

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 029 145**

51 Int. Cl.:

H04N 19/13 (2014.01)

H04N 19/184 (2014.01)

H04N 19/593 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **24.06.2020 PCT/GB2020/051531**

87 Fecha y número de publicación internacional: **30.12.2020 WO20260874**

96 Fecha de presentación y número de la solicitud europea: **24.06.2020 E 20735666 (8)**

97 Fecha y número de publicación de la concesión europea: **26.03.2025 EP 3991418**

54 Título: **Codificación y decodificación de datos de imagen**

30 Prioridad:

25.06.2019 GB 201909141

31.07.2019 GB 201910924

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
23.06.2025

73 Titular/es:

SONY GROUP CORPORATION (100.00%)
1-7-1, Konan Minato-ku
Tokyo 108-0075, JP

72 Inventor/es:

KEATING, STEPHEN MARK;
SHARMAN, KARL JAMES y
BROWNE, ADRIAN RICHARD

74 Agente/Representante:

DEL VALLE VALIENTE, Sonia

ES 3 029 145 T3

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

DESCRIPCIÓN

Codificación y decodificación de datos de imagen

5

Antecedentes**Campo**

10 Esta descripción se refiere a la codificación y decodificación de datos de imágenes.

Descripción de la técnica relacionada

15 La descripción de “antecedentes” proporcionada en la presente memoria tiene el propósito de presentar en general el contexto de la descripción. El trabajo de los inventores actualmente nombrados, en la medida en que se describe en esta sección de antecedentes, así como los aspectos de la descripción que de otro modo no se calificarían como estado de la técnica en el momento de la presentación, no se admiten expresa ni implícitamente como estado de la técnica frente a la presente descripción.

20 Hay varios sistemas de codificación y decodificación de datos de vídeo que implican transformar los datos de vídeo en una representación del dominio de la frecuencia, cuantificar los coeficientes del dominio de la frecuencia y después aplicar alguna forma de codificación por entropía a los coeficientes cuantificados. Esto puede lograr la compresión de los datos de vídeo. Se aplica una técnica de decodificación o descompresión correspondiente para recuperar una versión reconstruida de los datos de vídeo originales.

25

La codificación de vídeo de alta eficiencia (HEVC), también conocida como H.265 o MPEG-H Part 2, es una propuesta sucesora de H.264/MPEG-4 AVC. Se pretende que la HEVC mejore la calidad del vídeo y duplique la relación de compresión de datos en comparación con H.264, y para que sea escalable a partir de una resolución de 128×96 a 7680×4320 píxeles, aproximadamente el equivalente a tasas de bits que oscilan entre 128 kbit/s y 800 Mbit/s.

30

Resumen

La presente descripción aborda o mitiga los problemas que surgen de este procesamiento.

35

La presente invención proporciona un aparato de codificación de datos de imágenes, que comprende:

un codificador por entropía configurado para codificar selectivamente elementos de datos que representan datos de imagen mediante un sistema de codificación seleccionado entre un sistema de codificación de codificación aritmética binaria adaptativa al contexto, CABAC, y un sistema de codificación por derivación;

40

en donde los elementos de datos incluyen (i) un primer elemento de datos que indica si un modo de croma de modelo lineal multidireccional, MDLM, se aplica en una región de imagen actual; y (ii) un segundo elemento de datos que indica una dirección de procesamiento para el funcionamiento del modo de croma MDLM;

45

en donde el codificador por entropía está configurado para codificar instancias del primer elemento de datos mediante el sistema de codificación CABAC utilizando un primer contexto CABAC y para codificar instancias del segundo elemento de datos independientemente del primer contexto CABAC.

50

La presente invención también proporciona un método de codificación de datos de imagen, que comprende:

codificar selectivamente elementos de datos que representan datos de imagen mediante un sistema de codificación seleccionado entre un sistema de codificación de codificación aritmética binaria adaptativa al contexto, CABAC, y un sistema de codificación por derivación;

55

en donde los elementos de datos incluyen (i) un primer elemento de datos que indica si un modo de croma de modelo lineal multidireccional, MDLM, se aplica en una región de imagen actual; y (ii) un segundo elemento de datos que indica una dirección de procesamiento para el funcionamiento del modo de croma MDLM;

60

en donde el paso de codificación comprende codificar instancias del primer elemento de datos mediante el sistema de codificación CABAC utilizando un primer contexto CABAC y codificar instancias del segundo elemento de datos independientemente del primer contexto CABAC.

65

La presente invención también proporciona un aparato de decodificación de datos de imágenes, que comprende:

un decodificador por entropía configurado para decodificar selectivamente elementos de datos que representan datos de imagen mediante un sistema de decodificación seleccionado entre un sistema de decodificación de codificación

aritmética binaria adaptativa al contexto (CABAC, por sus siglas en inglés) y un sistema de decodificación por derivación;

5 en donde los elementos de datos incluyen (i) un primer elemento de datos que indica si un modo de croma de modelo lineal multidireccional, MDLM, se aplica en una región de imagen actual; y (ii) un segundo elemento de datos que indica una dirección de procesamiento para el funcionamiento del modo de croma MDLM;

10 en donde el decodificador por entropía está configurado para decodificar instancias del primer elemento de datos mediante el sistema de decodificación CABAC utilizando un primer contexto CABAC y para decodificar instancias del segundo elemento de datos independientemente del primer contexto CABAC.

La presente invención también proporciona un método de decodificación de datos de imagen, que comprende:

15 decodificar selectivamente elementos de datos que representan datos de imagen mediante un sistema de decodificación seleccionado entre un sistema de decodificación de codificación aritmética binaria adaptativa al contexto, CABAC, y un sistema de decodificación por derivación;

20 en donde los elementos de datos incluyen (i) un primer elemento de datos que indica si un modo de croma de modelo lineal multidireccional, MDLM, se aplica en una región de imagen actual; y (ii) un segundo elemento de datos que indica una dirección de procesamiento para el funcionamiento del modo de croma MDLM;

25 en donde el paso de decodificación comprende decodificar instancias de los primeros datos mediante el sistema de decodificación CABAC utilizando un primer contexto CABAC y decodificar instancias del segundo elemento de datos independientemente del primer contexto CABAC.

En las reivindicaciones adjuntas se definen aspectos y características respectivos adicionales de la presente descripción.

30 Debe entenderse que tanto la descripción general anterior como la siguiente descripción detallada son ilustrativas, pero no restrictivas, de la presente tecnología.

Breve descripción de los dibujos

35 Se obtendrá fácilmente una apreciación más completa de la descripción y muchas de las ventajas que la misma conlleva a medida que la misma se entienda mejor con referencia a la siguiente descripción detallada cuando se considera en conexión con los dibujos adjuntos, en donde:

40 la figura 1 ilustra esquemáticamente un sistema de transmisión y recepción de datos de audio/vídeo (A/V) que utiliza compresión y descompresión de datos de vídeo;

la figura 2 ilustra esquemáticamente un sistema de visualización de vídeo que utiliza a descompresión de datos de vídeo;

45 la figura 3 ilustra esquemáticamente un sistema de almacenamiento de audio/vídeo que utiliza compresión y descompresión de datos de vídeo;

la figura 4 ilustra esquemáticamente una cámara de vídeo que utiliza compresión de datos de vídeo;

50 las figuras 5 y 6 ilustran esquemáticamente unos medios de almacenamiento;

la figura 7 proporciona una visión general esquemática de un aparato de compresión y descompresión de datos de vídeo;

55 la figura 8 ilustra esquemáticamente un predictor;

la figura 9 ilustra esquemáticamente una imagen parcialmente codificada;

la figura 10 ilustra esquemáticamente un conjunto de posibles direcciones de intrapredicción;

60 la figura 11 ilustra esquemáticamente un conjunto de modos de predicción;

la figura 12 ilustra esquemáticamente otro conjunto de modos de predicción;

65 la figura 13 ilustra esquemáticamente un proceso de intrapredicción;

la figura 14 ilustra esquemáticamente un codificador CABAC;

las figuras 15 y 16 ilustran esquemáticamente las técnicas de codificación CABAC;

las figuras 17 y 18 ilustran esquemáticamente las técnicas de decodificación CABAC; y

5

Las figuras 19 y 20 son diagramas de flujo esquemáticos que ilustran los métodos respectivos.

Descripción de las realizaciones preferidas

10 Haciendo referencia ahora a los dibujos, se proporcionan las figuras 1 a 4 para aportar ilustraciones esquemáticas de aparatos o sistemas que utilizan el aparato de compresión y/o descompresión que se describirá a continuación en relación con las realizaciones de la presente tecnología.

15 Todos los aparatos de compresión y/o descompresión de datos que se describirán a continuación pueden implementarse en hardware, en software que se ejecuta en un aparato de procesamiento de datos de propósito general, tal como un ordenador de propósito general, como hardware programable, tal como un circuito integrado de aplicación específica (ASIC) o una matriz de puertas programables en campo (FPGA) o como combinaciones de estos. En los casos en los que las realizaciones se implementen mediante software y/o firmware, se apreciará que tal software y/o firmware, así como los medios de almacenamiento de datos no transitorios mediante los cuales dicho software y/o firmware se almacena o está provisto de otro modo, se consideran realizaciones de la presente tecnología.

20

La figura 1 ilustra esquemáticamente un sistema de transmisión y recepción de datos de audio/vídeo que utiliza compresión y descompresión de datos de vídeo.

25 Se suministra una señal de audio/vídeo de entrada 10 a un aparato de compresión de datos de vídeo 20 que comprime al menos el componente de vídeo de la señal de audio/vídeo 10 para su transmisión a lo largo de una ruta de transmisión 30, tal como un cable, una fibra óptica, un enlace inalámbrico o similar. La señal comprimida es procesada por un aparato de descompresión 40 para proporcionar una señal de audio/vídeo de salida 50. Para la ruta de retorno, un aparato de compresión 60 comprime una señal de audio/vídeo para su transmisión a lo largo de la ruta de transmisión 30 en un aparato de descompresión 70.

30

Por lo tanto, el aparato de compresión 20 y el aparato de descompresión 70 pueden formar un nodo de un enlace de transmisión. El aparato de descompresión 40 y el aparato de descompresión 60 pueden formar otro nodo del enlace de transmisión. Por supuesto, en los casos en los que el enlace de transmisión es unidireccional, solo uno de los nodos requeriría un aparato de compresión y el otro nodo solo requeriría un aparato de descompresión.

35

La figura 2 ilustra esquemáticamente un sistema de visualización de vídeo que utiliza descompresión de datos de vídeo. En particular, un aparato 110 de descompresión procesa una señal 100 de audio/vídeo comprimida para proporcionar una señal descomprimida que puede visualizarse en una pantalla 120. El aparato 110 de descompresión podría implementarse como parte integral de la pantalla 120, por ejemplo, proporcionándose dentro de la misma carcasa que el dispositivo de visualización. De forma alternativa, el aparato 110 de descompresión se puede proporcionar (por ejemplo) como un decodificador denominado (STB) de sobremesa, teniendo en cuenta que la expresión “de sobremesa” no implica el requisito de que el decodificador esté ubicado en una orientación o posición particular con respecto a la pantalla 120; sino que es simplemente un término utilizado en la técnica para indicar un dispositivo que se puede conectar a una pantalla como dispositivo periférico.

40

45

La figura 3 ilustra esquemáticamente un sistema de almacenamiento de audio/vídeo que utiliza compresión y descompresión de datos de vídeo. Se suministra una señal 130 de audio/vídeo de entrada a un aparato 140 de compresión que genera una señal comprimida para su almacenamiento mediante un dispositivo 150 de almacenamiento, tal como un dispositivo de disco magnético, un dispositivo de disco óptico, un dispositivo de cinta magnética, un dispositivo de almacenamiento de estado sólido, tal como una memoria de semiconductores u otro dispositivo de almacenamiento. Para la reproducción, los datos comprimidos se leen del dispositivo 150 de almacenamiento y se pasan a un aparato 160 de descompresión para su descompresión con el fin de proporcionar una señal 170 de audio/vídeo de salida.

50

55

Se apreciará que la señal comprimida o codificada y un medio de almacenamiento, tal como un medio de almacenamiento no transitorio legible por máquina, que almacena esa señal, se consideran realizaciones de la presente tecnología.

60

La figura 4 ilustra esquemáticamente una videocámara que utiliza compresión de datos de vídeo. En la figura 4, un dispositivo 180 de captura de imágenes, tal como un sensor de imagen de dispositivo de carga acoplada (CCD) y la electrónica de control y lectura asociada, genera una señal de vídeo que se pasa a un aparato 190 de compresión. Un micrófono 200 (o varios micrófonos) genera una señal de audio que se pasará al aparato 190 de compresión. El aparato de compresión 190 genera una señal de audio/vídeo comprimida 210 para ser almacenada y/o transmitida (mostrada genéricamente como una etapa esquemática 220).

