

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 602 100**

51 Int. Cl.:

H04N 19/117 (2014.01)

H04N 19/184 (2014.01)

H04N 19/80 (2014.01)

H04N 19/33 (2014.01)

H04N 19/36 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **16.04.2008** **E 14187185 (5)**

97 Fecha y número de publicación de la concesión europea: **21.09.2016** **EP 2835976**

54 Título: **Escalabilidad de profundidad de bits**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
17.02.2017

73 Titular/es:

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

72 Inventor/es:

WIEGAND, THOMAS y
WINKEN, MARTIN

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 602 100 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).

Escalabilidad de profundidad de bits

DESCRIPCIÓN

5 La presente invención está relacionada con la codificación de instantáneas y/o de vídeo, y en particular, a codificación de calidad escalable que posibilita la escalabilidad de profundidad de bits usando flujos de datos de calidad escalable.

10 El Equipo Mixto de Vídeo (JVT) del Grupo de Expertos en Imágenes en Movimiento (MPEG) de ISO/IEC y el Grupo de Expertos en Codificación de Vídeo (VCEG) de la ITU-T recientemente han finalizado una extensión escalable de la norma de codificación de vídeo del estado de la técnica H.264/AVC denominada Codificación de Vídeo Escalable (SVC). SVC soporta codificación temporal, espacial y escalable de SNR de secuencias de vídeo o cualquier combinación de las mismas.

15 H.264/AVC como se describe en el documento de la ITU-T Rec. & ISO/IEC 14496-10 AVC, "Advanced Video Coding for Generic Audiovisual Services", versión 3, 2005, especifica un códec de vídeo híbrido en el que las señales de predicción de macrobloques se generan en el dominio temporal mediante predicción de movimiento compensado, o en el dominio espacial mediante intra predicción, y ambas predicciones son seguidas por codificación residual. La
 20 codificación H.264/AVC sin la extensión de escalabilidad se denomina como codificación H.264/AVC de capa única. El rendimiento tasa-distorsión comparable a H.264/AVC de capa única significa que típicamente se consigue la misma calidad de reproducción visual a tasa de bits del 10 %. Dado lo anterior, se considera la escalabilidad como una funcionalidad para la eliminación de partes del flujo de bits mientras se consigue un rendimiento R-D en cualquier resolución espacial, temporal o de SNR soportada que es comparable a codificación H.264/AVC de capa
 25 única en esa resolución particular.

El diseño básico de la codificación de vídeo escalable (SVC) puede clasificarse como un códec de vídeo en capas. En cada capa, los conceptos básicos de predicción de movimiento compensado e intra predicción se emplean como en H.264/AVC. Sin embargo, se han integrado mecanismos de predicción entre capas adicionales para aprovechar
 30 la redundancia entre varias capas espaciales o de SNR. La escalabilidad de SNR básicamente se consigue mediante cuantificación residual, mientras que para escalabilidad espacial, se emplea una combinación de predicción de movimiento compensado y descomposición piramidal sobremuestreada. Se mantiene el enfoque de escalabilidad temporal de H.264/AVC.

35 En general, la estructura del codificador depende del espacio de escalabilidad que se requiere por una aplicación. Para ilustración, la Figura 8 muestra una estructura de codificador 900 típica con dos capas espaciales 902a, 902b. En cada capa, se emplea una estructura de predicción de movimiento compensado jerárquica independiente 904a, b con parámetros de movimiento de capa específica 906a, b. La redundancia entre capas consecutivas 902a, b se aprovecha por los conceptos de predicción entre capas 908 que incluye mecanismos de predicción para parámetros
 40 de movimiento 906a, b así como datos de textura 910a, b. Una representación de base 912a, b de las instantáneas de entrada 914a, b de cada capa 902a, b se obtiene por codificación de transformación 916a, b similar a la de H.264/AVC, las correspondientes unidades de NAL (NAL - Capa de Abstracción de Red) contienen información de movimiento y datos de textura; las unidades de NAL de la representación de base de la capa más baja, es decir 912a, son compatibles con H.264/AVC de capa única.

45 Los flujos de bits resultantes emitidos por la codificación de capa de base 916a, b y la codificación de textura de perfeccionamiento de SNR progresiva 918a, b de las respectivas capas 902a, b, respectivamente, se multiplexan por un multiplexor 920 para dar como resultado el flujo de bits escalable 922. Este flujo de bits 922 es escalable en tiempo, espacio y calidad de SNR.

50 Resumiendo, de acuerdo con la extensión escalable anterior de la Norma de Codificación de Vídeo H.264/AVC, se proporciona la escalabilidad temporal usando una estructura de predicción jerárquica. Para esta estructura de predicción jerárquica, una de las normas H.264/AVC de capa única puede usarse sin ningún cambio. Para escalabilidad espacial y de SNR, tienen que añadirse herramientas adicionales al H.264/MPEG4.AVC de capa única como se describe en la extensión de SVC de H.264/AVC. Pueden combinarse los tres tipos de escalabilidad para
 55 generar un flujo de bits que soporta un gran grado de escalabilidad combinada.

60 Surgen problemas cuando una señal de origen de vídeo tiene un rango dinámico diferente que el requerido por el decodificador o reproductor, respectivamente. En la norma de SVC actual anterior, las herramientas de escalabilidad únicamente se especifican para el caso en el que tanto la capa de base como la capa de mejora representan un origen de vídeo dado con la misma profundidad de bits de los conjuntos correspondientes de muestras de luma y/o croma. Por lo tanto, considerando diferentes decodificadores y reproductores, respectivamente, que requieren diferentes profundidades de bits, tendrían que proporcionarse varios flujos de codificación especializados para cada una de las profundidades de bits por separado. Sin embargo, en el sentido de tasa/distorsión, esto significa una tara

aumentada y eficacia reducida, respectivamente.

5 Ha habido ya propuestas para añadir una escalabilidad en términos de profundidad de bits a la norma SVC. Por ejemplo, Shan Liu et al. describe en el documento de entrada al JVT - en concreto JVT-X075 - la posibilidad de derivar una predicción entre capas a partir de una representación de profundidad de bits inferior de una capa de base mediante el uso de un mapeo de tono inverso de acuerdo con el que un valor de píxel predicho entre capas o inversamente mapeado en tono p' se calcula a partir de un valor de píxel de capa de base p_b por $p' = p_b \cdot escala + desplazamiento$ estableciendo que la predicción entre capas se realizaría en macro bloques o tamaños de bloque más pequeños. En JVT-Y067 Shan Liu, presenta resultados para este esquema de predicción entre capas. De manera similar, Andrew Segall et al., propone en JVT-X071 una predicción entre capas para escalabilidad de profundidad de bits de acuerdo con la que se usa una operación de ganancia más desplazamiento para el mapeo de tono inverso. Los parámetros de ganancia se indexan y transmiten en el flujo de bits de capa de mejora en una base bloque a bloque. La señalización de los factores de escala y factores de desplazamiento se consigue por una combinación de predicción y perfeccionamiento. Además, se describe que la sintaxis de alto nivel soporta granularidades más bastas que la transmisión en una base bloque a bloque. Se hace referencia también a Andrew Segall "Scalable Coding of High Dynamic Range Video" en el documento ICIP 2007, I-1 a I-4 y el documento JVT, JVT-X067 y JVT-W113, también procedentes de Andrew Segall.

20 Aunque las propuestas anteriormente mencionadas para usar un mapeo de tono inverso para obtener una predicción desde una capa de base de profundidad de bits inferior, eliminar alguna de la redundancia entre la información de profundidad de bits inferior y la información de profundidad de bits superior, sería favorable conseguir incluso una mejor eficacia al proporcionar un flujo de bits escalable de profundidad de bits de este tipo, especialmente en el sentido de rendimiento de tasa/distorsión.

25 Es el objeto de la presente invención proporcionar un esquema de codificación que posibilita una manera más eficaz de proporcionar una codificación de una instantánea o vídeo que es adecuada para diferentes profundidades de bits.

30 Este objeto se consigue mediante un codificador de acuerdo con la reivindicación 1, un decodificador de acuerdo con la reivindicación 11, un método de acuerdo con la reivindicación 22 o 23, o un flujo de datos de calidad escalable de acuerdo con la reivindicación 24.

35 La presente invención está basada en el hallazgo de que la eficacia de un flujo de datos escalable de profundidad de bits puede aumentarse cuando se obtiene una predicción entre capas mapeando muestras de la representación de los datos de origen de instantánea o de vídeo con una primera profundidad de bits de muestra de instantánea a partir de un primer rango dinámico que corresponde a la primera profundidad de bits de muestra de instantánea a un segundo rango dinámico mayor que el primer rango dinámico y que corresponde a una segunda profundidad de bits de muestra de instantánea mayor que es superior que la primera profundidad de bits de muestra de instantánea mediante el uso de una o más funciones de mapeo global que son constantes dentro de los datos de origen de instantánea o de vídeo o que varían a una primera granularidad, y una función de mapeo local que modifica localmente la una o más funciones de mapeo global y que varía a una segunda granularidad menor que la primera granularidad, formando el flujo de datos de calidad escalable basándose en la función de mapeo local de manera que la función de mapeo local puede derivarse a partir del flujo de datos de calidad escalable. Aunque la provisión de una o más funciones de mapeo global además de una función de mapeo local que, a su vez, modifica localmente la una o más funciones de mapeo global, a primera vista aumenta la cantidad de información secundaria dentro del flujo de datos escalable, este aumento es más que lo recompensado por el hecho de que esta sub-división en una función de mapeo global, por una parte, y una función de mapeo local, por la otra, posibilita que la función de mapeo local y sus parámetros para su parametrización puedan ser y son pequeños, por lo tanto, codificables de una manera altamente eficaz. La función de mapeo global puede codificarse dentro del flujo de datos de calidad escalable y, puesto que el mismo es constante dentro de los datos de origen de instantánea o de vídeo o varía a una granularidad mayor, la tara o flexibilidad para definir esta función de mapeo global puede aumentarse de modo que esta función de mapeo global puede adaptarse con precisión a las estadísticas promedio de los datos de origen de instantánea o de vídeo, reduciendo adicionalmente de esta manera la magnitud de la función de mapeo local.

55 A continuación, se describen realizaciones preferidas de la presente solicitud con referencia a las figuras. En particular, como se muestra en

- La Figura 1 un diagrama de bloques de un codificador de vídeo de acuerdo con una realización de la presente invención;
- 60 La Figura 2 un diagrama de bloques de un decodificador de vídeo de acuerdo con una realización de la presente invención;
- La Figura 3 un diagrama de flujo para una posible implementación de un modo de operación del módulo de predicción 134 de la Figura 1 de acuerdo con una realización;

- La Figura 4 un diagrama esquemático de un vídeo y su subdivisión en secuencias de instantánea, instantáneas, pares de macrobloques, macrobloques y bloques de transformación de acuerdo con una realización;
- 5 La Figura 5 un diagrama esquemático de una porción de una instantánea, subdividida en bloques de acuerdo con la granularidad precisa subyacente a la función de mapeo/adaptación local ilustrando de manera concurrente un esquema de codificación predictivo para codificar los parámetros de función de mapeo/adaptación local de acuerdo con una realización;
- 10 La Figura 6 un diagrama de flujo para ilustrar un proceso de mapeo de tono inverso en el codificador de acuerdo con una realización;
- La Figura 7 un diagrama de flujo de un proceso de mapeo de tono inverso en el decodificador que corresponde al de la Figura 6, de acuerdo con una realización; y
- 15 La Figura 8 un diagrama de bloques de una estructura de codificador convencional para codificación de vídeo escalable.

La Figura 1 muestra un codificador 100 que comprende un medio de codificación de base 102, un medio de predicción 104, un medio de codificación residual 106 y un medio de combinación 108 así como una entrada 110 y una salida. El codificador 100 de la Figura 1 es un codificador de vídeo que recibe una señal de vídeo de alta calidad en la entrada 110 y que emite un flujo de bits de calidad escalable en la salida 112. El medio de codificación de base 102 codifica los datos en la entrada 110 en un flujo de datos de codificación de base que representa el contenido de esta señal de vídeo en la entrada 110 con una profundidad de bits de muestra de instantánea reducida y, opcionalmente, una resolución espacial reducida en comparación con la señal de entrada en la entrada 110. El medio de predicción 104 está adaptado para, basándose en el flujo de datos de codificación de base emitido por el medio de codificación de base 102, proporcionar una señal de predicción con profundidad de bits de muestra de instantánea total o aumentada y, opcionalmente, resolución espacial total o aumentada para la señal de vídeo en la entrada 110. Un restador 114 también comprendido por el codificador 100 forma una predicción residual de la señal de predicción proporcionada por el medio 104 con relación a la señal de entrada de alta calidad en la entrada 110, estando codificada la señal residual por el medio de codificación residual 106 en un flujo de datos de capa de mejora de calidad. El medio de combinación 108 combina el flujo de datos de codificación de base del medio de codificación de base 102 y el flujo de datos de capa de mejora de calidad emitido por el medio de codificación residual 106 para formar un flujo de datos de calidad escalable 112 en la salida 112. El medio de escalabilidad de calidad del que está compuesto una parte del flujo de datos en la salida 112 es autónomo en que posibilita una reconstrucción de la señal de vídeo 110 con la profundidad de bits reducida y, opcionalmente, la resolución espacial reducida sin ninguna información adicional y olvidando el resto del flujo de datos 112, por una parte, y una parte adicional que posibilita, en combinación con la primera parte, una reconstrucción de la señal de vídeo en la entrada 110 en la profundidad de bits original y siendo la resolución espacial original superior a la profundidad de bits y/o resolución espacial de la primera parte.

