

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 767 932**

51 Int. Cl.:

H04N 19/103 (2014.01)

H04N 19/112 (2014.01)

H04N 19/16 (2014.01)

H04N 19/176 (2014.01)

H04N 19/70 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **03.09.2004 PCT/US2004/029032**

87 Fecha y número de publicación internacional: **24.03.2005 WO05027494**

96 Fecha de presentación y número de la solicitud europea: **03.09.2004 E 04783322 (3)**

97 Fecha y número de publicación de la concesión europea: **27.11.2019 EP 1665761**

54 Título: **Codificación y decodificación de plano de bits para el estado de predicción AC e información de tipo de codificación de campo/trama de macrobloque**

30 Prioridad:

07.09.2003 US 501081 P

02.09.2004 US 933959

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

19.06.2020

73 Titular/es:

MICROSOFT TECHNOLOGY LICENSING, LLC

(100.0%)

One Microsoft Way

Redmond, WA 98052, US

72 Inventor/es:

HSU, POHSIANG;

SRINIVASAN, SRIDHAR;

HOLCOMB, THOMAS, W;

MUKERJEE, KUNAL y

LIN, CHIH-LUNG

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 767 932 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 plano de bits para el estado de predicción AC e información de tipo de codificación de campo/trama de macrobloque

5 Una parte de la divulgación de este documento de patente contiene material sujeto a protección de derechos de autor. El propietario de los derechos de autor no tiene ninguna objeción a la reproducción facsímil por parte de cualquiera de la divulgación de la patente, tal como aparece en los archivos o registros de patentes de la Oficina de Patentes y Marcas, pero por lo demás se reserva todos los derechos de copyright.

10 Se describen técnicas y herramientas para la codificación y decodificación de vídeo. Por ejemplo, un codificador de plano de bits codifica información de estado de predicción AC. Como otro ejemplo, un codificador señala el tipo de transformación de campo/trama para macrobloques en una imagen codificada de trama entrelazada. Un decodificador realiza la decodificación correspondiente.

15 El vídeo digital consume grandes cantidades de almacenamiento y capacidad de transmisión. Una secuencia de vídeo digital en bruto típica incluye 15 o 30 imágenes por segundo. Cada imagen puede incluir decenas o cientos de miles de píxeles (también llamados pels). Cada píxel representa un pequeño elemento de la imagen. En bruto, una computadora comúnmente representa un píxel con 24 bits o más. Por lo tanto, la cantidad de bits por segundo, o tasa de bits, de una secuencia de vídeo digital en bruto típica puede ser de 5 millones de bits/segundo o más.

20 La mayoría de las computadoras y las redes de computadoras carecen de los recursos para procesar vídeo digital en bruto. Por esta razón, los ingenieros usan la compresión (también llamada codificación o cifrado) para reducir la tasa de bits del vídeo digital. La compresión puede ser sin pérdidas, en la cual la calidad del vídeo no sufre pero las disminuciones en la tasa de bits están limitadas por la complejidad del vídeo. O, la compresión puede ser con pérdida, en la cual la calidad del vídeo sufre pero las disminuciones en la tasa de bits son más dramáticas. La descompresión invierte la compresión.

25 En general, las técnicas de compresión de vídeo incluyen compresión "intra" y compresión "inter" o predictiva. Para tramas de vídeo, las técnicas de intracompresión comprimen tramas individuales, normalmente llamadas tramas I o tramas clave. Las técnicas de intercompresión comprimen tramas con referencia a tramas anteriores y/o siguientes, y las tramas intercomprimidas generalmente se llaman tramas previstas, tramas P o tramas B.

I. Inter e Intracompresión en Windows Media Video, Versiones 8 y 9

30 Windows Media Video de Microsoft Corporation, versión 8 ["WMV8"] incluye un codificador de vídeo y un decodificador de vídeo. El codificador WMV8 usa compresión intra e inter, y el decodificador WMV8 usa descompresión intra e inter. Windows Media Video, versión 9 ["WMV9"] usa una arquitectura similar para muchas operaciones.

