

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 671 596**

51 Int. Cl.:

**H04N 19/597** (2014.01)  
**H04N 13/00** (2008.01)  
**H04N 19/196** (2014.01)  
**H04N 19/139** (2014.01)  
**H04N 19/159** (2014.01)  
**H04N 19/61** (2014.01)  
**H04N 19/176** (2014.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **11.08.2011 PCT/EP2011/063852**
- 87 Fecha y número de publicación internacional: **16.02.2012 WO12020092**
- 96 Fecha de presentación y número de la solicitud europea: **11.08.2011 E 11741600 (8)**
- 97 Fecha y número de publicación de la concesión europea: **07.03.2018 EP 2604036**

54 Título: **Códec de señal de múltiples vistas**

30 Prioridad:

**11.08.2010 US 372785 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**07.06.2018**

73 Titular/es:

**GE VIDEO COMPRESSION, LLC (100.0%)**  
**8 Southwoods Boulevard**  
**Albany, NY 12211, US**

72 Inventor/es:

**WIEGAND, THOMAS;**  
**MARPE, DETLEV;**  
**MUELLER, KARSTEN;**  
**MERKLE, PHILIPP;**  
**TECH, GERHARD;**  
**RHEE, HUNN y**  
**SCHWARZ, HEIKO**

74 Agente/Representante:

**ARIZTI ACHA, Monica**

**ES 2 671 596 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).

Códec de señal de múltiples vistas

**DESCRIPCIÓN**

5 La presente invención se refiere a codificación de señales de múltiples vistas.

Las señales de múltiples vistas se ven implicadas en muchas aplicaciones, tales como aplicaciones de vídeo en 3D que incluyen, por ejemplo, pantallas estéreo y de múltiples vistas, aplicaciones de vídeo de punto de vista libre, etc. Para contenido de vídeo estéreo y de múltiples vistas, se ha especificado la norma MVC [1, 2]. Esta norma  
10 comprime secuencias de vídeo desde un número de cámaras adyacentes. El proceso de decodificación de MVC únicamente reproduce estas vistas de cámara en sus posiciones de cámara original. Para diferentes pantallas de múltiples vistas, sin embargo, se requiere un número diferente de las vistas con diferentes posiciones espaciales de manera que se requieren vistas adicionales, por ejemplo entre las posiciones de cámara original.

15 La dificultad al manejar señales de múltiples vistas es la enorme cantidad de datos necesarios para transportar información en las múltiples vistas incluidas en la señal de múltiples vistas. En caso del requisito recién mencionado para posibilitar extracción/síntesis de vista intermedia, la situación incluso empeora, puesto que en este caso los vídeos asociados con las vistas individuales pueden acompañarse por datos complementarios tales como datos de mapa de profundidad/disparidad que posibilitan volver a proyectar la respectiva vista en otra vista, tal como una vista  
20 intermedia. Debido a la enorme cantidad de datos, es muy importante maximizar la tasa de compresión del códec de la señal de múltiples vistas tanto como sea posible.

El artículo Merkle et al., "The effects of multi-view depth video compression on multi-view rendering", Signal processing: image Communication, vol. 24, N.º 1-2. P. 73-88, analiza estructuras de MVD que usan codificación en  
25 cuña para mapas de profundidad. El documento WO 2010/043773 A1 sugiere que puede posibilitarse la predicción de movimiento entre el vídeo de mapa de profundidad y el vídeo de textura de la misma vista.

Por lo tanto, es el objeto de la presente invención proporcionar un códec de señal de múltiples vistas que posibilite  
30 una compresión superior o mejor relación de tasa/distorsión.

Este objeto se consigue mediante la materia objeto de las reivindicaciones independientes pendientes.

La presente solicitud proporciona realizaciones que comprenden adicionalmente adoptar o predecir segundos  
35 parámetros de codificación usados para codificar una segunda vista de la señal de múltiples vistas desde primeros parámetros de codificación usados al codificar una primera vista de la señal de múltiples vistas. En otras palabras, las redundancias entre vistas de una señal de múltiples vistas no están restringidas a las mismas vistas, tal como la misma información de vídeo, pero los parámetros de codificación que codifican en paralelo estas vistas muestran similitudes que pueden aprovecharse para mejorar adicionalmente la tasa de codificación.

Algunas realizaciones de la presente solicitud adicionalmente aprovechan un hallazgo de acuerdo con el que la  
40 segmentación de un mapa de profundidad/disparidad asociado con un cierto fotograma de un vídeo de una cierta vista, usado al codificar el mapa de profundidad/disparidad, puede determinarse o predecirse usando un borde detectado en el fotograma de vídeo como un indicio, en concreto determinando una línea de separación en cuña para extenderse a lo largo del borde en el fotograma de vídeo. Aunque la detección de borde aumenta la complejidad en el lado del decodificador, la deficiencia puede ser aceptable en escenarios de aplicación donde son  
45 más importantes las tasas de transmisión bajas a calidad aceptable que los problemas de complejidad. Tales escenarios pueden implicar aplicaciones donde los decodificadores se implementan como dispositivos estáticos.

Además, algunas realizaciones de la presente solicitud aprovechan adicionalmente un hallazgo de acuerdo con el  
50 cual la vista de los parámetros de codificación de los cuales se adoptan/predicen desde información de codificación de otra vista, pueden codificarse, es decir predecirse y corregirse residualmente, a una resolución espacial inferior, de esta manera ahorrando bits codificados, si la adopción/predicción de los parámetros de codificación incluye escalamiento de estos parámetros de codificación de acuerdo con una relación entre las resoluciones espaciales.

Implementaciones ventajosas de realizaciones de los aspectos anteriormente señalados son el objeto de las  
55 reivindicaciones dependientes adjuntas. En particular, se describen a continuación realizaciones preferidas de la presente solicitud con respecto a las figuras entre las que

La Figura 1 muestra un diagrama de bloques de un codificador de acuerdo con una realización;

60 La Figura 2 muestra un diagrama esquemático de una porción de una señal de múltiples vistas para ilustración de reutilización de información a través de vistas y límites de profundidad/disparidad de vídeo;

La Figura 3 muestra un diagrama de bloques de un decodificador de acuerdo con una realización;

La Figura 4 estructura de predicción y vectores de movimiento/disparidad en codificación de múltiples vistas en

- La Figura 5 el ejemplo de dos vistas y dos instancias de tiempo; correspondencias de punto por vector de disparidad entre vistas adyacentes;
- La Figura 6 síntesis de vista intermedia por proyección de contenido de escena desde las vistas 1 y 2, usando vectores de disparidad escalados;
- La Figura 7 extracción de N vistas desde datos de color y complementarios decodificados de manera separada para generar vistas intermedias en posiciones de visualización arbitrarias; y
- La Figura 8 ejemplo de extracción de N vistas de un flujo de bits de dos vistas para una visualización de 9 vistas.

La Figura 1 muestra un codificador para codificar una señal de múltiples vistas de acuerdo con una realización. La señal de múltiples vistas de la Figura 1 se indica de manera ilustrativa en 10 comprendiendo dos vistas  $12_1$  y  $12_2$ , aunque la realización de la Figura 1 sería también factible con un número de vistas superior. Además, de acuerdo con la realización de la Figura 1, cada vista  $12_1$  y  $12_2$  comprende un vídeo 14 y datos de mapa de profundidad/disparidad 16, aunque muchos de los principios ventajosos de la realización descrita con respecto a la Figura 1 podrían ser también ventajosos si se usaran en relación con señales de múltiples vistas con vistas que no comprendieran ningún dato de mapa de profundidad/disparidad. Tal generalización de la presente realización se describe adicionalmente a continuación después de la descripción de las figuras 1 a 3.

El vídeo 14 de las respectiva vistas  $12_1$  y  $12_2$  representa un muestreo espacio-temporal de una proyección de una escena común a lo largo de diferentes direcciones de proyección/visualización. Preferentemente, las tasas de muestreo temporal de los vídeos 14 de las vistas  $12_1$  y  $12_2$  son iguales entre sí aunque esta restricción no tiene que cumplirse necesariamente. Como se muestra en la Figura 1, preferentemente cada vídeo 14 comprende una secuencia de fotogramas, estando asociado cada fotograma con una respectiva indicación de tiempo  $t$ ,  $t - 1$ ,  $t - 2$ . En la Figura 1 los fotogramas de vídeo se indican por  $V_{\text{número de vista, número de indicación de tiempo}}$ . Cada fotograma  $V_{i,t}$  representa un muestreo espacial de la escena a lo largo de la respectiva dirección de vista a la respectiva indicación de tiempo  $t$ , y por lo tanto comprende una o más series de muestras tales como, por ejemplo, una serie de muestras para muestras de luminancia y dos series de muestras con muestras de crominancia, o simplemente muestras de luminancia o series de muestras para otros componentes de color, tal como componentes de color de un espacio de color RGB o similares. La resolución espacial de la una o más series de muestras puede diferir tanto dentro de un vídeo 14 como dentro de los vídeos 14 de las diferentes vistas  $12_1$  y  $12_2$ .

De manera similar, los datos de mapa de profundidad/disparidad 16 representan un muestreo espacio-temporal de la profundidad de los objetos de escena de la escena común, medidos a lo largo de la respectiva dirección de visualización de las vistas  $12_1$  y  $12_2$ . La tasa de muestreo temporal de los datos de mapa de profundidad/disparidad 16 puede ser igual a la tasa de muestreo temporal del vídeo asociado de la misma vista como se representa en la Figura 1, o puede ser diferente de la misma. En el caso de la Figura 1, cada fotograma de vídeo  $v$  tiene asociado con el mismo un respectivo mapa de profundidad/disparidad  $d$  de los datos de mapa de profundidad/disparidad 16 de la respectiva vista  $12_1$  y  $12_2$ . En otras palabras, en el ejemplo de la Figura 1, cada fotograma de vídeo  $v_{i,t}$  de la vista  $i$  e indicación de tiempo  $t$  tiene un mapa de profundidad/disparidad  $d_{i,t}$  asociado con el mismo. Con respecto a la resolución espacial de los mapas de profundidad/disparidad  $d$ , lo mismo se aplica como se ha indicado anteriormente con respecto a los fotogramas de vídeo. Es decir, la resolución espacial puede ser diferente entre los mapas de profundidad/disparidad de diferentes vistas.

Para comprimir la señal de múltiples vistas 10 de manera eficaz, el codificador de la Figura 1 codifica en paralelo las vistas  $12_1$  y  $12_2$  en un flujo de datos 18, sin embargo, los parámetros de codificación usados para codificar la primera vista  $12_1$  se vuelven a usar para adoptar, o predecir, los mismos que los segundos parámetros de codificación a usarse al codificar la segunda vista  $12_2$ . Mediante esta medida, el codificador de la Figura 1 aprovecha el hecho descubierto por los inventores, de acuerdo con el cual la codificación paralela de las vistas  $12_1$  y  $12_2$  da como resultado que el codificador determine los parámetros de codificación para estas vistas de manera similar, de modo que pueden aprovecharse redundancias entre estos parámetros de codificación de manera eficaz para aumentar la tasa de compresión o relación de tasa/distorsión (con la distorsión medida, por ejemplo, como una distorsión media de ambas vistas y la tasa medida como una tasa de codificación de la totalidad del flujo de datos 18).

En particular, el codificador de la Figura 1 se indica en general por el signo de referencia 20 y comprende una entrada para recibir la señal de múltiples vistas 10 y una salida para emitir el flujo de datos 18. Como puede observarse en la Figura 2, el codificador 20 de la Figura 1 comprende dos ramales de codificación por vista  $12_1$  y  $12_2$ , en concreto uno para los datos de vídeo y el otro para los datos de mapa de profundidad/disparidad. Por consiguiente, el codificador 20 comprende un ramal de codificación  $22_{v,1}$  para los datos de vídeo de la vista 1, un ramal de codificación  $22_{d,1}$  para los datos de mapa de disparidad de profundidad de la vista 1, un ramal de codificación  $24_{v,2}$  para los datos de vídeo de la segunda vista y un ramal de codificación  $22_{d,2}$  para los datos de mapa de profundidad/disparidad de la segunda vista. Cada uno de estos ramales de codificación 22 se construye de

manera similar. Para describir la construcción y funcionalidad del codificador 20, la siguiente descripción empieza con la construcción y funcionalidad del ramal de codificación  $22_{v,1}$ . Esta funcionalidad es común a todos los ramales 22. Posteriormente, se analizan las características individuales de los ramales 22.

5 El ramal de codificación  $22_{v,1}$  es para codificar el vídeo  $14_1$  de la primera vista  $12_1$  de la señal de múltiples vistas  $12$ , y en consecuencia el ramal  $22_{v,1}$  tiene una entrada para recibir el vídeo  $14_1$ . Más allá de este, el ramal  $22_{v,1}$  comprende, conectados en serie entre sí en el orden mencionado, un restador 24, un módulo de cuantificación/transformada 26, un módulo de recuantificación/transformada inversa 28, un sumador 30, un módulo de procesamiento adicional 32, una memoria intermedia de instantánea decodificada 34, dos módulos de predicción 36 y 38 que, a su vez, están conectados en paralelo entre sí, y un combinador o selector 40 que está conectado  
10 entre las salidas de los módulos de predicción 36 y 38 por una parte y a la entrada de inversión del restador 24 por otra parte. La salida del combinador 40 está también conectada a una entrada adicional del sumador 30. La entrada de no inversión del restador 24 recibe el vídeo  $14_1$ .

15 Los elementos 24 a 40 del ramal de codificación  $22_{v,1}$  cooperan para codificar el vídeo  $14_1$ . La codificación codifica el vídeo  $14_1$  en unidades de ciertas porciones. Por ejemplo, al codificar el vídeo  $14_1$ , los fotogramas  $v_{1,k}$  se segmentan en segmentos tales como bloques u otros grupos de muestra. La segmentación puede ser constante con el tiempo o puede variar en el tiempo. Además, la segmentación puede ser conocida para el codificador y decodificador por defecto o puede señalizarse en el flujo de datos 18. La segmentación puede ser una segmentación regular de los fotogramas en bloques tales como una disposición no solapante de bloques en filas y columnas, o puede ser una  
20 segmentación basada en árbol cuádruple en bloques de tamaño variable. Un segmento de vídeo actualmente codificado  $14_1$  que entra en la entrada no de inversión del restador 24 se denomina una porción actual de vídeo  $14_1$  en la siguiente descripción.

25 Los módulos de predicción 36 y 38 son para predecir la porción actual y para este fin, los módulos de predicción 36 y 38 tienen sus entradas conectadas a la memoria intermedia de instantánea decodificada 34. En efecto, ambos módulos de predicción 36 y 38 usan porciones previamente reconstruidas de vídeo  $14_1$  que residen en la memoria intermedia de instantánea decodificada 34 para predecir la porción/segmento actual que entra en la entrada no de inversión del restador 24. En este sentido, el módulo de predicción 36 actúa como un intra predictor que predice espacialmente la porción actual de vídeo  $14_1$  de porciones ya reconstruidas espacialmente vecinas del mismo  
30 fotograma del vídeo  $14_1$ , mientras que el módulo de predicción 38 actúa como un inter predictor que predice temporalmente la porción actual de fotogramas del vídeo previamente reconstruidos 14. Ambos módulos 36 y 38 realizan sus predicciones de acuerdo con, o descritas por, ciertos parámetros de predicción. Para ser más precisos, los últimos parámetros se determinan para que estén en el codificador 20 en alguna estructura de optimización para optimizar algún objetivo de optimización tal como optimizar una relación de tasa/distorsión bajo alguna restricción tal  
35 como una tasa de bits máxima, o sin ninguna.

40 Por ejemplo, el módulo de intra predicción 36 puede determinar parámetros de predicción espacial para la porción actual tal como una dirección de predicción a lo largo de la cual se expande/copia contenido de porciones ya reconstruidas vecinas del mismo fotograma de vídeo  $14_1$  en la porción actual para predecir el último. El módulo de inter predicción 38 puede usar compensación de movimiento para predecir la porción actual de fotogramas previamente reconstruidos y los parámetros de inter predicción implicados con los mismos pueden comprender un vector de movimiento, un índice de trama de referencia, una información de subdivisión de predicción de movimiento con respecto a la porción actual, un número hipotético o cualquier combinación de los mismos. El combinador 40  
45 puede combinar una o más predicciones proporcionadas por los módulos 36 y 38 o seleccionar simplemente una de las mismas.

El combinador o selector 40 reenvía la predicción resultante de la porción actual para insertar la entrada del restador 24 y la entrada adicionalmente del sumador 30, respectivamente.

50 En la salida del restador 24, se emite el residuo de la predicción de la porción actual y el módulo de cuantificación/transformada 36 está configurado para transformar esta señal residual cuantificando los coeficientes de transformada. La transformada puede ser cualquier transformada de descomposición espectral tal como una DCT. Debido a la cuantificación, el resultado de procesamiento del módulo de cuantificación/transformada 26 es irreversible. Es decir, la codificación da como resultado pérdidas. La salida del módulo 26 es la señal residual  $42_1$  a  
55 transmitirse en el flujo de datos. La señal residual  $42_1$  se descuantifica y se transforma a la inversa en el módulo 28 para reconstruir la señal residual tanto como sea posible, es decir para corresponder a la señal residual como se emite por el restador 24 a pesar del ruido de cuantificación. El sumador 30 combina esta señal residual reconstruida con la predicción de la porción actual mediante suma. Serían factibles también otras combinaciones. Por ejemplo, el restador 24 podría operar como un divisor para medir el residuo en proporciones, y el sumador podría implementarse  
60 como un multiplicador para reconstruir la porción actual, de acuerdo con una alternativa. La salida del sumador 30, por lo tanto, representa una reconstrucción preliminar de la porción actual. El procesamiento adicional, sin embargo, en el módulo 32 puede usarse opcionalmente para mejorar la reconstrucción. Tal procesamiento adicional puede implicar, por ejemplo, desbloqueo, filtración adaptativa y similares. Todas las reconstrucciones disponibles hasta ahora se almacenan en memoria intermedia en la memoria intermedia de instantánea decodificada 34. Por lo tanto,

la memoria intermedia de instantánea decodificada 34 almacena en memoria intermedia fotogramas de vídeo previamente reconstruidos 14<sub>1</sub> y porciones previamente reconstruidas del fotograma actual al que pertenece la porción actual.