Después de haber descrito bastante en general la estructura y funcionalidad del codificador 100, se describe su estructura interna en más detalle a continuación. En particular, el medio de codificación de base 102 comprende un módulo de conversión descendente 116, un restador 118, un módulo de transformación 120 y un módulo de cuantificación 122 conectados en serie, en el orden mencionado, entre la entrada 110, y el medio de combinación 108 y el medio de predicción 104, respectivamente. El módulo de conversión descendente 116 es para reducir la profundidad de bits de las muestras de instantáneas de y, opcionalmente, la resolución espacial de las instantáneas de la señal de vídeo en la entrada 110. En otras palabras, el módulo de conversión descendente 116 convierte de manera descendente e irreversible la señal de vídeo de entrada de alta calidad en la entrada 110 a una señal de vídeo de calidad base. Como se describirá en más detalle a continuación, esta conversión descendente puede incluir reducir la profundidad de bits de las muestras de señal, es decir valores de píxel, en la señal de vídeo en la entrada 110 usando cualquier esquema de mapeo de tono, tal como redondeo de los valores de muestra, sub-muestreo de los componentes de croma en caso de que se proporcione la señal de vídeo en forma de componentes de luma más croma, filtrado de la señal de entrada en la entrada 110, tal como mediante una conversión de RGB a YCbCr, o cualquier combinación de las mismas. Se presentan más detalles sobre posibles mecanismos de predicción a continuación. En particular, es posible que el módulo de conversión descendente 116 use diferentes esquemas de conversión descendente para cada instantánea de la señal de vídeo o secuencia de instantáneas introducidas en la entrada 110 o use el mismo esquema para todas las instantáneas. Esto se analiza también en más detalle a continuación.

El restador 118, el módulo de transformación 120 y el módulo de cuantificación 122 cooperan para codificar la señal de calidad de base emitida por el módulo de conversión descendente 116 mediante el uso de, por ejemplo, un esquema de codificación de vídeo no escalable, tal como H.264/AVC. De acuerdo con el ejemplo de la Figura 1, el restador 118, el módulo de transformación 120 y el módulo de cuantificación 122 cooperan con un filtro de bucle de

predicción opcional 124, un módulo predictor 126, un módulo de transformación inverso 128 y un sumador 130 comprendidos comúnmente por el medio de codificación de base 102 y el medio de predicción 104 para formar la parte de reducción de irrelevancia de un codificador híbrido que codifica la señal de vídeo de calidad de base emitida por el módulo de conversión descendente 116 por predicción basada en compensación de movimiento y seguido de la compresión de la predicción residual. En particular, el restador 118 resta de una instantánea o macrobloque actual de la señal de vídeo de calidad de base una instantánea predicha o porción de macrobloque predicha reconstruida a partir de instantáneas previamente codificadas de la señal de vídeo de calidad de base mediante el uso de, por ejemplo, compensación de movimiento. El módulo de transformación 120 aplica una transformación en la predicción residual, tal como una DCT, FFT o transformada de ondícula. La señal residual transformada puede representar una representación espectral y sus coeficientes de transformación se cuantifican de manera irreversible en el módulo de cuantificación 122. La señal residual cuantificada resultante representa el residuo del flujo de datos de codificación de base emitido por el medio de codificación de base 102.

Además del filtro de bucle de predicción opcional 124 y el módulo predictor 126, el módulo de transformación inverso 128, y el sumador 130, el medio de predicción 104 comprende un filtro opcional para reducir artefactos de codificación 132 y un módulo de predicción 134. El módulo de transformación inverso 128, el sumador 130, el filtro de bucle de predicción opcional 124 y el módulo predictor 126 cooperan para reconstruir la señal de vídeo con una profundidad de bits reducida y, opcionalmente, una resolución espacial reducida, como se define por el módulo de conversión descendente 116. En otras palabras, crean una profundidad de bits baja y, opcionalmente, señal de vídeo de resolución espacial baja para el filtro opcional 132 que representa una representación de baja calidad de la señal de origen en la entrada 110 pudiéndose reconstruir también en el lado del decodificador. En particular, el módulo de transformación inverso 128 y el sumador 130 están conectados en serie entre el módulo de cuantificación 122 y el filtro opcional 132, mientras que el filtro de bucle de predicción opcional 124 y el módulo de predicción 126 están conectados en serie, en el orden mencionado, entre una salida del sumador 130 así como una entrada adicional del sumador 130. La salida del módulo predictor 126 está también conectada a una entrada de inversión del restador 118. El filtro opcional 132 está conectado entre la salida del sumador 130 y el módulo de predicción 134, que, a su vez, está conectado entre la salida del filtro opcional 132 y la inversión introducida al restador 114.

El módulo de transformación inverso 128 transforma a la inversa las instantáneas residuales codificadas de base emitidas por el medio de codificación de base 102 para conseguir baja profundidad de bits y, opcionalmente, instantáneas residuales de baja resolución espacial. Por consiguiente, el módulo de transformación inverso 128 realiza una transformación inversa que es una inversión de la transformación y cuantificación realizada por los módulos 120 y 122. Como alternativa, puede proporcionarse por separado un módulo de des-cuantificación en el lado de entrada del módulo de transformación inverso 128. El sumador 130 añade una predicción a las instantáneas residuales reconstruidas, basándose en la predicción en instantáneas previamente reconstruidas de la señal de vídeo. En particular, el sumador 130 emite una señal de vídeo reconstruida con una profundidad de bits reducida y, opcionalmente, resolución espacial reducida. Estas instantáneas reconstruidas se filtran por el filtro de bucle 124 para reducir artefactos, por ejemplo, y se usan posteriormente por el módulo predictor 126 para predecir la instantánea actualmente a reconstruir por medio de, por ejemplo, compensación de movimiento, a partir de instantáneas previamente reconstruidas. La señal de calidad de base obtenida de esta manera en la salida del sumador 130 se usa por la conexión serie del filtro opcional 132 y el módulo de predicción 134 para obtener una predicción de la señal de entrada de alta calidad en la entrada 110, la última predicción a usarse para formar la señal de mejora de alta calidad en la salida del medio de codificación residual 106. Esto se describe en mayor detalle a continuación.

En particular, la señal de baja calidad obtenida desde el sumador 130 se filtra opcionalmente por el filtro opcional 132 para reducir artefactos de codificación. Los filtros 124 y 132 pueden incluso operar de la misma manera y, por lo tanto, aunque los filtros 124 y 132 se muestran por separado en la Figura 1, ambos pueden sustituirse por solamente un filtro dispuesto entre la salida del sumador 130 y la entrada del módulo de predicción 126 y el módulo de predicción 134, respectivamente. Posteriormente, la señal de vídeo de baja calidad se usa por el módulo de predicción 134 para formar una señal de predicción para la señal de vídeo de alta calidad recibida en la entrada no de inversión del sumador 114 que está conectada a la entrada 110. Este proceso de formar la predicción de alta calidad puede incluir mapear las muestras de instantánea de señal de calidad de base decodificada mediante el uso de una función de mapeo combinada como se describe en más detalle a continuación, usando el respectivo valor de las muestras de señal de calidad de base para indexar una tabla de consulta que contiene los correspondientes valores de muestra de alta calidad, usando el valor de la muestra de señal de calidad de base para un proceso de interpolación para obtener el correspondiente valor de muestra de alta calidad, sobremuestreo de los componentes de croma, filtrado de la señal de calidad de base mediante el uso de, por ejemplo, conversión de YCbCr a RGB, o cualquier combinación de las mismas. Se describen otros ejemplos a continuación.

Por ejemplo, el módulo de predicción 134 puede mapear las muestras de la señal de vídeo de calidad de base desde un primer rango dinámico a un segundo rango dinámico que es superior al primer rango dinámico y, opcionalmente, mediante el uso de un filtro de interpolación espacial, interpolar espacialmente muestras de la señal de vídeo de calidad de base para aumentar la resolución espacial para corresponder con la resolución espacial de la

señal de vídeo en la entrada 110. De una manera similar a la descripción anterior del módulo de conversión descendente 116, es posible usar un proceso de predicción diferente para diferentes instantáneas de la secuencia de señal de vídeo de calidad de base así como usar el mismo proceso de predicción para todas las instantáneas.

5 El restador 114 resta la predicción de alta calidad recibida desde el módulo de predicción 134 de la señal de vídeo de alta calidad recibida desde la entrada 110 para emitir una señal de predicción residual de alta calidad, es decir con la profundidad de bits original y, opcionalmente, resolución espacial al medio de codificación residual 106. En el medio de codificación residual 106, la diferencia entre la señal de entrada de alta calidad original y la predicción derivada de la señal de calidad de base decodificada se codifica de manera ejemplar usando un esquema de
10 codificación de compresión tal como se especifica, por ejemplo, en H.264/AVC. Para este fin, el medio de codificación residual 106 de la Figura 1 comprende de manera ejemplar un módulo de transformación 136, un módulo de cuantificación 138 y un módulo de codificación por entropía 140 conectados en serie entre una salida del restador 114 y el medio de combinación 108 en el orden mencionado. El módulo de transformación 136 transforma la señal residual de las instantáneas de la misma, respectivamente, en un dominio de transformación o dominio
15 espectral, respectivamente, donde los componentes espectrales se cuantifican mediante el módulo de cuantificación 138 y codificándose por entropía los valores de transformación cuantificados mediante el módulo de codificación por entropía 140. El resultado de la codificación por entropía representa el flujo de datos de capa de mejora de alta calidad emitido por el medio de codificación residual 106. Si los módulos 136 a 140 implementan una codificación H.264/AVC, que soporta transformaciones con un tamaño de 4x4 o 8x8 muestras para codificar el contenido de
20 luma, el tamaño de transformación para transformar el componente de luma de la señal residual del restador 114 en el módulo de transformación 136 puede elegirse de manera arbitraria para cada macrobloque y no tiene que ser necesariamente el mismo que se usa para codificar la señal de calidad de base en el módulo de transformación 120. Para codificar los componentes de croma, la norma H.264/AVC, no proporciona elección. Cuando se cuantifican los coeficientes de transformación en el módulo de cuantificación 138, puede usarse el mismo esquema de
25 cuantificación que en H.264/AVC, que significa que el tamaño de paso del cuantificador puede controlarse por un parámetro de cuantificación QP, que puede tomar valores desde -6*(profundidad de bits de componente de señal de vídeo de alta calidad-8) a 51. El QP usado para codificar el macrobloque de representación de calidad de base en el módulo de cuantificación 122 y el QP usado para codificar el macrobloque de mejora de alta calidad en el módulo de cuantificación 138 no tienen que ser el mismo.

30 El medio de combinación 108 comprende un módulo de codificación por entropía 142 y el multiplexor 144. El módulo de codificación por entropía 142 está conectado entre una salida del módulo de cuantificación 122 y una primera entrada del multiplexor 144, mientras que una segunda entrada del multiplexor 144 está conectada a una salida del módulo de codificación por entropía 140. La salida del multiplexor 144 representa la salida 112 del codificador 100.

35 El módulo de codificación por entropía 142 codifica por entropía los valores de transformación cuantificados emitidos por el módulo de cuantificación 122 para formar un flujo de datos de capa de calidad de base desde el flujo de datos de codificación de base emitido mediante el módulo de cuantificación 122. Por lo tanto, como se ha mencionado anteriormente, los módulos 118, 120, 122, 124, 126, 128, 130 y 142 pueden diseñarse para cooperar de acuerdo
40 con H.264/AVC, y representan juntos un codificador híbrido con el codificador por entropía 142 que realiza una compresión sin pérdidas de la predicción residual cuantificada.

45 El multiplexor 144 recibe tanto el flujo de datos de capa de calidad de base como el flujo de datos de capa de alta calidad y los pone juntos para formar el flujo de datos de calidad escalable.

50 Como ya se ha descrito anteriormente, y como se muestra en la Figura 3, la manera en la que el módulo de predicción 134 realiza la predicción desde la señal de calidad de base reconstruida al dominio de señal de alta calidad puede incluir una expansión de profundidad de bits de muestra, también denominada un mapeo de tono inverso 150 y, opcionalmente, una operación de sobremuestreo espacial, es decir una operación de filtrado de sobremuestreo 152 caso en el que las señales de base y de alta calidad son de diferente resolución espacial. El orden en el que el módulo de predicción 134 realiza el mapeo de tono inverso 150 y la operación de sobremuestreo espacial opcional 152 puede fijarse, y, por consiguiente, ser conocido previamente para tanto los lados del codificador como del decodificador, o puede elegirse de manera adaptativa en una base bloque a bloque o instantánea a instantánea o alguna otra granularidad, caso en el que el módulo de predicción 134 señala
55 información en el orden entre las etapas 150 y 152 usadas, a alguna entidad, tal como el codificador 106, para que se introduzca como información secundaria en el flujo de bits 112 para señalizarse al lado de decodificador como parte de la información secundaria. La adaptación del orden entre las etapas 150 y 152 se ilustra en la Figura 3 por el uso de una flecha discontinua de doble punta 154 y la granularidad a la que el orden puede elegirse de manera adaptativa puede señalizarse también e incluso variarse dentro del vídeo.

60 Al realizar el mapeo de tono inverso 150, el módulo de predicción 134 usa dos partes, en concreto una o más funciones de mapeo de tono inverso global y una adaptación local de las mismas. En general, la una o más funciones de mapeo de tono inverso global están especializadas para contabilizar las características promedio generales de la secuencia de instantáneas del vídeo y, por consiguiente, del mapeo de tono que se ha aplicado

inicialmente a la señal de vídeo de entrada de alta calidad para obtener la señal de vídeo de calidad de base en el módulo de conversión descendente 116. En comparación a la misma, la adaptación local deberá contabilizar las desviaciones individuales desde el modelo de mapeo de tono inverso global para los bloques individuales de las instantáneas del vídeo.

