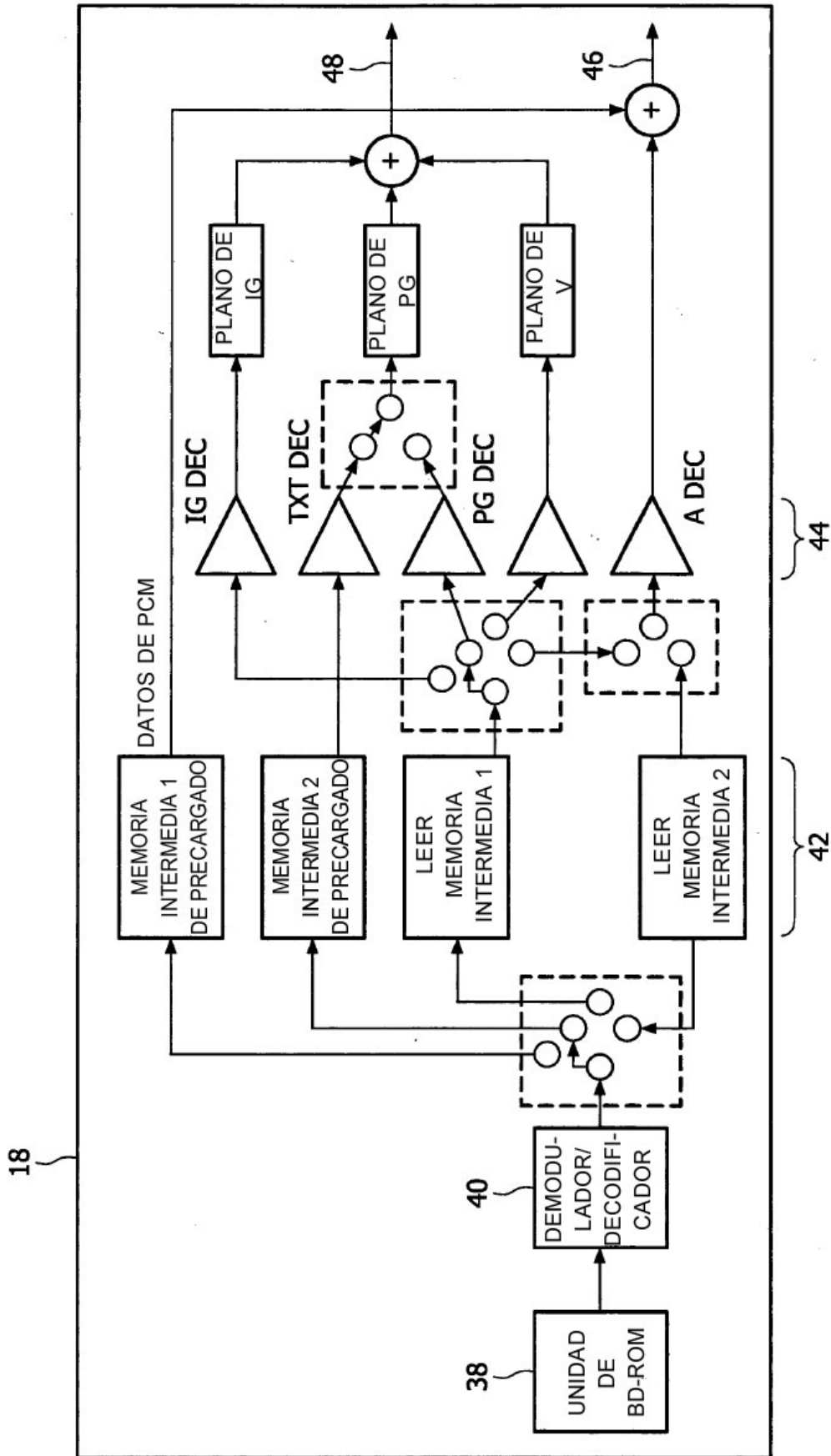


FIG. 4

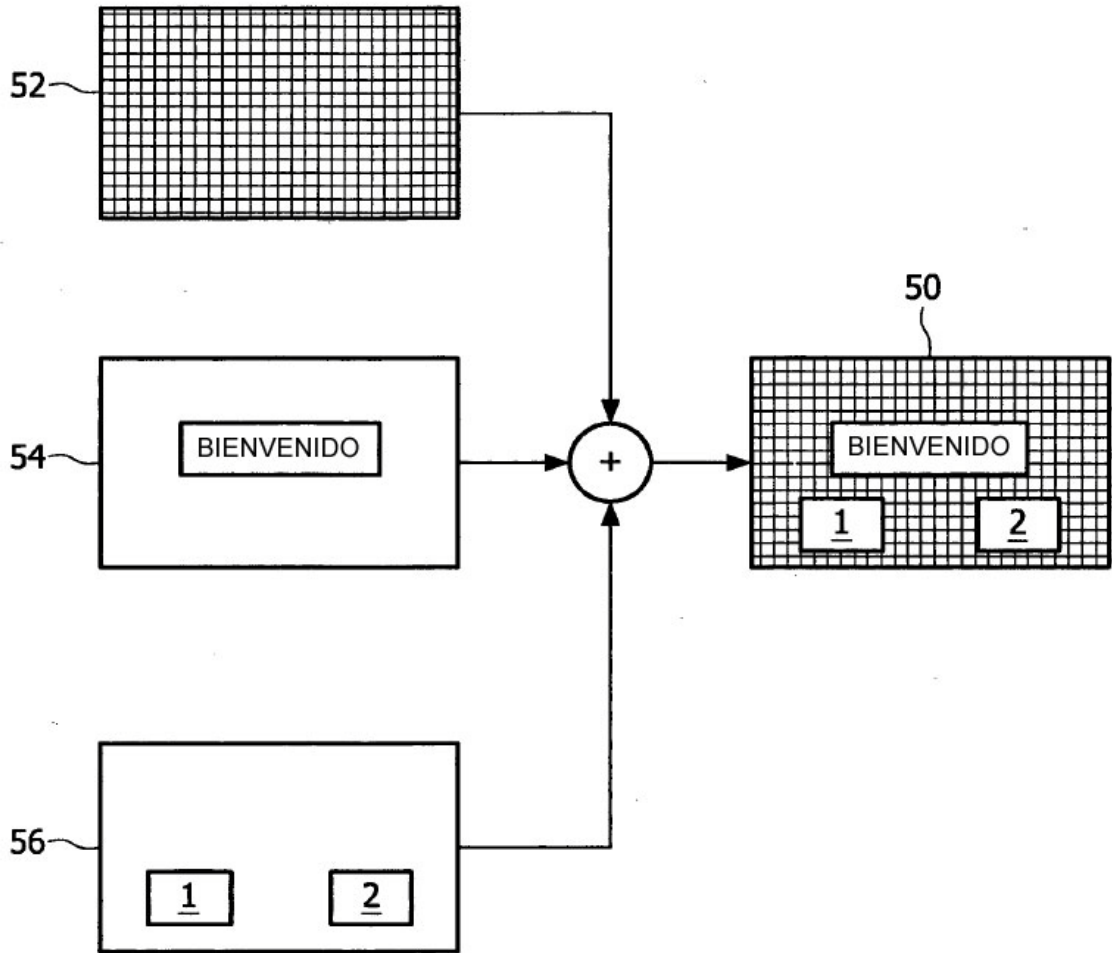