5 Para posibilitar que el decodificador reconstruya la señal de múltiples vistas desde el flujo de datos 18, el módulo de cuantificación/transformada 26 reenvía la señal residual 42<sub>1</sub> a un multiplexor 44 del codificador 20. De manera concurrente, el módulo de predicción 36 reenvía parámetros de intra predicción 46<sub>1</sub> al multiplexor 44, el módulo de inter predicción 38 reenvía parámetros de inter predicción 48<sub>1</sub> al multiplexor 44 y el módulo de procesamiento adicional 32 reenvía parámetros de procesamiento adicional 50<sub>1</sub> al multiplexor 44 que, a su vez, multiplexa o inserta toda esta información en el flujo de datos 18.

15 Como se hará evidente a partir del análisis anterior de acuerdo con la realización de la Figura 1, la codificación de vídeo 14<sub>1</sub> por el ramal de codificación 22<sub>v,1</sub> es autónoma en que la codificación es independiente de los datos de mapa de profundidad/disparidad 16<sub>1</sub> y los datos de cualquiera de las otras vistas 12<sub>2</sub>. Desde un punto de vista más general, el ramal de codificación 22<sub>v,1</sub> puede considerarse como codificar vídeo 14<sub>1</sub> en el flujo de datos 18 determinando parámetros de codificación y, de acuerdo con los primeros parámetros de codificación, predecir una porción actual del vídeo 14<sub>1</sub> desde una porción del vídeo previamente codificada 14<sub>1</sub>, codificada en el flujo de datos 18 por el codificador 20 antes de la codificación de la porción actual, y determinar un error de predicción de la predicción de la porción actual para obtener datos de corrección, en concreto la señal residual anteriormente mencionada 42<sub>1</sub>. Los parámetros de codificación y los datos de corrección se insertan en el flujo de datos 18.

20 Los parámetros de codificación recién mencionados insertados en el flujo de datos 18 por el ramal de codificación 22<sub>v,1</sub> definen/señalizan la segmentación de los fotogramas de vídeo 14<sub>1</sub> como se ha analizado anteriormente de manera breve.

- 25
- Además, los parámetros de codificación pueden comprender información de modo de codificación que indica para cada segmento o porción actual, el modo de codificación que se ha de usar para predecir el respectivo segmento tal como intra predicción, inter predicción, o una combinación de las mismas.
  - 30 - Los parámetros de codificación pueden comprender también los parámetros de predicción recién mencionados tal como los parámetros de intra predicción para porciones/segmentos previstos por intra predicción, y parámetros de inter predicción para porciones/segmentos inter previstos.
  - Los parámetros de codificación pueden comprender adicionalmente, sin embargo, parámetros de procesamiento adicional 50<sub>1</sub> que señalizan al lado de decodificación cómo procesar adicionalmente las porciones de vídeo ya reconstruidas 14<sub>1</sub> antes de usar las mismas para predecir las porciones de vídeo actuales o siguientes 14<sub>1</sub>. Estos parámetros de procesamiento adicional 50<sub>1</sub> pueden comprender índices que indexan respectivos filtros, coeficientes de filtro o similares.
  - 35 - Los parámetros de predicción 46<sub>1</sub>, 48<sub>1</sub> y los parámetros de procesamiento adicional 50<sub>1</sub> pueden comprender incluso adicionalmente datos de sub-segmentación para definir una sub-segmentación adicional con relación a la segmentación anteriormente mencionada que define la granularidad de la selección de modo, o que define una segmentación completamente independiente tal como para la aplicación de diferentes filtros adaptativos para diferentes porciones de los fotogramas en el procesamiento adicional.
  - 40 - Los parámetros de codificación pueden influenciar también la determinación de la señal residual y por lo tanto, ser parte de la señal residual 42<sub>1</sub>. Por ejemplo, los niveles de coeficiente de transformada espectral emitidos por el módulo de cuantificación/transformada 26 pueden considerarse como datos de corrección, mientras que el tamaño de etapa de cuantificación puede señalizarse en el flujo de datos 18 también, y el parámetro de tamaño de etapa de cuantificación puede considerarse como un parámetro de codificación en el sentido de la descripción presentada a continuación.
  - 45 - Los parámetros de codificación pueden definir adicionalmente parámetros de predicción que definen una predicción de segunda etapa de la predicción residual de la primera etapa de predicción anteriormente analizada.
  - 50 Puede usarse en este sentido intra/inter predicción.

Para aumentar la eficacia de codificación, el codificador 20 comprende un módulo de intercambio de información de codificación 52 que recibe todos los parámetros de codificación e información adicional que influencia, o que se ve influenciado por, el procesamiento en los módulos 36, 38 y 32. Por ejemplo, como se indica de manera ilustrativa, extendiendo verticalmente las flechas que apuntan desde los respectivos módulos hasta el módulo de intercambio de información de codificación 52. El módulo de intercambio de información de codificación 52 es responsable de compartir los parámetros de codificación que codifica adicionalmente de manera opcional información entre los ramales de codificación 22 de modo que los ramales pueden predecir o adoptar parámetros de codificación entre sí. En la realización de la Figura 1 se define un orden entre las entidades de datos, en concreto vídeo y datos de mapa de profundidad/disparidad, de las vistas 12<sub>1</sub> y 12<sub>2</sub> de la señal de múltiples vistas 10 para este fin. En particular, el vídeo 14<sub>1</sub> de la primera vista 12<sub>1</sub> precede los datos de mapa de profundidad/disparidad 16<sub>1</sub> de la primera vista seguido por el vídeo 14<sub>2</sub> y a continuación los datos de mapa de profundidad/disparidad 16<sub>2</sub> de la segunda vista 12<sub>2</sub> y así sucesivamente. Debería observarse en este punto que este orden estricto entre las entidades de datos de la señal de múltiples vistas 10 no necesita aplicarse estrictamente para la codificación de toda la señal de múltiples

vistas 10, pero por motivos de un análisis más fácil, se supone a continuación que este orden es constante. El orden entre las entidades de datos, de manera natural, también define un orden entre los ramales 22 que están asociados con el mismo.

5 Como ya se ha indicado anteriormente, los ramales de codificación adicionales 22 tal como el ramal de codificación  $22_{d,1}$ ,  $22_{v,2}$  y  $22_{d,2}$  actúan de manera similar al ramal de codificación  $22_{v,1}$  para codificar la respectiva entrada  $16_1$ ,  $14_2$  y  $16_2$ , respectivamente. Sin embargo, debido al orden recién mencionado entre los vídeos y datos de mapa de profundidad/disparidad de las vistas  $12_1$  y  $12_2$ , respectivamente, y el correspondiente orden definido entre los ramales de codificación 22, el ramal de codificación  $22_{d,1}$  tiene, por ejemplo, libertad adicional al predecir parámetros de codificación para usarse para codificar porciones actuales de los datos de mapa de profundidad/disparidad  $16_1$  de la primera vista  $12_1$ . Esto es debido al orden anteriormente mencionado entre vídeo y datos de mapa de profundidad/disparidad de las diferentes vistas: por ejemplo, cada una de estas entidades se permite que se codifique usando porciones reconstruidas de sí misma así como entidades de la misma precedentes en el orden anteriormente mencionado entre estas entidades de datos. Por consiguiente, al codificar los datos de mapa de profundidad/disparidad  $16_1$ , el ramal de codificación  $22_{d,1}$  se permite que use información conocida a partir de porciones previamente reconstruidas del correspondiente vídeo  $14_1$ . Cómo el ramal  $22_{d,1}$  aprovecha las porciones reconstruidas del vídeo  $14_1$  para predecir alguna propiedad de los datos de mapa de profundidad/disparidad  $16_1$ , que posibilita una mejor tasa de compresión de la compresión de los datos de mapa de profundidad/disparidad  $16_1$ , se describe en más detalle a continuación. Más allá de esto, sin embargo, el ramal de codificación  $22_{d,1}$  puede predecir/adoptar parámetros de codificación implicados al codificar vídeo  $14_1$  como se ha mencionado anteriormente, para obtener parámetros de codificación para codificar los datos de mapa de profundidad/disparidad  $16_1$ . En caso de adopción, la señalización de cualesquiera parámetros de codificación con respecto a los datos de mapa de profundidad/disparidad  $16_1$  en el flujo de datos 18 puede suprimirse. En caso de predicción, simplemente la predicción residual/datos de corrección con respecto a estos parámetros de codificación puede tener que señalizarse en el flujo de datos 18. Ejemplos para tal predicción/adopción de parámetros de codificación se describe adicionalmente a continuación, también.

30 Se presentan capacidades de predicción adicionales para las entidades de datos posteriores, en concreto el vídeo  $14_2$  y los datos de mapa de profundidad/disparidad  $16_2$  de la segunda vista  $12_2$ . Con respecto a estos ramales de codificación, el módulo de inter predicción de los mismos puede no únicamente realizar predicción temporal, sino también predicción inter-*vista*. Los correspondientes parámetros de inter predicción comprenden información similar en comparación con predicción temporal, en concreto segmento previsto por inter-*vista*, un vector de disparidad, un índice de vista, un índice de trama de referencia y/o una indicación de un número de hipótesis, es decir la indicación de un número de inter predicciones que participan al formar la inter predicción inter-*vista* por medio de suma, por ejemplo. Tal predicción inter-*vista* está disponible no únicamente para el ramal  $22_{v,2}$  con respecto al vídeo  $14_2$ , sino también para el módulo de inter predicción 38 del ramal  $22_{d,2}$  con respecto a los datos de mapa de profundidad/disparidad  $16_2$ . De manera natural, estos parámetros de predicción inter-*vista* también representan parámetros de codificación que pueden servir como una base para adopción/predicción para datos de vista posteriores de una posible tercera vista que, sin embargo, no se muestra en la Figura 1.

40 Debido a las medidas anteriores, la cantidad de datos a insertarse en el flujo de datos 18 por el multiplexor 44 se reduce adicionalmente. En particular, la cantidad de parámetros de codificación de los ramales de codificación  $22_{d,1}$ ,  $22_{v,2}$  y  $22_{d,2}$  puede reducirse enormemente adoptando parámetros de codificación de ramales de codificación anteriores o simplemente insertando residuos de predicción con relación a los mismos en el flujo de datos 28 mediante el multiplexor 44. Debido a la capacidad de elegir entre predicción temporal y de inter-*vista*, también puede reducirse la cantidad de datos residuales  $42_3$  y  $42_4$  de los ramales de codificación  $22_{v,2}$ , y  $22_{d,2}$ . La reducción en la cantidad de datos residuales sobre-compensa el esfuerzo de codificación adicional al diferenciar modos de predicción temporal y de inter-*vista*.

50 Para explicar los principios de adopción/predicción del parámetro de codificación en más detalle, se hace referencia a la Figura 2. La Figura 2 muestra una porción ejemplar de la señal de múltiples vistas 10. La Figura 2 ilustra el fotograma de vídeo  $v_{1,t}$  segmentándose en segmentos o porciones 60a, 60b y 60c. Por razones de simplificación, únicamente se muestran tres porciones de fotograma  $v_{1,t}$ , aunque la segmentación puede dividir sin interrupciones y sin huecos el fotograma en segmentos/porciones. Como se ha mencionado anteriormente, la segmentación de fotograma de vídeo  $v_{1,t}$  puede fijarse o variar en el tiempo, y la segmentación se señala en el flujo de datos. La Figura 2 ilustra que las porciones 60a y 60b se predicen temporalmente usando vectores de movimiento  $62a$  y  $62b$  a partir de una versión reconstruida de cualquier fotograma de vídeo de referencia  $14_1$ , que en el presente caso es ejemplarmente el fotograma  $v_{1,t-1}$ . Como es conocido en la técnica, el orden de codificación entre los fotogramas de vídeo  $14_1$  puede no coincidir con el orden de presentación entre estos fotogramas, y en consecuencia el fotograma de referencia puede suceder al fotograma actual  $v_{1,t}$  en orden de tiempo de presentación 64. La porción 60c es, por ejemplo, una porción intra prevista para la cual se insertan parámetros de intra predicción en el flujo de datos 18.

Al codificar el mapa de profundidad/disparidad  $d_{1,t}$  el ramal de codificación  $22_{d,1}$  usa la segmentación de fotograma de vídeo  $v_{1,t}$  como una base/predicción para la segmentación a usarse para el mapa de profundidad/disparidad  $d_{1,t}$

señalizando la desviación de la segmentación con relación al fotograma de vídeo  $v_{1,t}$  mediante el flujo de datos 18. La Figura 2 ilustra el caso de que el ramal de codificación  $22_{d,1}$  usa la segmentación de fotograma de vídeo  $v_1$  como una pre-segmentación de mapa de profundidad/disparidad  $d_{1,t}$ . Es decir, el ramal de codificación  $22_{d,1}$  adopta la pre-segmentación a partir de la segmentación de vídeo  $v_{1,t}$  o predice la pre-segmentación a partir del mismo.

- 5
- Además, el ramal de codificación  $22_{d,1}$  puede adoptar o predecir los modos de codificación de las porciones 66a, 66b y 66c del mapa de profundidad/disparidad  $d_{1,t}$  a partir de los modos de codificación asignados a la respectiva porción 60a, 60b y 60c en el fotograma de vídeo  $v_{1,t}$ . En caso de una segmentación diferente entre el fotograma de vídeo  $v_{1,t}$  y el mapa de profundidad/disparidad  $d_{1,t}$ , la adopción/predicción de modos de codificación del fotograma de vídeo  $v_{1,t}$  puede controlarse de manera que la adopción/predicción se obtiene a partir de porciones de la segmentación del fotograma de vídeo  $v_{1,t}$  co-localizadas. Una definición apropiada de co-localización podría ser como sigue. La porción co-localizada en el fotograma de vídeo  $v_{1,t}$  para una porción actual en el mapa de profundidad/disparidad  $d_{1,t}$ , puede ser, por ejemplo, la que comprende la posición co-localizada en la esquina superior izquierda del fotograma actual en el mapa de profundidad/disparidad  $d_{1,t}$ . En caso de predicción de los modos de codificación, el ramal de codificación  $22_{d,1}$  puede señalar las desviaciones de modo de codificación de las porciones 66a a 66c del mapa de profundidad/disparidad  $d_{1,t}$  con relación a los modos de codificación en el fotograma de vídeo  $v_{1,t}$  señalado explícitamente en el flujo de datos 18.
- 10
- En lo que respecta a los parámetros de predicción, el ramal de codificación  $22_{d,1}$  tiene la libertad de adoptar o predecir espacialmente parámetros de predicción usados para codificar porciones vecinas en el mismo mapa de profundidad/disparidad  $d_{1,t}$  o para adoptar/prededir algunos de los parámetros de predicción usados para codificar porciones co-localizadas 60a a 6c de fotograma de vídeo  $V_{1,t}$ . Por ejemplo, la Figura 2 ilustra que la porción 66a del mapa de profundidad/disparidad  $d_{1,t}$  es una porción inter prevista, y el correspondiente vector de movimiento 68a puede adoptarse o predecirse del vector de movimiento 62a de la porción co-localizada 60a de fotograma de vídeo  $v_{1,t}$ . En caso de predicción, simplemente la diferencia de vector de movimiento se ha de insertar en el flujo de datos 18 como parte de parámetros de inter predicción  $48_2$ .
- 15
- En términos de eficacia de codificación, puede ser favorable para el ramal de codificación  $22_{d,1}$  que tenga la capacidad de subdividir segmentos de la pre-segmentación del mapa de profundidad/disparidad  $d_{1,t}$  usando una denominada línea de separación en cuña 70 señalizando la localización de esta línea de separación en cuña 70 al lado de decodificación en el flujo de datos 18. Mediante esta medida, en el ejemplo de la Figura 2, la porción 66c de mapa de profundidad/disparidad  $d_{1,t}$  se subdivide en dos porciones con forma en cuña 72a y 72b. El ramal de codificación  $24_{d,1}$  puede configurarse para codificar estos sub-segmentos 72a y 72b de manera separada. En el caso de la Figura 2, ambos sub-segmentos 72a y 72b se inter predicen ejemplarmente usando respectivos vectores de movimiento 68c y 68d. En caso de usar intra predicción de ambos sub-segmentos 72a y 72b, un valor de CC para cada segmento puede derivarse por extrapolación de los valores de CC de segmentos causales vecinos con la opción de perfeccionar cada uno de estos valores de CC derivados transmitiendo un valor de CC de perfeccionamiento correspondiente al decodificador como un parámetro de intra predicción. Existen varias posibilidades para posibilitar que el decodificador determine la línea de separación en cuñas que se han usado para que el codificador subdivida la pre-segmentación del mapa de profundidad/disparidad. El ramal de codificación  $22_{d,1}$  puede configurarse para usar cualquiera de estas posibilidades de manera exclusiva. Como alternativa, el ramal de codificación  $12_{d,1}$  puede tener la libertad de elegir entre las siguientes opciones de codificación, y para señalar la elección al decodificador como información secundaria en el flujo de datos 18:
- 20
- 25
- 30
- 35
- 40
- 45
- 50
- 55
- 60
- La línea de separación en cuña 70 puede ser, por ejemplo, una línea recta. La señalización de la localización de esta línea 70 al lado de decodificación puede implicar la señalización de un punto de intersección a lo largo del borde del segmento 66c junto con una pendiente o información de gradiente o la indicación de los dos puntos de intersección de la línea de separación en cuña 70 con el borde del segmento 66c. En una realización, la línea de separación en cuña 70 puede señalizarse explícitamente en el flujo de datos por indicación de los dos puntos de intersección de la línea de separación en cuña 70 con el borde del segmento 66c, donde la granularidad de la cuadrícula que indica posibles puntos de intersección, es decir la granularidad o resolución de la indicación de los puntos de intersección, puede depender del tamaño del segmento 66c o parámetros de codificación como, por ejemplo, el parámetro de cuantificación.
- En una realización alternativa, donde se proporciona la pre-segmentación por, por ejemplo, un particionamiento de bloques basado en árbol cuádruple usando bloques cuadrados diádicos, el conjunto permisible de puntos de intersección para cada tamaño de bloque puede proporcionarse como una tabla de correspondencia (LUT) de manera que la señalización de cada punto de intersección implica la señalización de un índice de LUT correspondiente.
- De acuerdo con incluso otra posibilidad, sin embargo, el ramal de codificación  $22_{d,1}$  usa la porción reconstruida 60c del fotograma de vídeo  $v_{1,t}$  en la memoria intermedia de instantánea decodificada 34 para predecir la localización de la línea de separación en cuña 70 señalizando en el flujo de datos, si lo hubiera, una desviación de la línea de separación en cuña 70 realmente para usarse al codificar el segmento 66c, al

5 decodificador. En particular, el módulo 52 puede realizar una detección de borde en el vídeo  $v_{1,t}$  a una localización que corresponde a la localización de porción 66c en el mapa de profundidad/disparidad  $d_{1,t}$ . Por ejemplo, la detección puede ser sensible a bordes en el fotograma de vídeo  $v_{1,t}$  donde el gradiente espacial de alguna característica escalada de intervalo tal como el brillo, el componente de luminancia o un componente de croma o crominancia o similar, supera algún umbral mínimo. Basándose en la localización de este borde 72, el módulo 52 podría determinar la línea de separación en cuña 70 de manera que la misma se extienda a lo largo del borde 72. A medida que el decodificador también tiene acceso al fotograma de vídeo reconstruido  $v_{1,t}$ , el decodificador puede determinar de manera análoga la línea de separación en cuña 70 para subdividir la porción 66c en subporciones con forma en cuña 72a y 72b. Por lo tanto, se graba la capacidad de señalización para señalar la línea de separación en cuña 70. El aspecto de tener un tamaño de porción 66c que depende de la resolución para representar la localización de línea de separación en cuña podría aplicarse también para el presente aspecto de determinación de la localización de la línea 70 por detección de borde, y para transmitir la desviación opcional de la localización prevista.