A. Intracompresión

35 La Figura 1A ilustra la intracompresión 100 basada en bloques de un bloque 105 de píxeles en una trama clave en el codificador WMV8. Un bloque es un conjunto de píxeles, por ejemplo, una disposición de píxeles de 8x8. El codificador WMV8 divide una trama de vídeo clave en bloques de píxeles de 8x8 y aplica una Transformación discreta de coseno 8x8 ["DCT"] 110 a bloques individuales como el bloque 105. Una DCT es un tipo de transformación de frecuencia que convierte el bloque de píxeles 8x8 (información espacial) en un bloque 8x8 de coeficientes DCT 115, que son información de frecuencia. La operación DCT en sí es sin pérdidas o casi sin pérdidas. En comparación con los valores de píxeles originales, sin embargo, los coeficientes DCT son más eficientes para que el codificador comprima ya que la mayor parte de la información significativa se concentra en coeficientes de baja frecuencia (convencionalmente, la esquina superior izquierda del bloque 115) y muchos de los coeficientes de alta frecuencia (convencionalmente, la esquina inferior derecha del bloque 115) tiene valores de cero o cercanos a cero.

40 El codificador cuantifica 120 los coeficientes DCT, resultando en un bloque 8x8 de coeficientes 125 DCT cuantificados. Por ejemplo, el codificador aplica un tamaño de etapa uniforme, de cuantificación escalar a cada coeficiente. La cuantificación es con pérdida. Dado que los coeficientes DCT de baja frecuencia tienden a tener valores más altos, la cuantificación da como resultado una pérdida de precisión pero no una pérdida completa de la información de los coeficientes. Por otra parte, dado que los coeficientes DCT de alta frecuencia tienden a tener valores de cero o cercanos a cero, la cuantificación de los coeficientes de alta frecuencia típicamente da como resultado regiones contiguas de valores cero. Además, en algunos casos, los coeficientes DCT de alta frecuencia se cuantifican de manera más gruesa que los coeficientes DCT de baja frecuencia, resultando en una mayor pérdida de precisión/información para los coeficientes DCT de alta frecuencia.

45 El codificador luego prepara el bloque 8x8 de los coeficientes 125 DCT cuantificados para la codificación de entropía, que es una forma de compresión sin pérdidas. El tipo exacto de codificación de entropía puede variar dependiendo de si un coeficiente es un coeficiente DC (frecuencia más baja), un coeficiente AC (otras frecuencias) en la fila superior o columna izquierda, u otro coeficiente AC.

55 El codificador codifica el coeficiente DC 126 como un diferencial del coeficiente DC 136 de un bloque vecino de 8x8,

que es un vecino previamente codificado (por ejemplo, arriba o izquierda) del bloque que se codifica. (La Figura 1A muestra un bloque 135 vecino que está situado a la izquierda del bloque que se codifica en la trama.) El codificador codifica por entropía 140 el diferencial.

5 El codificador de entropía puede codificar la columna izquierda o la fila superior de coeficientes AC como diferencial de una columna izquierda o fila superior correspondiente del bloque vecino de 8x8. Este es un ejemplo de predicción del coeficiente AC. La Figura 1A muestra la columna izquierda 127 de coeficientes AC codificados como un diferencial 147 de la columna izquierda 137 del siguiente (en realidad, a la izquierda) bloque 135. La codificación diferencial aumenta la posibilidad de que los coeficientes diferenciales tengan valores cero. Los coeficientes AC restantes son del bloque 125 de coeficientes DCT cuantificados.

10 La Figura 1B muestra los candidatos de predicción AC para un bloque de 8x8 en una trama I. Para la predicción superior, la fila 177 superior de coeficientes AC en el bloque 175 adyacente superior se usa como el predictor para la fila 129 superior de coeficientes AC en el bloque 125 de coeficientes DCT cuantificados. Para la predicción izquierda, la columna 137 más a la izquierda de los coeficientes AC en el bloque 135 vecino izquierdo se usa como el predictor para la columna más a la izquierda de los coeficientes AC en el bloque 125.

15 En algunos modos, los predictores del coeficiente AC se escalan o se procesan antes de calcular o combinar con valores diferenciales.

20 Si un bloque vecino no existe en la dirección de predicción especificada, los valores pronosticados para los siete coeficientes AC en la columna más a la izquierda o en la fila superior se establecen en cero. Por ejemplo, si la dirección de predicción está hacia arriba y el bloque actual está en la fila superior, cada uno de los coeficientes AC previstos en la fila superior del bloque actual se establece en cero porque no hay un bloque adyacente en la dirección hacia arriba. Los coeficientes AC en la fila o columna pronosticada se agregan a los coeficientes AC decodificados correspondientes (que son valores diferenciales) en el bloque actual para producir el bloque de coeficiente de transformación cuantificado completamente reconstruido.