65

Las técnicas que se describirán a continuación se refieren principalmente a la compresión y descompresión de datos de vídeo. Se apreciará que se pueden utilizar muchas técnicas existentes para la compresión de datos de audio junto con las técnicas de compresión de datos de vídeo que se van a describir, para generar una señal de audio/vídeo comprimida. En consecuencia, no se proporcionará una discusión separada sobre la compresión de datos de audio.

5 También se apreciará que la velocidad de datos asociada con los datos de vídeo, en particular, los datos de vídeo con calidad de emisión, es generalmente mucho más alta que la velocidad de datos asociada con los datos de audio (ya estén comprimidos o no comprimidos). Por lo tanto, se apreciará que los datos de audio sin comprimir podrían acompañar a los datos de vídeo comprimidos para formar una señal de audio/vídeo comprimida. Se apreciará además que, aunque los presentes ejemplos (mostrados en las figuras 1 a 4) se refieren a datos de audio/vídeo, las técnicas

10 que se describirán a continuación pueden ser útiles en un sistema que simplemente trate (es decir, comprima, descomprima, almacene, visualice y/o transmita) datos de vídeo. Es decir, las realizaciones se pueden aplicar a una compresión de datos de vídeo sin tener que manipular necesariamente en modo alguno los datos de audio asociados.

15 Por lo tanto, la figura 4 proporciona un ejemplo de un aparato de captura de vídeos que comprende un sensor de imagen y un aparato de codificación del tipo que se expondrá a continuación. Por lo tanto, la figura 2 proporciona un ejemplo de un aparato de decodificación del tipo que se expondrá a continuación y una pantalla a la que se envían las imágenes decodificadas.

20 Una combinación de las figuras 2 y 4 puede proporcionar un aparato de captura de vídeos que comprende un sensor 180 de imagen y un aparato 190 de codificación, un aparato 110 de decodificación y una pantalla 120 a la que se envían las imágenes decodificadas.

25 Las figuras 5 y 6 ilustran esquemáticamente los medios de almacenamiento, que almacenan (por ejemplo) los datos comprimidos generados por el aparato 20, 60, los datos comprimidos introducidos en el aparato 110 o los medios o pasos 150, 220 de almacenamiento. La figura 5 ilustra esquemáticamente un medio de almacenamiento en disco, tal como un disco magnético u óptico, y la figura 6 ilustra esquemáticamente un medio de almacenamiento en estado sólido, tal como una memoria flash. Cabe destacar que las figuras 5 y 6 también pueden proporcionar ejemplos de medios de almacenamiento no transitorios legibles por máquina que almacenan software informático que, cuando lo ejecuta un ordenador, hace que el ordenador lleve a cabo uno o más de los métodos que se expondrán a continuación.

30 Por lo tanto, las disposiciones anteriores proporcionan ejemplos de aparatos de almacenamiento, captura, transmisión o recepción de vídeo que incorporan cualquiera de las técnicas actuales.

35 La figura 7 proporciona una visión general esquemática de un aparato de compresión y descompresión de datos de vídeo.

40 Un controlador 343 controla el funcionamiento general del aparato y, en particular en lo referente a un modo de compresión, controla un proceso de codificación de prueba actuando como un selector para seleccionar varios modos de funcionamiento, tales como tamaños y formas de bloques, y si los datos de vídeo se van a codificar sin pérdidas o de otro modo. Se considera que el controlador forma parte del codificador o decodificador de imágenes (según sea el caso). Se suministran las imágenes sucesivas de una señal 300 de vídeo de entrada a un sumador 310 y a un predictor 320 de imágenes. El predictor 320 de imágenes se describirá a continuación en más detalle con referencia a la figura 8. El codificador o decodificador de imágenes (según sea el caso) más el predictor dentro de la imagen de la figura 8 pueden utilizar características del aparato de la figura 7. Sin embargo, esto no significa que el codificador o

45 decodificador de imágenes requiera necesariamente todas las características de la figura 7.

De hecho, el sumador 310 realiza una operación de resta (suma negativa), en el sentido de que recibe la señal 300 de vídeo de entrada en una entrada “+” y la salida del predictor 320 de imágenes en una entrada “-”, de modo que la imagen predicha se resta de la imagen de entrada. El resultado es generar una denominada señal de imagen residual 330 que representa la diferencia entre las imágenes reales y proyectadas.

50

Una razón por la que se genera una señal de imagen residual es la siguiente. Las técnicas de codificación de datos que se van a describir, es decir, las técnicas que se aplicarán a la señal de imagen residual, tienden a trabajar de manera más eficiente cuando hay menos “energía” en la imagen que se va a codificar. En este caso, el término “eficientemente” se refiere a la generación de una pequeña cantidad de datos codificados; para un nivel de calidad de imagen particular, es deseable (y considerado “eficiente”) generar tan pocos datos como sea posible en la práctica. La referencia a la “energía” en la imagen residual se refiere a la cantidad de información contenida en la imagen residual. Si la imagen predicha fuera idéntica a la imagen real, la diferencia entre las dos (es decir, la imagen residual) contendría cero información (energía cero) y sería muy fácil de codificar en una pequeña cantidad de datos codificados.

55

60 En general, si se puede hacer que el proceso de predicción trabaje razonablemente bien de tal modo que el contenido de la imagen predicha sea similar al contenido de la imagen que se va a codificar, la expectativa es que los datos de imagen residuales contengan menos información (menos energía) que la imagen de entrada y, por lo tanto, sean más fáciles de codificar en una pequeña cantidad de datos codificados.

65 Ahora se describirá el resto del aparato que actúa como codificador (para codificar la imagen residual o de diferencia). Se suministran los datos 330 de imagen residuales a una unidad o circuitería 340 de transformada que genera una

representación de transformada de coseno discreta (DCT) de bloques o regiones de los datos de imagen residuales. La propia técnica de DCT es bien conocida y no se describirá en detalle aquí. Cabe destacar también que la utilización de la DCT es solo ilustrativa de un ejemplo de disposición. Otras transformadas que podrían utilizarse incluyen, por ejemplo, una transformada sinusoidal discreta (DST). Una transformada también podría comprender una secuencia o cascada de transformadas individuales, tal como una disposición en la que una transformada va seguida (ya sea directamente o no) de otra transformada. La elección de la transformada se puede determinar explícitamente y/o depender de la información secundaria utilizada para configurar el codificador y el decodificador.

La salida de la unidad 340 de transformada, es decir, un conjunto de coeficientes DCT para cada bloque transformado de datos de imagen, se suministra a un cuantificador 350. Se conocen diversas técnicas de cuantificación en el campo de la compresión de datos de vídeo, que van desde una simple multiplicación por un factor de escalado de cuantificación hasta la aplicación de tablas de búsqueda complicadas bajo el control de un parámetro de cuantificación. El objetivo general es doble. En primer lugar, el proceso de cuantificación reduce el número de valores posibles de los datos transformados. En segundo lugar, el proceso de cuantificación puede aumentar la probabilidad de que los valores de los datos transformados sean cero. Ambos pueden hacer que el proceso de codificación por entropía, que se describirá a continuación, funcione de manera más eficiente al generar pequeñas cantidades de datos de vídeo comprimidos.

Una unidad de escaneo 360 aplica un proceso de escaneo de datos. El propósito del proceso de escaneo es reordenar los datos transformados cuantificados para reunir tantos como sea posible de coeficientes transformados cuantificados distintos de cero y, por supuesto, reunir tantos coeficientes de valor cero como sea posible. Estas características pueden permitir una aplicación eficiente de la denominada codificación por longitud de series o técnicas similares. Por lo tanto, el proceso de escaneo implica seleccionar coeficientes de los datos transformados cuantificados y, en particular, de un bloque de coeficientes que corresponde a un bloque de datos de imagen que se ha transformado y cuantificado, según un "orden de escaneo", de modo que (a) todos los coeficientes se seleccionan una vez como parte del escaneo, y (b) el escaneo tiende a proporcionar el reordenamiento deseado. Un ejemplo de orden de escaneo que puede tender a dar resultados útiles es el denominado orden de escaneo diagonal arriba a la derecha.

Los coeficientes escaneados se pasan entonces a un codificador 370 por entropía (EE). De nuevo, se pueden utilizar varios tipos de codificación por entropía. Dos ejemplos son las variantes del denominado sistema CABAC (codificación aritmética binaria adaptativa al contexto) y las variantes del denominado sistema CAVLC (codificación de longitud variable adaptativa al contexto). En términos generales, se considera que la CABAC proporciona una mayor eficiencia y, en algunos estudios, se ha demostrado que proporciona una reducción del 10 al 20 % en la cantidad de datos de salida codificados para una calidad de imagen comparable en comparación con la CAVLC. Sin embargo, se considera que CAVLC representa un nivel de complejidad mucho menor (en términos de su implementación) que CABAC. Cabe destacar que el proceso de escaneo y el proceso de codificación por entropía se muestran como procesos separados, pero, de hecho, se pueden combinar o tratar juntos. Es decir, la lectura de datos en el codificador de entropía puede tener lugar en el orden de escaneo. Las consideraciones correspondientes se aplican a los respectivos procesos inversos que se describirán a continuación.

La salida del codificador 370 por entropía, junto con los datos adicionales (mencionados anteriormente y/o expuestos a continuación), que definen, por ejemplo, la manera en que el predictor 320 generó la imagen predicha, proporciona una señal 380 de vídeo de salida comprimida.

Sin embargo, también se proporciona una ruta de retorno porque el funcionamiento del propio predictor 320 depende de una versión descomprimida de los datos de salida comprimidos.

El motivo de esta función es el siguiente. En la etapa apropiada del proceso de descompresión (que se describirá más adelante) se genera una versión descomprimida de los datos residuales. Estos datos residuales descomprimidos deben añadirse a una imagen predicha para generar una imagen de salida (porque los datos residuales originales eran la diferencia entre la imagen de entrada y una imagen predicha). Para que este proceso sea comparable, como entre el lado de compresión y el lado de descompresión, las imágenes predichas generadas por el predictor 320 deberían ser las mismas durante el proceso de compresión y durante el proceso de descompresión. Por supuesto, en la descompresión, el aparato no tiene acceso a las imágenes de entrada originales, sino solo a las imágenes descomprimidas. Por lo tanto, en la compresión, el predictor 320 basa su predicción (al menos, para la codificación entre imágenes) en versiones descomprimidas de las imágenes comprimidas.

Se considera que el proceso de codificación por entropía llevado a cabo por el codificador 370 por entropía (al menos en algunos ejemplos) es "sin pérdidas", es decir, que puede invertirse para llegar exactamente a los mismos datos que se suministraron en primer lugar vez al codificador 370 por entropía. Por lo tanto, en estos ejemplos, la ruta de retorno se puede implementar antes de la etapa de codificación por entropía. De hecho, también se considera que el proceso de escaneo llevado a cabo por la unidad 360 de escaneo también es sin pérdidas, pero en la presente realización la ruta 390 de retorno va desde la salida del cuantificador 350 hasta la entrada de un cuantificador 420 inverso complementario. En los casos en los que una etapa introduce una pérdida o pérdida potencial, esa etapa puede incluirse en el bucle de retroalimentación formado por la ruta de retorno. Por ejemplo, la etapa de codificación por entropía puede, al menos en principio, hacerse con pérdidas, por ejemplo, mediante técnicas en las que los bits

se codifican dentro de la información de paridad. En tal caso, la codificación y decodificación por entropía deberían formar parte del bucle de retroalimentación.

5 En términos generales, un decodificador 410 por entropía, la unidad 400 de escaneo inverso, un cuantificador inverso 420 y una unidad o circuitería 430 de transformación inversa proporcionan las funciones inversas respectivas del codificador 370 por entropía, la unidad 360 de escaneo, el cuantificador 350 y la unidad 340 de transformación. Por ahora, la discusión continuará durante el proceso de compresión; el proceso para descomprimir una señal de vídeo comprimida de entrada se expondrá a continuación, por separado.

10 En el proceso de compresión, los coeficientes escaneados se pasan por la ruta 390 de retorno desde el cuantificador 350 hasta el cuantificador inverso 420 que lleva a cabo la operación inversa de la unidad 360 de escaneo. Las unidades 420, 430 llevan a cabo un proceso de cuantificación inversa y transformación inversa para generar una señal 440 de imagen residual comprimida y descomprimida.

15 La señal de imagen 440 se añade, en un sumador 450, a la salida del predictor 320 para generar una imagen de salida reconstruida 460. Esto forma una entrada al predictor 320 de imágenes, como se describirá a continuación.