15 Al codificar el vídeo 14<sub>2</sub>, el ramal de codificación 22<sub>v,2</sub> tiene, además de las opciones de modo de codificación disponibles para el ramal de codificación 22<sub>v,1</sub>, la opción de predicción de inter-vista.

20 La Figura 2 ilustra, por ejemplo, que una porción 64b de la segmentación del fotograma de vídeo  $v_{2,t}$  se predice inter-vista a partir del fotograma de vídeo temporalmente correspondiente  $v_{1,t}$  de la primera vista de vídeo 14<sub>1</sub> usando un vector de disparidad 76.

25 A pesar de esta diferencia, el ramal de codificación 22<sub>v,2</sub> puede aprovechar adicionalmente toda la información disponible a partir de la codificación de fotograma de vídeo  $v_{1,t}$  y mapa de profundidad/disparidad  $d_{1,t}$  tal como, en particular, los parámetros de codificación usados en estas codificaciones. Por consiguiente, el ramal de codificación 22<sub>v,2</sub> puede adoptar o predecir los parámetros de movimiento que incluyen el vector de movimiento 78 para una porción temporalmente inter prevista 74a de fotograma de vídeo  $v_{2,t}$  desde cualquiera, o una combinación de, los vectores de movimiento 62a y 68a de porciones co-localizadas 60a y 66a del fotograma de vídeo temporalmente alineado  $v_{1,t}$  y mapa de profundidad/disparidad  $d_{1,t}$ , respectivamente. Si alguna vez, una predicción residual puede señalizarse con respecto a los parámetros de inter predicción para la porción 74a. En este sentido, debería recalcarse que el vector de movimiento 68a pueda ya haberse sometido a predicción/adopción del mismo vector de movimiento 62a.

35 Las otras posibilidades de adoptar/predecir parámetros de codificación para codificar el fotograma de vídeo  $V_{2,t}$  como se ha descrito anteriormente con respecto a la codificación del mapa de profundidad/disparidad  $d_{1,t}$ , son aplicables para la codificación del fotograma de vídeo  $v_{2,t}$  por el ramal de codificación 22<sub>v,2</sub> también, con los datos comunes disponibles por el módulo 52 que se aumentan, sin embargo, puesto que los parámetros de codificación de tanto el fotograma de vídeo  $v_{1,t}$  y el correspondiente mapa de profundidad/disparidad  $d_{1,t}$  están disponibles.

40 A continuación, el ramal de codificación 22<sub>d,2</sub> codifica el mapa de profundidad/disparidad  $d_{2,t}$  de manera similar a la codificación del mapa de profundidad/disparidad  $d_{1,t}$  por el ramal de codificación 22<sub>d,1</sub>. Esto se cumple, por ejemplo, con respecto a todas las ocasiones de adopción/predicción de parámetro de codificación desde el fotograma de vídeo  $v_{2,t}$  de la misma vista 12<sub>2</sub>. Adicionalmente, sin embargo, el ramal de codificación 22<sub>d,2</sub> tiene la oportunidad también de adoptar/predecir parámetros de codificación desde parámetros de codificación que se han usado para codificar el mapa de profundidad/disparidad  $d_{1,t}$  de la vista precedente 12<sub>1</sub>. Adicionalmente, el ramal de codificación 22<sub>d,2</sub> puede usar predicción de inter-vista como se ha explicado con respecto al ramal de codificación 22<sub>v,2</sub>.

50 Con respecto a la adopción/predicción de parámetro de codificación, puede merecer la pena restringir la posibilidad del ramal de codificación 22<sub>d,2</sub> para adoptar/predecir sus parámetros de codificación desde los parámetros de codificación de entidades previamente codificadas de la señal de múltiples vistas 10 al vídeo 14<sub>2</sub> de la misma vista 12<sub>2</sub> y los datos de mapa de profundidad/disparidad 16<sub>1</sub> de la vista vecina previamente codificada 12<sub>1</sub> para reducir la sobrecarga de señalización que proviene de la necesidad de señalar al lado de decodificación en el flujo de datos 18 la fuente de adopción/predicción para las respectivas porciones del mapa de profundidad/disparidad  $d_{2,t}$ . Por ejemplo, el ramal de codificación 22<sub>d,2</sub> puede predecir los parámetros de predicción para una porción prevista inter-vista 80a de mapa de profundidad/disparidad  $d_{2,t}$  que incluye el vector de disparidad 82 desde el vector de disparidad 76 de la porción co-localizada 74b del fotograma de vídeo  $v_{2,t}$ . En este caso, una indicación de la entidad de datos desde la que se realiza la adopción/predicción, en concreto el vídeo 14<sub>2</sub> en el caso de la Figura 2, puede omitirse puesto que el vídeo 14<sub>2</sub> es la única fuente posible para adopción/predicción de vector de disparidad para el mapa de profundidad/disparidad  $d_{2,t}$ . Al adoptar/predecir los parámetros de inter predicción de una porción temporalmente inter prevista 80b, sin embargo, el ramal de codificación 22<sub>d,2</sub> puede adoptar/predecir el correspondiente vector de movimiento 84 desde uno cualquiera de los vectores de movimiento 78, 68a y 62a y por consiguiente, el ramal de codificación 22<sub>d,2</sub> puede configurarse para señalar en el flujo de datos 18 la fuente de adopción/predicción para el vector de movimiento 84. Restringir las posibles fuentes a vídeo 14<sub>2</sub> y mapa de profundidad/disparidad 16<sub>1</sub> reduce la sobrecarga en este sentido.



Con respecto a las líneas de separación, el ramal de codificación  $22_{d,2}$  tiene las siguientes opciones además de aquellas anteriormente analizadas:

5 Para codificar el mapa de profundidad/disparidad  $d_{2,t}$  de la vista  $12_2$  usando una línea de separación en cuña, pueden usarse las correspondientes porciones de señal compensadas de disparidad  $d_{1,t}$ , tal como por detección de borde y derivar implícitamente la correspondiente línea de separación en cuña. La compensación de disparidad se usa a continuación para transferir la línea detectada en el mapa de profundidad/disparidad  $d_{1,t}$  al mapa de profundidad/disparidad  $d_{2,t}$ . Para compensación de disparidad pueden usarse los valores de profundidad/disparidad de primer plano a lo largo del respectivo borde detectado en el mapa de profundidad/disparidad  $d_{1,t}$ .

10 Como alternativa, para codificar el mapa de profundidad/disparidad  $d_{2,t}$  de la vista  $12_2$  usando una línea de separación en cuña, pueden usarse las correspondientes porciones de señal compensadas en disparidad  $d_{1,t}$ , usando una línea de separación en cuña dada en la porción compensada en disparidad de  $d_{1,t}$ , es decir usando una línea de separación en cuña que se ha usado al codificar una porción co-localizada de la señal  $d_{1,t}$  como un predictor o adoptando el mismo.

20 Después de haber descrito el codificador 20 de la Figura 1, debería observarse que el mismo puede implementarse en software, hardware o firmware, es decir hardware programable. Aunque el diagrama de bloques de la Figura 1 sugiere que el codificador 20 comprende estructuralmente ramales de codificación paralelos, en concreto un ramal de codificación por vídeo y datos de profundidad/disparidad de la señal de múltiples vistas 10, este no es necesario que sea el caso. Por ejemplo, rutinas de software, porciones de circuito o porciones de lógica programable configuradas para realizar las tareas de elementos 24 a 40, respectivamente, pueden usarse secuencialmente para satisfacer las tareas para cada uno de los ramales de codificación. En procesamiento paralelo, los procesos de los ramales de codificación paralelos pueden realizarse en núcleos de procesador paralelos o en circuiterías de ejecución paralelas.

30 La Figura 3 muestra un ejemplo para un decodificador que puede decodificar el flujo de datos 18 para reconstruir uno o varios vídeos que corresponden a la escena representada por la señal de múltiples vistas desde el flujo de datos 18. En gran medida, la estructura y funcionalidad del decodificador de la Figura 3 es similar al codificador de la Figura 20 de modo que los signos de referencia de la Figura 1 se han vuelto a usar siempre que sea posible para indicar que la descripción de funcionalidad proporcionada anteriormente con respecto a la Figura 1 también se aplica a la Figura 3.

35 El decodificador de la Figura 3 se indica en general con el signo de referencia 100 y comprende una entrada para el flujo de datos 18 y una salida para emitir la reconstrucción de la una o varias vistas 102 anteriormente mencionadas. El decodificador 100 comprende un demultiplexor 104 y un par de ramales de decodificación 106 para cada una de las entidades de datos de la señal de múltiples vistas 10 (Figura 1) representadas por el flujo de datos 18 así como un extractor de vista 108 y un intercambiador de parámetro de codificación 110. Como era el caso con el codificador de la Figura 1, los ramales de decodificación 106 comprenden los mismos elementos de decodificación en una misma interconexión, que se describen, por consiguiente, de manera representativa con respecto al ramal de decodificación  $106_{v,1}$  responsable de la decodificación del vídeo  $14_1$  de la primera vista  $12_1$ . En particular, cada ramal de codificación 106 comprende una entrada conectada a una respectiva salida del multiplexor 104 y una salida conectada a una respectiva entrada de la vista extractor 108 para emitir al extractor de vista 108 la respectiva entidad de datos de la señal de múltiples vistas 10, es decir el vídeo  $14_1$  en caso de ramal de decodificación  $106_{v,1}$ . Mientras tanto, cada ramal de codificación 106 comprende un módulo de descuantificación/transformada inversa 28, un sumador 30, un módulo de procesamiento adicional 32 y una memoria intermedia de instantánea decodificada 34 conectada en serie entre el multiplexor 104 y el extractor de vista 108. El sumador 30, el módulo de procesamiento adicional 32 y la memoria intermedia de instantánea decodificada 34 forman un bucle junto con una conexión paralela de módulos de predicción 36 y 38 seguidos por un combinador/selector 40 que está, en el orden mencionado, conectado entre la memoria intermedia de instantánea decodificada 34 y la entrada adicional del sumador 30. Como se indica usando los mismos números de referencia como en el caso de la Figura 1, la estructura y funcionalidad de los elementos 28 a 40 de los ramales de decodificación 106 son similares a los correspondientes elementos de los ramales de codificación en la Figura 1 en que los elementos de los ramales de decodificación 106 emulan el procesamiento del proceso de codificación mediante el uso de la información transportada en el flujo de datos 18. De manera natural, los ramales de decodificación 106 simplemente invierten el procedimiento de codificación con respecto a los parámetros de codificación finalmente elegidos por el codificador 20, mientras que el codificador 20 de la Figura 1 tiene que hallar un conjunto óptimo de parámetros de codificación en algún sentido de optimización tal como parámetros de codificación que optimizan una función de coste de tasa/distorsión sometiéndose, opcionalmente, a ciertas restricciones tales como tasa máxima de bits o similar.

El demultiplexor 104 es para distribuir el flujo de datos 18 a los diversos ramales de decodificación 106. Por ejemplo, el demultiplexor 104 proporciona al módulo de descuantificación/transformada-inversa 28 con los datos residuales  $42_1$ , al módulo de procesamiento adicional 32 con los parámetros de procesamiento adicional  $50_1$ , al módulo de intra

predicción 36 con los parámetros de intra predicción 46<sub>1</sub> y al módulo de inter predicción 38 con los módulos de inter predicción 48<sub>1</sub>. El intercambiador de parámetro de codificación 110 actúa como el correspondiente módulo 52 en la Figura 1 para distribuir los parámetros de codificación comunes y otros datos comunes entre los diversos ramales de decodificación 106.

5 El extractor de vista 108 recibe la señal de múltiples vistas según se reconstruye por los ramales de decodificación paralelos 106 y extrae a partir de la misma una o varias vistas 102 que corresponden a los ángulos de vista o direcciones de vista prescritas por datos de control de extracción de vista intermedia proporcionados externamente 112.

10 Debido a la construcción similar del decodificador 100 con relación a la porción correspondiente del codificador 20, su funcionalidad hasta la interfaz al extractor de vista 108 se explica fácilmente de manera análoga a la descripción anterior.

15 De hecho, los ramales de decodificación 106<sub>v,1</sub> y 106<sub>d,1</sub> actúan juntos para reconstruir la primera vista 12<sub>1</sub> de la señal de múltiples vistas 10 desde el flujo de datos 18, de acuerdo con primeros parámetros de codificación contenidos en el flujo de datos 18 (tal como parámetros de escalamiento en 42<sub>1</sub>, los parámetros 46<sub>1</sub>, 48<sub>1</sub>, 50<sub>1</sub>, y los correspondientes no adoptados, y residuos de predicción, de los parámetros de codificación del segundo ramal 16<sub>d,1</sub>, en concreto 42<sub>2</sub>, los parámetros 46<sub>2</sub>, 48<sub>2</sub>, 50<sub>3</sub>), prediciendo una porción actual de la primera vista 12<sub>1</sub> desde una  
 20 porción previamente reconstruida de la señal de múltiples vistas 10, reconstruida desde el flujo de datos 18 anterior a la reconstrucción de la porción actual de la primera vista 12<sub>1</sub> y corrigiendo un error de predicción de la predicción de la porción actual de la primera vista 12<sub>1</sub> usando primeros datos de corrección, es decir en 42<sub>1</sub> y 42<sub>2</sub>, también contenidos en el flujo de datos 18. Mientras que el ramal de decodificación 106<sub>v,1</sub> es responsable de decodificar el vídeo 14<sub>1</sub>, un ramal de codificación 106<sub>d,1</sub> asume la responsabilidad de reconstruir los datos de mapa de  
 25 profundidad/disparidad 16<sub>1</sub>. Véase, por ejemplo, la Figura 2: el ramal de decodificación 106<sub>v,1</sub> reconstruye el vídeo 14<sub>1</sub> de la primera vista 12<sub>1</sub> desde el flujo de datos 18, de acuerdo con correspondientes parámetros de codificación leídos desde el flujo de datos 18, es decir parámetros de escalamiento en 42<sub>1</sub>, los parámetros 46<sub>1</sub>, 48<sub>1</sub>, 50<sub>1</sub>, prediciendo una porción actual del vídeo 14<sub>1</sub> tal como 60a, 60b o 60c desde una porción previamente reconstruida de la señal de múltiples vistas 10 y corrigiendo un error de predicción de esta predicción usando correspondientes  
 30 datos de corrección obtenidos desde el flujo de datos 18, es decir desde niveles de coeficiente de transformada en 42<sub>1</sub>. Por ejemplo, el ramal de decodificación 106<sub>v,1</sub> procesa el vídeo: 14<sub>1</sub> en unidades de los segmentos/porciones usando el orden de codificación entre los fotogramas de vídeo y, para codificar los segmentos en el fotograma, un orden de codificación entre los segmentos de estos fotogramas como lo hizo el correspondiente ramal de codificación del codificador. Por consiguiente, todas las porciones previamente reconstruidas de vídeo 14<sub>1</sub> están  
 35 disponibles para predicción para una porción actual. Los parámetros de codificación para una porción actual pueden incluir uno o más de los parámetros de intra predicción 50<sub>1</sub>, parámetros de inter predicción 48<sub>1</sub>, parámetros de filtro para el módulo de procesamiento adicional 32 y así sucesivamente. Los datos de corrección para corregir el error de predicción pueden representarse por los niveles de coeficiente espectral de transformada en los datos residuales 42<sub>1</sub>. No todos estos parámetros de codificación necesitan transmitirse en su totalidad. Algunos de ellos pueden haberse  
 40 previsto espacialmente a partir de parámetros de codificación de segmentos vecinos de vídeo 14<sub>1</sub>. Los vectores de movimiento para el vídeo 14<sub>1</sub>, por ejemplo, pueden transmitirse en el flujo de bits como la diferencia de vectores de movimiento entre vectores de movimiento de porciones/segmentos vecinos de vídeo 14<sub>1</sub>.