FIG. 5

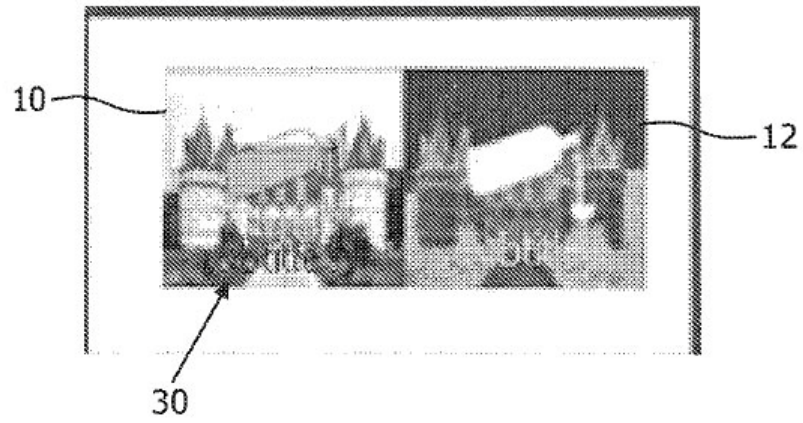


FIG. 6

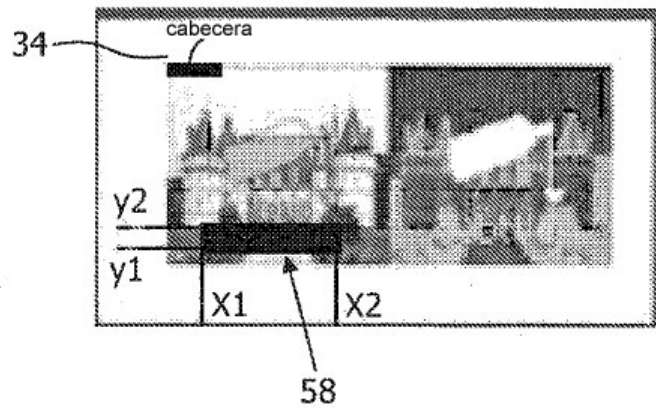