25 El codificador escanea 150 el bloque 145 de 8x8 de los coeficientes AC DCT cuantificados en una matriz 155 unidimensional y luego la codifica por entropía los coeficientes AC escaneados usando una variación de la codificación 160 de longitud de ejecución. El codificador selecciona un código de entropía de una o más tablas 165 de ejecución/nivel/último y genera el código de entropía.

B. Intercompresión

30 La intercompresión en el codificador WMV8 utiliza una codificación de predicción compensada por movimiento basada en bloques seguida de una codificación por transformación del error residual. Las Figuras 2 y 3 ilustran la intercompresión basada en bloques para una trama prevista en el codificador WMV8. En particular, la figura 2 ilustra la estimación de movimiento para una trama 210 prevista y la figura 3 ilustra la compresión de un residuo de predicción para un bloque compensado por movimiento de una trama prevista.

35 Por ejemplo, en la figura 2, el codificador WMV8 calcula un vector de movimiento para un macrobloque 215 en la trama 210 prevista. Para calcular el vector de movimiento, el codificador busca en un área 235 de búsqueda de una trama 230 de referencia. Dentro del área 235 de búsqueda, el codificador compara el macrobloque 215 de la trama 210 prevista con varios macrobloques candidatos para encontrar un macrobloque candidato que sea una buena coincidencia. El codificador emite información que especifica el vector de movimiento (codificado por entropía) para el macrobloque correspondiente. El vector de movimiento está codificado diferencialmente con respecto a un predictor de vector de movimiento.

40 Después de reconstruir el vector de movimiento agregando el diferencial al predictor del vector de movimiento, un decodificador usa el vector de movimiento para calcular un macrobloque de predicción para el macrobloque 215 usando información de la trama 230 de referencia, que es una trama previamente reconstruida disponible en el codificador y el decodificador. La predicción rara vez es perfecta, por lo que el codificador generalmente codifica bloques de diferencias de píxeles (también llamados bloques de error o residuales) entre el macrobloque de predicción y el macrobloque 215 en sí.

45 La Figura 3 ilustra un ejemplo de cálculo y codificación de un bloque 335 de error en el codificador WMV8. El bloque 335 de error es la diferencia entre el bloque 315 previsto y el bloque 325 actual original. El codificador aplica una DCT 340 al bloque 335 de error, resultando en un bloque 8x8 345 de coeficientes. El codificador cuantifica 350 los coeficientes DCT, resultando en un bloque 8x8 de coeficientes 355 DCT cuantificados. El codificador escanea 360 el bloque 8x8 355 en una matriz 365 unidimensional de modo que los coeficientes generalmente se ordenan de la frecuencia más baja a la más alta. El codificador codifica por entropía los coeficientes escaneados utilizando una variación de la codificación 370 de longitud de ejecución. El codificador selecciona un código de entropía de una o más tablas 375 de ejecución/nivel/último y genera el código de entropía.

55 La figura 4 muestra un ejemplo de un procedimiento 400 de decodificación correspondiente para un bloque intercodificado. En resumen de la Figura 4, un decodificador decodifica (410, 420) información codificada por entropía que representa un residuo de predicción utilizando la decodificación 410 de longitud variable con una o más tablas

415 de ejecución/nivel/último y decodificación 420 de longitud de ejecución. El decodificador escanea inversamente 430 una matriz 425 unidimensional que almacena la información decodificada por entropía en un bloque 435 bidimensional. El decodificador cuantifica a la inversa y hace DCT inversa (juntos, 440) de los datos, resultando en un bloque 445 de error reconstruido. En una ruta de compensación de movimiento separada, el decodificador calcula un 5 bloque 465 previsto usando la información 455 del vector de movimiento para el desplazamiento desde una trama de referencia. El decodificador combina 470 el bloque 465 previsto con el bloque 445 de error reconstruido para formar el bloque 475 reconstruido.

En software para un codificador WMV anterior y software para un decodificador WMV anterior, la información de predicción AC se señala en un bit por macrobloque a nivel de macrobloque en la secuencia de bits.