45 En lo que respecta al segundo ramal de decodificación 106<sub>d,1</sub>, el mismo tiene acceso: no únicamente a los datos residuales 42<sub>2</sub> y la correspondiente predicción y parámetros de filtro según se señalizan en el flujo de datos 18 y se distribuyen al respectivo ramal de decodificación 106<sub>d,1</sub> por el demultiplexor 104, es decir los parámetros de codificación no previstos a través de límites inter-vista, sino también indirectamente a los parámetros de codificación y datos de corrección proporcionados mediante el demultiplexor 104 al ramal de decodificación 106<sub>v,1</sub> o cualquier información derivable a partir de los mismos, según se distribuye mediante módulo de intercambio de información de  
 50 codificación 110. Por lo tanto, el ramal de decodificación 106<sub>d,1</sub> determina sus parámetros de codificación para reconstruir los datos de mapa de profundidad/disparidad 16<sub>1</sub> desde una porción de los parámetros de codificación reenviados mediante el demultiplexor 104 al par de ramales de decodificación 106<sub>v,1</sub> y 106<sub>d,1</sub> para la primera vista 12<sub>1</sub>, que solapa parcialmente la porción de estos parámetros de codificación especialmente dedicados y reenviados al ramal de decodificación 106<sub>v,1</sub>. Por ejemplo, el ramal de decodificación 106<sub>d,1</sub> determina el vector de movimiento  
 55 68a desde el vector de movimiento 62a explícitamente transmitido en 48<sub>1</sub>, por ejemplo, como una diferencia de vector de movimiento a otra porción vecina del fotograma v<sub>1,t</sub>, por un lado, y una diferencia de vector de movimiento explícitamente transmitida en 48<sub>2</sub>, por otro lado. Adicionalmente, o como alternativa, el ramal de decodificación 106<sub>d,1</sub> puede usar porciones reconstruidas del vídeo 14<sub>1</sub> como se ha descrito anteriormente con respecto a la predicción de la línea de separación en cuña para predecir parámetros de codificación para decodificar datos de  
 60 mapa de profundidad/disparidad 16<sub>1</sub>.

Para ser incluso más precisos, el ramal de decodificación 106<sub>d,1</sub> reconstruye los datos de mapa de profundidad/disparidad 14<sub>1</sub> de la primera vista 12<sub>1</sub> desde el flujo de datos por el uso de parámetros de codificación que se predicen al menos parcialmente desde los parámetros de codificación usados por el ramal de decodificación

106<sub>v,1</sub> (o adoptados a partir de los mismos) y/o se predicen desde las porciones reconstruidas de vídeo 14<sub>1</sub> en la memoria intermedia de instantánea decodificada 34 del ramal de decodificación 106<sub>v,1</sub>. Los residuos de predicción de los parámetros de codificación pueden obtenerse mediante el demultiplexor 104 desde el flujo de datos 18. Otros parámetros de codificación para el ramal de decodificación 106<sub>d,1</sub> pueden transmitirse en el flujo de datos 108 por completo o con respecto a otra base, en concreto haciendo referencia a un parámetro de codificación que se ha usado para codificar cualquiera de las porciones previamente reconstruidas de los mismos datos de mapa de profundidad/disparidad 16<sub>1</sub>. Basándose en estos parámetros de codificación, el ramal de decodificación 106<sub>d,1</sub> predice una porción actual de los datos de mapa de profundidad/disparidad 14<sub>1</sub> desde una porción previamente reconstruida de los datos de mapa de profundidad/disparidad 16<sub>1</sub>, reconstruidos desde el flujo de datos 18 por el ramal de decodificación 106<sub>d,1</sub> antes de la reconstrucción de la porción actual de los datos de mapa de profundidad/disparidad 16<sub>1</sub>, y corrigiendo un error de predicción de la predicción de la porción actual de los datos de mapa de profundidad/disparidad 16<sub>1</sub> usando los respectivos datos de corrección 42<sub>2</sub>.

Por lo tanto, el flujo de datos 18 puede comprender para una porción tal como la porción 66a de los datos de mapa de profundidad/disparidad 16<sub>1</sub>, lo siguiente:

una indicación en cuanto a si, o en cuanto a qué parte de, los parámetros de codificación para esa porción actual se ha de adoptar o predecir desde correspondientes parámetros de codificación, por ejemplo, de una porción de vídeo co-localizada y alineada en el tiempo 14<sub>1</sub> (o desde otros datos específicos de vídeo 14<sub>1</sub> tal como la versión reconstruida de los mismos para predecir la línea de separación en cuña),

- en caso afirmativo, en caso de predicción, el residuo del parámetro de codificación,
- si no, todos los parámetros de codificación para la porción actual, en el que los mismos pueden señalizarse como residuos de predicción en comparación con parámetros de codificación de porciones previamente reconstruidas de los datos de mapa de profundidad/disparidad 16<sub>1</sub>
- si no todos los parámetros de codificación se han de predecir/adaptar como se ha mencionado anteriormente, una parte restante de los parámetros de codificación para la porción actual, en el que los mismos pueden señalizarse como residuos de predicción en comparación con parámetros de codificación de porciones previamente reconstruidas de los datos de mapa de profundidad/disparidad 16<sub>1</sub>.

Por ejemplo, si la porción actual es una porción inter prevista tal como la porción 66a, el vector de movimiento 68a puede señalizarse en el flujo de datos 18 adoptándose o prediciéndose desde el vector de movimiento 62a. Además, el ramal de decodificación 106<sub>d,1</sub> puede predecir la localización de la línea de separación en cuña 70 dependiendo de los bordes detectados 72 en las porciones reconstruidas de vídeo 14<sub>1</sub> como se ha descrito anteriormente y aplicar esta línea de separación en cuña sin ninguna señalización en el flujo de datos 18 o dependiendo de una respectiva señalización en el flujo de datos 18. En otras palabras, la aplicación de la predicción de línea de separación en cuña para un fotograma actual puede suprimirse o permitirse por medio de señalización en el flujo de datos 18. En incluso otras palabras, el ramal de decodificación 106<sub>d,1</sub> puede predecir de manera eficaz la circunferencia de la porción actualmente reconstruida de los datos de mapa de profundidad/disparidad.

La funcionalidad del par de ramales de decodificación 106<sub>v,2</sub> y 106<sub>d,2</sub> para la segunda vista 12<sub>2</sub> es, como ya se ha descrito anteriormente con respecto a la codificación, similar que para la primera vista 12<sub>1</sub>. Ambos ramales cooperan para reconstruir la segunda vista 12<sub>2</sub> de la señal de múltiples vistas 10 desde el flujo de datos 18 mediante el uso de propios parámetros de codificación. Simplemente esa parte de estos parámetros de codificación necesita transmitirse y distribuirse mediante el demultiplexor 104 a cualquiera de estos dos los ramales de decodificación 106<sub>v,2</sub> y 106<sub>d,2</sub>, que no se adoptan/predicen a través del límite de vista entre las vistas 14<sub>1</sub> y 14<sub>2</sub>, y, opcionalmente, un residuo de la parte prevista inter-vista. Las porciones actuales de la segunda vista 12<sub>2</sub> se predicen desde porciones previamente reconstruidas de la señal de múltiples vistas 10, reconstruidas desde el flujo de datos 18 por cualquiera de los ramales de decodificación 106 antes de la reconstrucción de las respectivas porciones actuales de la segunda vista 12<sub>2</sub>, y corrigiendo el error de predicción en consecuencia usando los datos de corrección, es decir 42<sub>3</sub> y 42<sub>4</sub>, reenviados por el demultiplexor 104 hasta este par de ramales de decodificación 106<sub>v,2</sub> y 106<sub>d,2</sub>.

El ramal de decodificación 106<sub>v,2</sub> está configurado para al menos adoptar o predecir parcialmente sus parámetros de codificación desde los parámetros de codificación usados por cualquiera de los ramales de decodificación 106<sub>v,1</sub> y 106<sub>d,1</sub>. La siguiente información sobre parámetros de codificación puede estar presente para una porción actual del vídeo 14<sub>2</sub>:

- una indicación en cuanto a si, o en cuanto a qué parte de, los parámetros de codificación para esa porción actual se ha de adoptar o predecir, desde correspondientes parámetros de codificación, por ejemplo, de una porción co-localizada y alineada en el tiempo del vídeo 14<sub>1</sub> o datos de profundidad/disparidad 16<sub>1</sub>,
- en caso afirmativo, en caso de predicción, el residuo del parámetro de codificación,

- si no, todos los parámetros de codificación para la porción actual, en el que la misma puede señalizarse como residuos de predicción en comparación con parámetros de codificación de porciones previamente reconstruidas del vídeo 14<sub>2</sub>

5  
- si no todos los parámetros de codificación se han de predecir/adaptar como se ha mencionado anteriormente, una parte restante de los parámetros de codificación para la porción actual, en la que la misma puede señalizarse como residuos de predicción en comparación con parámetros de codificación de porciones previamente reconstruidas del vídeo 14<sub>2</sub>.

10  
- una señalización en el flujo de datos 18 puede señalar para una porción actual 74a si los correspondientes parámetros de codificación para esa porción, tal como el vector de movimiento 78, se han de leer desde el flujo de datos de manera completamente nueva, predecirse espacialmente o predecirse desde un vector de movimiento de una porción co-localizada del vídeo 14<sub>1</sub> o datos de mapa de profundidad/disparidad 16<sub>1</sub> de la primera vista 12<sub>1</sub> y el ramal de decodificación 106<sub>v,2</sub> pueden actuar en consecuencia, es decir extrayendo el vector de movimiento 78 desde el flujo de datos 18 en su totalidad, adoptando o prediciendo el mismo con, en el último caso, extrayendo datos de error de predicción con respecto a los parámetros de codificación para la porción actual 74a desde el flujo de datos 18.

20 El ramal de decodificación 106<sub>d,2</sub> puede actuar de manera similar. Es decir, el ramal de decodificación 106<sub>d,2</sub> puede determinar sus parámetros de codificación al menos parcialmente por adopción/predicción desde parámetros de codificación usados por cualquiera de los ramales de decodificación 106<sub>v,1</sub>, 106<sub>d,1</sub> y 106<sub>v,2</sub>, desde el vídeo reconstruido 14<sub>2</sub> y/o desde los datos de mapa de profundidad/disparidad 16<sub>1</sub> reconstruidos de la primera vista 12<sub>1</sub>. Por ejemplo, el flujo de datos 18 puede señalar para una porción actual 80b de los datos de mapa de profundidad/disparidad 16<sub>2</sub> en cuanto a si, y en cuanto a qué parte de, los parámetros de codificación para esta porción actual 80b se han de adoptar o predecir desde una porción co-localizada de cualquiera del vídeo 14<sub>1</sub>, datos de mapa de profundidad/disparidad 16<sub>1</sub> y vídeo 14<sub>2</sub> o un subconjunto apropiado del mismo. La parte de interés de estos parámetros de codificación puede implicar, por ejemplo, un vector de movimiento tal como 84, o un vector de disparidad tal como el vector de disparidad 82. Además, otros parámetros de codificación, tal como con respecto a la línea de separación en cuñas, pueden derivarse por el ramal de decodificación 106<sub>d,2</sub> mediante uso de la detección de borde en el vídeo 14<sub>2</sub>. Como alternativa, la detección de borde puede incluso aplicarse a los datos de mapa de profundidad/disparidad 16<sub>1</sub> reconstruidos aplicando una re-proyección predeterminada para transferir la localización del borde detectado en el mapa de profundidad/disparidad d<sub>1,t</sub> al mapa de profundidad/disparidad d<sub>2,t</sub> para servir como una base para una predicción de la localización de una línea de separación en cuña.

30 En cualquier caso, las porciones reconstruidas de los datos de múltiples vistas 10 llegan al extractor de vista 108 donde las vistas contenidas en el mismo son la base para una extracción de vista de nuevas vistas, es decir los vídeos asociados con estas nuevas vistas, por ejemplo. Esta extracción de vista puede comprender o implicar una re-proyección de los vídeos 14<sub>1</sub> y 14<sub>2</sub> usando los datos de mapa de profundidad/disparidad asociados con los mismos. Hablando con franqueza, al re-proyectar un vídeo en otra vista intermedia, las porciones del vídeo que corresponden a porciones de escena situadas más cerca del espectador que se desplazan a lo largo de la dirección de disparidad, es decir la dirección del vector de diferencia de dirección de visualización, más de las porciones del vídeo que corresponden a porciones de escena localizadas más alejadas de la posición de espectador. Un ejemplo para la extracción de vista realizada por el extractor de vista 108 se destaca a continuación con respecto a la Figura 4-6 y 8. El manejo de la desobstrucción puede realizarse por el extractor también. Sin embargo, antes de describir realizaciones adicionales a continuación, debería observarse que pueden realizarse varias modificaciones con respecto a las realizaciones señaladas anteriormente. Por ejemplo, la señal de múltiples vistas 10 no tiene que comprender necesariamente los datos de mapa de profundidad/disparidad para cada vista. Sin embargo, la reutilización de parámetro de codificación y compartición entre las múltiples vistas como se ha señalado anteriormente produce un aumento de eficacia de codificación. Además, para algunas vistas, los datos de mapa de profundidad/disparidad pueden estar restringidos a transmitirse en el flujo de datos a áreas de desobstrucción, es decir áreas que han de rellenar áreas desobstruidas en vistas re-proyectadas desde otras vistas de la señal de múltiples vistas estableciéndose a un valor que no importa en las áreas restantes de los mapas.

Como ya se ha indicado anteriormente, las vistas 12<sub>1</sub> y 12<sub>2</sub> de la señal de múltiples vistas 10 pueden tener diferentes resoluciones espaciales. Es decir, pueden transmitirse en el flujo de datos 18 usando diferentes resoluciones. Incluso en otras palabras, la resolución espacial a la que los ramales de codificación 21<sub>v,1</sub> y 22<sub>d,1</sub> realizan la codificación predictiva pueden ser más alta que la resolución espacial a la que los ramales de codificación 22<sub>v,2</sub> y 22<sub>d,2</sub> realizan la codificación predictiva de la vista posterior 12<sub>2</sub> que sigue a la vista 12<sub>1</sub> en el orden anteriormente mencionado entre las vistas. Los inventores de la presente invención hallan que esta medida mejora adicionalmente la relación de tasa/distorsión cuando se considera la calidad de las vistas sintetizadas 102. Por ejemplo, el codificador de la Figura 1 podría recibir la vista 12<sub>1</sub> y la vista 12<sub>2</sub> inicialmente a la misma resolución espacial, sin embargo, submuestreando el vídeo 14<sub>2</sub> y los datos de mapa de profundidad/disparidad 16<sub>2</sub> de la segunda vista 12<sub>2</sub> a una resolución espacial inferior antes de someter la misma al procedimiento de codificación predictivo realizado por los módulos 24 a 40. Sin embargo, las medidas anteriormente mencionadas de adopción y predicción de parámetros de codificación a través de límites de vista podrían aún realizarse escalando los parámetros de codificación que

forman la base de adopción o predicción de acuerdo con la relación entre las diferentes resoluciones de vista de origen y destino. Véase, por ejemplo, la Figura 2. Si el ramal de codificación  $22_{v,2}$  pretende adoptar o predecir el vector de movimiento 78 desde cualquiera de los vectores de movimiento 62a y 68a, a continuación el ramal de codificación  $22_{v,2}$  escalaría descendentemente por un valor que corresponde a la relación entre la alta resolución espacial de la vista  $12_1$ , es decir la vista de origen, y la baja resolución espacial de la vista  $12_2$ , es decir la vista de destino. De manera natural, lo mismo se aplica con respecto al decodificador y los ramales de decodificación 106. Los ramales de decodificación  $106_{v,2}$  y  $106_{d,2}$  realizarían la decodificación predictiva a la resolución espacial inferior con relación a los ramales de decodificación  $106_{v,1}$  y  $106_{d,1}$ . Después de la reconstrucción, se usaría sobre-muestreo para transferir las instantáneas reconstruidas y mapas de profundidad/disparidad emitidos por las memorias intermedias de instantánea de decodificación 34 de ramales de decodificación  $106_{v,2}$  y  $106_{d,2}$  desde la resolución espacial inferior a la resolución espacial superior antes de que la última alcance el extractor de vista 108. Un respectivo sobre-muestreador se posicionaría entre la respectiva memoria intermedia de instantánea decodificada y la respectiva entrada de la vista extractor 108. Como se ha mencionado anteriormente, en una vista  $12_1$  o  $12_2$ , vídeo y datos de mapa de profundidad/disparidad asociados pueden tener la misma resolución espacial. Sin embargo, adicionalmente o como alternativa, estos pares tienen diferente resolución espacial y las medidas recién descritas anteriormente se realizan a través de límites de resolución espacial, es decir entre datos de mapa de profundidad/disparidad y vídeo. Además, de acuerdo con otra realización, habría tres vistas que incluyen la vista  $12_3$ , no mostrada en las Figuras 1 a 3 para fines de ilustración, y mientras que la primera y segunda vistas tendrían la misma resolución espacial, la tercera vista  $12_3$  tendría la resolución espacial inferior. Por lo tanto, de acuerdo con las realizaciones recién descritas, algunas vistas posteriores, tal como la vista  $12_2$ , se submuestran antes de la codificación y se sobremuestran después de la decodificación. Este sub- y sobremuestreo, respectivamente, representa una clase de pre- o post-procesamiento de los ramales de de/codificación, en el que los parámetros de codificación usados para adopción/predicción de parámetros de codificación de cualquier vista posterior (destino) se escalan de acuerdo con la respectiva relación de las resoluciones espaciales de las vistas de origen y destino. Como ya se ha mencionado anteriormente, la calidad inferior a la que se transmiten las vistas posteriores, tal como la vista  $12_2$ , y se codifican de manera predictiva, no afecta significativamente a la calidad de la salida de la vista intermedia 102 del extractor de vista intermedia 108 debido al procesamiento en el extractor de vista intermedia 108. El extractor de vista 108 realiza una clase de interpolación/filtración de paso bajo en los vídeos  $14_1$  y  $14_2$  de cualquier manera debido a la re-proyección en la vista o vistas intermedias y el remuestreo necesario de los valores de muestra de vídeo re-proyectados en la misma cuadrícula de la vista o vistas intermedias. Para aprovechar el hecho de que la primera vista  $12_1$  se ha transmitido a una resolución espacial aumentada con relación a la vista vecina  $12_2$ , las vistas intermedias entre las mismas pueden obtenerse principalmente desde la vista  $12_1$ , usando la vista de baja resolución espacial  $12_2$  y su vídeo  $14_2$  simplemente como una vista auxiliar tal como, por ejemplo, simplemente para rellenar las áreas de desobstrucción de la versión reproyectada de vídeo  $14_1$ , o simplemente participando a un factor de ponderación reducido cuando se realiza algún promedio entre la versión reproyectada de vídeos de la vista  $12_1$  por una parte y  $12_2$  por otra parte. Mediante esta medida, la resolución espacial inferior de la vista  $12_2$  se compensa aunque la tasa de codificación de la segunda vista  $12_2$  se haya reducido significativamente debido a la transmisión a la resolución espacial inferior.