5 Para ilustrar esto, la Figura 4 muestra una porción de un vídeo 160 ejemplar que consiste en cuatro instantáneas consecutivas 162a a 162d del vídeo 160. En otras palabras, el vídeo 160 comprende una secuencia de instantáneas 162 entre las que se muestran ejemplarmente cuatro en la Figura 4. El vídeo 160 puede dividirse en secuencias no solapantes de instantáneas consecutivas para las que se transmiten parámetros globales o elementos de sintaxis dentro del flujo de datos 112. Para fines de ilustración únicamente, se supone que las cuatro instantáneas consecutivas 162a a 162d mostradas en la Figura 4 deberán formar tal secuencia 164 de instantáneas. Cada instantánea, a su vez, se subdivide en una pluralidad de macrobloques 166 como se ilustra en la esquina izquierda inferior de la instantánea 162d. Un macrobloque es un contenedor dentro del que los coeficientes de transformación junto con otros elementos de sintaxis de control que pertenecen a la clase de codificación del macrobloque se transmiten dentro del flujo de bits 112. Un par 168 de macrobloques 166 cubre una porción continua de la respectiva instantánea 162d. Dependiendo de un modo de par de macrobloques del respectivo par de macrobloques 168, el macrobloque superior 162 de este par 168 cubre cualquiera de las muestras de la mitad superior del par de macrobloques 168 o las muestras de cada línea con número impar dentro del par de macrobloques 168, relacionando el macrobloque inferior a las otras muestras en el mismo, respectivamente. Cada macrobloque 160, a su vez, puede subdividirse en bloques de transformación como se ilustra en 170, formando estos bloques de transformación las bases de bloque en las que el módulo de transformación 120 realiza la transformación y el módulo de transformación inversa 128 realiza la transformación inversa.

Haciendo referencia de vuelta de nuevo a la función de mapeo de tono inverso global recién mencionada, el módulo de predicción 134 puede configurarse para usar una o más tales funciones de mapeo de tono inverso global de manera constante para la totalidad del vídeo 160 o, como alternativa, para una subporción del mismo, tal como la secuencia 164 de instantáneas consecutivas o una misma instantánea 162. Las últimas opciones implicarían que el módulo de predicción 134 varíe la función de mapeo de tono inverso global a una granularidad que corresponde a un tamaño de secuencia de instantánea o un tamaño de instantánea. Se proporcionan ejemplos para funciones de mapeo de tono inverso global a continuación. Si el módulo de predicción 134 adapta la función de mapeo de tono inverso global a las estadísticas del vídeo 160, el módulo de predicción 134 emite información al módulo de codificación por entropía 140 o al multiplexor 144 de modo que el flujo de bits 112 contiene información sobre la función o funciones de mapeo de tono inverso global y la variación de las mismas dentro del vídeo 160. En caso de que la una o más funciones de mapeo de tono inverso global usadas por el módulo de predicción 134 se apliquen de manera constante a la totalidad del vídeo 160, lo mismo puede ser conocido previamente para el decodificador o transmitirse como información secundaria dentro del flujo de bits 112.

A una granularidad incluso menor, una función de mapeo de tono inverso local usada por el módulo de predicción 134 varía dentro del vídeo 160. Por ejemplo, tal función de mapeo de tono inverso local varía a una granularidad menor que un tamaño de instantánea tal como, por ejemplo, el tamaño de un macrobloque, un par de macrobloques o un tamaño de bloque de transformación.

Para tanto la función o funciones de mapeo de tono inverso global como la función de mapeo de tono inverso local, la granularidad a la que varía la respectiva función o a la que se definen las funciones en el flujo de bits 112 puede variarse dentro del vídeo 160. La variación de la granularidad, a su vez, puede señalizarse dentro del flujo de bits 112.

Durante el mapeo de tono inverso 150, el módulo de predicción 134 mapea una muestra predeterminada del vídeo 160 desde la profundidad de bits de calidad de base a la profundidad de bits de alta calidad mediante el uso de una combinación de una función de mapeo de tono inverso global que se aplica a la respectiva instantánea y a la función de mapeo de tono inverso local como se define en el respectivo bloque al que pertenece la muestra predeterminada.

Por ejemplo, la combinación puede ser una combinación aritmética y, en particular, una adición. El módulo de predicción puede configurarse para obtener un valor de muestra de alta profundidad de bits predicho S_{high} desde el correspondiente valor de muestra de baja profundidad de bits reconstruido S_{low} mediante el uso de $S_{high} = f_k(S_{low}) + m \cdot S_{low} + n$.

En esta fórmula, la función f_k representa un operador de mapeo de tono inverso global en el que el índice k selecciona qué operador de mapeo de tono inverso global se elige en caso de que se use más de un único esquema o más de una función de mapeo de tono inverso global. La parte restante de esta fórmula constituye la adaptación local o la función de mapeo de tono inverso local siendo n un valor de compensación y siendo m un factor de escala. Los valores de k , m y n pueden especificarse en una base bloque a bloque dentro del flujo de bits 112. En otras palabras, el flujo de bits 112 podría posibilitar revelar las ternas $\{k, m, n\}$ para todos los bloques del vídeo 160 con un tamaño de bloque de estos bloques que depende de la granularidad de la adaptación local de la función de mapeo

de tono inverso global con esta granularidad, a su vez, variando posiblemente dentro del video 160.

Los siguientes mecanismos de mapeo pueden usarse para el proceso de predicción siempre que se refieran a la función de mapeo de tiempo inverso global $f(x)$. Por ejemplo, puede usarse mapeo lineal a nivel de pieza donde puede especificarse un número arbitrario de puntos de interpolación. Por ejemplo, para la muestra de calidad de base con valor x y dos puntos de interpolación dados (x_n, y_n) y (x_{n+1}, y_{n+1}) la muestra de predicción correspondiente y se obtiene mediante el módulo 134 de acuerdo con la siguiente fórmula

$$f(x) = y_n + \frac{x - x_n}{x_{n+1} - x_n} (y_{n+1} - y_n)$$

Esta interpolación lineal puede realizarse con poca complejidad computacional usando únicamente un desplazamiento de bit en lugar de operaciones de división si $x_{n+1} - x_n$ se restringe para que sea una potencia de dos.

Un mecanismo de mapeo global posible adicional representa un mapeo de tabla de consulta en el que, por medio de los valores de muestra de calidad de base, se realiza una consulta de tabla en una tabla de consulta en la que se especifica para cada posible valor de muestra de calidad de base siempre que la función de mapeo de tono inverso global haga referencia al valor de muestra de predicción global (x) correspondiente. La tabla de consulta puede proporcionarse en el lado del decodificador como información secundaria o puede conocerse en el lado del decodificador por defecto.

Además, escalar con un desplazamiento constante puede usarse para el mapeo global. De acuerdo con esta alternativa, para conseguir la muestra de predicción global de alta calidad correspondiente (x) que tiene profundidad de bits superior, el módulo 134 multiplica las muestras de calidad de base x por un factor constante 2^{M-N-K} , y posteriormente se añade un desplazamiento constante $2^{M-1} - 2^{M-1-K}$, de acuerdo con, por ejemplo, una de las siguientes fórmulas:

$$f(x) = 2^{M-N-K} x + 2^{M-1} - 2^{M-1-K}$$

o

$$f(x) = \min(2^{M-N-K} x + 2^{M-1} - 2^{M-1-K}, 2^M - 1),$$

respectivamente, en las que M es la profundidad de bits de la señal de alta calidad y N es la profundidad de bits de la señal de calidad de base.

Mediante esta medida, el rango dinámico de baja calidad $[0; 2^N - 1]$ se mapea al segundo rango dinámico $[0; 2^M - 1]$ de una manera de acuerdo con la que los valores mapeados de x se distribuyen de una manera centralizada con respecto al rango dinámico posible $[0; 2^M - 1]$ de la calidad superior dentro de una extensión que se determina por K . El valor de K podría ser un valor entero o valor real, y podría transmitirse como información secundaria al decodificador dentro de, por ejemplo, el flujo de datos de calidad escalable de modo que en el decodificador algún medio de predicción pueda actuar de la misma manera que el módulo de predicción 134 como se describirá a continuación. Puede usarse una operación de redondeo para obtener los valores con valor entero $f(x)$.

Otra posibilidad para el escalamiento global es escalar con desplazamiento variable: las muestras de calidad de base x se multiplican por un factor constante, y posteriormente se añade un desplazamiento variable, de acuerdo con, por ejemplo, una de las siguientes fórmulas:

$$f(x) = 2^{M-N-K} x + D$$

o

$$f(x) = \min(2^{M-N-K} x + D, 2^M - 1)$$

Mediante esta medida, el rango dinámico de baja calidad se mapea globalmente al segundo rango dinámico de una manera de acuerdo con la que los valores mapeados de x se distribuyen dentro de una porción del posible rango dinámico de las muestras de alta calidad, la extensión de la cual se determina por K , y el desplazamiento del cual con respecto al límite inferior se determina por D . D puede ser entero o real. El resultado $f(x)$ representa un valor de muestra de instantánea mapeada globalmente de la señal de predicción de alta profundidad de bits. Los valores de

K y D podrían transmitirse como información secundaria al decodificador dentro de, por ejemplo, el flujo de datos de calidad escalable. De nuevo, puede usarse una operación de redondeo para obtener valores con valor entero $f(x)$, cumpliéndose lo último también para los otros ejemplos dados en la presente solicitud para los mapeos de profundidad de bits globales sin establecerlo de manera explícita repetitivamente.

Una posibilidad incluso adicional para el mapeo global es escalar con superposición: las muestras de predicción de alta profundidad de bits mapeadas globalmente $f(x)$ se obtienen desde la respectiva muestra de calidad de base x de acuerdo con, por ejemplo, una de las siguientes fórmulas, donde $\text{floor}(a)$ redondea uno por debajo del entero más cercano:

$$f(x) = \text{floor}(2^{M-N}x + 2^{M-2N}x)$$

$$f(x) = \min(\text{floor}(2^{M-N}x + 2^{M-2N}x), 2^M - 1)$$

Las posibilidades recién mencionadas pueden combinarse. Por ejemplo, puede usarse escalamiento global con superposición y desplazamiento constante: las muestras de predicción de alta profundidad de bits mapeadas globalmente $f(x)$ se obtienen de acuerdo con, por ejemplo, una de las siguientes fórmulas, donde $\text{floor}(a)$ redondea uno por debajo del entero más cercano:

$$f(x) = \text{floor}(2^{M-N-K}x + 2^{M-2N-K}x + 2^{M-1} - 2^{M-1-K})$$

$$f(x) = \min(\text{floor}(2^{M-N-K}x + 2^{M-2N-K}x + 2^{M-1} - 2^{M-1-K}), 2^M - 1)$$

El valor de K puede especificarse como información secundaria al decodificador.

De manera similar, puede usarse escalamiento global con superposición y desplazamiento variable: las muestras de predicción de alta profundidad de bits mapeadas globalmente $f(x)$ se obtienen de acuerdo con la siguiente fórmula, donde $\text{floor}(a)$ redondea uno por debajo del entero más cercano:

$$f(x) = \text{floor}(2^{M-N-K}x + 2^{M-2N-K}x + D)$$

$$f(x) = \min(\text{floor}(2^{M-N-K}x + 2^{M-2N-K}x + D), 2^M - 1)$$

Los valores de D y K pueden especificarse como información secundaria al decodificador.

Transfiriendo los ejemplos recién mencionados para la función de mapeo de tono inverso global a la Figura 4, los parámetros allí mencionados para definir la función de mapeo de tono inverso global, en concreto (x_1, y_1) , K y D , pueden ser conocidos para el decodificador, pueden transmitirse dentro del flujo de bits 112 con respecto al vídeo total 160 en caso de que la función de mapeo de tono inverso global sea constante dentro del vídeo 160 o estos parámetros se transmitan dentro del flujo de bits para diferentes porciones del mismo, tal como, por ejemplo, para secuencias de instantáneas 164 o instantáneas 162 dependiendo de la granularidad basta subyacente a la función de mapeo global. En caso de que el módulo de predicción 134 use más de una función de mapeo de tono inverso global, los parámetros anteriormente mencionados (x_1, y_1) , K y D pueden pensarse como que se proporcionan con un índice k con estos parámetros (x_1, y_1) , K_k y D_k que definen la k -ésima función de mapeo de tono inverso global f_k .

Por lo tanto, sería posible especificar el uso de diferentes mecanismos de mapeo de tono inverso global para cada bloque de cada instantánea señalizando un valor correspondiente k dentro del flujo de bits 112 así como, como alternativa, usando el mismo mecanismo para la secuencia completa del vídeo 160.

Además, es posible especificar diferentes mecanismos de mapeo global para los componentes de luma y croma de la señal de calidad de base para tener en cuenta que las estadísticas pueden ser diferentes, tales como su función de densidad de probabilidad.