20 Volviendo ahora al proceso aplicado para descomprimir una señal 470 de vídeo comprimida recibida, se suministra la señal al decodificador 410 por entropía y desde allí a la cadena de la unidad 400 de escaneo inverso, el cuantificador inverso 420 y la unidad 430 de transformada inversa antes de añadirse a la salida del predictor 320 de imágenes mediante el sumador 450. Por lo tanto, en el lado del decodificador, el decodificador reconstruye una versión de la imagen residual y después la aplica (mediante el sumador 450) a la versión predicha de la imagen (bloque a bloque) para decodificar cada bloque. En términos sencillos, la salida 460 del sumador 450 forma la señal 480 de vídeo descomprimida de salida. En la práctica, se puede aplicar opcionalmente un filtrado adicional (por ejemplo, mediante un filtro 560 mostrado en la figura 8, pero omitido en la figura 7 para mayor claridad del diagrama de nivel superior de la figura 7) antes de emitir la señal.

30 El aparato de las figuras 7 y 8 puede actuar como un aparato de compresión (codificación) o un aparato de descompresión (decodificación). Las funciones de los dos tipos de aparatos se superponen sustancialmente. La unidad 360 de escaneo y el codificador 370 por entropía no se utilizan en un modo de descompresión, y el funcionamiento del predictor 320 (que se describirá en detalle a continuación) y de otras unidades siguen la información de modo y parámetro contenida en el flujo de bits comprimido recibido en lugar de generar tal información por sí mismas.

35 La figura 8 ilustra esquemáticamente la generación de imágenes predichas y, en particular, el funcionamiento del predictor 320 de imágenes.

40 Hay dos modos básicos de predicción que lleva a cabo el predictor 320 de imágenes: la denominada predicción dentro de la imagen y la denominada predicción entre imágenes o con compensación de movimiento (CM). En el lado del codificador, cada una implica detectar una dirección de predicción con respecto a un bloque actual que se va a predecir y generar un bloque predicho de muestras según otras muestras (en la misma imagen (intra) u otra (inter)). En virtud de las unidades 310 o 450, la diferencia entre el bloque predicho y el bloque real se codifica o aplica para codificar o decodificar el bloque, respectivamente.

45 (En el decodificador, o en el lado de decodificación inversa del codificador, la detección de una dirección de predicción puede ser en respuesta a los datos asociados con los datos codificados por el codificador, lo que indica qué dirección se utilizó en el codificador. O la detección puede ser en respuesta a los mismos factores que aquellos sobre los que se tomó la decisión en el codificador).

50 La predicción entre imágenes basa la predicción del contenido de un bloque o región de la imagen en los datos de dentro de la misma imagen. Esto corresponde a la denominada codificación de fotogramas en I en otras técnicas de compresión de vídeo. Sin embargo, a diferencia de la codificación de fotogramas en I, que implica codificar toda la imagen por intracodificación, en las presentes realizaciones la elección entre intracodificación e intercodificación se puede hacer bloque por bloque, aunque en otras realizaciones la elección se sigue haciendo imagen por imagen.

55 La predicción con compensación de movimiento es un ejemplo de predicción entre imágenes y utiliza información de movimiento que intenta definir la fuente, en otra imagen adyacente o cercana, del detalle de la imagen que se codificará en la imagen actual. Por consiguiente, en un ejemplo ideal, el contenido de un bloque de datos de imagen en la imagen predicha se puede codificar de manera muy simple como una referencia (un vector de movimiento) que apunta a un bloque correspondiente en la misma posición o en una posición ligeramente diferente de una imagen adyacente.

60 Una técnica conocida como predicción de "copia en bloque" es, en algunos aspectos, un híbrido de las dos, ya que utiliza un vector para indicar un bloque de muestras en una posición desplazada del bloque predicho actualmente dentro de la misma imagen, que debe copiarse para formar el bloque predicho actualmente.

65 Volviendo a la figura 8, se muestran dos disposiciones de predicción de imágenes (correspondientes a la predicción dentro de la imagen y entre imágenes), cuyos resultados son seleccionados por un multiplexor 500 bajo el control de

una señal 510 de modo (por ejemplo, desde el controlador 343) para proporcionar bloques de la imagen predicha para su suministro a los sumadores 310 y 450. La elección se hace en función de qué selección produzca la “energía” más baja (lo que, como se ha expuesto anteriormente, puede considerarse como contenido de información que requiere codificación), y la elección se señala al decodificador dentro del flujo de datos de salida codificados. La energía de la imagen, en este contexto, puede detectarse, por ejemplo, realizando una sustracción de prueba de un área de las dos versiones de la imagen predicha a partir de la imagen de entrada, elevando al cuadrado cada valor de píxel de la imagen de diferencia, sumando los valores cuadrados e identificando cuál de las dos versiones da lugar al valor cuadrático medio inferior de la imagen de diferencia en relación con esa área de la imagen. En otros ejemplos, se puede llevar a cabo una codificación de prueba para cada selección o selección potencial, realizándose después una elección según el coste de cada selección potencial en términos de uno o ambos de entre el número de bits requeridos para la codificación y la distorsión de foto.

La predicción real, en el sistema de intracodificación, se realiza basándose en bloques de imágenes recibidos como parte de la señal 460, es decir, la predicción se basa en bloques de imágenes codificadas y decodificadas para que pueda hacerse exactamente la misma predicción en un aparato de descompresión. Sin embargo, los datos pueden derivarse de la señal de vídeo de entrada 300 mediante un selector 520 de modo intra para controlar el funcionamiento del predictor 530 dentro de la imagen.

Para la predicción entre imágenes, un predictor con compensación de movimiento (CM) 540 utiliza información de movimiento, tal como vectores de movimiento derivados por un estimador de movimiento 550 a partir de la señal 300 de vídeo de entrada. Esos vectores de movimiento se aplican a una versión procesada de la imagen reconstruida 460 mediante el predictor 540 con compensación de movimiento para generar bloques de la predicción entre imágenes.

En consecuencia, cada una de las unidades 530 y 540 (que funcionan con el estimador 550) actúa como detectores para detectar una dirección de predicción con respecto a un bloque actual que se va a predecir, y como un generador para generar un bloque predicho de muestras (que forma parte de la predicción que se pasa a las unidades 310 y 450) según otras muestras definidas por la dirección de predicción.

Ahora se describirá el procesamiento aplicado a la señal 460. En primer lugar, la señal se filtra opcionalmente mediante una unidad de filtro 560, que se describirá con mayor detalle a continuación. Esto implica aplicar un filtro de “desbloqueo” para eliminar o al menos tender a reducir los efectos del procesamiento basado en bloques llevado a cabo por la unidad 340 de transformada y las operaciones posteriores. También se puede utilizar un filtro de compensación adaptativa de muestras (SAO). También, se aplica opcionalmente un filtro de bucle adaptativo utilizando coeficientes derivados al procesar la señal reconstruida 460 y la señal 300 de vídeo de entrada. El filtro de bucle adaptativo es un tipo de filtro que, utilizando técnicas conocidas, aplica coeficientes de filtro adaptativo a los datos que se van a filtrar. Es decir, los coeficientes de filtro pueden variar en función de varios factores. Los datos que definen qué coeficientes de filtro utilizar se incluyen como parte del flujo de datos de salida codificado.

La salida filtrada de la unidad de filtro 560 forma, de hecho, la señal 480 de vídeo de salida cuando el aparato funciona como un aparato de descompresión. También se almacena en memoria intermedia en uno o más almacenes 570 de imágenes o fotogramas; el almacenamiento de imágenes sucesivas es un requisito del procesamiento de predicción con compensación de movimiento y, en particular, de la generación de vectores de movimiento. Para ahorrar en los requisitos de almacenamiento, las imágenes almacenadas en los almacenes 570 de imágenes se pueden mantener en forma comprimida y después descomprimirse para utilizarse en la generación de vectores de movimiento. Para este propósito particular, se puede utilizar cualquier sistema de compresión/descompresión conocido. Las imágenes almacenadas se pasan a un filtro de interpolación 580 que genera una versión de mayor resolución de las imágenes almacenadas; en este ejemplo, las muestras intermedias (submuestras) se generan de tal modo que la resolución de la imagen interpolada emitida por el filtro 580 de interpolación es 4 veces (en cada dimensión) la de las imágenes almacenadas en los almacenes 570 de imágenes para el canal de luminancia de 4:2:0 y 8 veces (en cada dimensión) la de las imágenes almacenadas en los almacenes 570 de imágenes para los canales de crominancia de 4:2:0. Las imágenes interpoladas se pasan como una entrada al estimador 550 de movimiento y también al predictor 540 con movimiento compensado.

Ahora se describirá la forma en la que se divide una imagen para el procesamiento por compresión. En un nivel básico, una imagen que se va a comprimir se considera una matriz de bloques o regiones de muestras. La división de una imagen en dichos bloques o regiones puede llevarse a cabo mediante un árbol de decisiones, tal como el descrito en Bross y col.: “High Efficiency Video Coding (HEVC) text specification draft 6”, JCTVC-H1003_d0 (noviembre de 2011) y “Versatile Video Coding (Draft 5)”, JVET-N1001v8.

En algunos ejemplos, los bloques o regiones resultantes tienen tamaños y, en algunos casos, formas que, en virtud del árbol de decisiones, generalmente pueden seguir la disposición de las características de la imagen dentro de la imagen. Esto en sí mismo puede permitir una eficiencia de codificación mejorada porque las muestras que representan o siguen características de imagen similares tenderían a agruparse entre sí mediante una disposición de este tipo. En algunos ejemplos, se pueden seleccionar regiones o bloques cuadrados de diferentes tamaños (como muestras de 4 x 4 hasta, por ejemplo, bloques de 64 x 64 o más). En otros ejemplos de disposición, se pueden utilizar regiones o bloques de diferentes formas, tales como bloques rectangulares (por ejemplo, orientados vertical u horizontalmente).

Se contemplan otros bloques no cuadrados y no rectangulares. El resultado de la división de la imagen en tales bloques o regiones es (al menos en los presentes ejemplos) que cada muestra de una imagen se asigna a uno, y solo a uno, de dichos bloques o regiones.

5 Ahora se analizará el proceso de intrapredicción. En términos generales, la intrapredicción implica generar una predicción de un bloque actual de muestras a partir de muestras previamente codificadas y decodificadas en la misma imagen.

10 La figura 9 ilustra esquemáticamente una imagen 800 parcialmente codificada. En este caso, la imagen se codifica de arriba a la izquierda a abajo a la derecha bloque a bloque. Un ejemplo de bloque codificado a la mitad de la manipulación de la imagen completa se muestra como bloque 810. Ya se ha codificado una región sombreada 820 por encima y a la izquierda del bloque 810. La predicción dentro de la imagen del contenido del bloque 810 puede hacer uso de cualquiera de las áreas sombreadas 820, pero no puede hacer uso del área no sombreada por debajo de ella.

15 En algunos ejemplos, la imagen se codifica bloque a bloque, de tal modo que los bloques más grandes (denominados unidades de codificación o CU) se codifican en un orden tal como el orden descrito con referencia a la figura 9. Dentro de cada CU, existe la posible opción (dependiendo del proceso de división de bloques que haya tenido lugar) de que la CU se manipule como un conjunto de dos o más bloques o unidades de transformada (TU) más pequeños. Esto puede proporcionar un orden jerárquico de codificación para que la imagen se codifique CU a CU, y cada CU se codifique potencialmente TU a TU. Sin embargo, cabe señalar que para una TU individual dentro de la unidad de árbol de codificación actual (el nodo más grande de la estructura de árbol de la división de bloques), el orden jerárquico de codificación (CU a CU y después TU a TU) expuesto anteriormente significa que puede haber muestras previamente codificadas en la CU actual y disponibles para la codificación de esa TU que estén, por ejemplo, por encima de la derecha o por debajo de la izquierda de esa TU.

20 El bloque 810 representa una CU; como se ha expuesto anteriormente, para los propósitos del procesamiento de predicción dentro de la imagen, esto puede subdividirse en un conjunto de unidades más pequeñas. Un ejemplo de una TU 830 actual se muestra dentro de la CU 810. De manera más general, la foto se divide en regiones o grupos de muestras para permitir una codificación eficiente de la información de señalización y de los datos transformados. La señalización de la información puede requerir una estructura de árbol de subdivisiones diferente a la de la transformada y, de hecho, a la de la información de predicción o la propia predicción. Por esta razón, las unidades de codificación pueden tener una estructura de árbol diferente a la de los bloques o regiones de transformada, los bloques o regiones de predicción y la información de predicción. En algunos ejemplos, tal como el HEVC, la estructura puede ser el denominado árbol cuádruple de unidades de codificación, cuyos nodos hoja contienen una o más unidades de predicción y una o más unidades de transformada; las unidades de transformada pueden contener múltiples bloques de transformada correspondientes a las representaciones de luma y croma de la foto, y la predicción podría considerarse aplicable a nivel de bloque de transformación. En los ejemplos, se puede considerar que los parámetros aplicados a un grupo particular de muestras están definidos predominantemente a nivel de bloque, que potencialmente no tiene la misma granularidad que la estructura de transformada.