Debería mencionarse también que las realizaciones pueden modificarse en términos de la estructura interna de los ramales de codificación/decodificación. Por ejemplo, los modos de infra-predicción pueden no estar presentes, es decir pueden no estar disponibles modos de predicción espacial. De manera similar, puede abandonarse cualquiera de los modos de predicción inter-vida y temporal. Además, todas las opciones de procesamiento adicionales son opcionales. Por otra parte, pueden estar presentes módulos de post-procesamiento fuera de bucle en las salidas de ramales de decodificación 106 para realizar, por ejemplo, filtración adaptativa u otras medidas de mejora de calidad y/o el sobre-muestreo anteriormente mencionado. Además, puede no realizarse transformación del residuo. En su lugar, el residuo puede transmitirse en el dominio espacial en lugar del dominio de la frecuencia. En un sentido más general, los diseños de codificación/decodificación híbrida mostrados en las Figuras 1 y 3 pueden sustituirse por otros conceptos de codificación/decodificación tal como los basados en transformada de ondícula.

Debería mencionarse también que el decodificador no comprende necesariamente el extractor de vista 108. En su lugar, el extractor de vista 108 puede no estar presente. En este caso, el decodificador 100 es simplemente para reconstruir cualquiera de las vistas  $12_1$  y  $12_2$ , tal como una o varias de ellas. En caso de que no estén presentes datos de profundidad/disparidad para dos vistas individuales  $12_1$  y  $12_2$ , un extractor de vista 108 puede realizar, sin embargo, una extracción de vista intermedia aprovechando los vectores de disparidad relacionados con correspondientes porciones de vistas vecinas entre sí. Usando estos vectores de disparidad como vectores de disparidad de soporte de un campo de vector de disparidad asociado con vídeos de vistas vecinas, el extractor de vista 108 puede crear un vídeo de vista intermedia a partir de tales vídeos de vistas vecinas  $12_1$  y  $12_2$  aplicando este campo de vector de disparidad. Imagínese, por ejemplo, que el fotograma de vídeo  $v_{2,t}$  tenía el 50 % de sus porciones/segmentos previstos por inter-vida. Es decir, para el 50 % de las porciones/segmentos, existirían vectores de disparidad. Para las porciones restantes, podrían determinarse vectores de disparidad por el extractor de vista 108 por medio de interpolación/extrapolación en el sentido espacial. Puede usarse también la interpolación temporal usando los vectores de disparidad para porciones/segmentos de fotogramas de vídeo previamente reconstruidos  $14_2$ . El fotograma de vídeo  $v_{2,t}$  y/o el fotograma de vídeo de referencia  $v_{1,t}$  pueden distorsionarse a continuación de acuerdo con estos vectores de disparidad para producir una vista intermedia. Para este fin, los vectores de disparidad se escalan de acuerdo con la posición de la vista intermedia de la vista intermedia entre posiciones de

vista de la primera vista  $12_1$  y una segunda vista  $12_2$ . Los detalles con respecto a este procedimiento se señalan en más detalle a continuación.

- Una ganancia de eficacia de codificación se obtiene usando la opción anteriormente mencionada de determinación de la línea de separación en cuñas para extenderse a lo largo de bordes detectados en un fotograma del vídeo actual reconstruido. Por lo tanto, como se ha explicado anteriormente la predicción de posición de línea de separación en cuña anteriormente descrita puede usarse para cada una de las vistas, es decir todas ellas o simplemente un subconjunto apropiado de las mismas.
- Hasta ahora, el análisis anterior de la Figura 3 también revela un decodificador que tiene un ramal de decodificación  $106_{c,1}$  configurado para reconstruir un fotograma actual  $v_{1,t}$  de un vídeo  $14_1$  desde un flujo de datos 18, y un ramal de decodificación  $106_{d,1}$  configurado para detectar un borde 72 en el fotograma actual reconstruido  $v_{1,t}$ , determinar una línea de separación en cuña 70 para extenderse a lo largo del borde 72, y reconstruir un mapa de profundidad/disparidad  $d_{1,t}$  asociado con el fotograma actual  $v_{1,t}$  en unidades de segmentos 66a, 66b, 72a, 72 b de una segmentación del mapa de profundidad/disparidad  $d_{1,t}$  en el que se separan dos segmentos vecinos 72a, 72b de la segmentación entre sí por la línea de separación en cuña 70, desde el flujo de datos 18. El decodificador puede estar configurado para predecir el mapa de profundidad/disparidad  $d_{1,t}$  por segmentos usando distintos conjuntos de parámetros previstos para los segmentos, de segmentos previamente reconstruidos del mapa de profundidad/disparidad  $d_{1,t}$  asociados con el fotograma actual  $v_{1,t}$  o un mapa de profundidad/disparidad  $d_{1,t-1}$  asociado con cualquiera de los fotogramas previamente decodificados  $v_{1,t-1}$  del vídeo. El decodificador puede estar configurado de manera que la línea de separación en cuña 70 es una línea recta y el decodificador está configurado para determinar la segmentación de una pre-segmentación basada en bloques del mapa de profundidad/disparidad  $d_{1,t}$  dividiendo un bloque 66c de la pre-segmentación a lo largo de la línea de separación en cuña 70 de modo que los dos segmentos vecinos 72a, 72b son segmentos con forma en cuña que juntos forman el bloque 66c de la pre-segmentación.
- Resumiendo algunas de las realizaciones anteriores, estas realizaciones posibilitan la extracción de vista de decodificación comúnmente de vídeo de múltiples vistas y datos complementarios. La expresión "datos complementarios" se usa a continuación para indicar datos de mapa de profundidad/disparidad. De acuerdo con estas realizaciones, el vídeo de múltiples vistas y los datos complementarios están embebidos en una representación comprimida. Los datos complementarios pueden consistir en mapas de profundidad por píxel o disparidad. Las vistas extraídas 102 pueden ser diferentes de las vistas  $12_1$ ,  $12_2$  contenidas en la representación comprimida o flujo de bits 18 en términos de número de vista y posición espacial. La representación comprimida 18 se ha generado antes por un codificador 20, que puede usar los datos complementarios para mejorar también la codificación de los datos de vídeo.
- En contraste a los métodos del estado de la técnica actual, se lleva a cabo una decodificación conjunta, donde la decodificación de vídeo y los datos complementarios pueden soportarse y controlarse por información común. Ejemplos son un conjunto común de vectores de movimiento o de disparidad, que se usan para decodificar el vídeo así como los datos complementarios. Finalmente, se extraen vistas a partir de los datos de vídeo decodificados, datos complementarios y posibles datos combinados, donde el número y posición de vistas extraídas se controla por un control de extracción en el dispositivo de recepción.
- Además, el concepto de compresión de múltiples vistas anteriormente descrito puede usarse en relación con la síntesis de vista basada en disparidad. La síntesis de vista basada en disparidad significa lo siguiente. Si el contenido de escena se captura con múltiples cámaras, tal como los vídeos  $14_1$  y  $14_2$ , una percepción en 3D de este contenido puede presentarse al espectador. Para esto, se han de proporcionar pares estéreo con dirección de visualización ligeramente diferente para el ojo izquierdo y derecho. El desplazamiento del mismo contenido en ambas vistas para instancias de tiempo iguales se representa por el vector de disparidad. Similar a esto, el desplazamiento de contenido en una secuencia entre diferentes instancias de tiempo es el vector de movimiento, como se muestra en la Figura 4 para dos vistas en dos instancias de tiempo.
- Normalmente, se estima la disparidad directamente o como profundidad de escena, proporcionada externamente o grabada con sensores o cámaras especiales. La estimación de movimiento ya se ha llevado a cabo por un codificador convencional. Si se codifican múltiples vistas juntas, la dirección temporal y de inter-vista se tratan de manera similar, de manera que se lleva a cabo estimación de movimiento en dirección temporal así como en inter-vista durante la codificación. Esto ya se ha descrito anteriormente con respecto a las Figuras 1 y 2. Los vectores de movimiento estimados en dirección de inter-vista son los vectores de disparidad. Tales vectores de disparidad se mostraron en la Figura 2 de manera ejemplar en 82 y 76. Por lo tanto, el codificador 20 también lleva a cabo estimación de disparidad implícitamente y los vectores de disparidad están incluidos en el flujo de bits codificado 18. Estos vectores pueden usarse para síntesis de vista intermedia adicional en el decodificador, en concreto en el extractor de vista 108.
- Considérese un píxel  $p_1(x_1, y_1)$  en la vista 1 en la posición  $(x_1, y_1)$  y un píxel  $p_2(x_2, y_2)$  en la vista 2 en la posición  $(x_2, y_2)$ , que tienen valores de luminancia idénticos. A continuación,

$$p_1(x_1, y_1) = p_2(x_2, y_2). \quad (1)$$

se conectan sus posiciones  $(x_1, y_1)$  y  $(x_2, y_2)$  por el vector de disparidad en 2D, por ejemplo desde la vista 2 a la vista 1, que es  $d_{21}(x_2, y_2)$  con los componentes  $d_{x,21}(x_2, y_2)$  y  $d_{y,21}(x_2, y_2)$ . Por lo tanto, se mantiene la siguiente ecuación:

$$(x_1, y_1) = (x_2 + d_{x,21}(x_2, y_2), y_2 + d_{y,21}(x_2, y_2)). \quad (2)$$

Combinando (1) y (2),

$$p_1(x_2 + d_{x,21}(x_2, y_2), y_2 + d_{y,21}(x_2, y_2)) = p_2(x_2, y_2). \quad (3)$$

Como se muestra en la Figura 5, en la parte inferior derecha, pueden conectarse dos puntos con contenido idéntico con un vector de disparidad: añadir este vector a las coordenadas de  $p_2$ , proporciona la posición de  $p_1$  en coordenadas de imagen. Si el vector de disparidad  $d_{21}(x_2, y_2)$  se escala ahora por un factor  $k = [0...1]$ , puede direccionarse cualquier posición intermedia entre  $(x_1, y_1)$  y  $(x_2, y_2)$ . Por lo tanto, pueden generarse vistas intermedias desplazando el contenido de imagen de la vista 1 y/o la vista 2 por los vectores de disparidad escalados. Un ejemplo se muestra en la Figura 6 para una vista intermedia.

Por lo tanto, pueden generarse nuevas vistas intermedias con cualquier posición entre la vista 1 y la vista 2.

Más allá de esto, puede conseguirse también extrapolación de vista usando los factores de escalamiento  $k < 0$  y  $k > 1$  para las disparidades.

Estos métodos de escalamiento pueden aplicarse también en dirección temporal, de manera que pueden extraerse nuevos fotogramas escalando los vectores de movimiento, que conduce a la generación de secuencias de vídeo de velocidad de fotograma más alta.

Ahora, volviendo a las realizaciones anteriormente descritas con respecto las Figuras 1 y 3, estas realizaciones describen, entre otros, una estructura de decodificación paralela con decodificadores para vídeo y datos complementarios tales como mapas de profundidad, que contienen un módulo de información común, en concreto el módulo 110. Este módulo usa información espacial de ambas señales que se generaron por un codificador.

Ejemplos comunes son un conjunto de vectores de movimiento o de disparidad, por ejemplo extraídos en el proceso de codificación de datos de vídeo, que también se usa para datos de profundidad, por ejemplo. En el decodificador, esta información común se usa para dirigir la decodificación de vídeo y datos de profundidad y proporcionar la información requerida a cada ramal de decodificador también, opcionalmente, para extraer nuevas vistas. Con esta información, todas las vistas requeridas, por ejemplo para una vista de N vistas, pueden extraerse en paralelo desde los datos de vídeo. Ejemplos para información común o parámetros de codificación para compartirse entre los ramales de codificación/decodificación individual son:

- Los vectores de movimiento y disparidad comunes, por ejemplo desde los datos de vídeo que se usan también para los datos complementarios
- La estructura de particionamiento de bloque común, por ejemplo desde el particionamiento de datos de vídeo que también se usa para los datos complementarios
- Modos de predicción
- Datos de borde y contorno en información de luminancia y/o crominancia, por ejemplo una línea recta en un bloque de luminancia. Esto se usa para particionamiento de bloque no rectangular de datos complementarios. Este particionamiento se denomina en cuña y separa un bloque en dos regiones por una línea recta con cierto ángulo y posición.

La información común puede usarse también como un predictor desde un ramal de decodificación (por ejemplo para vídeo) para perfeccionarse en el otro ramal (por ejemplo datos complementarios) y viceversa. Esto puede incluir por ejemplo perfeccionamiento de vectores de movimiento o disparidad, inicialización de la estructura de bloque en datos complementarios por la estructura de bloque de datos de vídeo, extraer una línea recta desde la información de borde o contorno de luminancia o crominancia desde un bloque de vídeo y usar esta línea para una predicción de línea de separación en cuña (con el mismo ángulo pero posiblemente diferente posición en el correspondiente bloque de profundidad que mantiene el ángulo. El módulo de información común también transfiere datos parcialmente reconstruidos desde un ramal de decodificación al otro. Finalmente, los datos desde este módulo pueden traspasarse también al módulo de extracción de vista, donde se extraen todas las vistas necesarias, por ejemplo para una visualización (las pantallas pueden ser 2D, estereoscópica con dos vistas, autoestereoscópicas con N vistas).

Un aspecto importante es que si se codifica/decodifica más de un único par de la vista y la señal de profundidad/complementaria usando la estructura de co-/decodificación anteriormente descrita, puede considerarse un escenario de aplicación donde tenemos que transmitir cada instante de tiempo  $t$  un par de vistas de color  $v_{Color\_1}(t)$ ,  $v_{Color\_2}(t)$  junto con los datos de profundidad correspondientes  $v_{Profundidad\_1}(t)$  y  $v_{Profundidad\_2}(t)$ . Las realizaciones

anteriores sugieren codificar/decodificar en primer lugar la señal  $v_{\text{Color}_1}(t)$ , por ejemplo, usando predicción de movimiento compensado convencional. A continuación, en una segunda etapa, para codificación/decodificación de la correspondiente señal de profundidad  $v_{\text{Profundidad}_1}(t)$  puede reutilizarse la información de la señal codificada/decodificada  $v_{\text{Color}_1}(t)$ , como se ha señalado anteriormente. Posteriormente, la información acumulada desde  $v_{\text{Color}_1}(t)$  y  $v_{\text{Profundidad}_1}(t)$  puede utilizarse adicionalmente para codificación/decodificación de  $v_{\text{Color}_2}(t)$  y/o  $v_{\text{Profundidad}_2}(t)$ . Por lo tanto, puede aprovecharse en gran medida compartir y reutilizar información común entre las diferentes vistas y/o redundancias.

La estructura de extracción de decodificación y vista de la Figura 3 puede ilustrarse como alternativa como se muestra en la Figura 7.

Como se muestra, la estructura del decodificador de la Figura 7 está basada en dos estructuras de decodificación de vídeo clásicas paralelizadas para color y datos complementarios. Además, contiene un módulo de información común. Este módulo puede enviar, procesar y recibir cualquier información compartida desde y a cualquier módulo de ambas estructuras de decodificación. El vídeo decodificado: y los datos complementarios se combinan finalmente en el módulo de extracción de vista para extraer el número necesario de las vistas. En este punto, puede usarse también la información común desde el nuevo módulo. Los nuevos módulos del método de decodificación y extracción de vista nuevamente propuesto se destacan por el recuadro gris en la Figura 7.

El proceso de decodificación empieza recibiendo una representación comprimida común o flujo de bits, que contiene datos de vídeo, datos complementarios así como información, común para ambas, por ejemplo vectores de movimiento o vectores de disparidad, información de control, información de particionamiento de bloque, modos de predicción, datos de contorno, etc., de una o más vistas.

En primer lugar, se aplica una decodificación por entropía al flujo de bits para extraer los coeficientes de transformada cuantificados para vídeo y datos complementarios, que se alimentan en los dos ramales de codificación separados, destacados por los recuadros en gris con línea discontinua de puntos en la Figura 7, etiquetados "procesamiento de Datos de Vídeo" y "procesamiento de Datos Complementarios". Adicionalmente, la decodificación por entropía también extrae datos compartidos o comunes y los realimenta en el nuevo módulo de información común.

Ambos ramales de decodificación operan de manera similar después de la decodificación por entropía. Los coeficientes de transformada cuantificados recibidos se escalan y se aplica una transformada inversa para obtener la señal de diferencia. Para esto, se añaden datos previamente decodificados desde vistas temporales o vecinas. El tipo de información a añadirse se controla por datos de control especial: en el caso de datos complementarios o de vídeo intra codificados, no estaba disponible información anterior o vecina, de manera que se aplica reconstrucción de intra fotograma. Para datos de vídeo codificado o complementarios, están disponibles datos previamente decodificados de vistas temporalmente anteriores o vecinas (ajuste de conmutación actual en la Figura 7). Los datos previamente codificados se desplazan por los vectores de movimiento asociados en el bloque de compensación de movimiento y se añaden a la señal de diferencia para generar fotogramas iniciales. Si los datos previamente decodificados pertenecen a una vista vecina, los datos de movimiento representan los datos de disparidad. Estos fotogramas o vistas iniciales se procesan adicionalmente por los filtros de desbloqueo y posiblemente métodos de mejora, por ejemplo suavizado de borde, etc., para mejorar la calidad visual.

Después de esta etapa de mejora, los datos reconstruidos se transfieren a la memoria intermedia de instantánea decodificada. Esta memoria intermedia ordena los datos decodificados y emite las instantáneas decodificadas en el orden temporal correcto para cada instancia de tiempo. Los datos almacenados también se usan para el siguiente ciclo de procesamiento para servir como entrada a la compensación de movimiento/disparidad escalable.