Como ya se ha indicado anteriormente, el módulo de predicción 134 usa una combinación de la función de mapeo de tono inverso global y una función de mapeo de tono inverso local para adaptar localmente la global. En otras palabras, se realiza una adaptación local en cada muestra de predicción de alta profundidad de bits mapeada globalmente $f(x)$ obtenida de esta manera. La adaptación local de la función de mapeo de tono inverso global se realiza mediante el uso de una función de mapeo de tono inverso local que está localmente adaptada, por ejemplo, adaptando localmente algunos parámetros de la misma. En el ejemplo proporcionado anteriormente, estos parámetros son el factor de escalamiento m y el valor de desplazamiento n . Los factores de escalamiento m y el valor de desplazamiento n pueden especificarse en una base bloque a bloque donde un bloque puede corresponder

a un tamaño de bloque de transformación que es, por ejemplo, en caso de H.264/AVC, 4x4 o 8x8 muestras, o un tamaño de macrobloque que es, en caso de H.264/AVC, por ejemplo, 16x16 muestras. Qué significado de "bloque" se está usando realmente puede fijarse, y, por lo tanto, conocerse previamente para tanto el codificador como el decodificador o puede elegirse de manera adaptativa por instantánea o por secuencia mediante el módulo de predicción 134, caso en el que se ha de señalar al decodificador como parte de información secundaria dentro del flujo de bits 112. En el caso de H.264/AVC, el conjunto de parámetros de secuencia y/o el conjunto de parámetros de instantánea podrían usarse para este fin. Adicionalmente, podría especificarse en la información secundaria que cualquiera del factor de escalamiento m o el valor de desplazamiento n o ambos valores se establecen igual a cero para una secuencia de vídeo completa o para un conjunto bien definido de instantáneas dentro de la secuencia de vídeo 160. En cualquier caso el factor de escalamiento m o el valor de desplazamiento n o ambos de ellos se especifican en una base bloque a bloque para una instantánea dada, para reducir la tasa de bits requerida para codificación de estos valores, únicamente pueden codificarse los valores de diferencia Δm , Δn a valores predichos correspondientes m_{pred} , n_{pred} de manera que los valores reales para m , n son obtenibles como sigue: $m = n_{pred} + \Delta m$, $n = n_{pred} + \Delta n$.

En otras palabras y como se ilustra en la Figura 5, que muestra una porción de una instantánea 162 dividida en bloques 170 que forman la base de la granularidad de la adaptación local de la función de mapeo de tono inverso global, el módulo de predicción 134, el módulo de codificación por entropía 140 y el multiplexor 144 están configurados de manera que los parámetros para adaptar localmente la función de mapeo de tono inverso global, en concreto el factor de escalamiento m y el valor de desplazamiento n como se usan para los bloques individuales 170 no están directamente codificados en el flujo de bits 112, sino meramente como una predicción residual a una predicción obtenida desde factores de escalamiento y valores de desplazamiento de bloques vecinos 170. Por lo tanto, $\{\Delta m, \Delta n, k\}$ se transmiten para cada bloque 170 caso en el que se use más de una función de mapeo de tono inverso global como se muestra en la Figura 5. Suponiendo, por ejemplo, que el módulo de predicción 134 usó parámetros $m_{i,j}$ y $n_{i,j}$ para mapeo de tono inverso las muestras de un cierto bloque i,j , en concreto el del medio de la Figura 5. A continuación, el módulo de predicción 134 está configurado para calcular valores de predicción $m_{i,j,pred}$ y $n_{i,j,pred}$ desde los factores de escalamiento y valores de desplazamiento de bloques vecinos 170, tal como, por ejemplo, $m_{i,j-1}$ y $n_{i,j-1}$. En este caso, el módulo de predicción 134 provocaría la diferencia entre los parámetros reales $m_{i,j}$ y $n_{i,j}$ y los predichos $m_{i,j} - n_{i,j,pred}$ y $n_{i,j} - n_{i,j,pred}$ para que se inserten en el flujo de bits 112. Estas diferencias se indican en la Figura 5 como $\Delta m_{i,j}$ y $\Delta n_{i,j}$ indicando los índices i, j el j -ésimo bloque desde la parte superior y el i -ésimo bloque desde el lado izquierdo de la instantánea 162, por ejemplo. Como alternativa, la predicción del factor de escalamiento m y el valor de desplazamiento n puede derivarse de los bloques ya transmitidos en lugar de los bloques vecinos de la misma instantánea. Por ejemplo, los valores de predicción pueden derivarse del bloque 170 de la instantánea precedente que radica en la misma o una localización espacial correspondiente. En particular, el valor predicho puede ser

- un valor fijo, que se transmite en la información secundaria o ya se conoce para tanto el codificador como el decodificador,
- el valor de la variable correspondiente y el bloque precedente,
- el valor de la mediana de las variables correspondientes en los bloques vecinos,
- el valor medio de las variables correspondientes en los bloques vecinos,
- un valor interpolado o extrapolado derivado de los valores de variables correspondientes en los bloques vecinos.

Cuál de estos mecanismos de predicción se usa realmente para un bloque particular puede conocerse para tanto el codificador como el decodificador o puede depender de valores de m, n en los propios bloques vecinos, si los hubiera.

Dentro de la señal de mejora de alta calidad codificada emitida por el módulo de codificación por entropía 140, podría transmitirse la siguiente información para cada macrobloque en caso de que los módulos 136, 138 y 140 implementen una codificación conforme a H.264/AVC. Podría incluirse una información de patrón de bloque codificado (CBP) que indica cuál de los cuatro bloques de transformación de luma 8x8 dentro del macrobloque y cuál de los bloques de transformación de croma asociados del macrobloque puede contener coeficientes de transformación distintos de cero. Si no hubiera coeficientes de transformación distintos de cero, no se transmitiría información adicional para el macrobloque particular. Además la información podría relacionar el tamaño de transformación usado para codificar el componente de luma, es decir el tamaño de los bloques de transformación en los que el macrobloque que consiste en 16x16 muestras de luma se transforma en el módulo de transformación 136, es decir en bloques de transformación de 4x4 o de 8x8. Además, el flujo de datos de capa de mejora de alta calidad podría incluir el parámetro de cuantificación QP usado en el módulo de cuantificación 138 para controlar el tamaño de paso del cuantificador. Además, los coeficientes de transformación cuantificados, es decir los niveles de coeficientes de transformación, podrían incluirse para cada macrobloque en el flujo de datos de capa de mejora de alta calidad emitido por el módulo de codificación por entropía 140.

Además de la información anterior, la siguiente información debería contenerse dentro del flujo de datos 112. Por ejemplo, en caso de que se use más de un esquema de mapeo de tono inverso global para la instantánea actual,

también tiene que transmitirse el correspondiente valor de índice k para cada bloque de la señal de mejora de alta calidad. Siempre que se haga referencia a la adaptación local, las variables Δm y Δn pueden señalizarse para cada bloque de aquellas instantáneas donde la transmisión de los valores correspondientes se indique en la información secundaria por medio de, por ejemplo, el conjunto de parámetros de secuencia y/o el conjunto de parámetros de instantánea en H.264/AVC. Para las tres nuevas variables k, Δm y Δn , tienen que introducirse nuevos elementos de sintaxis con unos esquemas de binarización correspondientes. Puede usarse una binarización unaria sencilla para preparar las tres variables para el esquema de codificación aritmético binario usado en H.264/AVC. Puesto que las tres variables tienen típicamente pequeñas magnitudes, el esquema de binarización unario sencillo es muy adecuado. Puesto que Δm y Δn son valores enteros con signo, podrían convertirse a valores sin signo como se describe en la siguiente Tabla:

valor con signo de $\Delta m, \Delta n$	valor sin signo (a binarizar)
0	0
1	1
2	-1
3	2
4	-2
5	3
6	-3
k	$(-1)^{k+1} \text{Ceil}(k/2)$

Por lo tanto, resumiendo alguna de las realizaciones anteriores, el medio de predicción 134 puede realizar las siguientes etapas durante su modo de operación. En particular, como se muestra en la Figura 6, el medio de predicción 134 establece, en la etapa 180, una o más función o funciones de mapeo global $f(x)$ o, en caso de más de una función de mapeo global, $f_k(x)$ indicando k el valor de índice correspondiente que apunta a la respectiva función de mapeo global. Como se ha descrito anteriormente, la una o más función o funciones de mapeo global pueden establecerse o definirse para que sean constantes a lo largo del vídeo 160 o pueden definirse o establecerse en una granularidad basta tal como, por ejemplo, una granularidad del tamaño de una instantánea 162 o una secuencia 164 de instantáneas. La una o más función o funciones de mapeo global pueden definirse como se ha descrito anteriormente. En general, la función o funciones de mapeo global pueden ser una función no trivial, es decir desigual a $f(x)=\text{constante}$, y pueden ser especialmente una función no lineal. En cualquier caso, cuando se usa cualquiera de los ejemplos anteriores para una función de mapeo global, la etapa 180 da como resultado un parámetro de función de mapeo global correspondiente, tal como K,D o (x_n, y_n) estableciéndose n E (1, ... N) para cada sección en la que se subdivide el vídeo de acuerdo con la granularidad basta.

Además, el módulo de predicción 134 establece una función de mapeo/adaptación local, tal como la anteriormente indicada, en concreto una función lineal de acuerdo con $m \cdot x + n$. Sin embargo, otra función de mapeo/adaptación local es también factible tal como una función constante parametrizándose simplemente por el uso de un valor de desplazamiento. El ajuste 182 se realiza a una granularidad más precisa tal como, por ejemplo, la granularidad de un tamaño menor que una instantánea tal como un macrobloque, bloque de transformación o par de macrobloques o incluso un segmento dentro de una instantánea en el que un segmento es un subconjunto de macrobloques o pares de macrobloques de una instantánea. Por lo tanto, la etapa 182 da como resultado, en el caso de la realización anterior para una función de mapeo/adaptación local, un par de valores Δm y Δn que se establecen o definen para cada uno de los bloques en los que se subdividen las instantáneas 162 del vídeo de acuerdo con la granularidad más precisa.

Opcionalmente, en concreto en el caso de que se use más de una función de mapeo global en la etapa 180, el medio de predicción 134 establece, en la etapa 184, el índice k para cada bloque de la granularidad precisa como se usa en la etapa 182.

Aunque sería posible que el módulo de predicción 134 estableciera la una o más función o funciones de mapeo global en la etapa 180 solamente dependiendo de la función de mapeo usada por el módulo de conversión descendente 116 para reducir la profundidad de bits de muestra de las muestras de instantánea original, tal como, por ejemplo, usando la función de mapeo inversa a la misma como una función de mapeo global en la etapa 180, debería observarse que es posible también que el medio de predicción 134 realice todos los ajustes dentro de las etapas 180, 182 y 184 de manera que se extreme un cierto criterio de optimización, tal como la proporción de

tasa/distorsión del flujo de bits resultante 112 tal como maximizado o minimizado. Mediante esta medida, tanto la función o funciones de mapeo global como la adaptación local a las mismas, en concreto la función de mapeo/adaptación local está adaptada a las estadísticas de tono de valor de muestra del vídeo determinando el mejor compromiso entre la tara de codificar la información secundaria necesaria para el mapeo global por una parte,
 5 y conseguir la mejor adaptación de la función de mapeo global a las estadísticas de tono necesitando meramente una pequeña función de mapeo/adaptación local por otra.

Al realizar el mapeo de tono inverso real en la etapa 186, el módulo de predicción 134 usa, para cada muestra de la señal de baja calidad reconstruida, una combinación de una función de mapeo global en caso de que solo haya una
 10 función de mapeo global o una de las funciones de mapeo global en caso de que haya más de una por una parte, y la función de mapeo/adaptación local por otra parte, ambas como se definen en el bloque o la sección a la que pertenece la muestra actual. En el ejemplo anterior, la combinación ha sido una adición. Sin embargo, sería posible también cualquier combinación aritmética. Por ejemplo, la combinación podría ser una aplicación en serie de ambas funciones.

Además, para informar al lado del decodificador acerca de la función de mapeo combinada usada para el mapeo de tono inverso, el módulo de predicción 134 provoca que se proporcione al menos información sobre la función de mapeo/adaptación local en el lado del decodificador mediante el flujo de bits 112. Por ejemplo, el módulo de predicción 134 provoca, en la etapa 188, que se codifiquen los valores de m y n en el flujo de bits 112 para cada
 20 bloque de la granularidad precisa. Como se ha descrito anteriormente, la codificación en la etapa 188 puede ser una codificación predictiva de acuerdo con la que los residuos de predicción de m y n se codifican en el flujo de bits en lugar de los propios valores reales, derivando la predicción de los valores que se derivan de los valores de m y n en los bloques vecinos 170 o de un bloque correspondiente en una instantánea anterior. En otras palabras, puede usarse una predicción local o temporal junto con codificación residual para codificar el parámetro m y n .

De manera similar, para el caso de que se use más de una función de mapeo global en la etapa 180, el módulo de predicción 134 puede provocar que se codifique el índice k en el flujo de bits para cada bloque en la etapa 190. Además, el módulo de predicción 134 puede provocar que se codifique la información en $f(x)$ o $f_k(x)$ en el flujo de bits en caso de que la misma información no sea conocida a priori para el decodificador. Esto se hace en la etapa 192.
 30 Además, el módulo de predicción 134 puede provocar que se codifique la información sobre la granularidad y el cambio de la granularidad dentro del vídeo para la función de mapeo/adaptación local y/o la función o funciones de mapeo global en el flujo de bits en la etapa 194.

Debería observarse que todas las etapas 180 a 194 no es necesario que se realicen en el orden mencionado. Incluso una realización secuencial estricta de estas etapas no es necesaria. En su lugar, las etapas 180 a 194 se muestran en un orden secuencial meramente para fines de ilustración y estas etapas se realizarán preferentemente de una manera solapante.

Aunque no se establece explícitamente en la descripción anterior, se observa que la información secundaria generada en las etapas 190 a 194 puede introducirse en la señal de capa de mejora de alta calidad o en la porción de alta calidad del flujo de bits 112 en lugar de en la porción de calidad de base que procede del módulo de codificación por entropía 142.

Después de haber descrito una realización para un codificador, con respecto a la Figura 2, se describe una realización de un decodificador. El decodificador de la Figura 2 se indica por el signo de referencia 200 y comprende un medio de demultiplexación 202, un medio de decodificación de base 204, un medio de predicción 206, un medio de decodificación residual 208 y un medio de reconstrucción 210 así como una entrada 212, una primera salida 214 y una segunda salida 216. El decodificador 200 recibe, en su entrada 212, el flujo de datos de calidad escalable, que se ha emitido, por ejemplo, por el codificador 100 de la Figura 1. Como se ha descrito anteriormente, la escalabilidad de la calidad puede hacer referencia a la profundidad de bits y, opcionalmente, a la reducción espacial. En otras palabras, el flujo de datos en la entrada 212 puede tener una parte independiente que es usable de manera aislada para reconstruir la señal de vídeo con una profundidad de bits reducida y, opcionalmente, resolución espacial reducida, así como una parte adicional que, en combinación con la primera parte, posibilita reconstruir la señal de vídeo con una profundidad de bits superior y, opcionalmente, resolución espacial superior. La señal de vídeo de reconstrucción de calidad inferior se emite en la salida 216, mientras que la señal de vídeo de reconstrucción superior se emite en la salida 214.