30 La predicción dentro de la imagen tiene en cuenta las muestras codificadas antes de que se considere la TU actual, tales como las que están por encima y/o a la izquierda de la TU actual. Las muestras fuente, a partir de las cuales se predicen las muestras requeridas, pueden estar ubicadas en diferentes posiciones o direcciones con respecto a la TU actual. Para decidir qué dirección es apropiada para una unidad de predicción actual, el selector 520 de modo de un ejemplo de codificador puede probar todas las combinaciones de estructuras TU disponibles para cada dirección candidata y seleccionar la dirección de predicción y la estructura TU con la mejor eficiencia de compresión.

35 La foto también puede codificarse en forma de "corte". En un ejemplo, un corte es un grupo de CU horizontalmente adyacentes. Pero en términos más generales, toda la imagen residual podría formar un corte, o un corte podría ser una única CU, o un corte podría ser una fila de CU, y así sucesivamente. Los cortes pueden ofrecer cierta resistencia a los errores, ya que se codifican como unidades independientes. Los estados del codificador y del decodificador se restablecen por completo en un límite del corte. Por ejemplo, la intrapredicción no se lleva a cabo a través de los límites de los cortes; para este propósito, los límites de los cortes se tratan como límites de imagen.

40 La figura 10 ilustra esquemáticamente un conjunto de posibles direcciones de predicción (candidatas). El conjunto completo de direcciones candidatas está disponible para una unidad de predicción. Las direcciones se determinan mediante un desplazamiento horizontal y vertical en relación con una posición de bloque actual, pero se codifican como "modos" de predicción, de los cuales se muestra un conjunto en la figura 11. Cabe destacar que el denominado modo DC representa una media aritmética simple de las muestras superiores e izquierdas circundantes. También cabe destacar que el conjunto de direcciones mostrado en la figura 10 es simplemente un ejemplo; en otros ejemplos, un conjunto de (por ejemplo) 65 modos angulares más DC y plano (un conjunto completo de 67 modos), como se muestra esquemáticamente en la figura 12, constituye el conjunto completo. Podrían utilizarse otros números de modos.

65 En términos generales, después de detectar una dirección de predicción, los sistemas son operativos para generar un bloque predicho de muestras según otras muestras definidas por la dirección de predicción. En unos ejemplos, el

codificador de imágenes está configurado para codificar datos que identifican la dirección de predicción seleccionada para cada muestra o región de la imagen (y el decodificador de imágenes está configurado para detectar dichos datos).

La figura 13 ilustra esquemáticamente un proceso de intrapredicción en el que una muestra 900 de un bloque o región 910 de muestras se deriva de otras muestras 920 de referencia de la misma imagen según una dirección 930 definida por el modo de intrapredicción asociado a esa muestra. Las muestras 920 de referencia de este ejemplo proceden de los bloques situados por encima y a la izquierda del bloque 910 en cuestión y el valor predicho de la muestra 900 se obtiene mediante un seguimiento a lo largo de la dirección 930 hasta las muestras 920 de referencia. La dirección 930 puede apuntar a una única muestra de referencia individual, pero en un caso más general se utiliza un valor interpolado entre las muestras de referencia circundantes como valor de predicción. Cabe destacar que el bloque 910 podría ser cuadrado, como se muestra en la figura 13, o podría tener otra forma, tal como rectangular.

Las técnicas son aplicables a tamaños de bloques variables y, de hecho, a formas de bloques.

Como se ha expuesto anteriormente con referencia a la figura 13, las matrices de muestras de referencia pueden extenderse más allá de la extensión del bloque que se va a predecir, con el fin de proporcionar modos o direcciones de predicción dentro del rango indicado en las figuras 10 a 12. Cuando sea necesario, si no hay muestras previamente decodificadas disponibles para su uso como muestras de referencia en posiciones particulares de muestra de referencia, se pueden reutilizar otras muestras de referencia en esas posiciones faltantes. Los procesos de filtrado de muestras de referencia se pueden utilizar en las muestras de referencia.

La figura 14 ilustra esquemáticamente el funcionamiento de un codificador por entropía CABAC.

El codificador CABAC funciona con respecto a datos binarios, es decir, datos representados solo por los dos símbolos 0 y 1. El codificador hace uso del denominado proceso de modelado de contexto que selecciona un "contexto" o modelo de probabilidad para los datos posteriores basándose en datos previamente codificados. La selección del contexto se lleva a cabo de forma determinista para que la misma determinación, sobre la base de datos previamente decodificados, pueda realizarse en el decodificador sin la necesidad de añadir más datos (especificando el contexto) al flujo de datos codificado que se pasa al decodificador.

Haciendo referencia a la figura 14, los datos de entrada que se van a codificar se pueden pasar a un convertidor binario 1400 si aún no están en forma binaria; si los datos ya están en forma binaria, se circunvala el convertidor 1400 (mediante un conmutador esquemático 1410). En las presentes realizaciones, la conversión a una forma binaria se lleva a cabo realmente expresando los datos del coeficiente DCT cuantificado como una serie de "mapas" binarios, que se describirán más adelante.

Los datos binarios pueden manipularse entonces mediante una de las dos rutas de procesamiento, una ruta "normal" y una ruta de "derivación" (que se muestran esquemáticamente como rutas separadas pero que, en las realizaciones de la invención que se analizan a continuación, de hecho, podrían implementarse mediante las mismas etapas de procesamiento, simplemente utilizando parámetros ligeramente diferentes). La ruta de derivación emplea un denominado codificador 1420 de derivación que no hace uso necesariamente del modelado de contexto de la misma forma que la ruta normal. En algunos ejemplos de codificación CABAC, esta ruta de derivación se puede seleccionar si se necesita un procesamiento particularmente rápido de un lote de datos, pero en las presentes realizaciones se observan dos características de los denominados datos de "derivación": en primer lugar, los datos de derivación son manipulados por el codificador CABAC (950, 1460), simplemente utilizando un modelo de contexto fijo que representa una probabilidad fija, por ejemplo, del 50 % (también denominado de contenedores "equiprobables" o "EP"); y en segundo lugar, los datos de derivación se refieren a determinadas categorías de datos, siendo un ejemplo particular los datos de signos de coeficiente. De otro modo, la ruta normal es seleccionada por los conmutadores esquemáticos 1430, 1440 que funcionan bajo el control del circuito de control 1435. Esto implica que los datos sean procesados por un modelador 1450 de contexto seguido de un motor 1460 de codificación.

El codificador por entropía mostrado en la figura 14 codifica un bloque de datos (es decir, por ejemplo, datos correspondientes a un bloque de coeficientes relacionados con un bloque de la imagen residual) como un valor único si el bloque está formado completamente por datos de valor cero. Para cada bloque que no entra en esta categoría, es decir, un bloque que contiene al menos algunos datos distintos de cero, se prepara un "mapa de significancia". El mapa de significancia indica si, para cada posición en un bloque de datos que se va a codificar, el coeficiente correspondiente en el bloque es distinto de cero. Los datos del mapa de significancia, al estar en forma binaria, están codificados con CABAC en sí mismos. El uso del mapa de significancia ayuda a la compresión porque no es necesario codificar ningún dato para un coeficiente con una magnitud que el mapa de significancia indica que es cero. Además, el mapa de significancia puede incluir un código especial para indicar el coeficiente final distinto de cero en el bloque, de modo que todos los coeficientes finales de alta frecuencia/ceros al final puedan omitirse de la codificación. El mapa de significancia va seguido, en el flujo de bits codificado, por datos que definen los valores de los coeficientes distintos de cero especificados por el mapa de significancia.

También se preparan y codifican niveles adicionales de datos de mapa. Un ejemplo es un mapa que define, como valor binario (1 = sí, 0 = no) si los datos del coeficiente en una posición del mapa, que el mapa de significancia ha

indicado que es “distinta de cero”, tienen realmente el valor de “uno”. Otro mapa especifica si los datos de coeficientes en una posición del mapa, que el mapa de significancia ha indicado que es “distinta de cero”, tienen realmente el valor de “dos”. Otro mapa indica, para aquellas posiciones del mapa en las que el mapa de significancia ha indicado que los datos del coeficiente son “distintos de cero”, si los datos tienen un valor “superior a dos”. Otro mapa indica, también para los datos identificados como “distintos de cero”, el signo del valor de los datos (utilizando una notación binaria predeterminada, tal como 1 para +, 0 para - o, por supuesto, al revés).

En las realizaciones de la invención, los mapas de significancia y los demás mapas se asignan de manera predeterminada al codificador CABAC o bien al codificador de derivación, y todos son representativos de diferentes atributos o rangos de valores respectivos de los mismos elementos de datos iniciales. En un ejemplo, al menos el mapa de significancia está codificado con CABAC y al menos algunos de los mapas restantes (como los datos de signos) están codificados por derivación. En consecuencia, cada elemento de datos se divide en subconjuntos de datos respectivos y los subconjuntos respectivos se codifican mediante un primer sistema de codificación (por ejemplo, CABAC) y un segundo (por ejemplo, de derivación). La naturaleza de los datos y de la codificación CABAC y de derivación es tal que, para una cantidad predeterminada de datos codificados con CABAC, se genera una cantidad variable de cero o más datos de derivación con respecto a los mismos elementos de datos iniciales. Así, por ejemplo, si los datos de DCT cuantificados y reordenados contienen sustancialmente todos los valores cero, entonces puede ser que no se generen datos de derivación o que se genere una cantidad muy pequeña de datos de derivación, porque los datos de derivación se refieren únicamente a las posiciones del mapa para las que el mapa de significancia ha indicado que el valor es distinto de cero. En otro ejemplo, en los datos DCT reordenados cuantificados que tienen muchos coeficientes de valor alto, se puede generar una cantidad significativa de datos de derivación.

En algunas realizaciones de la invención, el mapa de significancia y otros mapas se generan a partir de los coeficientes DCT cuantificados, por ejemplo, mediante la unidad 360 de escaneo, y se someten a un proceso de escaneo en zigzag (o a un proceso de escaneo seleccionado entre escaneo en zigzag, de rasterizado horizontal y de rasterizado vertical según el modo de intrapredicción) antes de someterse a una codificación CABAC.

En términos generales, la codificación CABAC implica predecir un contexto, o un modelo de probabilidad, para el siguiente bit que se va a codificar, basándose en otros datos previamente codificados. Si el siguiente bit es el mismo que el bit identificado como “más probable” por el modelo de probabilidad, entonces la codificación de la información de que “el siguiente bit concuerda con el modelo de probabilidad” se puede codificar con gran eficiencia. Es menos eficiente codificar cuando “el siguiente bit no concuerda con el modelo de probabilidad”, por lo que la derivación de los datos de contexto es importante para el buen funcionamiento del codificador. El término “adaptativo” significa que los modelos de contexto o probabilidad se adaptan, o varían durante la codificación, en un intento por proporcionar una buena coincidencia con los siguientes datos (aún no codificados).

Utilizando una analogía simple, en el idioma inglés escrito, la letra “U” es relativamente poco común. Pero en cuanto a la posición de la letra inmediatamente después de la letra “Q”, es muy común. Por lo tanto, un modelo de probabilidad puede establecer la probabilidad de una “U” como un valor muy bajo, pero si la letra actual es una “Q”, el modelo de probabilidad de una “U” como la siguiente letra podría establecerse como un valor de probabilidad muy alto.

La codificación CABAC se utiliza, en las presentes disposiciones, para al menos el mapa de significancia y los mapas que indican si los valores distintos de cero son uno o dos. El procesamiento de derivación, que en estas realizaciones es idéntico a la codificación CABAC, excepto por el hecho de que el modelo de probabilidad está fijado en una distribución de probabilidad igual (0,5:0,5) de 1 y 0, se utiliza para al menos los datos de signo y el mapa que indica si un valor es >2. Para aquellas posiciones de datos identificadas como >2, se puede utilizar una codificación de datos denominada de escape separada para codificar el valor real de los datos. Esto puede incluir una técnica de codificación de Golomb-Rice.

El proceso de codificación y modelado de contexto CABAC se describe con más detalle en WD4: Working Draft 4 of High-Efficiency Video Coding, JCTVC-F803_d5, Draft ISO/IEC 23008-HEVC; 201x(E) 2011-10-28.