10 El campo ACPRED es un elemento de un bit, de secuencia de bits a nivel de macrobloque que especifica si la predicción AC se utiliza para decodificar los coeficientes AC para todos los bloques en un macrobloque. ACPRED está presente en tramas I y en intra macrobloques de 1MV en tramas previstas. ACPRED = 0 generalmente indica que la predicción AC no se usa en el macrobloque, y ACPRED = 1 generalmente indica que la predicción AC se usa en el macrobloque. El bloque predictor es el bloque inmediatamente superior o a la izquierda del bloque actual. Sin embargo, 15 en una trama prevista (por ejemplo, una trama P o una trama B), si el bloque de predicción superior y el bloque de predicción izquierdo no están intracodificados, la predicción AC no se usa incluso si ACPRED = 1.

El codificador y el decodificador también usan señalización de predicción AC para tramas entrelazadas. El indicador ACPREDMB es un valor de un bit presente a nivel de macrobloque para macrobloques codificados por trama que especifica si se usa la predicción AC para todos los bloques en el macrobloque. Los indicadores ACPREDTFIELD y 20 ACPREDBFIELD son valores de un bit presentes a nivel de macrobloque para macrobloques codificados en campo que especifican si la predicción AC se usa para bloques en el campo superior e inferior de un macrobloque actual, respectivamente.

II. Vídeo entrelazado y vídeo progresivo

25 Una trama de vídeo contiene líneas de información espacial de una señal de vídeo. Para vídeo progresivo, estas líneas contienen muestras que comienzan desde un instante y continúan a través de líneas sucesivas hasta la parte inferior de la trama. Una trama I progresiva es una trama de vídeo progresivo intracodificado. Una trama P progresiva es una trama de vídeo progresivo codificado usando predicción directa, y una trama B progresiva es una trama de vídeo progresivo codificado usando predicción bidireccional.

30 Una trama de vídeo entrelazado típico que consiste en dos campos escaneados que comienzan en diferentes momentos. Por ejemplo, en referencia a la figura 5, una trama 500 de vídeo entrelazado incluye el campo 510 superior y el campo 520 inferior. Típicamente, las líneas pares (campo superior) se escanean comenzando al mismo tiempo (p. ej., tiempo t) y las líneas impares (campo inferior) se escanean comenzando en un momento diferente (generalmente más tarde) (por ejemplo, tiempo t + 1). Este tiempo puede crear características dentadas como dientes 35 en regiones de una trama de vídeo entrelazado en que hay movimiento cuando los dos campos se escanean comenzando en diferentes momentos. Por esta razón, las tramas de vídeo entrelazado se pueden reorganizar de acuerdo con una estructura de campo, con las líneas impares agrupadas en un campo y las líneas pares agrupadas en otro campo. Esta disposición, conocida como codificación de campo, es útil en imágenes de movimiento alto para reducir tales artefactos de bordes irregulares. Por otra parte, en regiones estacionarias, el detalle de la imagen en la trama de vídeo entrelazado puede conservarse de manera más eficiente sin dicha reorganización. Por consiguiente, 40 la codificación de trama se usa a menudo en tramas de vídeo entrelazado estacionarias o de bajo movimiento, en el que se conserva la disposición original de la línea de campo alterna.

El software para un codificador WMV anterior y el software para un decodificador anterior usan macrobloques que están dispuestos de acuerdo con una estructura de campo (macrobloques codificados en campo) o una estructura de trama (macrobloques codificados por trama) en tramas de vídeo entrelazado. La Figura 6 muestra una estructura para 45 macrobloques codificados en campo en el codificador y decodificador. Un macrobloque 610 entrelazado está permutado de modo que todas las líneas de campo superior (por ejemplo, líneas pares 0, 2,... 14) se colocan en la mitad superior del macrobloque 620 codificado en campo y todas las líneas de campo inferior (por ejemplo, líneas impares 1, 3,... 15) se colocan en la mitad inferior del macrobloque codificado en campo. Para un macrobloque codificado por trama, las líneas de campo superior y las líneas de campo inferior se alternan en todo el macrobloque, 50 como en el macrobloque 610 entrelazado.