El medio de demultiplexación 202 divide el flujo de datos de calidad escalable entrante en la entrada 212 en el flujo de datos de codificación de base y el flujo de datos de capa de mejora de alta calidad, ambos de los cuales se han mencionado con respecto a la Figura 1. El medio de decodificación de base 204 es para decodificar el flujo de datos de codificación de base en la representación de calidad de base de la señal de vídeo, que es directamente, como es el caso en el ejemplo de la Figura 2, o indirectamente mediante un filtro de reducción de artefactos (no mostrado), opcionalmente emisible en la salida 216. Basándose en la señal de vídeo de representación de calidad de base, el medio de predicción 206 forma una señal de predicción que tiene la profundidad de bits de muestra de instantánea

y/o la resolución de muestreo de croma aumentadas. El medio de decodificación 208 decodifica el flujo de datos de capa de mejora para obtener la predicción residual que tiene la profundidad de bits aumentada y, opcionalmente, resolución espacial aumentada. El medio de reconstrucción 210 obtiene la señal de vídeo de alta calidad desde la predicción y la predicción residual y emite la misma a la salida 214 mediante un filtro de reducción de artefacto opcional.

Internamente, el medio de demultiplexación 202 comprende un demultiplexor 218 y un módulo de decodificación por entropía 220. Una entrada del demultiplexor 218 está conectada a la entrada 212 y una primera salida del demultiplexor 218 está conectada al medio de decodificación residual 208. El módulo de decodificación por entropía 220 está conectado entre otra salida del demultiplexor 218 y el medio de decodificación de base 204. El demultiplexor 218 divide el flujo de datos de calidad escalable en el flujo de datos de capa de base y el flujo de datos de capa de mejora como han sido introducidos por separado en el multiplexor 144, como se ha descrito anteriormente. El módulo de decodificación por entropía 220 realiza, por ejemplo, una decodificación Huffman o algoritmo de decodificación aritmética para obtener los niveles de coeficientes de transformación, vectores de movimiento, información de tamaño de transformación y otros elementos de sintaxis necesarios para derivar la representación de base de la señal de vídeo desde los mismos. En la salida del módulo de decodificación por entropía 220, da como resultado el flujo de datos de codificación de base.

El medio de decodificación de base 204 comprende un módulo de transformación inverso 222, un sumador 224, un filtro de bucle opcional 226 y un módulo predictor 228. Los módulos 222 a 228 del medio de decodificación de base 204 corresponden, con respecto a funcionalidad e inter-conexión, a los elementos 124 a 130 de la Figura 1. Para ser más precisos, el módulo de transformación inverso 222 y el sumador 224 están conectados en serie en el orden mencionado entre el medio de demultiplexación 202 por una parte y el medio de predicción 206 y la salida de calidad de base, respectivamente, por otra parte, y el filtro de bucle opcional 226 y el módulo predictor 228 están conectados en serie en el orden mencionado entre la salida del sumador 224 y otra entrada del sumador 224. Mediante esta medida, el sumador 224 emite la señal de vídeo de representación de base con la profundidad de bits reducida y, opcionalmente, la resolución espacial reducida que puede recibirse desde el exterior en la salida 216.

El medio de predicción 206 comprende un filtro de reducción de artefactos opcional 230 y un módulo de información de predicción 232, ambos módulos funcionando de una manera síncrona con relación a los elementos 132 y 134 de la Figura 1. En otras palabras, el filtro de reducción de artefactos opcional 230 filtra opcionalmente la señal de vídeo de calidad de base para reducir artefactos en la misma y el módulo de formación de predicción 232 recupera instantáneas predichas con profundidades de bits aumentadas y, opcionalmente, resolución espacial aumentada de una manera ya descrita anteriormente con respecto al módulo de predicción 134. Es decir, el módulo de información de predicción 232 puede mapear, por medio de información secundaria contenida en el flujo de datos de calidad escalable, las muestras de instantáneas entrantes a un rango dinámico superior y, opcionalmente, aplicar un filtro de interpolación espacial al contenido de las instantáneas para aumentar la resolución espacial.

El medio de decodificación residual 208 comprende un módulo de decodificación por entropía 234 y un módulo de transformación inverso 236, que están conectados en serie entre el demultiplexor 218 y el medio de reconstrucción 210 en el orden recién mencionado. El módulo de decodificación por entropía 234 y el módulo de transformación inverso 236 cooperan para invertir la codificación realizada por los módulos 136, 138 y 140 de la Figura 1. En particular, el módulo de decodificación por entropía 234 realiza, por ejemplo, un algoritmo de decodificación Huffman o de decodificación aritmética para obtener elementos de sintaxis que comprenden, entre otros, niveles de coeficientes de transformación, que se transforman a la inversa, mediante el módulo de transformación inverso 236, para obtener una señal de predicción residual o una secuencia de instantáneas residuales. Además, el módulo de decodificación por entropía 234 revela la información secundaria generada en las etapas 190 a 194 en el lado del codificador de modo que el módulo de formación de predicción 232 puede emular el procedimiento de mapeo inverso realizado en el lado del codificador mediante el módulo de predicción 134, como ya se ha indicado anteriormente.

Similar a la Figura 6, que hace referencia al codificador, la Figura 7 muestra en más detalle el modo de operación del módulo de formación de predicción 232 y, parcialmente, el módulo de decodificación por entropía 234. Como se muestra en la misma, el proceso de obtener la predicción desde la señal de capa de base reconstruida en el lado del decodificador empieza con una cooperación del demultiplexor 218 y el decodificador por entropía 234 para derivar la información de granularidad codificada en la etapa 194 en la etapa 280, derivar la información sobre la función o funciones de mapeo global que se han codificado en la etapa 192 en la etapa 282, derivar el valor de índice k para cada bloque de la granularidad precisa como ha sido codificado en la etapa 190 en la etapa 284, y derivar el parámetro de la función de mapeo/adaptación local m y n para cada bloque de granularidad precisa como ha sido codificado en la etapa 188 en la etapa 286. Como se ilustra por medio de las líneas discontinuas, las etapas 280 a 284 son opcionales con su aplicación dependiendo de la realización actualmente usada.

En la etapa 288, el módulo de formación de predicción 232 realiza el mapeo de tono inverso basándose en la

información obtenida en las etapas 280 a 286, emulando de esta manera exactamente el mapeo de tono inverso que se ha realizado en el lado del codificador en la etapa 186. Similar a la descripción con respecto a la etapa 188, la derivación de los parámetros m , n puede comprender una decodificación predictiva donde se deriva un valor residual de predicción desde la parte de alta calidad del flujo de datos que entra al demultiplexor 218 mediante el uso de, por ejemplo, codificación por entropía según se realiza por el decodificador por entropía 234, y obtener los valores reales de m y n añadiendo estos valores de predicción residual a un valor derivado por predicción local y/o temporal.

El medio de reconstrucción 210 comprende un sumador 238 las entradas del cual están conectadas a la salida del módulo de información de predicción 232, y a la salida del módulo de transformación inverso 236, respectivamente. El sumador 238 añade la predicción residual y la señal de predicción para obtener la señal de vídeo de alta calidad que tiene la profundidad de bits aumentada y, opcionalmente, resolución espacial aumentada que se alimenta mediante un filtro de reducción de artefactos opcional 240 a la salida 214.

Por lo tanto, como puede derivarse a partir de la Figura 2, un decodificador de calidad de base puede reconstruir una señal de vídeo de calidad de base a partir del flujo de datos de calidad escalable en la entrada 212 y puede, para hacer esto, no incluir los elementos 230, 232, 238, 234, 236 y 240. Por otra parte, un decodificador de alta calidad puede no incluir la salida 216.

En otras palabras, en el proceso de decodificación, la decodificación de la representación de la calidad de base es sencilla. Para la decodificación de la señal de alta calidad, en primer lugar ha de decodificarse la señal de calidad de base, que se realiza mediante los módulos 218 a 228. Posteriormente, el proceso de predicción anteriormente descrito con respecto al módulo 232 y al módulo opcional 230 se emplea usando la representación de base decodificada. Los coeficientes de transformación cuantificados de la señal de mejora de alta calidad se escalan y transforman a la inversa mediante el módulo de transformación inverso 236, por ejemplo, como se especifica en H.264/AVC para obtener las muestras de señal residuales o de diferencia, que se añaden a la predicción derivada de las muestras de representación de base decodificadas mediante el módulo de predicción 232. Como una etapa final en el proceso de decodificación de la señal de vídeo de alta calidad a emitirse en la salida 214, puede emplearse un filtro opcional para eliminar o reducir artefactos de codificación visualmente molestos. Se ha de observar que el bucle de predicción de movimiento compensado que implica los módulos 226 y 228 es completamente independiente usando únicamente la representación de calidad de base. Por lo tanto, la complejidad de decodificación es moderada y no hay necesidad de un filtro de interpolación, que opera en los datos de alta profundidad de bits y, opcionalmente, de imagen de alta resolución espacial en el proceso de predicción de movimiento compensado del módulo de predictor 228.

Con respecto a las realizaciones anteriores, debería mencionarse que los filtros de reducción de artefactos 132 y 230 son opcionales y podrían eliminarse. Lo mismo se aplica para los filtros de bucle 124 y 226, respectivamente, y el filtro 240. Por ejemplo, con respecto a la Figura 1, ha de observarse que los filtros 124 y 132 pueden sustituirse por solamente un filtro común tal como, por ejemplo, un filtro de desbloqueo, la salida del cual está conectada tanto a la entrada del módulo de predicción de movimiento 126 así como la entrada del módulo de mapeo de tono inverso 134. De manera similar, los filtros 226 y 230 pueden sustituirse por un filtro común tal como, un filtro de desbloqueo, la salida del cual está conectada a la salida 216, la entrada del módulo de formación de predicción 232 y la entrada del predictor 228. Además, por motivos de completitud, se observa que los módulos de predicción 228 y 126, respectivamente, no predicen necesariamente de manera temporal muestras dentro de los macrobloques de una instantánea actual. En su lugar, puede usarse también una predicción espacial o intra-predicción que usa muestras de la misma instantánea. En particular, el tipo de predicción puede elegirse en una base de macrobloque a macrobloque, por ejemplo, o alguna otra granularidad. Además, la presente invención no está restringida a codificación de vídeo. En su lugar, la descripción anterior es también aplicable a codificación de imagen fija. Por consiguiente, el bucle de predicción de movimiento compensado que implica los elementos 118, 128, 130, 126 y 124 y los elementos 224, 228 y 226, respectivamente, puede eliminarse también. De manera similar, la codificación por entropía mencionada no necesita realizarse necesariamente.

Incluso más precisos, en las realizaciones anteriores, la codificación de capa de base 118-130, 142 se basó en la predicción de movimiento compensado basándose en una reconstrucción de instantáneas ya codificadas con pérdidas. En este caso, la reconstrucción del proceso de codificación de base puede verse también como una parte del proceso de formación de predicción de alta calidad como ya se ha realizado en la descripción anterior. Sin embargo, en caso de una codificación sin pérdidas de la representación de base, una reconstrucción no sería necesaria y la señal convertida de manera descendente podría reenviarse directamente a los medios 132, 134, respectivamente. En el caso de predicción no basada en compensación de movimiento en una codificación de capa de base con pérdidas, la reconstrucción para reconstruir la señal de calidad de base en el lado del codificador estaría especialmente especializada para la formación de predicción de alta calidad en 104. En otras palabras, la asociación anterior de los elementos 116-134 y 142 a los medios 102, 104 y 108, respectivamente, podría realizarse de otra manera. En particular, el módulo de codificación por entropía 142 podría verse como una parte del medio de codificación de base 102, comprendiendo simplemente el medio de predicción los módulos 132 y 134 y comprendiendo simplemente el medio de combinación 108 el multiplexor 144. Esta vista se correlaciona con la asociación módulo/medio usada en la Figura 2 en que el medio de predicción 206 no comprende la predicción

basada en compensación de movimiento. Adicionalmente, sin embargo, el medio de demultiplexación 202 podría verse como que no incluye el módulo de entropía 220 de modo que el medio de decodificación de base también comprende el módulo de decodificación por entropía 220. Sin embargo, ambas vistas conducen al mismo resultado de que la predicción en 104 se realiza basándose en una representación del material de origen con la profundidad de bits reducida y, opcionalmente, la resolución espacial reducida que se codifica de manera sin pérdidas y puede derivarse de manera sin pérdidas a partir del flujo de bits de calidad escalable y del flujo de datos de capa de base, respectivamente. De acuerdo con la vista subyacente a la Figura 1, la predicción 134 está basada en una reconstrucción del flujo de datos de codificación de base, mientras que en el caso de la vista alternativa, la reconstrucción empezaría desde una versión codificada intermedia o versión codificada a la mitad de la señal de calidad de base que pierde la codificación sin pérdidas de acuerdo con el módulo 142 para codificarse por completo en el flujo de datos de capa de base. En este sentido, debería observarse adicionalmente que la conversión descendente en el módulo 116 no tiene que realizarse por el codificador 100. En su lugar, el codificador 100 puede tener dos entradas, una para recibir la señal de alta calidad y la otra para recibir la versión convertida de manera descendente, desde el exterior.