Haciendo referencia ahora a las figuras 15 y 16, un codificador por entropía que forma parte de un aparato de codificación de vídeo comprende un primer sistema de codificación (por ejemplo, un sistema de codificación aritmética tal como un codificador CABAC 1500) y un segundo sistema de codificación (tal como un codificador de derivación 1510), dispuestos de modo que una palabra o valor de datos particular se codifique en el flujo de datos de salida final mediante el codificador CABAC o bien el codificador de derivación, pero no ambos. En algunas realizaciones de la invención, los valores de datos pasados al codificador CABAC y al codificador de derivación son subconjuntos respectivos de valores de datos ordenados, divididos o derivados de los datos de entrada iniciales (los datos DCT cuantificados reordenados en este ejemplo), que representan valores diferentes del conjunto de “mapas” generados a partir de los datos de entrada.

La representación esquemática de la figura 15 trata el codificador CABAC y el codificador de derivación como disposiciones separadas. Este puede ser el caso en la práctica, pero según otra posibilidad, que se muestra esquemáticamente en la figura 16, se utiliza el único codificador CABAC 1620 tanto como el codificador CABAC 1500 como el codificador de derivación 1510 de la figura 15. El codificador 1620 funciona bajo el control de una señal 1630

de selección de modo de codificación, para funcionar con un modelo de contexto adaptativo (como se ha descrito anteriormente) cuando está en el modo del codificador CABAC 1500, y para funcionar con un modelo de contexto de probabilidad fijo del 50 % cuando está en el modo del codificador 1510 de derivación.

5 Una tercera posibilidad combina estas dos, ya que dos codificadores CABAC sustancialmente idénticos pueden funcionar en paralelo (similar a la disposición en paralelo de la figura 15) con la diferencia de que el codificador CABAC que funciona como el codificador 1510 de derivación tiene su modelo de contexto fijado en un en un modelo de contexto de probabilidad del 50 %.

10 Las salidas del proceso de codificación CABAC y del proceso de codificación por derivación se pueden almacenar (al menos temporalmente) en las respectivas memorias intermedias 1540, 1550. En el caso de la figura 16, un conmutador o demultiplexor 1660 actúa bajo el control de la señal 1630 de modo para enrutar los datos codificados con CABAC a la memoria intermedia 1550 y derivar los datos codificados a la memoria intermedia 1540.

15 Por lo tanto, las figuras 14 a 16 proporcionan ejemplos de un codificador por entropía configurado para codificar selectivamente elementos de datos que representan datos de imagen mediante un sistema de codificación seleccionado entre un sistema CABAC y un sistema de derivación.

20 Las figuras 17 y 18 ilustran esquemáticamente ejemplos de un decodificador por entropía que forma parte de un aparato de decodificación de vídeo. Haciendo referencia a la figura 17, las respectivas memorias intermedias 1710, 1700 pasan datos a un decodificador CABAC 1730 y a un decodificador 1720 de derivación, dispuestos de modo que una palabra o valor de datos codificados particular sea decodificado por el decodificador CABAC o bien el decodificador de derivación, pero no por ambos. Los datos descodificados se reordenan mediante la lógica 1740 en el orden apropiado para las etapas de descodificación posteriores.

25 La representación esquemática de la figura 17 trata el decodificador CABAC y el decodificador de derivación como disposiciones separadas. Este puede ser muy bien el caso en la práctica, pero según otra posibilidad, que se muestra esquemáticamente en la figura 18, se utiliza un único decodificador CABAC 1850 tanto como decodificador CABAC 1730 como decodificador 1720 de derivación de la figura 17. El decodificador 1850 funciona bajo el control de una señal 1860 de selección de modo de decodificación, para funcionar con un modelo de contexto adaptativo (como se ha descrito anteriormente) cuando está en modo de decodificador CABAC 1730, y para funcionar con un modelo de contexto de probabilidad fijo del 50 % cuando está en modo de codificador 1720 de derivación.

30 Como antes, una tercera posibilidad combina estas dos, ya que dos descodificadores CABAC sustancialmente idénticos pueden funcionar en paralelo (similar a la disposición en paralelo de la figura 17) con la diferencia de que el decodificador CABAC que funciona como decodificador de derivación 1720 tiene su modelo de contexto fijo en un modelo de contexto de probabilidad del 50 %.

35 En el caso de la figura 18, un conmutador o multiplexor 1870 actúa bajo el control de la señal 1860 de modo para enrutar los datos codificados con CABAC al decodificador 1850 desde la memoria intermedia 1700 o la memoria intermedia 1710, según sea apropiado.

40 Por lo tanto, las figuras 14 a 16 proporcionan ejemplos de un decodificador por entropía configurado para decodificar selectivamente elementos de datos que representan datos de imagen mediante un sistema de decodificación seleccionado entre un sistema CABAC y un sistema de derivación.

45 Entre los elementos de datos que pueden generarse, codificarse y decodificarse en un sistema de manipulación de datos de imagen hay instancias de (i) un primer elemento de datos que indica si un modo de croma de modelo lineal multidireccional (MDLM, por sus siglas en inglés) se aplica a una región de imagen actual; (ii) un segundo elemento de datos que indica una dirección de procesamiento para el funcionamiento del modo de croma MDLM (tal como la utilización de muestras por encima o a la izquierda); y (iii) un indicador que indica un modo de intrapredicción plano.

50 En las disposiciones propuestas anteriormente (tal como las definidas en la última versión de especificación disponible en la fecha de prioridad de la presente solicitud, a saber, JVET-N1001-v8 (borrador 5), Bross y col., estos elementos (i) - (iii) se manipulan como datos codificados CABAC. En particular, con respecto a los elementos (i) y (ii), los dos elementos de datos se manipulan utilizando un contexto CABAC común (véase 9.5.4.2.1 y, en particular, la tabla 9.17 del documento JVET citado anteriormente).

55 La utilización de un contexto CABAC común para los elementos de datos (i) y (ii) puede conllevar posibles ineficiencias.

60 Por ejemplo, la utilización eficiente de un contexto CABAC se basa en el seguimiento por contexto de la verosimilitud de resultados particulares de la variable codificada, lo que se logra actualizando el contexto en respuesta a la codificación de la variable. En la situación de utilización del modo croma MDLM, si bien la utilización real de este modo puede tener una probabilidad particular para una imagen o región de imagen determinada, una vez que el codificador ha decidido utilizar el modo croma MDLM (elemento de datos (i)), la probabilidad de elegir una dirección de operación particular (elemento de datos (ii)) tiende a ser aproximadamente 50:50. Por lo tanto, al utilizar un contexto común para

65

5 los elementos de datos (i) y (ii), cualquier beneficio potencial de eficiencia de codificación logrado mediante el seguimiento de la probabilidad de utilización del modo croma MDLM se deshace potencialmente o, al menos, se contrarresta parcialmente siguiendo también con el mismo contexto la utilización sustancialmente igualmente probable de las diferentes direcciones de procesamiento (elemento de datos (ii)). O bien la variable de contexto resultante es menos eficiente de lo que podría ser para codificar el elemento de datos (i), o es inapropiada para codificar el elemento de datos que es sustancialmente igualmente probable (ii).

10 En un ejemplo de disposición, los elementos de datos primero y segundo se pueden expresar como sigue, con referencia a los denominados valores 4-6 de `intra_chroma_pred_mode`, tal como se definen en JVET-N1001-v8 (borrador 5):

intra_chroma_pred_mode	Primer elemento de datos		Segundo elemento de datos
	Indicador que indica Croma LM o no	indicador que indica MDLM o no	
4 (modo Croma LM = sí, pero MDLM = no)	Contexto CABAC 1	Método de codificación 2	(no está presente)
5 (modo Croma LM = sí y dirección MDLM = "izquierda")	Contexto CABAC 1	Método de codificación 2	Método de codificación 3
6 (croma LM = sí y dirección MDLM = "arriba")	Contexto CABAC 1	Método de codificación 2	Método de codificación 3

25 En algunos ejemplos de disposición, el método 2 de codificación es una codificación por derivación que utiliza contenedores EP. Por lo tanto, el elemento de datos (i) que abarca una indicación de que el modo de croma LM está en uso (teniendo en cuenta que el MDLM no es aplicable cuando el modo de croma LM no está en uso) y una indicación de si el MDLM también está en uso, utiliza al menos el contexto CABAC 1.

30 En algunos otros ejemplos de disposición, el método 2 de codificación es la codificación CABAC utilizando un segundo contexto CABAC 2. Por lo tanto, el primer elemento de datos que abarca una indicación de que el modo de croma LM está en uso (teniendo en cuenta que el MDLM no es aplicable cuando el modo de croma LM no está en uso) y una indicación de si el MDLM también está en uso, utiliza al menos el contexto CABAC 1 y también utiliza el contexto CABAC 2. En términos de la expresión de la presente técnica de tal modo que el primer elemento de datos se codifique utilizando al menos un primer contexto CABAC, en estos ejemplos de disposición, el primer contexto CABAC puede tomarse como Contexto 1 o bien como Contexto 2.

40 En algunos ejemplos de disposición, el método 3 de codificación utiliza un contexto CABAC adicional e independiente, pero en otros ejemplos el método 3 de codificación representa la codificación por derivación utilizando contenedores EP.

45 Se apreciará que la expresión "instancias de codificación de un elemento de datos" puede implicar comenzar con una representación inicial del elemento de datos y generar una representación codificada. De manera similar, el término "instancias de decodificación de un elemento de datos" puede implicar comenzar con una representación codificada del elemento de datos y generar una representación de salida.

50 Por lo tanto, en el lado de la codificación, el codificador por entropía está configurado para codificar instancias del primer elemento de datos y para codificar instancias del segundo elemento de datos independientemente de la codificación de las instancias del primer elemento de datos. En el lado de la decodificación, el decodificador por entropía está configurado para decodificar instancias del primer elemento de datos y para decodificar instancias del segundo elemento de datos independientemente de la decodificación de instancias del primer elemento de datos.

55 En al menos algunos de los ejemplos dados anteriormente, el codificador por entropía está configurado para codificar instancias del primer elemento de datos mediante el sistema de codificación CABAC utilizando al menos un primer contexto CABAC (por ejemplo, incluyendo la utilización del contexto CABAC 1) y para codificar instancias del segundo elemento de datos independientemente del primer contexto CABAC (por ejemplo, mediante el método 3 de codificación).

60 En al menos algunos ejemplos, el primer elemento de datos comprende datos indicativos de si un modo de predicción de croma en modo lineal se aplica a una región de imagen actual. En al menos algunos ejemplos, el codificador por entropía está configurado para codificar, utilizando el primer contexto CABAC, al menos los datos indicativos de si un modo de predicción de croma en modo lineal se aplica a la región de imagen actual, incluso si (como en algunos ejemplos) el indicador que indica si el MDLM está en uso se codifica utilizando un contexto CABAC adicional o un método de derivación.

65 Las consideraciones correspondientes se aplican en el lado de la decodificación.

En algunas realizaciones ilustrativas, el codificador por entropía de la figura 14 está configurado para codificar instancias del elemento de datos (i) mediante el sistema de codificación CABAC utilizando un primer contexto CABAC y para codificar instancias del elemento de datos (ii) independientemente del primer contexto CABAC.

5 En las realizaciones ilustrativas, esto se logra de varios modos:

En algunas realizaciones de la presente descripción, las instancias de los elementos de datos (i) y (ii) se codifican utilizando CABAC, pero con contextos separados.

10 En otras realizaciones de la presente descripción, las instancias del elemento de datos (i) se codifican utilizando CABAC y el primer contexto, pero las instancias del elemento de datos (ii) se codifican como datos de derivación, por ejemplo, mediante un sistema de codificación aritmética binaria (por ejemplo, 1420) utilizando un modelo fijo de probabilidad (EP) del 50 %

15 También se propone que, en las realizaciones ilustrativas, las instancias del elemento de datos (iii), donde el indicador indica un modo de intrapredicción plano, se manipulen como contenedores EP (derivación). Esta propuesta es independiente de los arreglos discutidos anteriormente en relación con los elementos de datos (i) y (ii).

20 La figura 19 es un diagrama de flujo esquemático que ilustra un método de codificación de datos de imagen, que comprende:

25 codificar selectivamente (en el paso 1900) elementos de datos que representan datos de imagen mediante un sistema de codificación seleccionado entre un sistema de codificación de codificación aritmética binaria adaptativa al contexto (CABAC) y un sistema de codificación por derivación;

30 en donde los elementos de datos incluyen (i) un primer elemento de datos que indica si un modo de croma de modelo lineal multidireccional (MDLM) se aplica en una región de imagen actual; y (ii) un segundo elemento de datos que indica una dirección de procesamiento para el funcionamiento del modo de croma MDLM;

en donde el paso de codificación comprende codificar instancias del primer elemento de datos y codificar instancias del segundo elemento de datos independientemente de la codificación de instancias del primer elemento de datos.