El codificador y decodificador anteriores utilizan un formato de macrobloque 4:1:1 en tramas entrelazadas. Un macrobloque 4:1:1 está compuesto por cuatro bloques de luminancia de 8x8 y dos bloques de 4x8 de cada canal de crominancia. En un macrobloque 4:1:1 codificado en campo, el macrobloque permutado se subdivide de modo que los dos bloques de luminancia superiores de 8x8 y el bloque de crominancia superior de 4x8 en cada canal de crominancia contienen solo líneas de campo superior, mientras que los dos bloques de luminancia inferiores de 8x8 y el bloque de crominancia inferior de 4x8 en cada canal de crominancia contienen solo líneas de campo inferior. 55

Una trama de vídeo progresivo típica consiste en una trama de contenido con líneas no alternas. A diferencia del vídeo entrelazado, el vídeo progresivo no divide las tramas de vídeo en campos separados, y se escanea una trama completa

de izquierda a derecha, de arriba a abajo a partir de una sola vez.

III. Modo de trama/campo de señalización para macrobloques entrelazados

En el software para un codificador y decodificador WMV anterior, el campo INTRLCF es un elemento de un bit, de capa de trama utilizado para indicar si los macrobloques se codifican solo en modo de trama, o en modo de campo o de trama. Si INTRLCF = 0, todos los macrobloques en la trama están codificados en modo de trama. Si INTRLCF = 1, los macrobloques en la trama pueden codificarse en modo de campo o trama, y el campo INTRLCMB sigue en la secuencia de bits para indicar el estado de codificación de trama/campo para cada macrobloque. INTRLCMB es un plano de bits presente en tramas I progresivas, tramas I entrelazadas, tramas P entrelazadas y tramas B entrelazadas. El plano de bits INTRLCMB decodificado representa el estado entrelazado para cada macrobloque como un campo de valores de un bit en el orden de escaneo de ráster desde la parte superior izquierda a la inferior derecha. Un valor de 0 indica que el macrobloque correspondiente está codificado en modo trama. Un valor de 1 indica que el macrobloque correspondiente está codificado en modo de campo.

El modo de codificación de campo/trama se indica para cada macrobloque en tramas I progresivas, tramas I entrelazadas, tramas P entrelazadas y tramas B entrelazadas, y el modo de codificación de campo/trama se señala solo a nivel de trama mediante un plano de bits. No hay una opción de señalización de capa de macrobloque disponible para señalar el modo de codificación de campo/trama, lo que limita la flexibilidad de la señalización.

IV. Codificación de plano de bits en software para un codificador y decodificador WMV anterior

En el software para un codificador y decodificador WMV anterior, cierta información binaria para macrobloques en una trama se codifica como una matriz bidimensional en uno de los siete modos de codificación de plano de bits, y se transmite en una cabecera de trama.

El codificador y el decodificador utilizan la codificación de plano de bits para señalar cuatro tipos diferentes de información binaria a nivel de trama para macrobloques en una trama: (1) macrobloques omitidos/no omitidos, (2) modo de codificación de campo o trama en imágenes entrelazadas, (3) un modo de codificación de un vector de movimiento ["1MV"] o cuatro vectores de movimiento ["4MV"]; y (4) modo de predicción directa/no directa en tramas B. Los siguientes elementos de sintaxis se utilizan en el esquema de codificación de plano de bits.

INVERT

El campo INVERT es un código de un bit que indica si el plano de bits tiene más bits iguales a 0 o más bits iguales a 1. Dependiendo de INVERT y el modo de codificación de plano de bits, el decodificador puede invertir el plano de bits decodificado para recrear el original.

IMODE

El campo IMODE es un código de longitud variable ["VLC"] que representa el modo de codificación del plano de bits. En general, los códigos más cortos se utilizan para codificar modos de codificación más frecuentes.

DATABITS

El campo DATABITS es una secuencia de símbolos codificada por entropía basada en el modo de codificación indicado en el campo IMODE. El tamaño de cada matriz bidimensional es $rowMB \times colMB$, en que $rowMB$ y $colMB$ son el número de filas y columnas de macrobloques, respectivamente, en la trama. Dentro de la secuencia de bits, cada matriz se codifica como un conjunto de bits consecutivos en uno de los siete modos de codificación de plano de bits. Los siete modos de codificación de plano de bits se describen a continuación.