En las realizaciones anteriormente descritas, la escalabilidad de calidad relacionó simplemente a la profundidad de bits y, opcionalmente, la resolución espacial. Sin embargo, las realizaciones anteriores pueden ampliarse fácilmente para incluir escalabilidad temporal, escalabilidad de formato de croma y escalabilidad de calidad granular.

Por consiguiente, las realizaciones anteriores de la presente invención forman un concepto para codificación escalable de contenido de instantánea o de vídeo con diferentes granularidades en términos de profundidad de bits de muestra y, opcionalmente, resolución espacial mediante el uso de mapeo de tono inverso localmente adaptativo. De acuerdo con realizaciones de la presente invención, tanto los procesos de predicción temporal y espacial como se especifican en la extensión de codificación de vídeo escalable H.264/AVC se amplían de tal manera que incluyen mapeos desde la fidelidad de profundidad de bits de muestra inferior a la superior así como, opcionalmente, desde la resolución espacial inferior a la superior. La extensión anteriormente descrita de SVC hacia la escalabilidad en términos de profundidad de bits y, opcionalmente, resolución espacial posibilita que el codificador almacene una representación de calidad base de una secuencia de vídeo, que puede decodificarse por cualquier decodificador de vídeo heredado junto con una señal de mejora para profundidad de bits superior y opcionalmente, resolución espacial superior, que se ignora por decodificadores de vídeo heredados. Por ejemplo, la representación de calidad de base podría contener una versión de 8 bits de la secuencia de vídeo en resolución CIF, en concreto muestras de 352x288, mientras que la señal de mejora de alta calidad contiene un "perfeccionamiento" a una versión de 10 bits en resolución 4CIF, es decir muestras de 704x476 de la misma secuencia. En una configuración diferente, es posible también usar la misma resolución espacial para tanto representaciones de base y de calidad de mejora, de manera que la señal de mejora de alta calidad únicamente contiene un perfeccionamiento de la profundidad de bits de muestra, por ejemplo desde 8 a 10 bits.

En otras palabras, las realizaciones anteriormente señaladas posibilitan formar un codificador-decodificador de vídeo, es decir un codificador o decodificador, para codificar-decodificar, es decir codificar o decodificar, una representación en capas de una señal de vídeo que comprende un método de codificación de vídeo normalizado para codificar una capa de calidad de base, un método de predicción para realizar una predicción de la señal de capa de mejora de alta calidad usando la señal de calidad base reconstruida y un método de codificación residual para codificación de la predicción residual de la señal de capa de mejora de alta calidad. En este sentido, la predicción puede realizarse usando una función de mapeo desde el rango dinámico asociado con la capa de calidad de base al rango dinámico asociado con la capa de mejora de alta calidad. Además, la función de mapeo puede crearse como la suma de una función de mapeo global, que sigue los esquemas de mapeo de tono inverso anteriormente descritos y una adaptación local. La adaptación local, a su vez, puede realizarse escalando los valores de muestra x de la capa de calidad de base y añadir un valor de compensación de acuerdo con $m \cdot x \cdot n$. En cualquier caso, la codificación residual puede realizarse de acuerdo con H.264/AVC.

Dependiendo de una implementación real, el esquema de codificación inventivo puede implementarse en hardware o en software. Por lo tanto, la presente invención también se refiere a un programa informático, que puede almacenarse en un medio legible por ordenador tal como un CD, un disco o cualquier otro soporte de datos. La presente invención es, por lo tanto, un programa informático que tiene un código de programa que, cuando se ejecuta en un ordenador, realiza el método inventivo descrito en relación con las figuras anteriores. En particular, las implementaciones de los medios y módulos en la Figura 1 y 2 pueden comprender sub-rutinas que se ejecutan en una CPU, partes de circuito de un ASIC o similares, por ejemplo.

REIVINDICACIONES

1. Método para decodificar un flujo de datos de calidad escalable en el que se codifican datos de origen de instantánea o de vídeo, comprendiendo el flujo de datos de calidad escalable un flujo de datos de capa de base que representa los datos de origen de instantánea o de vídeo con una primera profundidad de bits de muestra de instantánea y una primera resolución espacial, un flujo de datos de capa de mejora que representa una predicción residual siendo una segunda profundidad de bits de muestra de instantánea superior que la primera profundidad de bits de muestra de instantánea y una segunda resolución espacial superior que la primera resolución espacial, y una función de mapeo local definida a una segunda granularidad, comprendiendo el método:
- 5
10
15
20
25
- decodificar el flujo de datos de capa de base en unos datos de instantánea o de vídeo reconstruidos de profundidad de bits inferior;
 decodificar el flujo de datos de mejora en la predicción residual;
 interpolar espacialmente, usando un filtro de interpolación espacial, muestras de los datos de instantánea o de vídeo reconstruidos de profundidad de bits inferior para aumentar la primera resolución espacial para corresponder con la segunda resolución espacial;
 mapear muestras de los datos de instantánea o de vídeo reconstruidos de profundidad de bits inferior con la primera profundidad de bits de muestra de instantánea a partir de un primer rango dinámico que corresponde a la primera profundidad de bits de muestra de instantánea a un segundo rango dinámico mayor que el primer rango dinámico y que corresponde a la segunda profundidad de bits de muestra de instantánea, mediante el uso de una o más funciones de mapeo global que son constantes dentro del vídeo o que varían a una primera granularidad, y una función de mapeo local que modifica localmente la una o más funciones de mapeo global a la segunda granularidad que es menor que la primera granularidad, para obtener una predicción de los datos de origen de instantánea o de vídeo que tienen la segunda profundidad de bits de muestra de instantánea; y
 reconstruir la instantánea con la segunda profundidad de bits de muestra de instantánea basándose en la predicción y en la predicción residual.
2. Método de acuerdo con la reivindicación 1, en el que un orden en el que se realiza la interpolación espacial y el mapeo, se elige de manera adaptativa en una base bloque a bloque.
- 30
3. Método de acuerdo con la reivindicación 1, en el que un orden en el que se realiza la interpolación espacial y el mapeo, se elige de manera adaptativa en una base instantánea a instantánea.
- 35
4. Método de acuerdo con la reivindicación 1, en el que un orden en el que se realiza la interpolación espacial y el mapeo, se elige de manera adaptativa a una granularidad que se señala en el flujo de datos de calidad escalable.
5. Método de acuerdo con la reivindicación 1, en el que un orden en el que se realiza la interpolación espacial y el mapeo, se señala como información secundaria en el flujo de datos de calidad escalable.
- 40
6. Decodificador para decodificar un flujo de datos de calidad escalable en el que se codifican datos de origen de instantánea o de vídeo, comprendiendo el flujo de datos de calidad escalable un flujo de datos de capa de base que representa los datos de origen de instantánea o de vídeo con una primera profundidad de bits de muestra de instantánea y una primera resolución espacial, un flujo de datos de capa de mejora que representa una predicción residual siendo una segunda profundidad de bits de muestra de instantánea superior que la primera profundidad de bits de muestra de instantánea y una segunda resolución espacial superior que la primera resolución espacial, y una función de mapeo local definida a una segunda granularidad, comprendiendo el decodificador:
- 45
- medio (204) para decodificar el flujo de datos de capa de base en unos datos de instantánea o de vídeo reconstruidos de profundidad de bits inferior;
 medio (208) para decodificar el flujo de datos de mejora en la predicción residual;
 medio (206) para interpolar espacialmente, usando un filtro de interpolación espacial, muestras de los datos de instantánea o de vídeo reconstruidos de profundidad de bits inferior para aumentar la primera resolución espacial para corresponder con la segunda resolución espacial y mapear muestras de los datos de instantánea o de vídeo reconstruidos de profundidad de bits inferior con la primera profundidad de bits de muestra de instantánea a partir de un primer rango dinámico que corresponde a la primera profundidad de bits de muestra de instantánea a un segundo rango dinámico mayor que el primer rango dinámico y que corresponde a la segunda profundidad de bits de muestra de instantánea, mediante el uso de una o más funciones de mapeo global que son constantes dentro del vídeo o que varían a una primera granularidad, y una función de mapeo local que modifica localmente la una o más funciones de mapeo global a la segunda granularidad que es menor que la primera granularidad, para obtener una predicción de los datos de origen de instantánea o de vídeo que tienen la segunda profundidad de bits de muestra de instantánea; y
 medio (210) para reconstruir la instantánea con la segunda profundidad de bits de muestra de instantánea basándose en la predicción y en la predicción residual.
- 50
55
60

7. Decodificador de acuerdo con la reivindicación 6, en el que un orden en el que se realiza la interpolación espacial y el mapeo, se elige de manera adaptativa en una base bloque a bloque.
8. Decodificador de acuerdo con la reivindicación 6, en el que un orden en el que se realiza la interpolación espacial y el mapeo, se elige de manera adaptativa en una base instantánea a instantánea.
- 5 9. Decodificador de acuerdo con la reivindicación 6, en el que un orden en el que se realiza la interpolación espacial y el mapeo, se elige de manera adaptativa a una granularidad que se señala en el flujo de datos de calidad escalable.
- 10 10. Decodificador de acuerdo con la reivindicación 6, en el que un orden en el que se realiza la interpolación espacial y el mapeo, se señala como información secundaria en el flujo de datos de calidad escalable.
11. Flujo de datos de calidad escalable en el que se codifican datos de origen de instantánea o de vídeo, comprendiendo el flujo de datos de calidad escalable
- 15 un flujo de datos de capa de base que representa los datos de origen de instantánea o de vídeo con una primera profundidad de bits de muestra de instantánea y una primera resolución espacial, un flujo de datos de capa de mejora de profundidad de bits que representa una predicción residual siendo una segunda profundidad de bits de muestra de instantánea superior que la primera profundidad de bits de muestra de instantánea y una segunda resolución espacial superior que la primera resolución espacial, y
- 20 una función de mapeo local definida a una segunda granularidad, en el que una reconstrucción de la instantánea con la segunda profundidad de bits de muestra de instantánea se puede derivar a partir de la predicción residual y una predicción obtenida interpolando espacialmente, usando un filtro de interpolación espacial, muestras de los datos de instantánea o de vídeo reconstruidos de profundidad de bits inferior para aumentar la primera resolución espacial para corresponder con la segunda resolución espacial y
- 25 mapear muestras de los datos de instantánea o de vídeo reconstruidos de profundidad de bits inferior con la primera profundidad de bits de muestra de instantánea a partir de un primer rango dinámico que corresponde a la primera profundidad de bits de muestra de instantánea a un segundo rango dinámico mayor que el primer rango dinámico y que corresponde a la segunda profundidad de bits de muestra de instantánea, mediante el uso de una o más funciones de mapeo global que son constantes dentro del vídeo o que varían a una primera granularidad, y la
- 30 función de mapeo local que modifica localmente la una o más funciones de mapeo global a la segunda granularidad que es menor que la primera granularidad.
12. Método de acuerdo con la reivindicación 1, en el que un orden en el que se ha de realizar la interpolación espacial y el mapeo, se señala como información secundaria en el flujo de datos de calidad escalable.
- 35 13. Método para codificar unos datos de origen de instantánea o de vídeo (160) en un flujo de datos de calidad escalable (112), que comprende:
- 40 codificar los datos de origen de instantánea o de vídeo (160) en un flujo de datos de codificación de base que representa una representación de los datos de origen de instantánea o de vídeo con una primera profundidad de bits de muestra de instantánea y una primera resolución espacial; interpolar espacialmente, usando un filtro de interpolación espacial, muestras de los datos de instantánea o de vídeo reconstruidos de profundidad de bits inferior para aumentar la primera resolución espacial para corresponder con una segunda resolución espacial superior que la primera resolución espacial; y
- 45 mapear muestras de la representación de los datos de origen de instantánea o de vídeo (160) con la primera profundidad de bits de muestra de instantánea a partir de un primer rango dinámico que corresponde a la primera profundidad de bits de muestra de instantánea a un segundo rango dinámico mayor que el primer rango dinámico y que corresponde a una segunda profundidad de bits de muestra de instantánea que es superior a la primera profundidad de bits de muestra de instantánea, mediante el uso de una o más funciones de mapeo global que son constantes dentro de los datos de origen de instantánea o de vídeo (160) o que varían a una primera granularidad, y una función de mapeo local que modifica localmente la una o más funciones de mapeo global a una segunda granularidad más precisa que la primera granularidad para obtener una predicción de los datos de origen de instantánea o de vídeo que tienen la segunda profundidad de bits de muestra de instantánea;
- 50 codificar una predicción residual de la predicción en un flujo de datos de capa de mejora de profundidad de bits; y formar el flujo de datos de calidad escalable basándose en el flujo de datos de codificación de base, la función de mapeo local y el flujo de datos de capa de mejora de profundidad de bits de modo que la función de mapeo local puede derivarse del flujo de datos de calidad escalable.
- 55 14. Codificador para codificar unos datos de origen de instantánea o de vídeo (160) en un flujo de datos de calidad escalable (112), que comprende:
- 60 medio de codificación de base (102) para codificar los datos de origen de instantánea o de vídeo (160) en un flujo de datos de codificación de base que representan una representación de los datos de origen de instantánea o de vídeo con una primera profundidad de bits de muestra de instantánea y una primera resolución espacial;

- 5 medio de mapeo (104) para interpolar espacialmente, usando un filtro de interpolación espacial, muestras de los datos de instantánea o de vídeo reconstruidos de profundidad de bits inferior para aumentar la primera resolución espacial para corresponder con una segunda resolución espacial superior que la primera resolución espacial y mapear muestras de la representación de los datos de origen de instantánea o de vídeo (160) con la primera profundidad de bits de muestra de instantánea a partir de un primer rango dinámico que corresponde a la primera profundidad de bits de muestra de instantánea a un segundo rango dinámico mayor que el primer rango dinámico y que corresponde a una segunda profundidad de bits de muestra de instantánea que es superior a la primera profundidad de bits de muestra de instantánea, mediante el uso de una o más funciones de mapeo global que son constantes dentro de
- 10 los datos de origen de instantánea o de vídeo (160) o que varían a una primera granularidad, y una función de mapeo local que modifica localmente la una o más funciones de mapeo global a una segunda granularidad más precisa que la primera granularidad para obtener una predicción de los datos de origen de instantánea o de vídeo que tienen la segunda profundidad de bits de muestra de instantánea;
- 15 medio de codificación residual (106) para codificar una predicción residual de la predicción en un flujo de datos de capa de mejora de profundidad de bits; y
medio de combinación (108) para formar el flujo de datos de calidad escalable basándose en el flujo de datos de codificación de base, la función de mapeo local y el flujo de datos de capa de mejora de profundidad de bits de modo que la función de mapeo local puede derivarse a partir del flujo de datos de calidad escalable.
- 20 15. Programa informático que tiene código de programa para realizar, cuando se ejecuta en un ordenador, un método de acuerdo con la reivindicación 1 o 13.