35 La figura 20 es un diagrama de flujo esquemático que ilustra un método de decodificación de datos de imagen, que comprende:

40 decodificar selectivamente (en un paso 2000) elementos de datos que representan datos de imagen mediante un sistema de decodificación seleccionado entre un sistema de decodificación aritmética binaria adaptativa al contexto (CABAC) y un sistema de decodificación por derivación, en donde los elementos de datos incluyen (i) un primer elemento de datos que indica si un modo de croma de modelo lineal multidireccional (MDLM) se aplica a una región de imagen actual; y (ii) un segundo elemento de datos que indica una dirección de procesamiento para el funcionamiento del modo de croma MDLM;

45 en donde el paso de decodificación comprende decodificar instancias de los primeros datos y decodificar instancias del segundo elemento de datos independientemente de la decodificación de instancias del primer elemento de datos.

50 En cada caso, las realizaciones de la descripción están representadas por un software informático que, cuando es ejecutado por un ordenador, hace que el ordenador lleve a cabo el método respectivo y mediante un medio de almacenamiento no transitorio legible por máquina que almacena dicho software informático. En el caso de los métodos de codificación, las realizaciones de la descripción se representan mediante una señal de datos que comprende datos codificados generados según el método respectivo.

55 En la medida en que se ha descrito que algunas realizaciones de la descripción se implementan, al menos en parte, mediante un aparato de procesamiento de datos controlado por software, se apreciará que un medio no transitorio legible por máquina que contenga dicho software, tal como un disco óptico, un disco magnético, una memoria de semiconductores o similar, también se considera que representa una realización de la presente descripción. De manera similar, también se considera que una señal de datos que comprende datos codificados generados según los métodos expuestos anteriormente (estén o no incorporados en un medio legible por máquina no transitorio) representa una realización de la presente descripción.

60 Será evidente que son posibles numerosas modificaciones y variaciones de la presente descripción a la luz de las enseñanzas anteriores. Por lo tanto, se debe entender que, dentro del alcance de las cláusulas adjuntas, la tecnología se puede poner en práctica de otro modo que no sea la descrita específicamente en la presente memoria.

65

REIVINDICACIONES

1. Aparato de codificación de datos de imagen, que comprende:
 - 5 un codificador por entropía configurado para codificar selectivamente elementos de datos que representan datos de imagen mediante un sistema de codificación seleccionado entre un sistema de codificación de codificación aritmética binaria adaptativa al contexto, CABAC, y un sistema de codificación por derivación;
 - 10 en donde los elementos de datos incluyen (i) un primer elemento de datos que indica si un modo de croma de modelo lineal multidireccional, MDLM, se aplica en una región de imagen actual; y (ii) un segundo elemento de datos que indica una dirección de procesamiento para el funcionamiento del modo de croma MDLM;
 - caracterizado porque**
 - 15 el codificador por entropía está configurado para codificar instancias del primer elemento de datos mediante el sistema de codificación CABAC utilizando un primer contexto CABAC y para codificar instancias del segundo elemento de datos independientemente del primer contexto CABAC.

2. El aparato de codificación de datos de imagen de la reivindicación 1, en donde el codificador por entropía está configurado para codificar instancias del segundo elemento de datos utilizando el sistema de codificación por derivación.

3. El aparato de codificación de datos de imagen de la reivindicación 2, en donde el sistema de codificación por derivación es un sistema de codificación aritmética binaria que utiliza un modelo de contexto de probabilidad fijo del 50 %.

4. El aparato de codificación de datos de imagen de la reivindicación 1, en donde el codificador por entropía está configurado para codificar instancias del segundo elemento de datos mediante el sistema de codificación CABAC utilizando un segundo contexto CABAC independiente del primer contexto CABAC.

5. El aparato de codificación de datos de imagen de la reivindicación 1, en donde el primer elemento de datos comprende datos indicativos de si un modo de predicción de croma en modo lineal se aplica a una región de imagen actual.

6. El aparato de codificación de datos de imagen de la reivindicación 1, en donde:
 - 40 el primer elemento de datos comprende datos indicativos de si un modo de predicción de croma en modo lineal se aplica a una región de imagen actual; y
 - el codificador por entropía está configurado para codificar utilizando el primer contexto CABAC al menos los datos indicativos de si un modo de predicción de croma en modo lineal se aplica a la región de imagen actual.

7. Un método de codificación de datos de imagen, que comprende:
 - 45 codificar selectivamente elementos de datos que representan datos de imagen mediante un sistema de codificación seleccionado entre un sistema de codificación de codificación aritmética binaria adaptativa al contexto, CABAC, y un sistema de codificación por derivación;
 - 50 en donde los elementos de datos incluyen (i) un primer elemento de datos que indica si un modo de croma de modelo lineal multidireccional, MDLM, se aplica en una región de imagen actual; y (ii) un segundo elemento de datos que indica una dirección de procesamiento para el funcionamiento del modo de croma MDLM;
 - caracterizado porque**
 - 55 el paso de codificación comprende codificar instancias del primer elemento de datos mediante el sistema de codificación CABAC utilizando un primer contexto CABAC y codificar instancias del segundo elemento de datos independientemente del primer contexto CABAC.

8. Una señal de datos que comprende datos codificados generados según el método de la reivindicación 7.

9. Aparato de decodificación de datos de imágenes, que comprende:
 - 60 un decodificador por entropía configurado para decodificar selectivamente elementos de datos que representan datos de imagen mediante un sistema de decodificación seleccionado de un sistema de decodificación de codificación aritmética binaria adaptativa al contexto, CABAC, y un sistema de decodificación por derivación;
 - 65 en donde los elementos de datos incluyen (i) un primer elemento de datos que indica si un modo de croma de modelo lineal multidireccional, MDLM, se aplica en una región de imagen

actual; y (ii) un segundo elemento de datos que indica una dirección de procesamiento para el funcionamiento del modo de croma MDLM;

caracterizado porque

el decodificador por entropía está configurado para decodificar instancias del primer elemento de datos mediante el sistema de decodificación CABAC utilizando un primer contexto CABAC y para decodificar instancias del segundo elemento de datos independientemente del primer contexto CABAC.

- 5
10. Aparato de almacenamiento, captura, transmisión o recepción de vídeo que comprende un aparato según la reivindicación 1 o 9.
- 10
11. Un método de decodificación de datos de imagen, que comprende:
- 15
- decodificar selectivamente elementos de datos que representan datos de imagen mediante un sistema de decodificación seleccionado entre un sistema de decodificación de codificación aritmética binaria adaptativa al contexto, CABAC, y un sistema de decodificación por derivación; en donde los elementos de datos incluyen (i) un primer elemento de datos que indica si un modo de croma de modelo lineal multidireccional, MDLM, se aplica en una región de imagen actual; y (ii) un segundo elemento de datos que indica una dirección de procesamiento para el funcionamiento del modo de croma MDLM;
- 20
- caracterizado porque**
- el paso de decodificación comprende decodificar instancias de los primeros datos mediante el sistema de decodificación CABAC utilizando un primer contexto CABAC y decodificar instancias del segundo elemento de datos independientemente del primer contexto CABAC.
- 25
12. Software informático que, cuando lo ejecuta un ordenador, hace que el ordenador lleve a cabo el método de la reivindicación 7 u 11.
- 30
13. Un medio de almacenamiento no transitorio legible por máquina que almacena el software informático de la reivindicación 12.
- 35
- 40
- 45
- 50
- 55
- 60
- 65