1. Modo en bruto

En modo en bruto, el plano de bits se codifica como un bit por píxel escaneado en el orden de escaneo natural. DATABITS tiene $rowMB \times colMB$ bits de longitud.

2. Modo saltar fila

En modo Saltar fila, el campo ROWSKIP indica si el campo ROWBITS está presente para cada fila en el plano de bits. Si una fila completa de valores en el plano de bits es cero, ROWSKIP = 0 y ROWBITS se omite. Si al menos un valor en la fila no es cero, ROWSKIP = 1 y ROWBITS contiene un bit para cada valor en la fila. Las filas se escanean de arriba a abajo de la trama.

3. Modo saltar columna

En modo saltar columna, el campo COLUMNSKIP indica si el campo COLUMNBITS está presente para cada columna en el plano de bits. Si una columna completa de valores en el plano de bits es cero, COLUMNSKIP = 0 y COLUMNBITS se omite. Si al menos un valor en la columna no es cero, COLUMNSKIP = 1 y COLUMNBITS contiene un bit para cada valor en la columna. Las columnas se escanean de izquierda a derecha de la trama.

4. Modo normal-2

En modo Normal-2, Si $rowMB \times colMB$ es impar, el primer símbolo se representa simplemente con un bit que coincide con su valor, y los símbolos posteriores se codifican en pares en orden de escaneo natural utilizando una tabla VLC binaria.

5 **5. Modo normal-6**

En modo Normal-6, el plano de bits está codificado en grupos de seis píxeles. Estos píxeles se agrupan en mosaicos de 2x3 o 3x2. El plano de bits se divide en mosaico al máximo utilizando un conjunto de reglas de mosaico, y los píxeles restantes se codifican utilizando una variante de los modos Saltar fila y Saltar columna. Se utilizan mosaicos "verticales" de 3x2 si y solo si $rowMB$ es un múltiplo de 3 y $colMB$ no lo es. En caso contrario, se utilizan mosaicos "horizontales" 2x3.

Los mosaicos de seis elementos se codifican primero, seguido por los mosaicos lineales codificados de Saltar columna y Saltar fila. Si el tamaño de la matriz es un múltiplo de 3x2 o de 2x3, los últimos mosaicos lineales no existen y el plano de bits está en mosaico con solo mosaicos rectangulares de seis elementos.

6, 7. Modos Diff-2 y Diff-6

15 Si se utiliza cualquiera de los modos diferenciales (Diff-2 o Diff-6), se decodifica un plano de bits de "bits diferenciales" utilizando los modos normales correspondientes (Normal-2 o Normal-6, respectivamente). Los bits diferenciales se utilizan para regenerar el plano de bits original.

Para obtener más información sobre la codificación de plano de bits, véase la Solicitud de Patente de los Estados Unidos N.º 10/321.415, titulada "Saltar codificación de macrobloque" presentada el 16 de diciembre de 2002.

20 **V. Normas para la compresión y descompresión de vídeo**

Varias normas internacionales se relacionan con la compresión y descompresión de vídeo. Estas normas incluyen las normas de Motion Picture Experts Group ["MPEG"] 1, 2 y 4 y las normas H.261, H.262 (otro título para MPEG-2), H.263 y H.264 (también llamadas JVT/AVC) de la Unión Internacional de Telecomunicaciones ["UIT"]. Estas normas especifican aspectos de decodificadores de vídeo y formatos para información de vídeo comprimido. Directamente o por implicación, también especifican ciertos detalles del codificador, pero no se especifican otros detalles del codificador. Estas normas usan (o apoyan el uso de) diferentes combinaciones de descompresión y compresión dentro de la trama y entre tramas.

A. Señalización de macrobloques codificados por campo o trama en las normas

30 Algunas normas internacionales describen la señalización de la codificación por campo o la codificación por trama para macrobloques en imágenes entrelazadas.

El borrador JVT-d157 de la norma JVT/AVC describe el elemento de sintaxis `mb_field_decoding_flag`, que se utiliza para indicar si un par de macrobloques se decodifica en modo trama o modo campo en tramas P entrelazadas. La sección 7.3.4 describe una sintaxis de secuencia de bits en que `mb_field_decoding_flag` se envía como un elemento de datos de corte en los casos en que un parámetro de secuencia (`mb_frame_field_adaptive_flag`) indica el cambio entre la decodificación de trama y campo en macrobloques y un elemento de encabezado de corte (`pic_structure`) identifica la estructura de la imagen como una imagen de trama entrelazada.