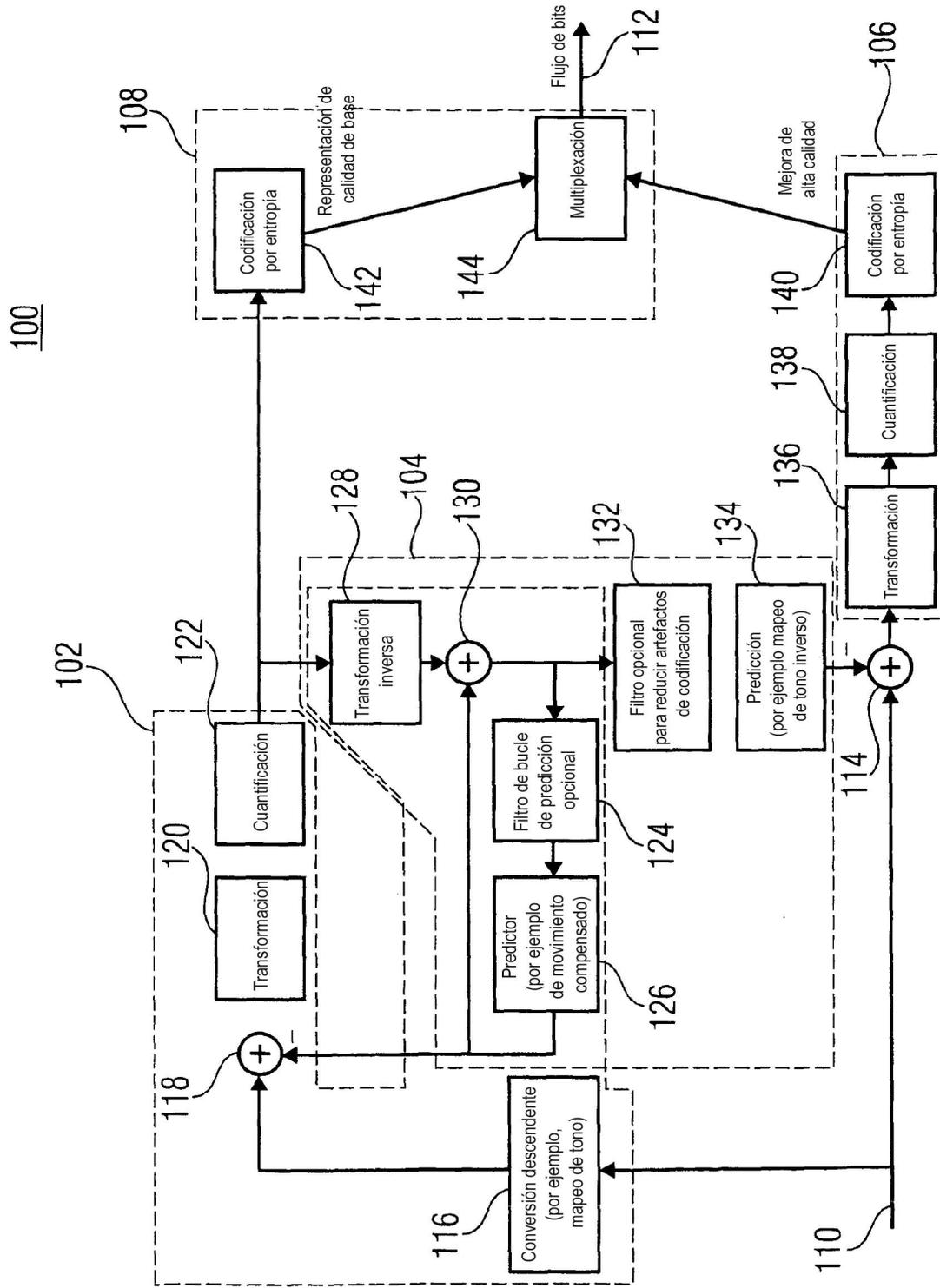


FIG 1

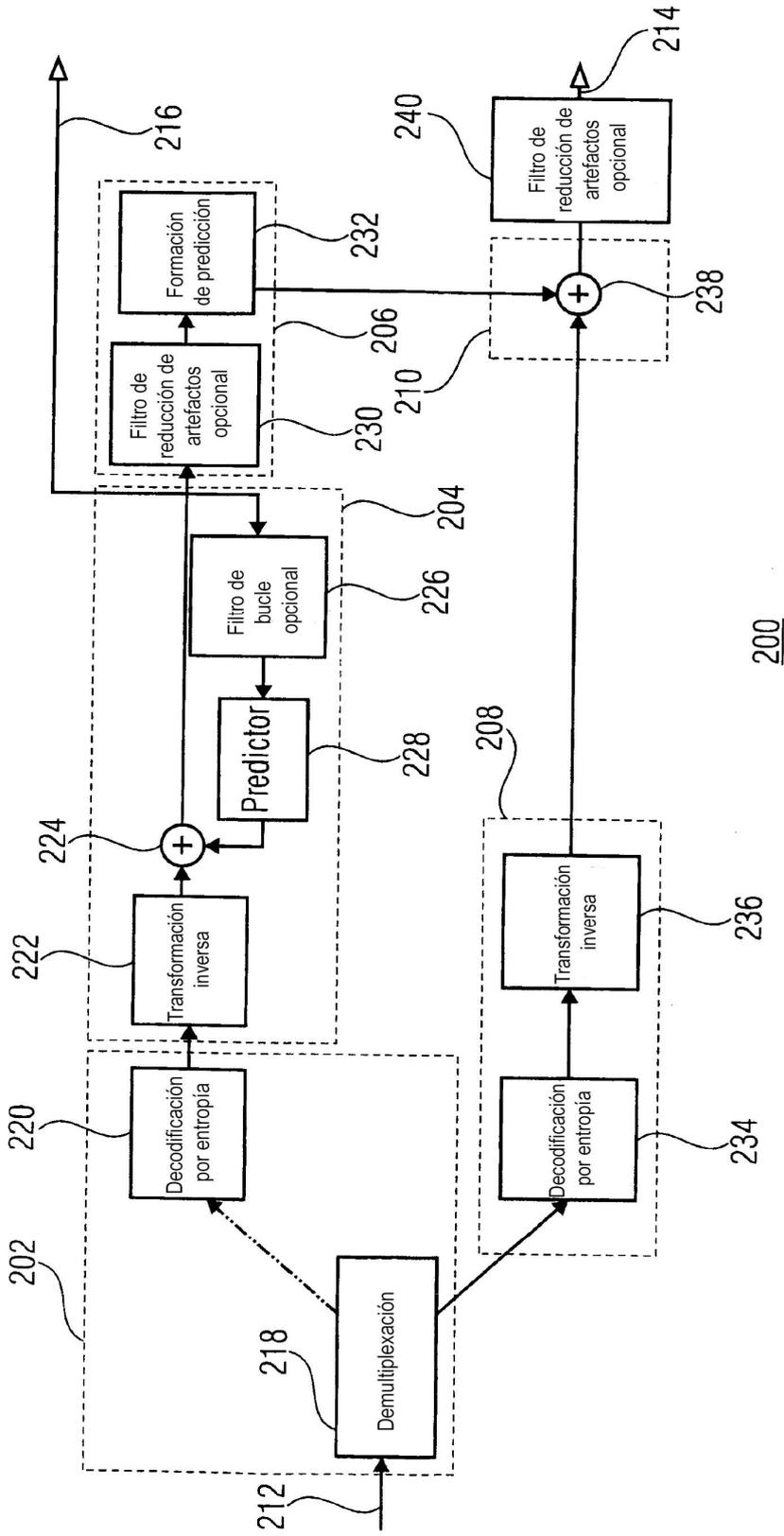


FIG 2

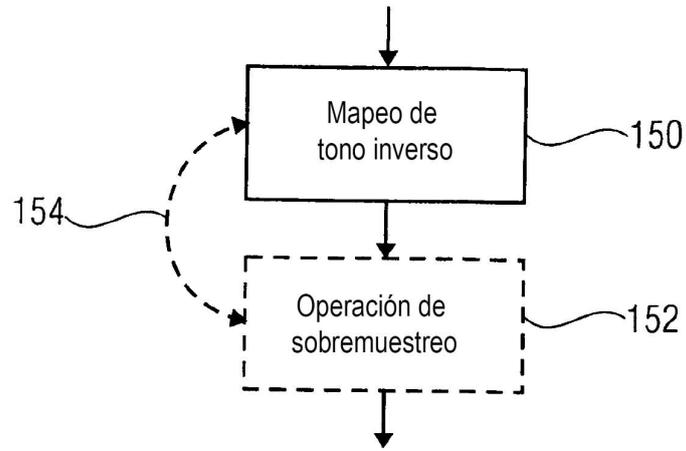


FIG 3

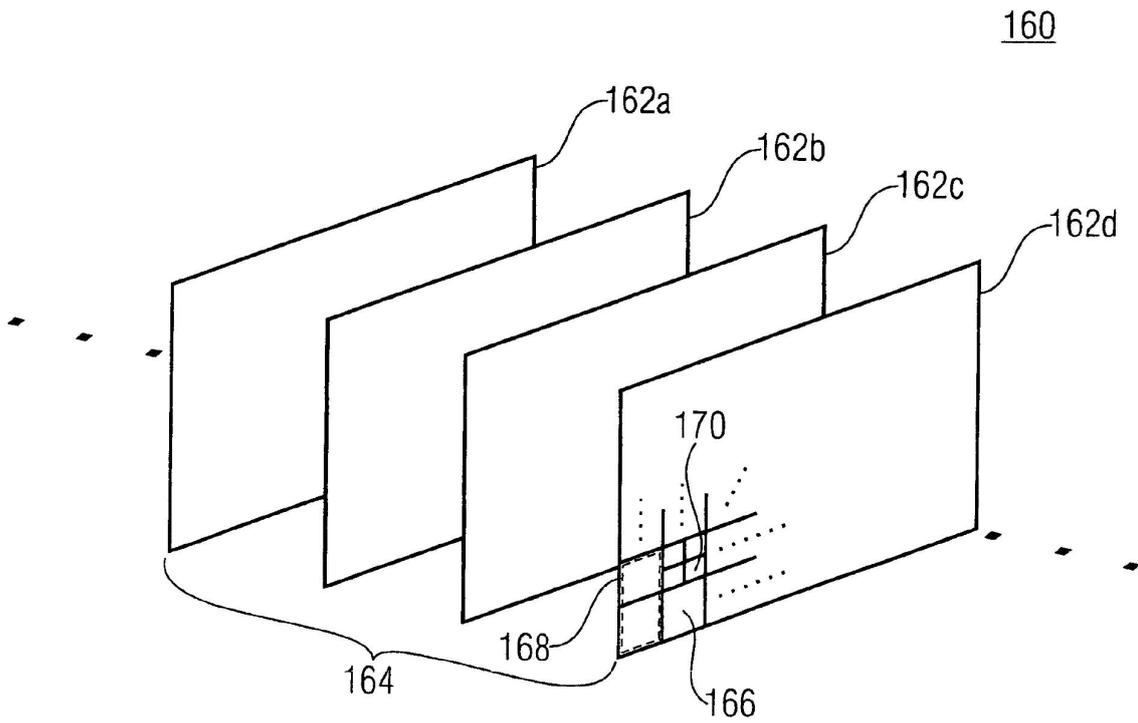


FIG 4