Además de esta decodificación de vídeo separada y complementaria, se usa el nuevo módulo de información común, que procesa cualquier dato, que es común a vídeo y datos complementarios. Ejemplos de información común incluyen vectores de movimiento/disparidad compartidos, información de particionamiento de bloque, modos de predicción, datos de contorno, datos de control, pero también coeficientes de transformación o modos comunes, datos de mejora de vista, etc. Cualquier dato, que se procese en el vídeo individual y módulos complementarios, puede ser también parte del módulo común. Por lo tanto, pueden existir las conexiones a y desde el módulo común a todas las partes de los ramales de decodificación individuales. También, el módulo de información común puede contener suficientes datos, que únicamente sea necesario un ramal de decodificación separado y el módulo común para decodificar todos los datos de vídeo y complementarios. Un ejemplo para esto es una representación comprimida, donde algunas partes únicamente contienen datos de vídeo y todas las otras partes contienen datos de vídeo común y complementarios. En este punto, los datos de vídeo se decodifican en el ramal de decodificación de vídeo, mientras que todos los datos complementarios se procesan en el módulo común y se emiten a la síntesis de vista. Por lo tanto, en este ejemplo, no se usa el ramal complementario separado. También, los datos individuales desde los módulos de los ramales de decodificación separados pueden enviar información de vuelta al módulo de procesamiento de información común, por ejemplo en forma de datos parcialmente decodificados, para usarse allí o transferirse al otro ramal de decodificación. Un ejemplo son los datos de vídeo decodificados, como coeficientes de



transformada, vectores de movimiento, modos o ajustes, que se transfieren a los módulos de decodificación complementarios apropiados.

5 Después de la decodificación, los datos de vídeo y complementarios reconstruidos se transfieren para la extracción de vista ya sea desde los ramales de decodificación separados o desde el módulo de información común. En el módulo de extracción de vista, tal como 110 en la Figura 3, se extraen las vistas requeridas para un dispositivo de recepción, por ejemplo visualización de múltiples vistas. Este proceso se controla por el control de extracción de vista intermedia, que establece el número requerido y posición de las secuencias de vista. Un ejemplo para extracción de vista es la síntesis de vista: si se ha de sintetizar una nueva vista entre dos vistas originales 1 y 2, como se muestra en la Figura 6, los datos desde la vista 1 pueden desplazarse a la nueva posición en primer lugar. Sin embargo, este desplazamiento de disparidad es diferente para objetos de primer plano y de segundo plano, ya que el desplazamiento es inversamente proporcional a la profundidad de escena original (distancia frontal desde la cámara). Por lo tanto, se vuelven visibles nuevas áreas de segundo plano en la vista sintetizada, que no eran visibles en la vista 1. En este punto, la vista 2 puede usarse para rellenar esta información. También, pueden usarse datos espacialmente vecinos, por ejemplo información de segundo plano adyacente.

20 Como un ejemplo, considérese el ajuste en la Figura 8. En este punto, los datos decodificados consisten en 2 secuencias de vista con datos de color  $V_{Color\ 1}$  y  $V_{Color\ 2}$ , así como datos de profundidad  $V_{Profundidad\ 1}$  y  $V_{Profundidad\ 2}$ . A partir de estos datos, deberán extraerse las vistas para una visualización de 9 vistas con las vistas  $V_D\ 1$ ,  $V_D\ 2$  y  $V_D\ 9$ . La visualización señala el número y posición espacial de las vistas mediante el control de extracción de vista intermedia. En este punto, se requieren 9 vistas con una distancia espacial de 0,25, de manera que las vistas de visualización vecinas (por ejemplo  $V_D\ 2$  y  $V_D\ 3$ ) están 4 veces más cerca juntas en términos de posición espacial y percepción estereoscópica que las vistas en el flujo de bits. Por lo tanto, el conjunto de factores de extracción de vista  $\{K_1, K_2, K_9\}$  se establece a  $\{-0,5, -0,25, 0, 0,25, 0,5, 0,75, 1, 1,25, 1,5\}$ . Esto indica que las vistas de color decodificadas  $V_{Color\ 1}$  y  $V_{Color\ 2}$  coinciden en su posición espacial con las vistas de visualización  $V_D\ 3$  y  $V_D\ 7$  (como  $K_3 = 0$  y  $K_7 = 1$ ). Adicionalmente,  $V_D\ 4$  y  $V_D\ 5$  se interpolan entre  $V_{Color\ 1}$  y  $V_{Color\ 2}$ . Finalmente,  $V_D\ 1$  y  $V_D\ 2$  así como  $V_D\ 8$  y  $V_D\ 9$  se extrapolan en cada lado del par de flujo de bits  $V_{Color\ 1}$ ,  $V_{Color\ 2}$ . Con el conjunto de factores de extracción de vista, los datos de profundidad  $V_{Profundidad\ 1}$  y  $V_{Profundidad\ 2}$  se transforman en información de desplazamiento por píxel y se escalan en consecuencia en la etapa de extracción de vista para obtener 9 versiones diferentemente desplazadas de los datos de color decodificados.

35 Aunque se han descrito algunos aspectos en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del método correspondiente, donde un bloque o dispositivo corresponde a una etapa de método o una característica de una etapa de método. De manera análoga, los aspectos descritos en el contexto de una etapa de método también representan una descripción de un bloque o elemento correspondiente o características de un correspondiente aparato. Algunas o todas las etapas de método pueden ejecutarse por (o usando) un aparato de hardware, como por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, alguno o más de la mayoría de las etapas de método importantes pueden ejecutarse por un aparato de este tipo.

40 La señal de múltiples vistas codificada inventiva puede almacenarse en un medio de almacenamiento digital o puede transmitirse en un medio de transmisión tal como un medio de transmisión inalámbrica o un medio de transmisión cableado tal como internet.

45 Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención pueden implementarse en hardware o en software. La implementación puede realizarse usando un medio de almacenamiento digital, por ejemplo un disco flexible, un DVD, un Blu-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM, o una memoria FLASH, que tiene señales de control electrónicamente legibles almacenadas en el mismo, que cooperan (o pueden cooperar) con un sistema informático programable de manera que se realiza el respectivo método. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.

55 Algunas realizaciones de acuerdo con la invención comprenden un soporte de datos que tiene señales de control electrónicamente legibles, que pueden cooperar con un sistema informático programable, de manera que se realiza uno de los métodos descritos en el presente documento.

En general, las realizaciones de la presente invención pueden implementarse como un producto de programa informático con un código de programa, siendo el código de programa operativo para realizar uno de los métodos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede almacenarse, por ejemplo, en un soporte legible por máquina.

60 Otras realizaciones comprenden el programa informático para realizar uno de los métodos descritos en el presente documento, almacenado en soporte legible por máquina.

En otras palabras, una realización del método inventivo es, por lo tanto, un programa informático que tiene un código

de programa para realizar uno de los métodos descritos en el presente documento, cuando el programa informático se ejecuta en un ordenador.

5 Una realización adicional de los métodos inventivos es, por lo tanto, un soporte de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, grabado en el mismo, el programa informático para realizar uno de los métodos descritos en el presente documento. El soporte de datos, el medio de almacenamiento digital o el medio grabado son típicamente tangibles y/o no transitorios.

10 Una realización adicional del método inventivo es, por lo tanto, un flujo de datos o una secuencia de señales que representan el programa informático para realizar uno de los métodos descritos en el presente documento. El flujo de datos o la secuencia de señales pueden configurarse, por ejemplo, para transferirse mediante una conexión de comunicación de datos, por ejemplo mediante Internet.

15 Una realización adicional comprende un medio de procesamiento, por ejemplo un ordenador, o un dispositivo de lógica programable, configurado para o adaptado para realizar uno de los métodos descritos en el presente documento.

20 Una realización adicional comprende un ordenador que tiene instalado en el mismo el programa informático para realizar uno de los métodos descritos en el presente documento.

25 Una realización adicional de acuerdo con la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, electrónicamente u ópticamente) un programa informático para realizar uno de los métodos descritos en el presente documento a un receptor. El receptor puede ser, por ejemplo, un ordenador, un dispositivo móvil, un dispositivo de memoria o similares. El aparato o sistema puede comprender, por ejemplo, un servidor de ficheros para transferir el programa informático al receptor.

30 En algunas realizaciones, un dispositivo de lógica programable (por ejemplo un campo de matriz de puertas programables) puede usarse para realizar algunas o todas las funcionalidades de los métodos descritos en el presente documento. En algunas realizaciones, un campo de matriz de puertas programables puede cooperar con un microprocesador para realizar uno de los métodos descritos en el presente documento. En general, los métodos se realizan preferentemente por cualquier aparato de hardware.

35 Las realizaciones anteriormente descritas son meramente ilustrativas de los principios de la presente invención. Se entiende que serán evidentes modificaciones y variaciones de las disposiciones y los detalles descritos en el presente documento para los expertos en la materia. Se pretende, por lo tanto, estar limitado únicamente por el alcance de las reivindicaciones de patente siguientes y no por los detalles específicos presentados por medio de descripción y explicación de las realizaciones del presente documento.

REIVINDICACIONES

1. Decodificador configurado para  
 5 reconstruir (106<sub>v,1</sub>, 106<sub>d,1</sub>) una primera vista (12<sub>1</sub>) de una señal de múltiples vistas (10) desde un flujo de datos (18),  
 de acuerdo con primeros parámetros de codificación (46<sub>1</sub>, 48<sub>1</sub>, 50<sub>1</sub>) obtenidos desde el flujo de datos (18),  
 prediciendo una porción actual de la primera vista (12<sub>1</sub>) desde una primera porción previamente reconstruida de la  
 señal de múltiples vistas (10), reconstruida desde el flujo de datos (18) por el decodificador antes de la  
 10 reconstrucción de la porción actual de la primera vista (12<sub>1</sub>) y corrigiendo un error de predicción de la predicción de  
 la porción actual de la primera vista (12<sub>1</sub>) usando primeros datos de corrección (42<sub>1,2</sub>) contenidos en el flujo de datos  
 (18);  
 adoptar o predecir al menos parcialmente segundos parámetros de codificación desde los primeros parámetros de  
 codificación; y  
 reconstruir (106<sub>v,2</sub>, 106<sub>d,2</sub>) una segunda vista (12<sub>2</sub>) de la señal de múltiples vistas (10) desde el flujo de datos (18), de  
 acuerdo con los segundos parámetros de codificación, prediciendo una porción actual de la segunda vista desde una  
 15 segunda porción previamente reconstruida de la señal de múltiples vistas (10), reconstruida desde el flujo de datos  
 (18) por el decodificador antes de la reconstrucción de la porción actual de la segunda vista (12<sub>2</sub>) y corrigiendo un  
 error de predicción de la predicción de la porción actual de la segunda vista (12<sub>2</sub>) usando segundos datos de  
 corrección (42<sub>3,4</sub>) contenidos en el flujo de datos (18), en el que cada una de la primera vista (12<sub>1</sub>) y la segunda vista  
 (12<sub>2</sub>) comprende un vídeo (14<sub>1,2</sub>) capturado desde una respectiva posición de cámara, y datos de mapa de  
 20 profundidad/disparidad asociados (16<sub>1,2</sub>).  
 en el que el decodificador está configurado para  
 reconstruir (106<sub>v,1</sub>) el vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) desde el flujo de datos (18), de acuerdo con una primera  
 porción (46<sub>1</sub>, 48<sub>1</sub>, 50<sub>1</sub>) de los primeros parámetros de codificación, prediciendo una porción actual del vídeo (14<sub>1</sub>) de  
 la primera vista (12<sub>1</sub>) desde una tercera porción previamente reconstruida de la señal de múltiples vistas (10),  
 25 reconstruida desde el flujo de datos por el decodificador antes de la reconstrucción de la porción actual del vídeo  
 (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) y corrigiendo un error de predicción de la predicción de la porción actual del vídeo  
 (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) usando un primer subconjunto (42<sub>1</sub>) de los primeros datos de corrección (42<sub>1,2</sub>)  
 contenidos en el flujo de datos (18),  
 adoptar o predecir al menos parcialmente una segunda porción de los primeros parámetros de codificación desde la  
 30 primera porción de los primeros parámetros de codificación,  
 reconstruir (106<sub>d,1</sub>) los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) desde el flujo de  
 datos (18), de acuerdo con la segunda porción de los primeros parámetros de codificación, prediciendo una porción  
 actual de los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) desde una cuarta porción  
 previamente reconstruida de la señal de múltiples vistas (10), reconstruida desde el flujo de datos (18) por el  
 35 decodificador antes de la reconstrucción de la porción actual de los datos de mapa de profundidad/disparidad (16<sub>1</sub>)  
 de la primera vista (12<sub>1</sub>) y corrigiendo un error de predicción de la predicción de la porción actual de los datos de  
 mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) usando un segundo subconjunto (42<sub>2</sub>) de los primeros  
 datos de corrección,  
 en el que los primeros parámetros de codificación definen una segmentación de los fotogramas del vídeo (14<sub>1</sub>) de la  
 40 primera vista (12<sub>1</sub>), y el decodificador está **caracterizado porque** está configurado para, al reconstruir los datos de  
 mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>), usar la segmentación de los fotogramas del vídeo  
 (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) como una pre-segmentación de los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de  
 la primera vista (12<sub>1</sub>) junto con una señalización de una desviación de la segmentación para los datos de mapa de  
 45 profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) desde la segmentación de los fotogramas del vídeo (14<sub>1</sub>) de  
 la primera vista (12<sub>1</sub>) mediante el flujo de datos.
2. Decodificador de acuerdo con la reivindicación 1, en el que el decodificador está configurado adicionalmente para  
 50 extraer (108) una vista intermedia (102) desde la primera y segunda vistas.
3. Decodificador de acuerdo con las reivindicaciones 1 o 2, en el que el decodificador está configurado para predecir  
 la porción actual (60a) del vídeo de la primera vista (12<sub>1</sub>) desde una primera porción previamente reconstruida (v<sub>1,t-1</sub>)  
 del vídeo de la primera vista (12<sub>1</sub>), reconstruida desde el flujo de datos (18) por el decodificador antes de la  
 reconstrucción de la porción actual del vídeo de la primera vista, y  
 predecir la porción actual (66a) de los datos de mapa de profundidad/disparidad de la primera vista (12<sub>1</sub>) desde una  
 55 primera porción previamente reconstruida (d<sub>1,t-1</sub>) de los datos de mapa de profundidad/disparidad de la primera vista  
 (12<sub>1</sub>), reconstruida desde el flujo de datos (18) por el decodificador antes de la reconstrucción de la porción actual  
 (66a) de los datos de mapa de profundidad/disparidad de la primera vista (12<sub>1</sub>).
4. Decodificador de acuerdo con cualquiera de las reivindicaciones 1 a 3, en el que el decodificador está configurado  
 60 para adoptar o predecir al menos parcialmente una primera porción de los segundos parámetros de codificación  
 desde los primeros parámetros de codificación, reconstruir (106<sub>v,2</sub>) el vídeo (14<sub>2</sub>) de la segunda vista (12<sub>2</sub>) desde el  
 flujo de datos (18), de acuerdo con la primera porción de los segundos parámetros de codificación, prediciendo una  
 porción actual del vídeo (14<sub>2</sub>) de la segunda vista (12<sub>2</sub>) desde una quinta porción previamente reconstruida de la  
 señal de múltiples vistas (10), reconstruida desde el flujo de datos (18) por el decodificador antes de la

- reconstrucción de la porción actual del vídeo (14<sub>2</sub>) de la segunda vista (12<sub>2</sub>) y corregir un error de predicción de la predicción de la porción actual del vídeo (14<sub>2</sub>) de la segunda vista (12<sub>2</sub>) usando un primer subconjunto (42<sub>3</sub>) de los segundos datos de corrección contenidos en el flujo de datos (18),
- 5 adoptar o predecir al menos parcialmente una segunda porción de los segundos parámetros de codificación desde los primeros parámetros de codificación y/o la primera porción de los segundos parámetros de codificación, y reconstruir (106<sub>d,2</sub>) los datos de mapa de profundidad/disparidad (16<sub>2</sub>) de la segunda vista (12<sub>2</sub>) desde el flujo de datos (18), de acuerdo con la segunda porción de los segundos parámetros de codificación, prediciendo una porción actual de los datos de mapa de profundidad/disparidad (16<sub>2</sub>) de la segunda vista (12<sub>2</sub>) desde una sexta porción previamente reconstruida de la señal de múltiples vistas (10), reconstruida desde el flujo de datos (18) por el
- 10 decodificador antes de la reconstrucción de la porción actual de los datos de mapa de profundidad/disparidad (16<sub>2</sub>) de la segunda vista (12<sub>2</sub>) y corrigiendo un error de predicción de la predicción de la porción actual de los datos de mapa de profundidad/disparidad (16<sub>2</sub>) de la segunda vista (12<sub>2</sub>) usando un segundo subconjunto (42<sub>4</sub>) de los segundos datos de corrección.
- 15 5. Decodificador de acuerdo con la reivindicación 4, en el que el decodificador está configurado para predecir la porción actual (74a; 74b) del vídeo (14<sub>2</sub>) de la segunda vista (12<sub>2</sub>) desde una porción previamente reconstruida (v<sub>2,t-1</sub>) del vídeo (14<sub>2</sub>) de la segunda vista (12<sub>2</sub>), reconstruida desde el flujo de datos (18) por el decodificador antes de la reconstrucción de la porción actual del vídeo (14<sub>2</sub>) de la segunda vista (12<sub>2</sub>), o desde una segunda porción previamente reconstruida (v<sub>1,t</sub>) del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>), reconstruida desde el flujo de datos (18) por
- 20 el decodificador antes de la reconstrucción de la porción actual del vídeo (14<sub>2</sub>) de la segunda vista (12<sub>2</sub>), y predecir la porción actual (80a; 80b) de los datos de mapa de profundidad/disparidad (16<sub>2</sub>) de la segunda vista (12<sub>2</sub>) desde una porción previamente reconstruida (d<sub>2,t-1</sub>) de los datos de mapa de profundidad/disparidad (16<sub>2</sub>) de la segunda vista (12<sub>2</sub>), reconstruida desde el flujo de datos (18) por el decodificador antes de la reconstrucción de la porción actual de los datos de mapa de profundidad/disparidad (16<sub>2</sub>) de la segunda vista (12<sub>2</sub>), o desde una segunda porción
- 25 previamente reconstruida (d<sub>1,t</sub>) de los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>), reconstruida desde el flujo de datos (18) por el decodificador antes de la reconstrucción de la porción actual de los datos de mapa de profundidad/disparidad de la segunda vista.
- 30 6. Decodificador de acuerdo con la reivindicación 4 o 5, en el que el decodificador está configurado para detectar un borde (72) en un fotograma actual (v<sub>2,t</sub>) del vídeo de la segunda vista (12<sub>2</sub>) y determinar una línea de separación en cuña (70) para extenderse a lo largo del borde (72), y al reconstruir los datos de mapa de profundidad/disparidad (16<sub>2</sub>) de la segunda vista desde el flujo de datos (18), establecer un borde de la porción actual de un mapa de profundidad/disparidad (d<sub>2,t</sub>) de los datos de mapa de
- 35 profundidad/disparidad (16<sub>2</sub>) de la segunda vista, asociados con el fotograma actual (v<sub>2,t</sub>) del vídeo la segunda vista, para coincidir con la línea de separación en cuña.
7. Decodificador de acuerdo con la reivindicación 6, en el que el decodificador está configurado para realizar la predicción en la reconstrucción de los datos de mapa de profundidad/disparidad (16<sub>2</sub>) de la segunda vista desde el flujo de datos (18) por segmentos en unidades de segmentos a los que pertenece la porción actual, usando distintos
- 40 conjuntos de parámetros de predicción para los segmentos.
8. Decodificador de acuerdo con las reivindicaciones 6 o 7, en el que el decodificador está configurado de manera que la línea de separación en cuña (70) es una línea recta y el decodificador está configurado para dividir un bloque (66c) de una pre-segmentación de los datos de mapa de profundidad/disparidad (16<sub>2</sub>) de la segunda vista a lo largo
- 45 de la línea de separación en cuña (70) de modo que los dos segmentos vecinos (72a, 72b) son segmentos con forma en cuña que juntos forman el bloque (66c) de la pre-segmentación.
9. Decodificador de acuerdo con cualquiera de las reivindicaciones anteriores, en el que el decodificador está configurado de manera que los primeros y segundos parámetros de codificación comprenden primeros y segundos
- 50 parámetros de predicción, respectivamente, que controlan la predicción de la porción actual de la primera vista y la predicción de la porción actual de la segunda vista, respectivamente.
10. Decodificador de acuerdo con cualquiera de las reivindicaciones anteriores, en el que las porciones actuales son
- 55 segmentos de una segmentación de los fotogramas del vídeo de la primera y segunda vista, respectivamente.
11. Decodificador de acuerdo con cualquiera de las reivindicaciones anteriores, en el que el decodificador está configurado para reconstruir la primera vista (12<sub>1</sub>) de la señal de múltiples vistas (10) desde el flujo de datos realizando la predicción y corrección de la porción actual de la misma a una primera resolución espacial, y para reconstruir la segunda vista (12<sub>2</sub>) de la señal de múltiples vistas (10) desde el flujo de datos realizando la predicción
- 60 y corrección de la porción actual de la misma a una segunda resolución espacial inferior a la primera resolución espacial, sobre-muestreando a continuación la porción actual reconstruida de la segunda vista desde la segunda resolución espacial a la primera resolución espacial, en el que el decodificador está configurado para, adoptar o predecir al menos parcialmente los segundos parámetros de codificación desde los primeros parámetros de codificación, escalando los primeros parámetros de codificación