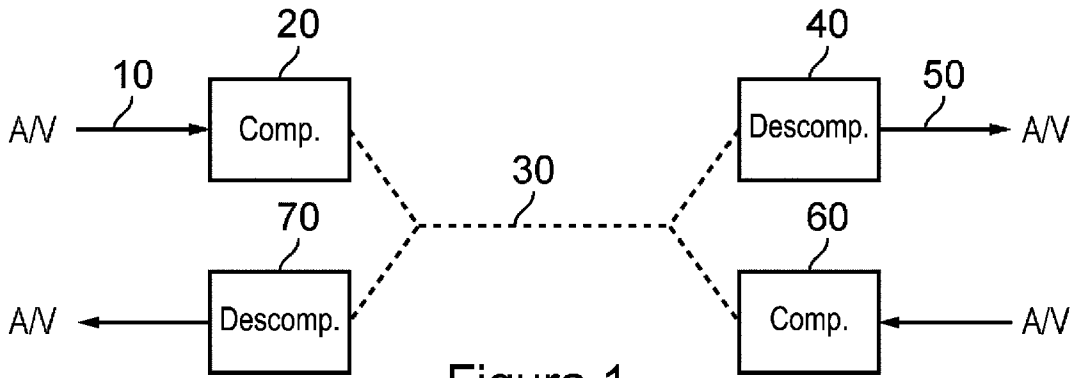


Figura 1

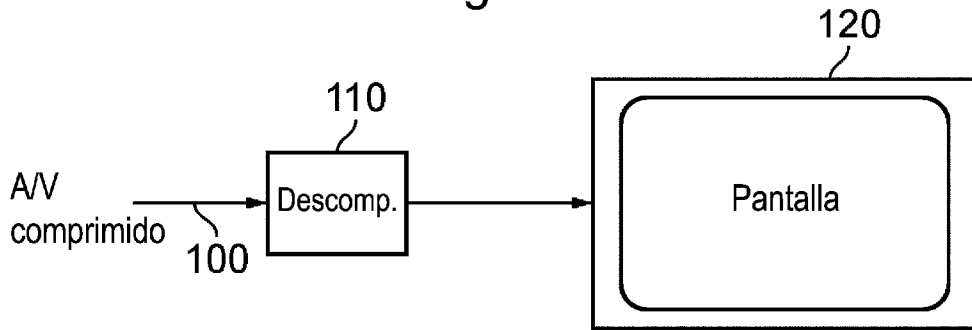


Figura 2

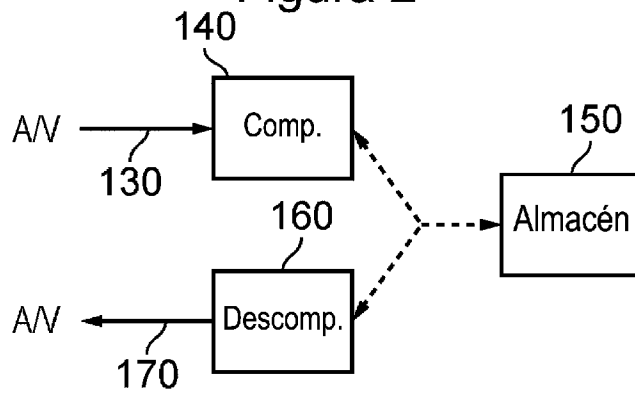


Figura 3

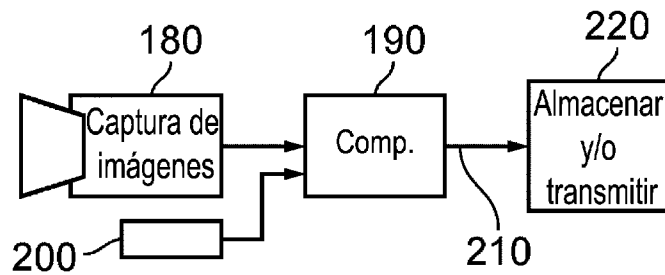


Figura 4

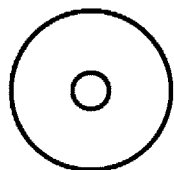


Figura 5

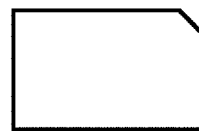


Figura 6

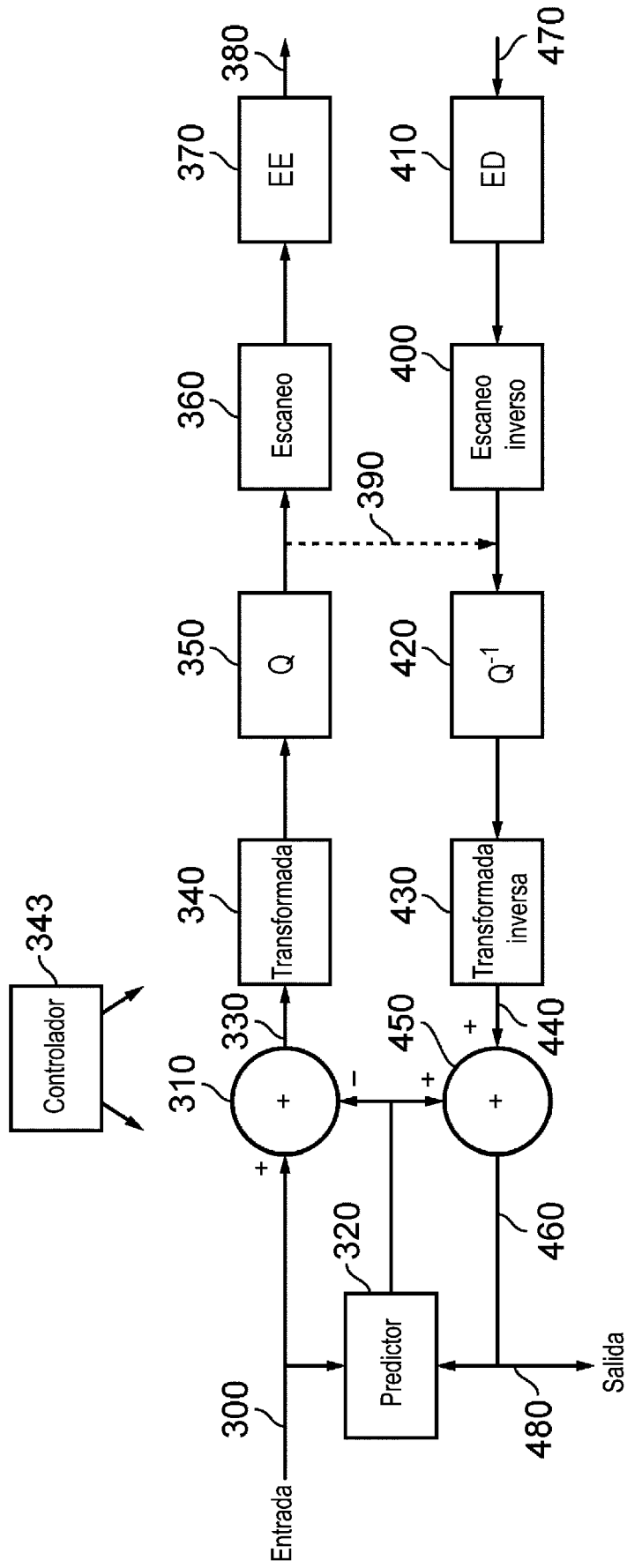


Figura 7

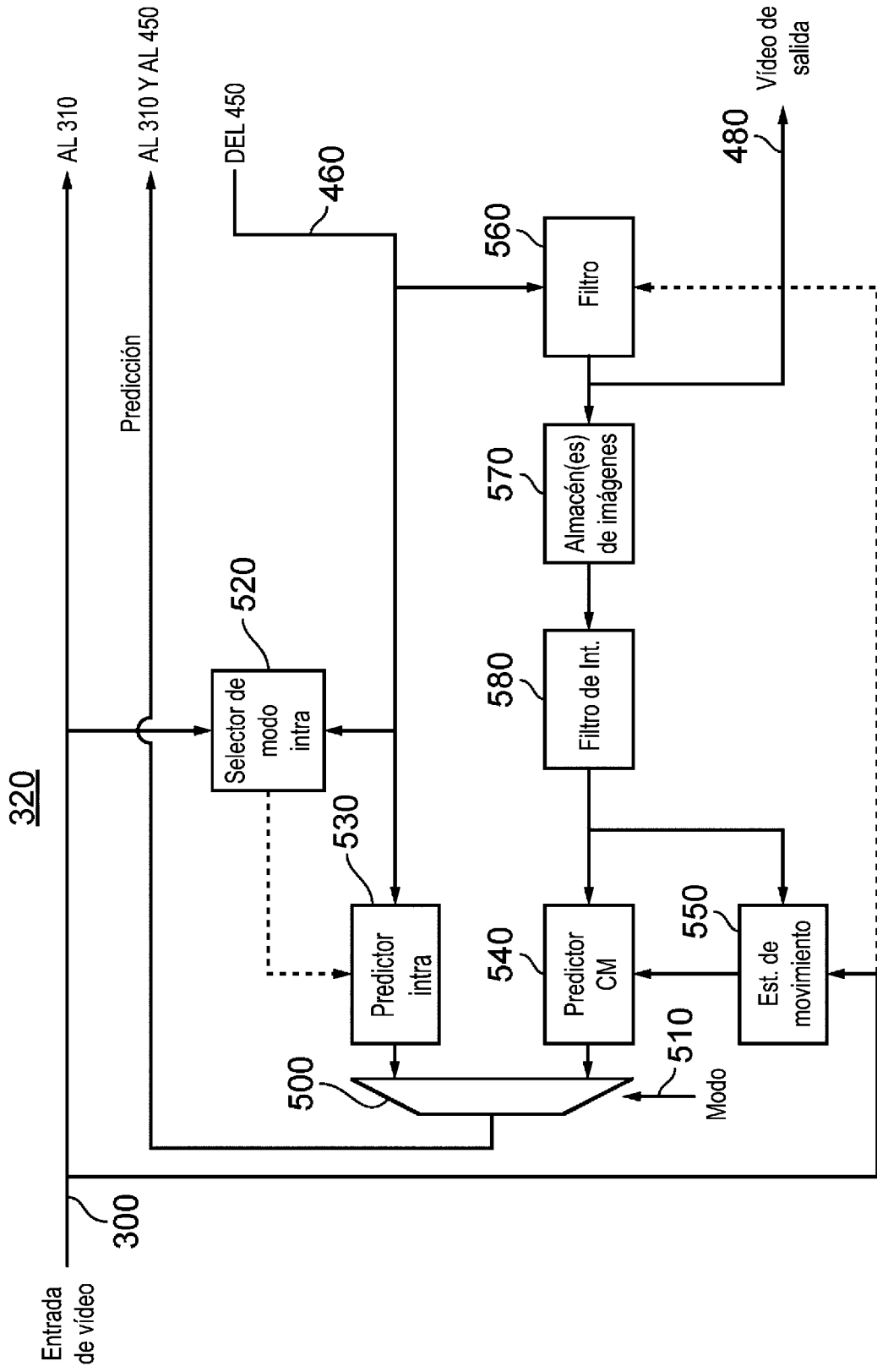


Figura 8

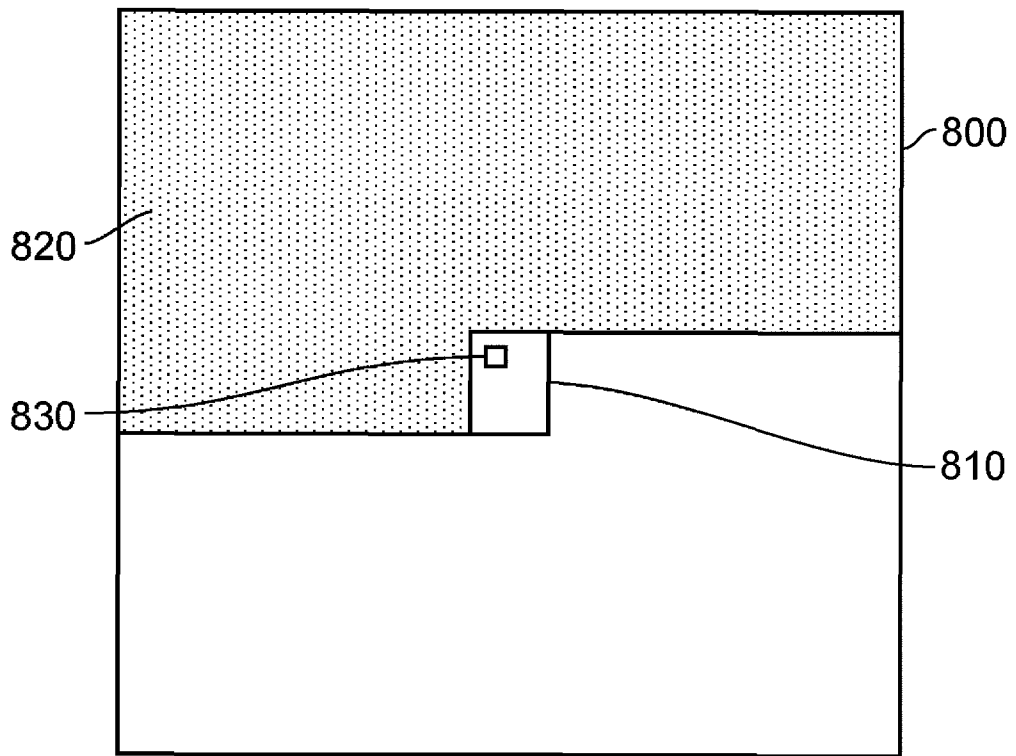


Figura 9