40 El borrador del comité del 28 de mayo de 1998 de MPEG-4 describe el elemento de sintaxis `dct_type`, que se utiliza para indicar si un macrobloque tiene codificación DCT de trama o codificación DCT de campo. De acuerdo con las Secciones 6.2.7.3 y 6.3.7.3, `dct_type` es un elemento de capa de macrobloque que solo está presente en la secuencia de bits MPEG-4 en contenido entrelazado en que el macrobloque tiene un patrón de bloque codificado distinto de cero o está intracodificado.

45 En MPEG-2, el elemento `dct_type` también es un elemento de capa de macrobloque que indica si un macrobloque está codificado con DCT de trama o codificado con DCT de campo. MPEG-2 también describe un elemento de extensión de codificación de imagen `frame_pred_frame_dct`. Cuando `frame_pred_frame_dct` se establece en '1', solo la codificación DCT de trama se usa en tramas entrelazadas. La condición `dct_type = 0` se "deriva" cuando `frame_pred_frame_dct = 1` y el elemento `dct_type` no está presente en la secuencia de bits.

B. Señalización de predicción de coeficiente AC en las normas

Algunas normas internacionales describen la señalización de diferentes modos de predicción del coeficiente AC espacial para macrobloques.

50 El borrador del comité del 28 de mayo de 1998 de MPEG-4 describe el elemento de sintaxis `ac_pred_flag`, que es un indicador de un bit para indicar si los coeficientes AC en la primera fila o columna de un intra macrobloque están codificados diferencialmente. En la secuencia de bits MPEG-4, `ac_pred_flag` se envía como un bit por macrobloque

en una estructura de datos de partición de datos de un plano de objeto de vídeo (por ejemplo, `data_partitioned_I_VOP()`, `data_partitioned_P_VOP()`) o en una estructura de datos de capa de macrobloque (`macrobloque()`).

- 5 En la norma H.263, el Anexo I describe un modo avanzado de intracodificación que opcionalmente usa predicción AC. El elemento de capa de macrobloque `INTRA_MODE` es un código de longitud variable que indica si un macrobloque está codificado en un modo que utiliza predicción AC.

C. Limitaciones de las normas

Estas normas internacionales están limitadas de varias maneras importantes. Por ejemplo, aunque las normas proporcionan la señalización de la información del tipo de campo/trama y la predicción AC, la señalización se realiza típicamente en un bit por macrobloque.

- 10 Dada la importancia crítica de la compresión y descompresión de vídeo para vídeo digital, no es sorprendente que la compresión y descompresión de vídeo sean campos muy desarrollados. Cualesquiera que sean los beneficios de las técnicas anteriores de compresión y descompresión de vídeo, sin embargo, no tienen las ventajas de las siguientes técnicas y herramientas.

- 15 El documento WO 03/053066 A1 se refiere a omitir la codificación de macrobloques. Según una primera implementación, la información omitida del macrobloque se señala en la capa de imagen en la secuencia de bits de vídeo, mediante lo cual el codificador y el decodificador seleccionan entre cuatro modos de codificación para codificar y decodificar la información omitida del macrobloque, a saber, sin codificación de bits omitida, codificación normal, fila y columna. El codificador selecciona los modos de codificación basado en imagen por imagen, o alternativamente, el codificador selecciona los modos de codificación de alguna otra manera. Según una segunda implementación, la información omitida del macrobloque y/u otros datos binarios 2D se codifican con una eficiencia mejorada. Un bloque omitido es una condición en una secuencia de bits en que no se necesita transmitir más información a ese nivel de granularidad. Un macrobloque omitido es un macrobloque que tiene un tipo predeterminado, movimiento predeterminado y error residual predeterminado. Si la trama actual es una trama I, no se omiten macrobloques en la trama actual y termina la omisión de la codificación de macrobloques para la trama. Se pueden usar siete modos de codificación para tramas P, a saber, salto de fila, salto de columna, modo normal-2, modo normal-6, modo diff-2 y diff-6, así como el modo en bruto. Todos los modos, excepto el modo en bruto, codifican un plano de bits en el nivel de trama, que requiere un segundo paso a través de la trama durante la codificación. Sin embargo, de acuerdo con el modo en bruto, el codificador codifica un bit para cada macrobloque.