de acuerdo con una relación entre la primera y segunda resoluciones espaciales.

12. Codificador configurado para codificar una primera vista de una señal de múltiples vistas en un flujo de datos

5 determinando primeros parámetros de codificación,  
de acuerdo con los primeros parámetros de codificación, prediciendo una porción actual de la primera vista desde una primera porción previamente codificada de la señal de múltiples vistas, codificada en el flujo de datos por el codificador antes de la codificación de la porción actual de la primera vista, y determinando un error de predicción de la predicción de la porción actual de la primera vista para obtener primeros datos de corrección, e  
10 insertar los primeros parámetros de codificación y los primeros datos de corrección en el flujo de datos;

codificar una segunda vista de la señal de múltiples vistas en el flujo de datos,

15 determinando segundos parámetros de codificación adoptando o prediciendo los segundos parámetros de codificación desde los primeros parámetros de codificación;  
de acuerdo con los segundos parámetros de codificación, predecir una porción actual de la segunda vista desde una segunda porción previamente codificada de la señal de múltiples vistas, codificada en el flujo de datos por el codificador antes de la codificación de la porción actual de la segunda vista y determinar un error de predicción de la predicción de la porción actual de la segunda vista para obtener segundos datos de corrección contenidos en el flujo de datos, insertar los segundos datos de corrección en el flujo de datos.

en el que cada una de la primera vista (12<sub>1</sub>) y la segunda vista (12<sub>2</sub>) comprende un vídeo (14<sub>1,2</sub>) capturado desde una respectiva posición de cámara, y datos de mapa de profundidad/disparidad asociados (16<sub>1,2</sub>).

25 en el que el codificador está configurado para  
codificar (106<sub>v,1</sub>) el vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) en el flujo de datos (18), de acuerdo con una primera porción (46<sub>1</sub>, 48<sub>1</sub>, 50<sub>1</sub>) de los primeros parámetros de codificación, prediciendo una porción actual del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) desde una tercera porción previamente codificada de la señal de múltiples vistas (10), codificada en el flujo de datos por el codificador antes de la codificación de la porción actual del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) y corrigiendo un error de predicción de la predicción de la porción actual del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) usando un primer subconjunto (42<sub>1</sub>) de los primeros datos de corrección (42<sub>1,2</sub>) e insertando el error de predicción en el flujo de datos (18),

30 adoptar o predecir al menos parcialmente una segunda porción de los primeros parámetros de codificación desde la primera porción de los primeros parámetros de codificación,  
codificar los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) en el flujo de datos (18), de acuerdo con la segunda porción de los primeros parámetros de codificación, prediciendo una porción actual de los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) desde una cuarta porción previamente codificada de la señal de múltiples vistas (10), codificada en el flujo de datos (18) por el codificador antes de la codificación de la porción actual de los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) y corrigiendo un error de predicción de la predicción de la porción actual de los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) usando un segundo subconjunto (42<sub>2</sub>) de los primeros datos de corrección,

40 en el que los primeros parámetros de codificación definen una segmentación de los fotogramas del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>), y el codificador está **caracterizado porque** está configurado para, al codificar los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>), usar la segmentación de los fotogramas del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) como una pre-segmentación de los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) junto con una señalización de una desviación de la segmentación para los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) desde la segmentación de los fotogramas del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) mediante el flujo de datos.

50 13. Método de decodificación que comprende  
reconstruir (106<sub>v,1</sub>, 106<sub>d,1</sub>) una primera vista (12<sub>1</sub>) de una señal de múltiples vistas (10) desde un flujo de datos (18), de acuerdo con primeros parámetros de codificación (46<sub>1</sub>, 48<sub>1</sub>, 50<sub>1</sub>) obtenidos desde el flujo de datos (18), prediciendo una porción actual de la primera vista (12<sub>1</sub>) desde una primera porción previamente reconstruida de la señal de múltiples vistas (10), reconstruida desde el flujo de datos (18) por el decodificador antes de la  
55 reconstrucción de la porción actual de la primera vista (12<sub>1</sub>) y corrigiendo un error de predicción de la predicción de la porción actual de la primera vista (12<sub>1</sub>) usando primeros datos de corrección (42<sub>1,2</sub>) contenidos en el flujo de datos (18);

adoptar o predecir al menos parcialmente segundos parámetros de codificación desde los primeros parámetros de codificación; y

60 reconstruir (106<sub>v,2</sub>, 106<sub>d,2</sub>) una segunda vista (12<sub>2</sub>) de la señal de múltiples vistas (10) desde el flujo de datos (18), de acuerdo con los segundos parámetros de codificación, prediciendo una porción actual de la segunda vista desde una segunda porción previamente reconstruida de la señal de múltiples vistas (10), reconstruida desde el flujo de datos (18) por el decodificador antes de la reconstrucción de la porción actual de la segunda vista (12<sub>2</sub>) y corrigiendo un error de predicción de la predicción de la porción actual de la segunda vista (12<sub>2</sub>) usando segundos datos de

corrección (42<sub>3,4</sub>) contenidos en el flujo de datos (18), en el que cada una de la primera vista (12<sub>1</sub>) y la segunda vista (12<sub>2</sub>) comprende un vídeo (14<sub>1,2</sub>) capturado desde una respectiva posición de cámara, y datos de mapa de profundidad/disparidad asociados (16<sub>1,2</sub>).

en el que

5 reconstruir (106<sub>v,1</sub>) el vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) desde el flujo de datos (18) se realiza, de acuerdo con una primera porción (46<sub>1</sub>, 48<sub>1</sub>, 50<sub>1</sub>) de los primeros parámetros de codificación, prediciendo una porción actual del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) desde una tercera porción previamente reconstruida de la señal de múltiples vistas (10), reconstruida desde el flujo de datos por el decodificador antes de la reconstrucción de la porción actual del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) y corrigiendo un error de predicción de la predicción de la porción actual del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) usando un primer subconjunto (42<sub>1</sub>) de los primeros datos de corrección (42<sub>1,2</sub>) contenidos en el flujo de datos (18),

adoptar o predecir al menos parcialmente una segunda porción de los primeros parámetros de codificación desde la primera porción de los primeros parámetros de codificación,

15 reconstruir (106<sub>d,1</sub>) los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) desde el flujo de datos (18) se realiza, de acuerdo con la segunda porción de los primeros parámetros de codificación, prediciendo una porción actual de los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) desde una cuarta porción previamente reconstruida de la señal de múltiples vistas (10), reconstruida desde el flujo de datos (18) por el decodificador antes de la reconstrucción de la porción actual de los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) y corrigiendo un error de predicción de la predicción de la porción actual de los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) usando un segundo subconjunto (42<sub>2</sub>) de los primeros datos de corrección,

en el que los primeros parámetros de codificación definen una segmentación de los fotogramas del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>), y la reconstrucción de los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) está **caracterizada por** comprender usar la segmentación de los fotogramas del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) como una pre-segmentación de los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) junto con una señalización de una desviación de la segmentación para los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) desde la segmentación de los fotogramas del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) mediante el flujo de datos.

30 14. Método de codificación que comprende codificar una primera vista de una señal de múltiples vistas en un flujo de datos

determinando primeros parámetros de codificación,

35 de acuerdo con los primeros parámetros de codificación, prediciendo una porción actual de la primera vista desde una primera porción previamente codificada de la señal de múltiples vistas, codificada en el flujo de datos por el codificador antes de la codificación de la porción actual de la primera vista, y determinando un error de predicción de la predicción de la porción actual de la primera vista para obtener primeros datos de corrección, e insertar los primeros parámetros de codificación y los primeros datos de corrección en el flujo de datos; y

40 codificar una segunda vista de la señal de múltiples vistas en el flujo de datos, determinando segundos parámetros de codificación adoptando o prediciendo los segundos parámetros de codificación desde los primeros parámetros de codificación;

45 de acuerdo con los segundos parámetros de codificación, prediciendo una porción actual de la segunda vista desde una segunda porción previamente codificada de la señal de múltiples vistas, codificada en el flujo de datos por el codificador antes de la codificación de la porción actual de la segunda vista y determinando un error de predicción de la predicción de la porción actual de la segunda vista para obtener segundos datos de corrección contenidos en el flujo de datos, insertando los segundos datos de corrección en el flujo de datos,

50 en el que cada una de la primera vista (12<sub>1</sub>) y la segunda vista (12<sub>2</sub>) comprende un vídeo (14<sub>1,2</sub>) capturado desde una respectiva posición de cámara, y datos de mapa de profundidad/disparidad asociados (16<sub>1,2</sub>).

en el que

55 codificar (106<sub>v,1</sub>) el vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) en el flujo de datos (18) se realiza, de acuerdo con una primera porción (46<sub>1</sub>, 48<sub>1</sub>, 50<sub>1</sub>) de los primeros parámetros de codificación, prediciendo una porción actual del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) desde una tercera porción previamente codificada de la señal de múltiples vistas (10), codificada en el flujo de datos por el codificador antes de la codificación de la porción actual del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) y corrigiendo un error de predicción de la predicción de la porción actual del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) usando un primer subconjunto (42<sub>1</sub>) de los primeros datos de corrección (42<sub>1,2</sub>) e insertando el error de predicción en el flujo de datos (18),

adoptar o predecir al menos parcialmente una segunda porción de los primeros parámetros de codificación desde la primera porción de los primeros parámetros de codificación,

60 codificar los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) en el flujo de datos (18) se

- 5 realiza, de acuerdo con la segunda porción de los primeros parámetros de codificación, prediciendo una porción actual de los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) desde una cuarta porción previamente codificada de la señal de múltiples vistas (10), codificada en el flujo de datos (18) por el codificador antes de la codificación de la porción actual de los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) y corrigiendo un error de predicción de la predicción de la porción actual de los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) usando un segundo subconjunto (42<sub>2</sub>) de los primeros datos de corrección,
- 10 en el que los primeros parámetros de codificación definen una segmentación de los fotogramas del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>), y la codificación de los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>), está **caracterizada por** comprender usar la segmentación de los fotogramas del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) como una pre-segmentación de los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) junto con una señalización de una desviación de la segmentación para los datos de mapa de profundidad/disparidad (16<sub>1</sub>) de la primera vista (12<sub>1</sub>) desde la segmentación de los fotogramas del vídeo (14<sub>1</sub>) de la primera vista (12<sub>1</sub>) mediante el flujo de datos.
- 15 15. Programa informático que comprende instrucciones que, cuando el programa se ejecuta por un ordenador, provocan que el ordenador lleve a cabo las etapas del método de acuerdo con la reivindicación 13 o 14.