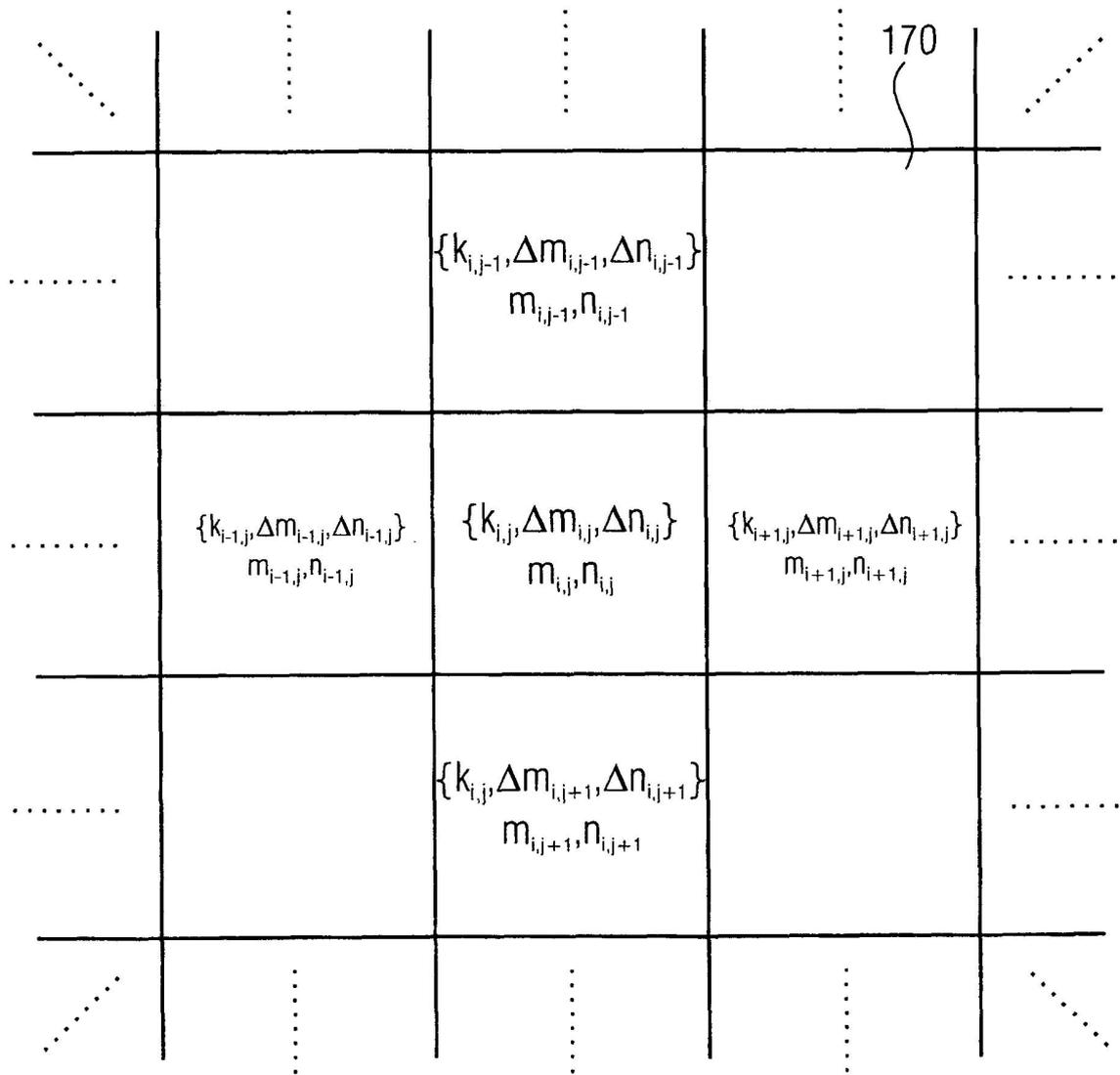


FIG 5

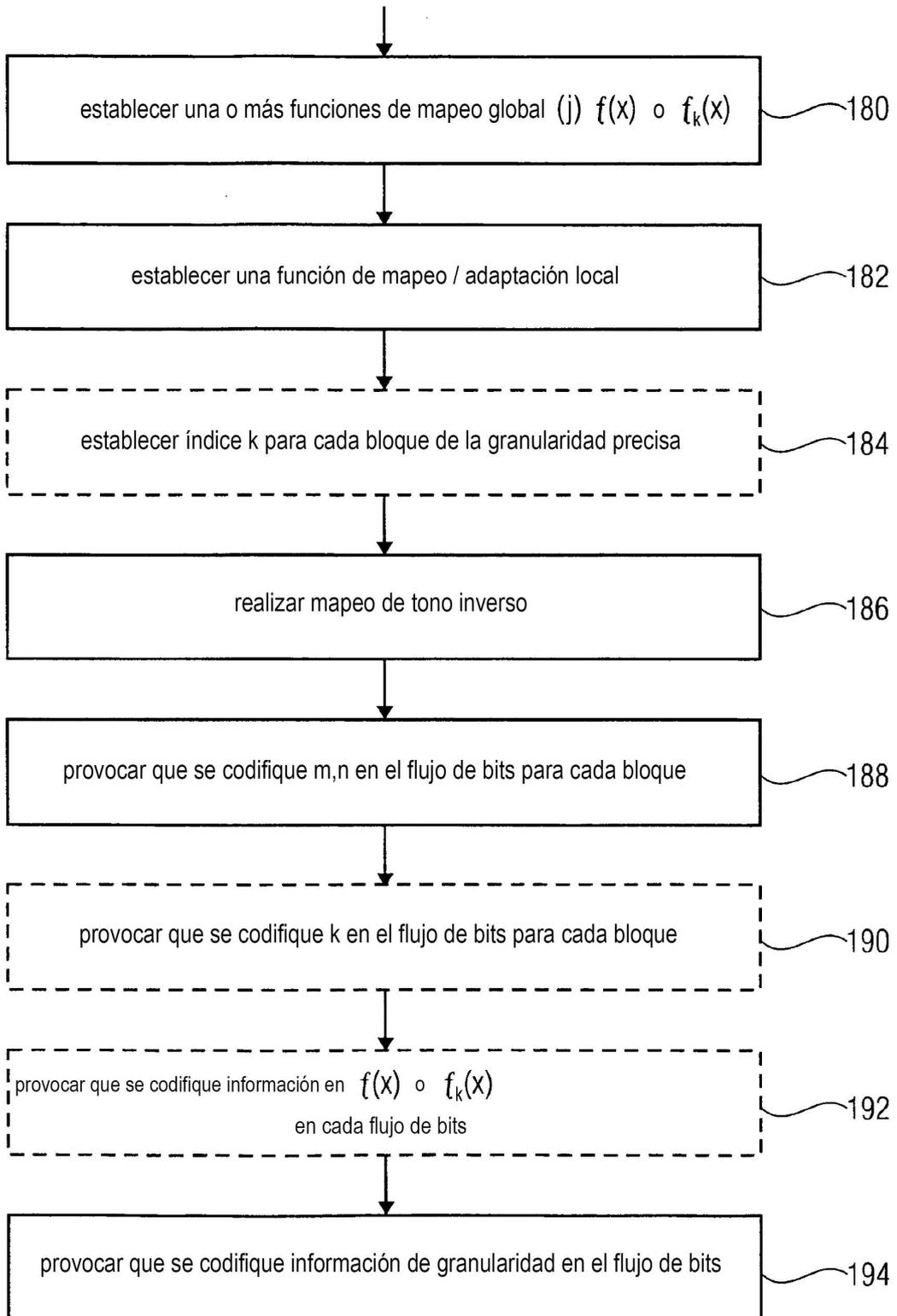


FIG 6

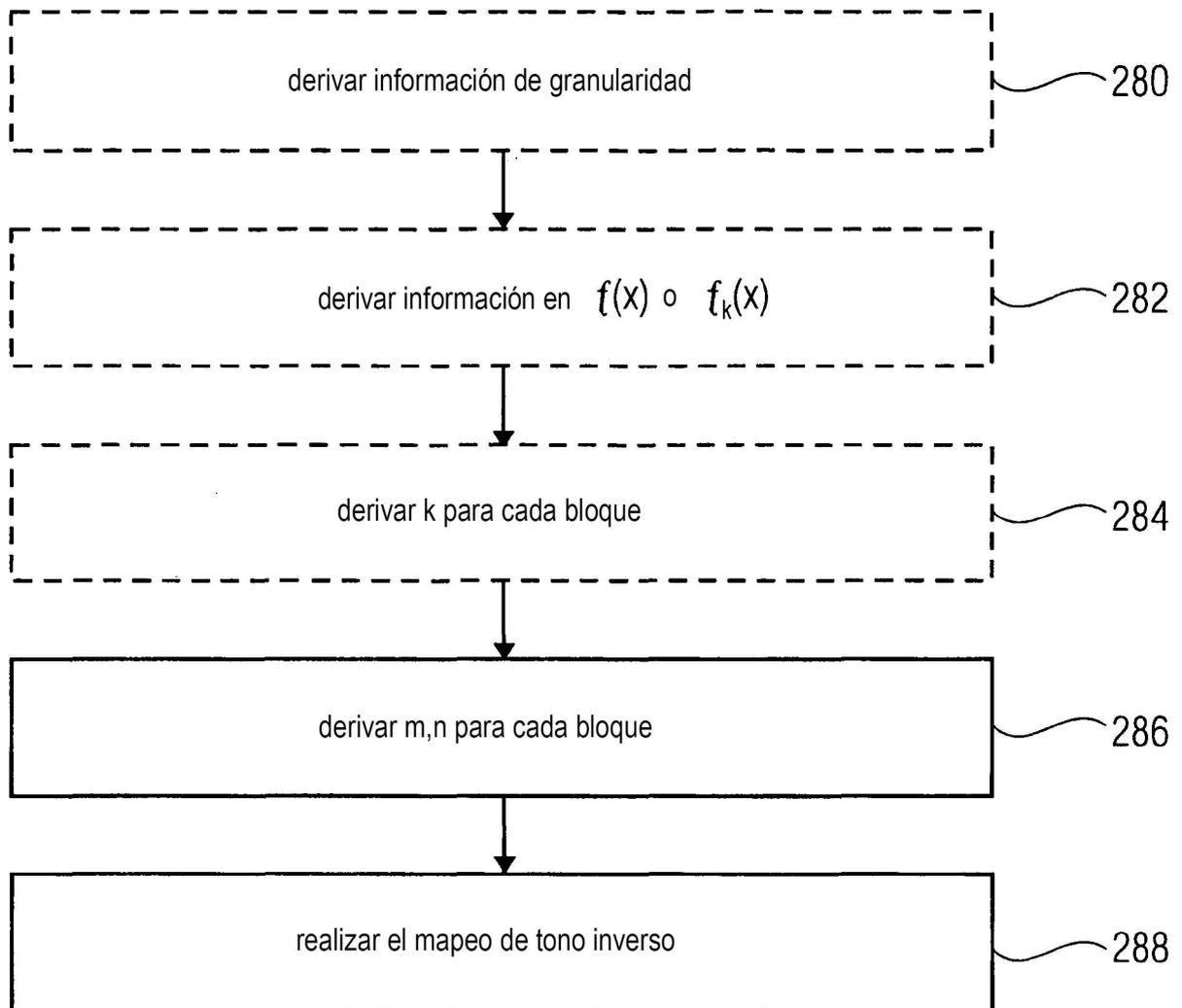


FIG 7

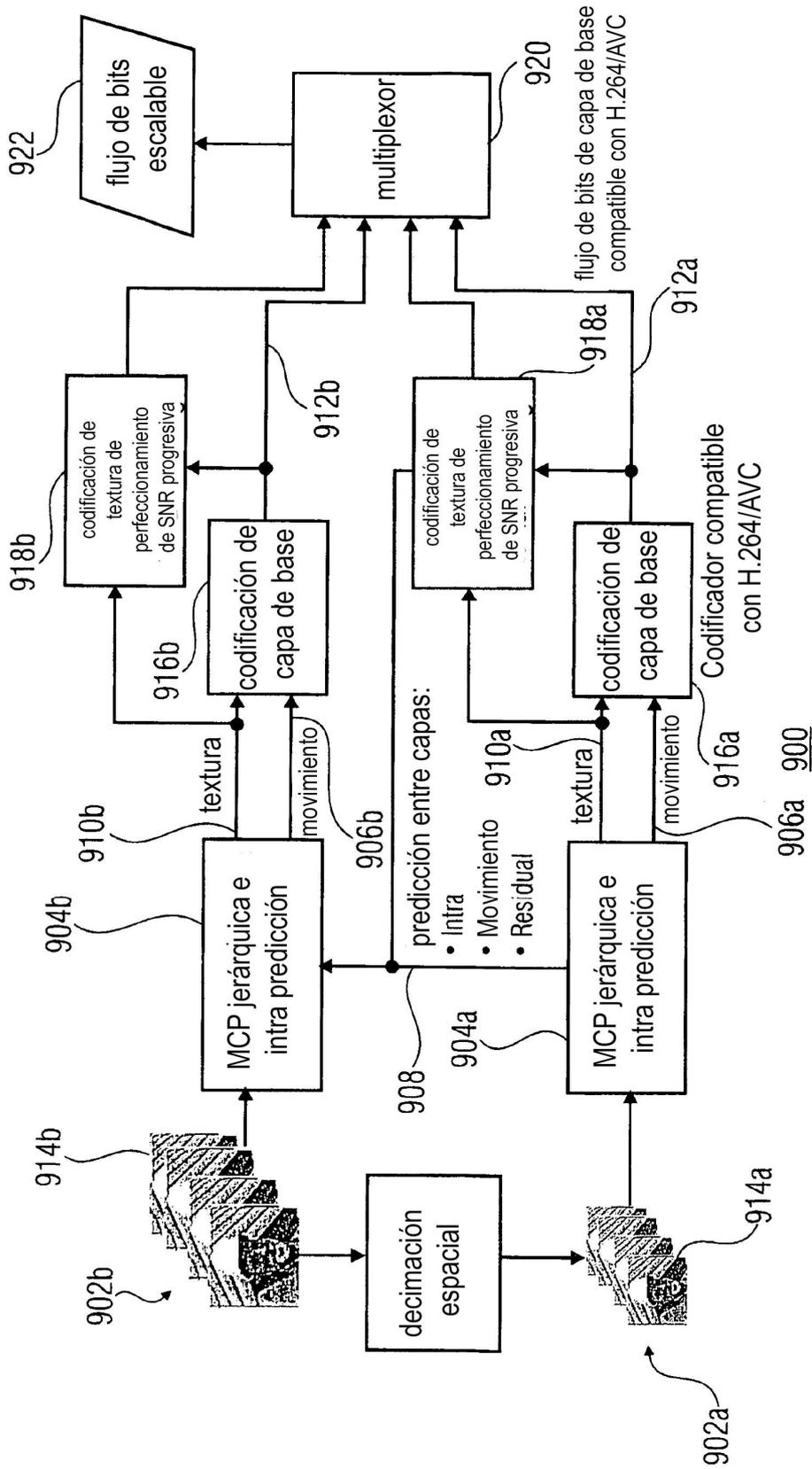


FIG 8