FIG. 7

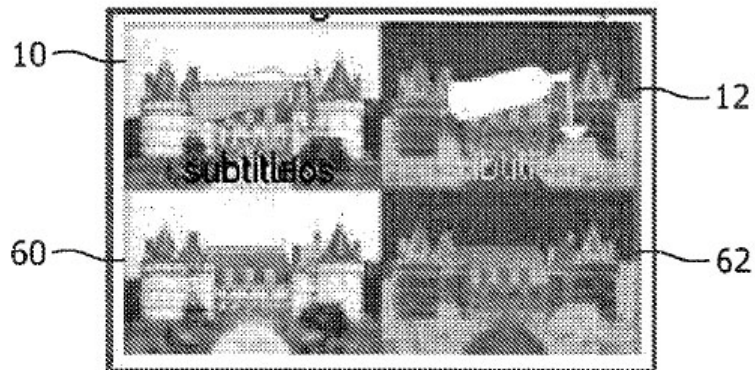


FIG. 8

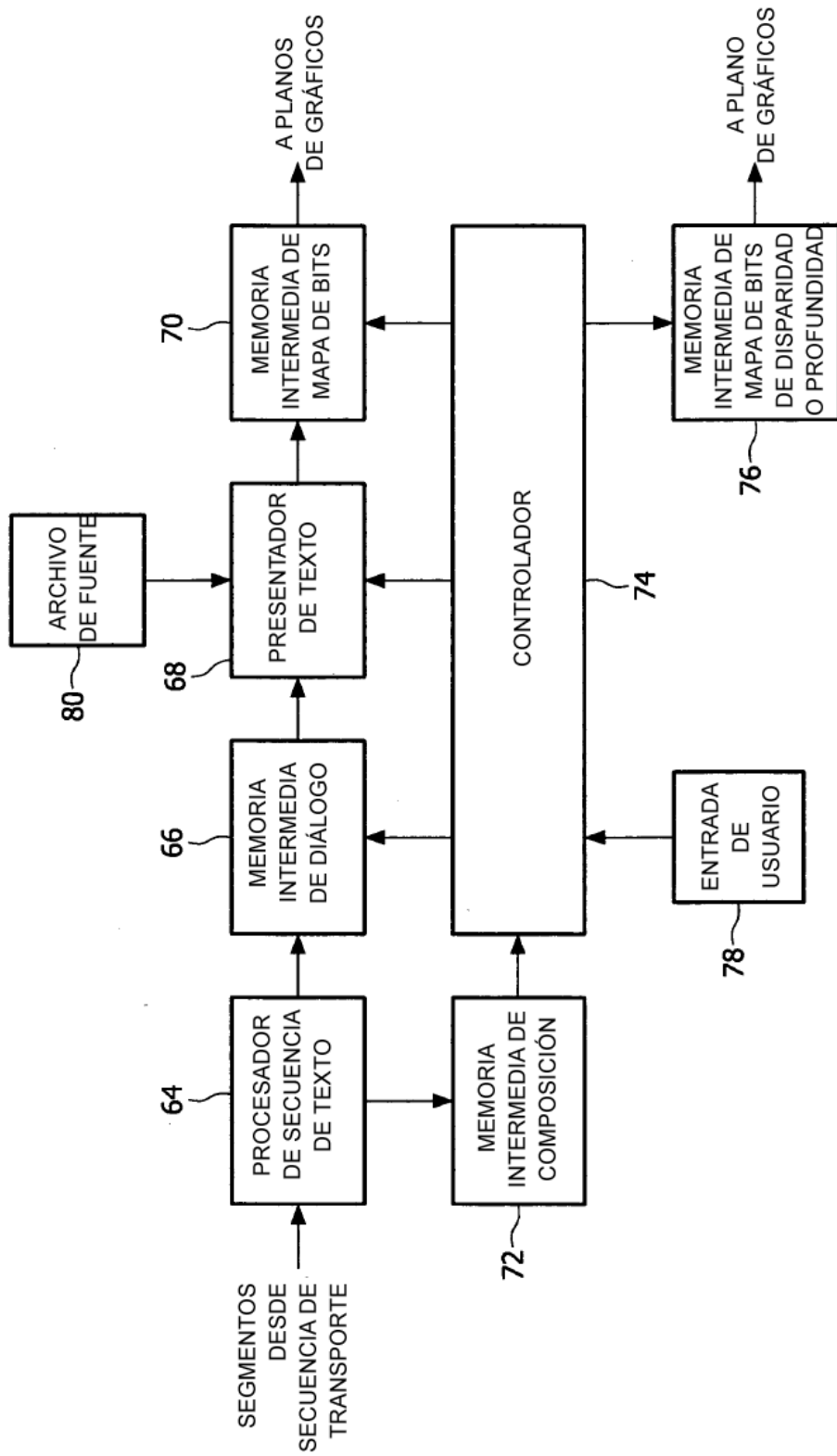


FIG. 9

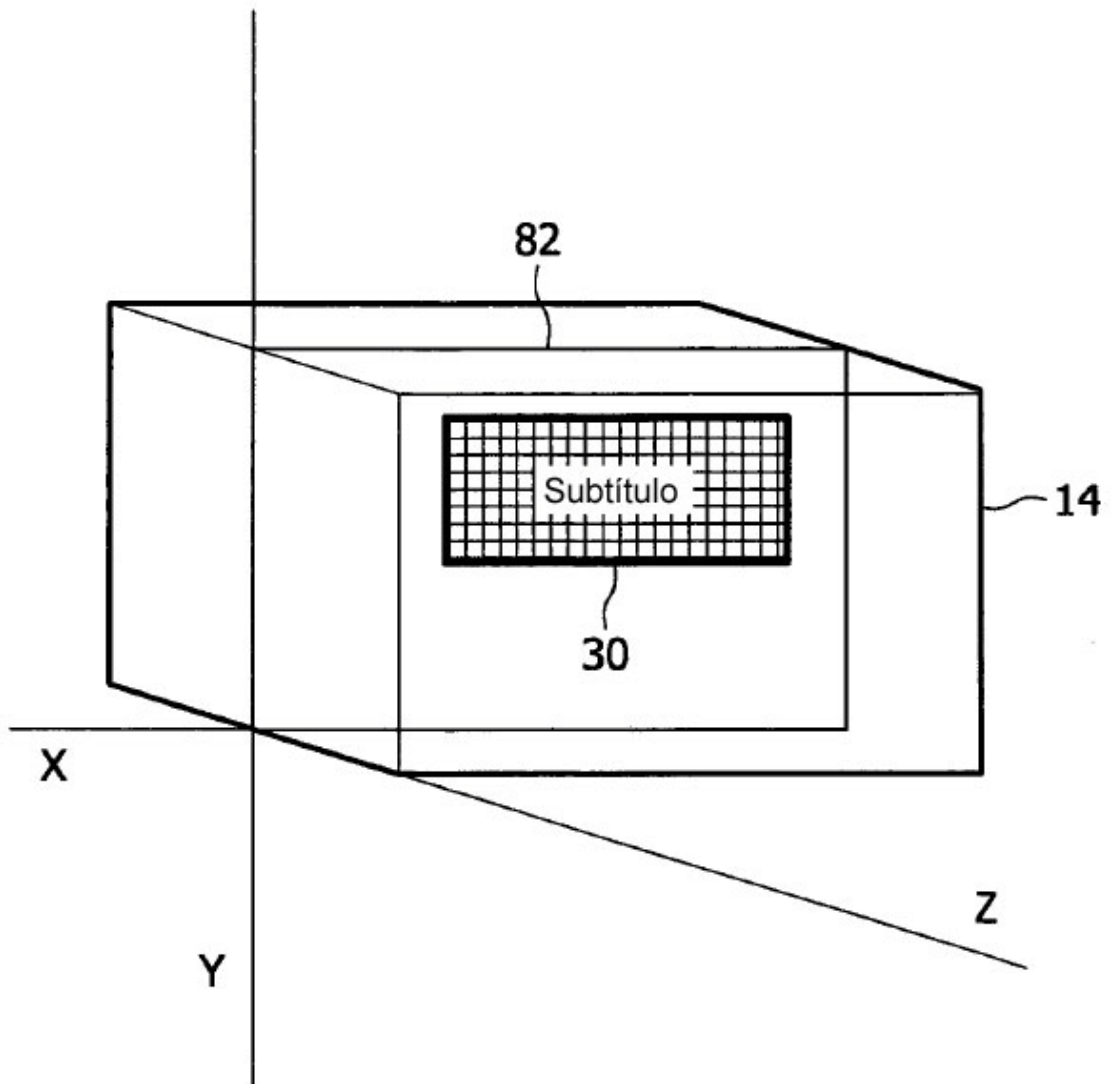


FIG. 10