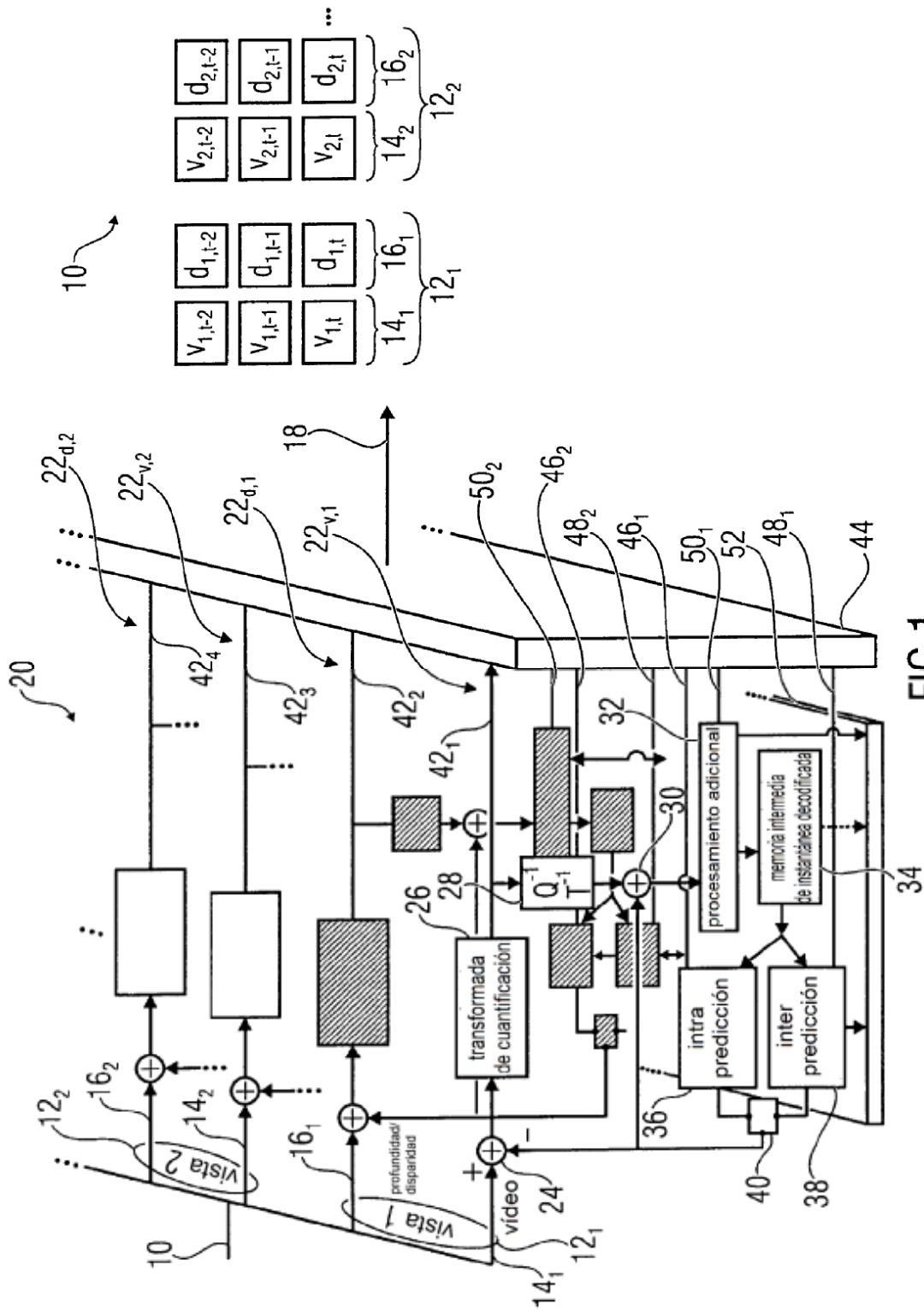


FIG 1



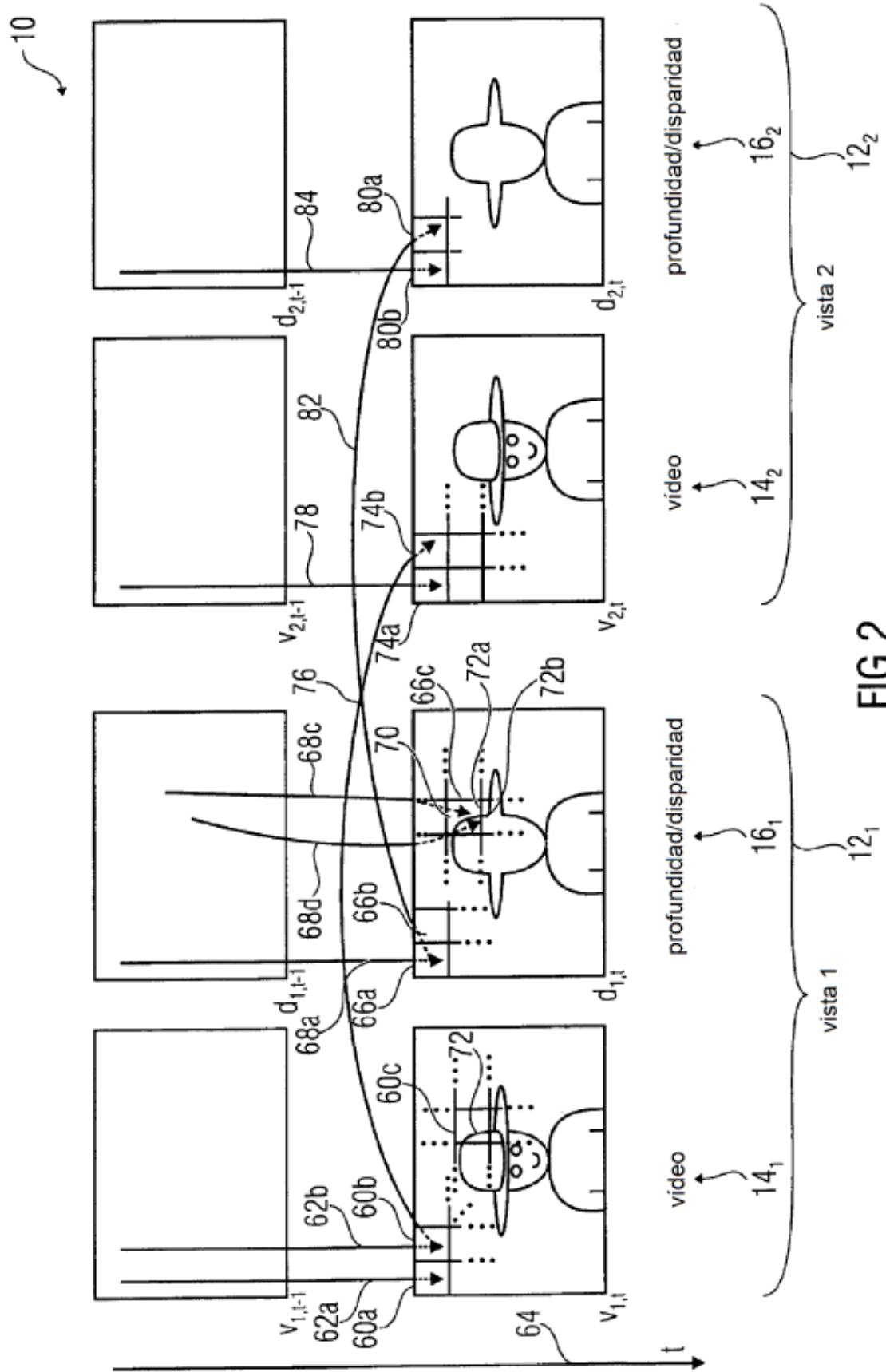


FIG 2

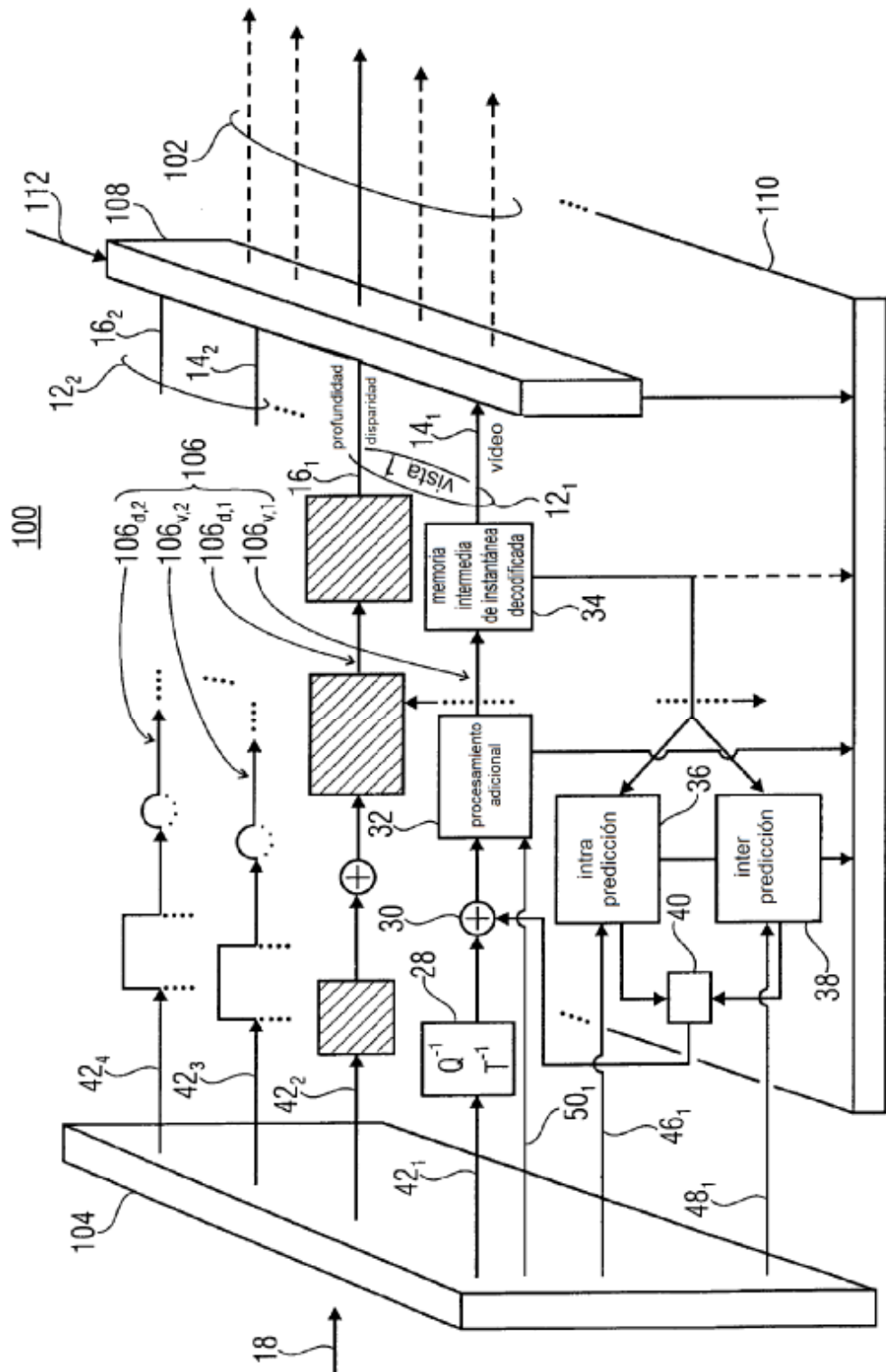


FIG 3

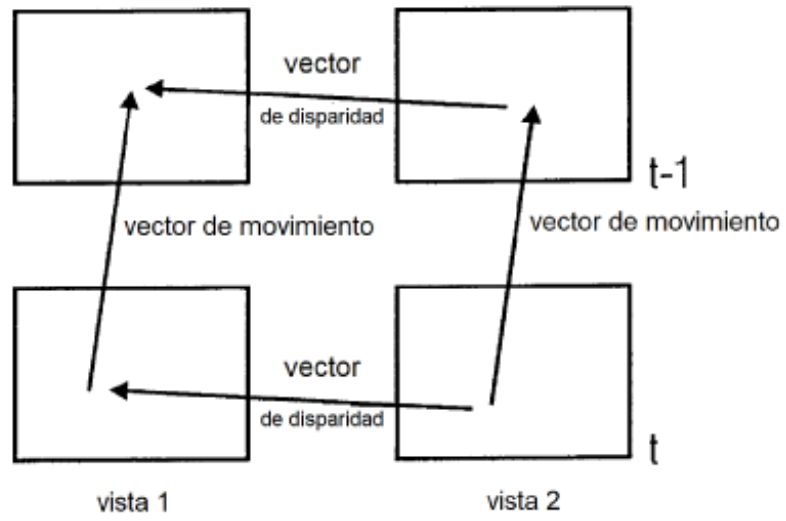


FIG 4

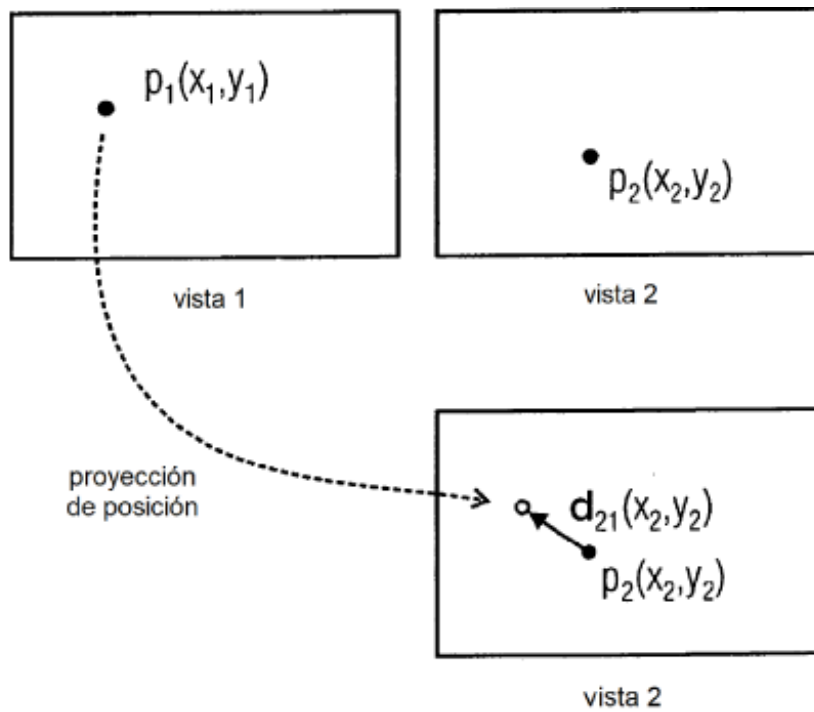


FIG 5

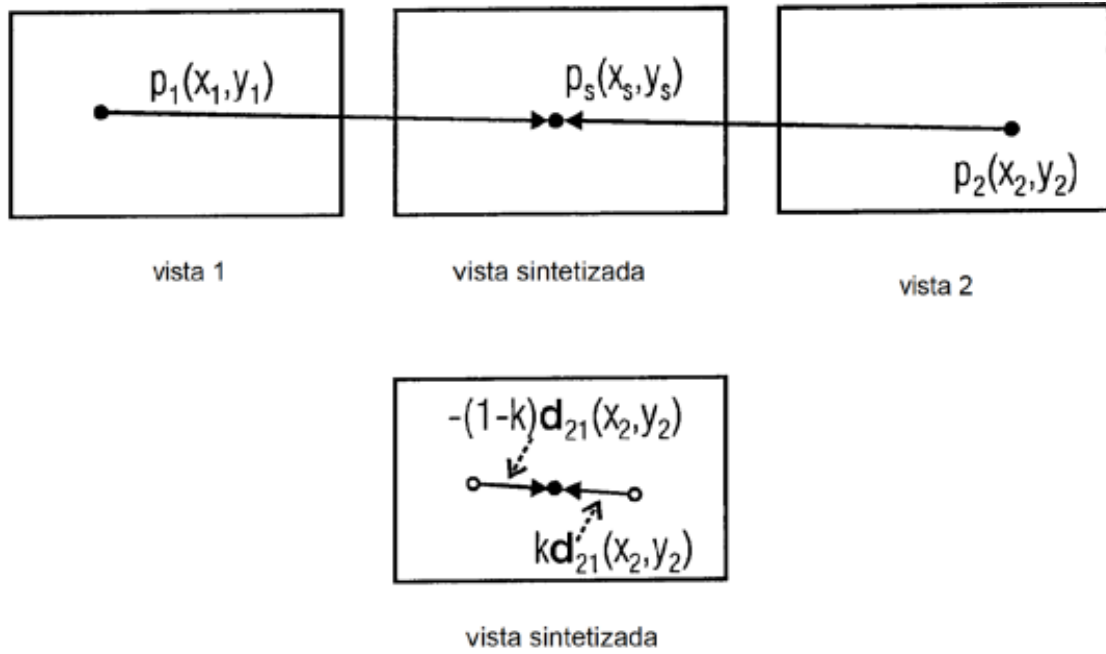


FIG 6

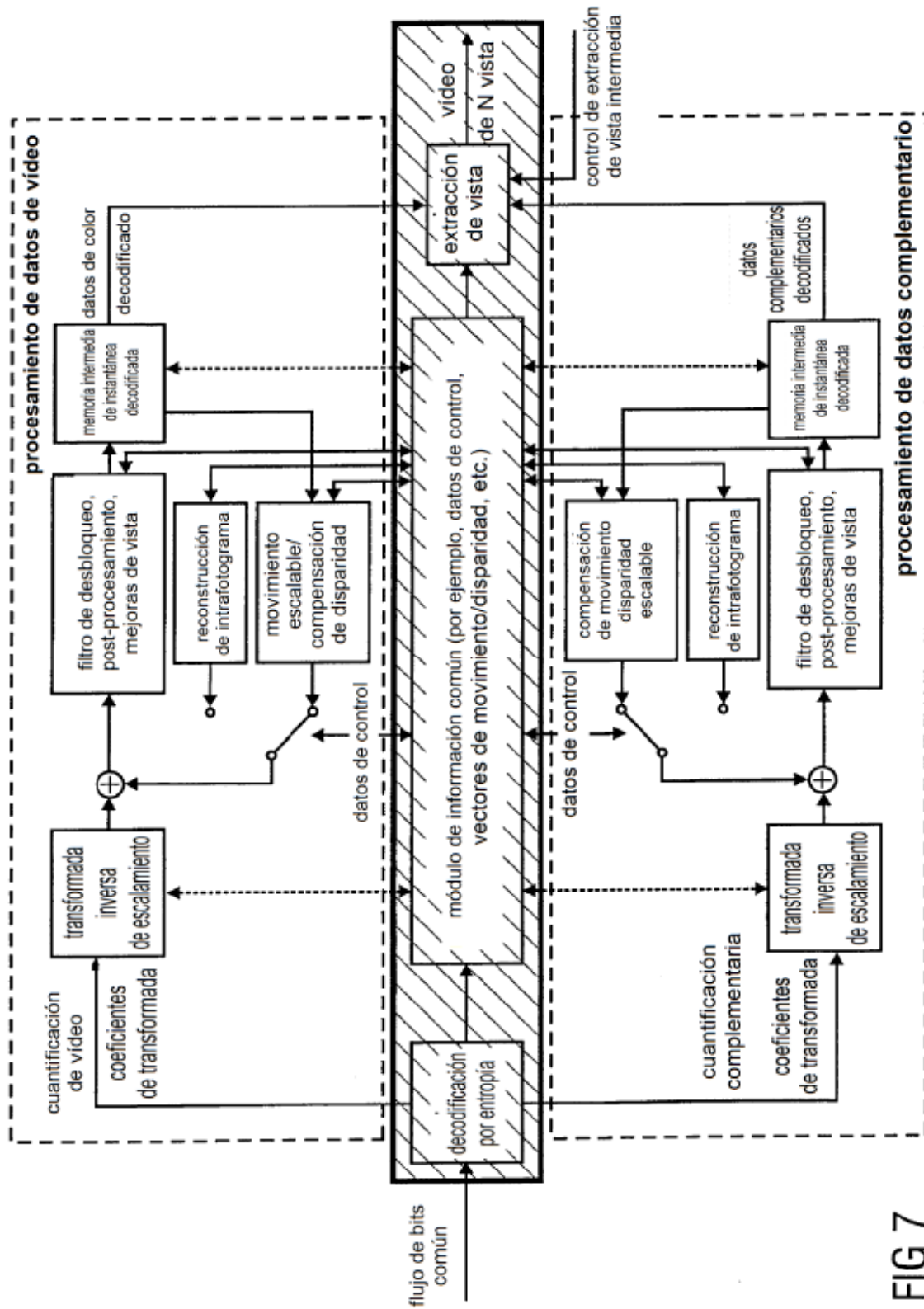


FIG 7

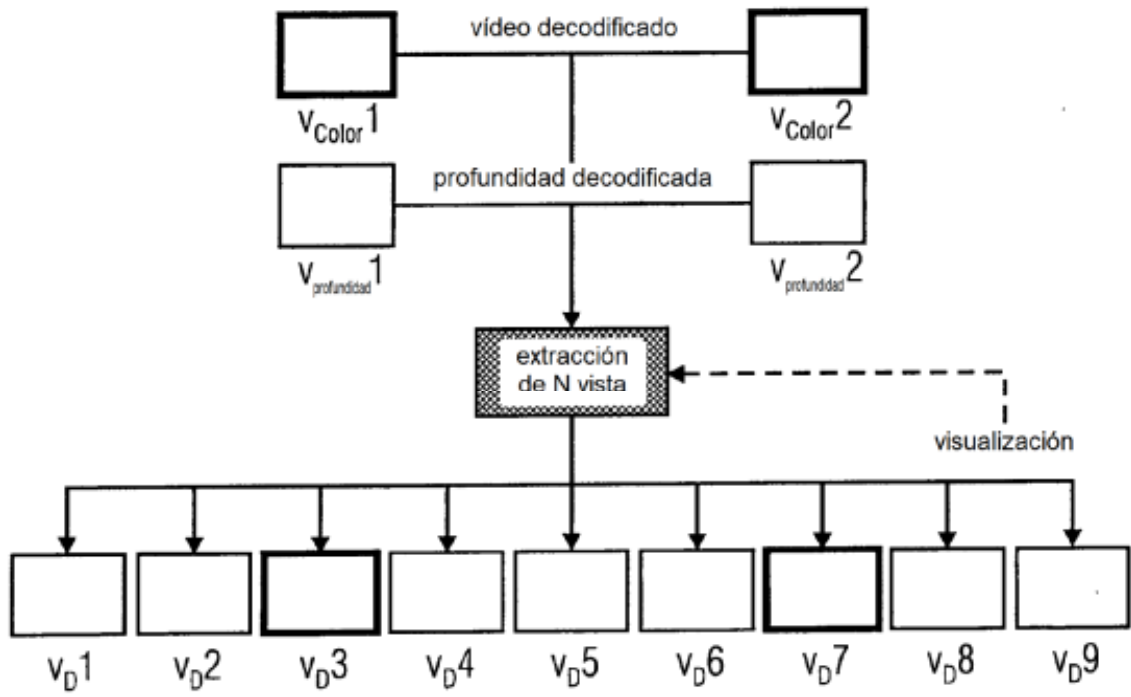


FIG 8