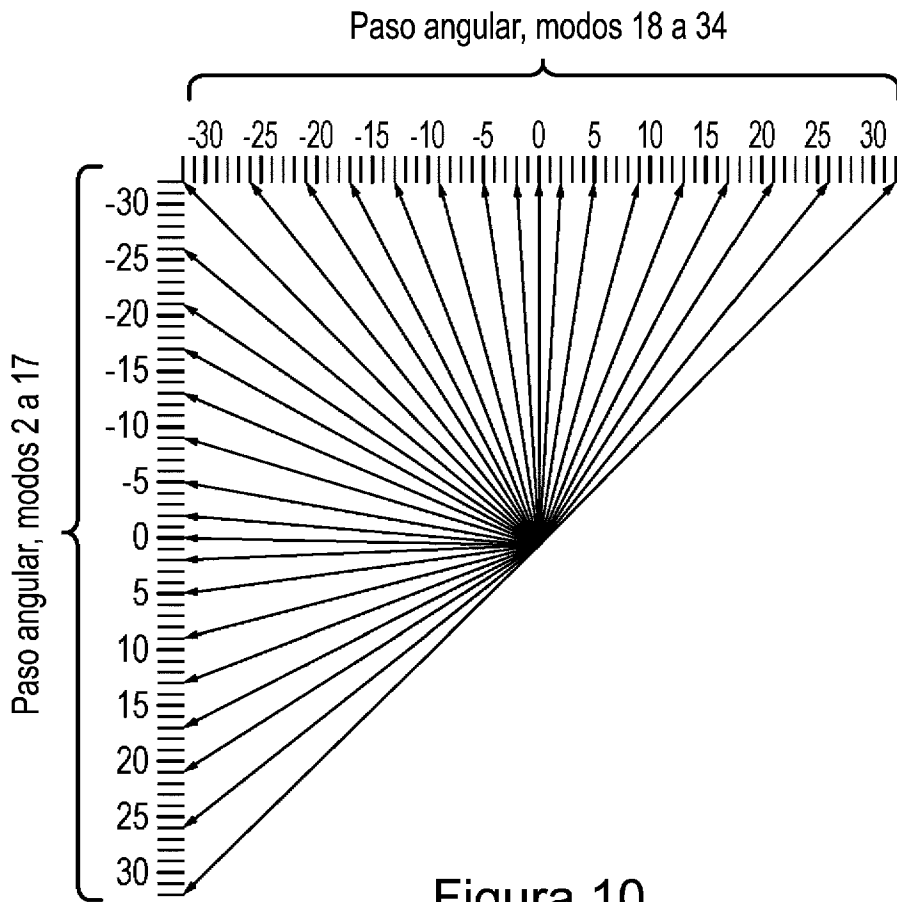


Figura 10

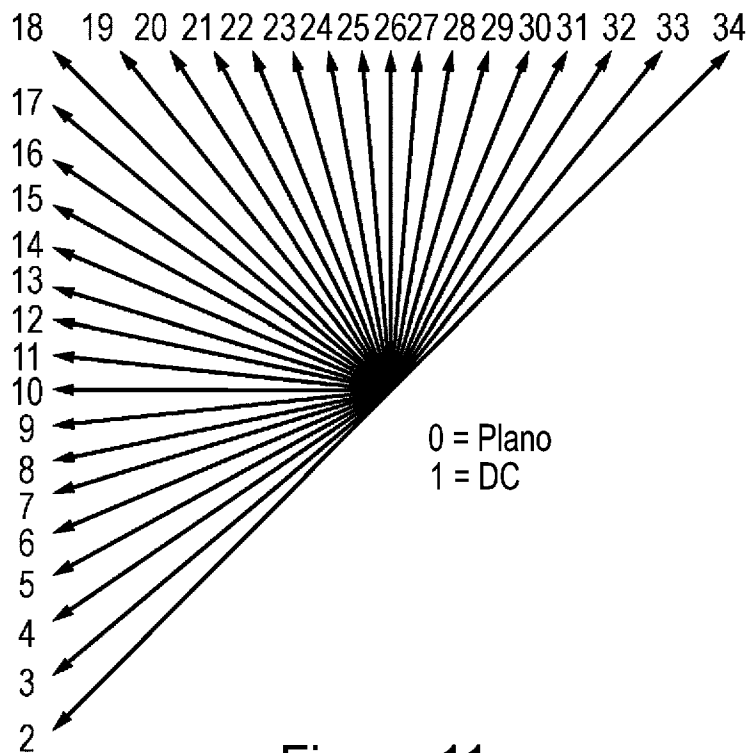


Figura 11

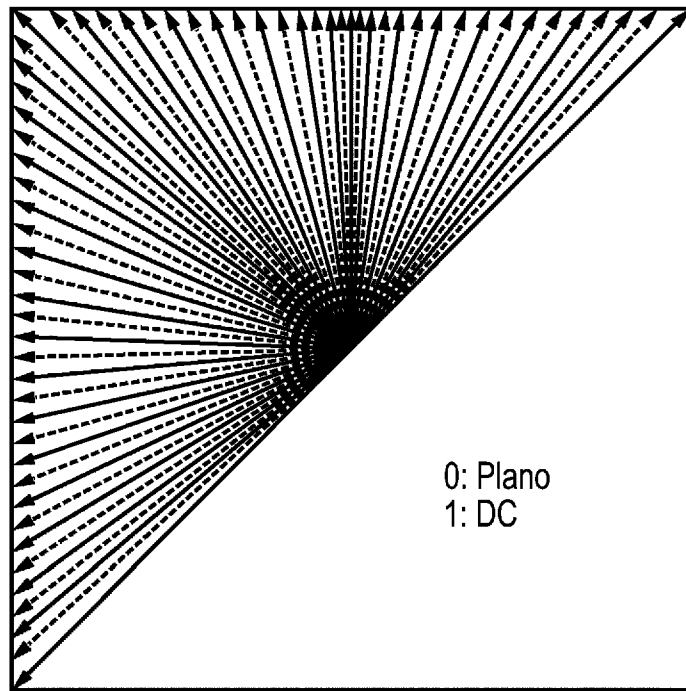


Figura 12

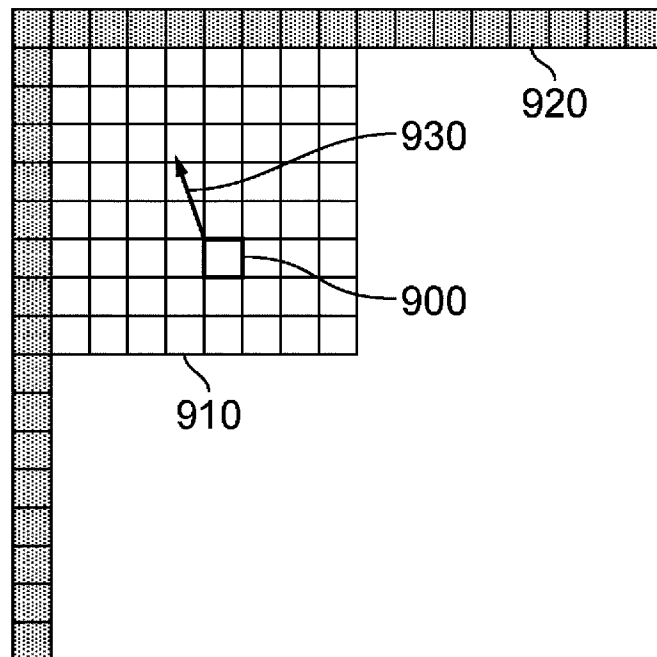


Figura 13

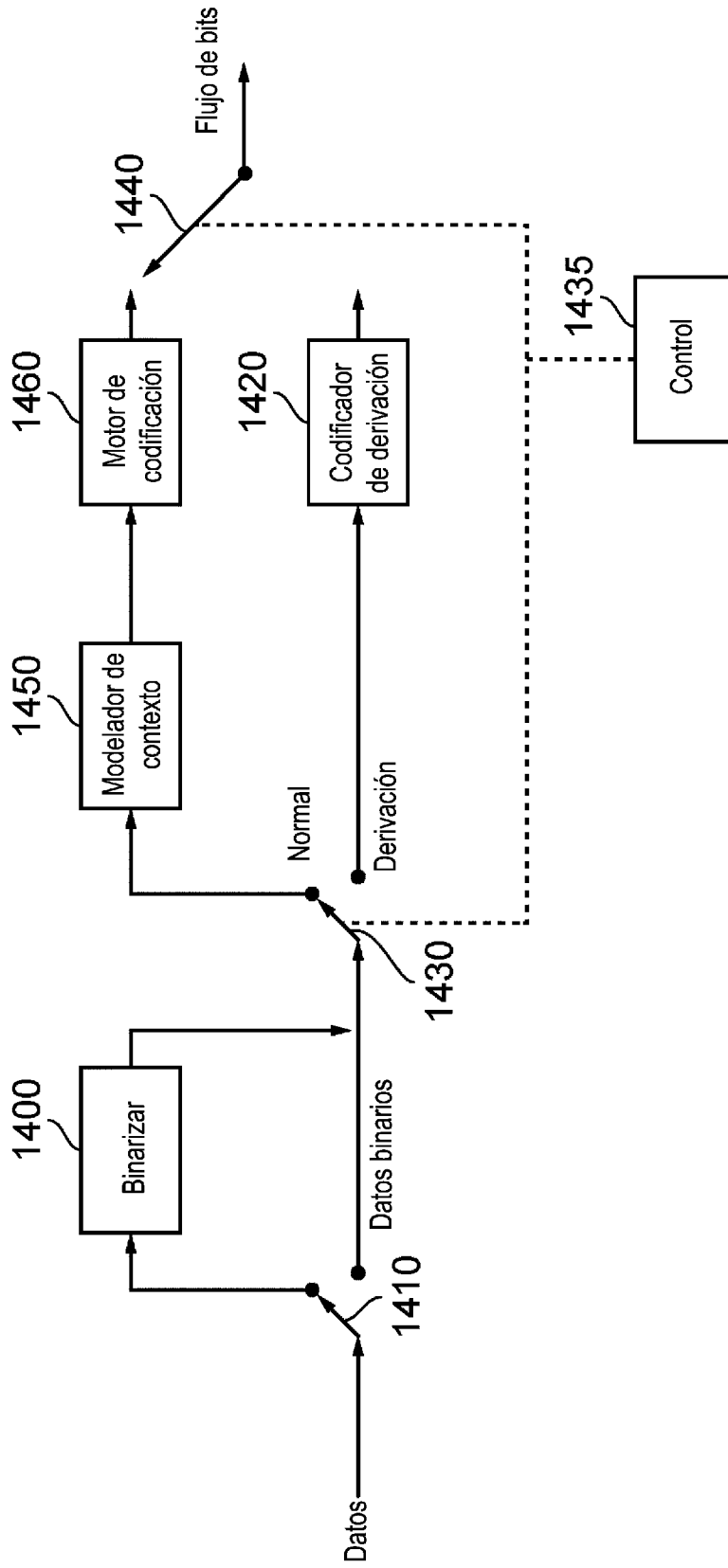


Figura 14

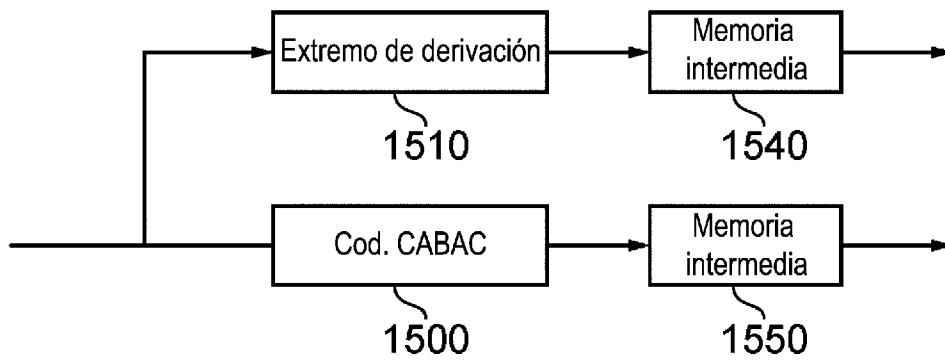


Figura 15

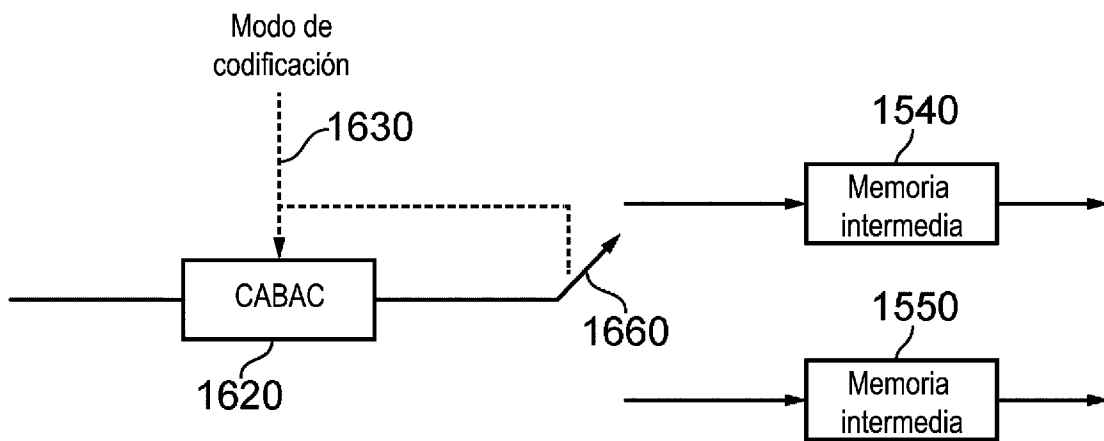


Figura 16

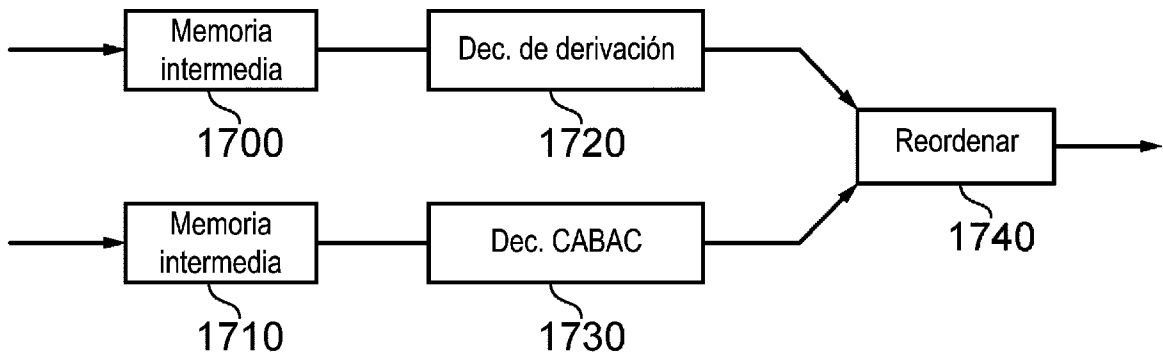


Figura 17

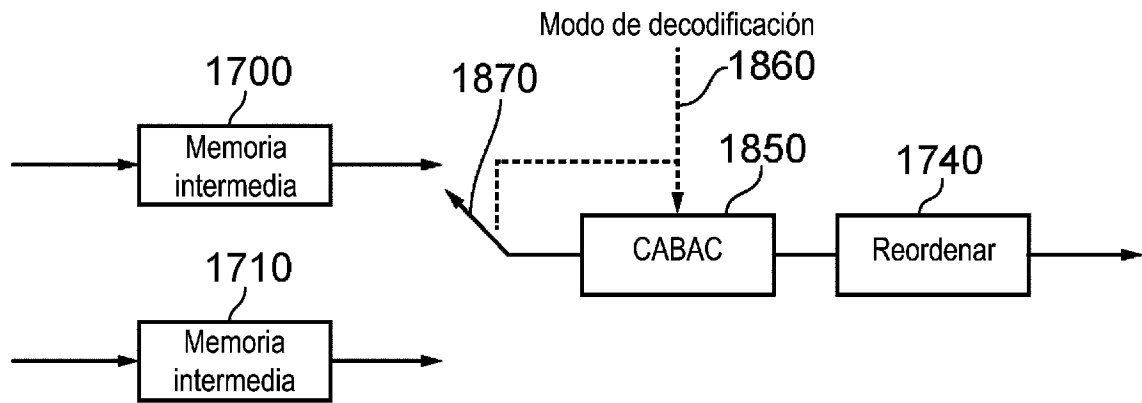


Figura 18



Figura 19



Figura 20