- 30 El "Modelo de verificación de vídeo MPEG4 versión 7.0", Organización Internacional de Normalización - Organization Internationale de Normalization, vol. 1642, 1 de abril de 1997, páginas 1 a 271 se refiere a la codificación de imágenes en movimiento y la información de audio asociada (MPEG4).

- 35 Puri A. y. col. "Codificación de vídeo adaptativa de movimiento de trama/campo compensado", Procesamiento de señal. Comunicación de imagen, Elsevier Science Publishers, Ámsterdam, NL, vol. 5, n.º 1-2, 1 de febrero de 1993, páginas 39-58 se refiere a un esquema de codificación de vídeo adaptativa de movimiento de campo/trama compensado.

Es el objeto de la presente invención proporcionar un procedimiento mejorado para la codificación y decodificación de plano de bits para información de tipo de codificación de campo/trama de macrobloque, así como un sistema correspondiente.

Este objeto se soluciona mediante la materia objeto de las reivindicaciones independientes.

- 40 Se definen realizaciones preferidas en las reivindicaciones dependientes.

- 45 En resumen, la descripción detallada se dirige a varias técnicas y herramientas para la codificación y decodificación de vídeo. Por ejemplo, un codificador de plano de bits codifica información de estado de predicción AC. Como otro ejemplo, un codificador señala el tipo de transformación de campo/trama para macrobloques en una imagen codificada de trama entrelazada. Un decodificador realiza la decodificación correspondiente. Las realizaciones descritas implementan una o más de las técnicas y herramientas descritas que incluyen, pero sin limitarse a, lo siguiente: En un aspecto, un codificador/decodificador selecciona un modo de plano de bits de un grupo de múltiples modos de plano de bits disponibles, y procesa un plano de bits de acuerdo con el modo de plano de bits seleccionado, en el que el plano de bits indica información de estado de predicción AC para la pluralidad de macrobloques de una imagen de vídeo.

- 50 En otro aspecto, un codificador codifica un plano de bits que indica información de estado de predicción AC para la pluralidad de macrobloques de una imagen de vídeo y señala el plano de bits codificado.

En otro aspecto, un decodificador recibe un plano de bits codificado y decodifica el plano de bits, en el que el plano de bits indica información de estado de predicción AC para la pluralidad de macrobloques de una imagen de vídeo.

En otro aspecto, para una primera trama de vídeo entrelazado en una secuencia de vídeo, un decodificador decodifica

5 un plano de bits señalado en la capa de trama para la primera trama de vídeo entrelazado. El plano de bits representa los tipos de transformación de campo/trama para la pluralidad de macrobloques de la primera trama de vídeo entrelazado. Para una segunda trama de vídeo entrelazado en la secuencia de vídeo, para cada uno de al menos uno pero no toda la pluralidad de macrobloques de la segunda trama de vídeo entrelazado, el decodificador procesa un bit de tipo de transformación de trama/campo por macrobloque señalado en la capa de macrobloque.

10 En otro aspecto, para una primera trama de vídeo entrelazado en una secuencia de vídeo, un codificador codifica un plano de bits y señala el plano de bits en la capa de trama para la primera trama de vídeo entrelazado. El plano de bits representa los tipos de transformación de campo/trama para la pluralidad de macrobloques de la primera trama de vídeo entrelazado. Para una segunda trama de vídeo entrelazado en la secuencia de vídeo, para cada uno de al menos uno pero no toda la pluralidad de macrobloques de la segunda trama de vídeo entrelazado, el codificador señala un bit de tipo de transformación de trama/campo por macrobloque en la capa de macrobloque.

15 En otro aspecto, un decodificador comprende medios para decodificar planos de bits señalados en la capa de trama, en el que los planos de bits incluyen un primer plano de bits para una primera trama de vídeo entrelazado, y en el que el primer plano de bits representa tipos de transformación de campo/trama para la pluralidad de macrobloques de la primera trama de vídeo entrelazado, y medios para procesar bits de tipo de transformación de campo/trama por macrobloque señalados en la capa de macrobloque, en el que los bits de tipo de transformación de trama/campo por macrobloque incluyen un bit de tipo de transformación de trama/campo por macrobloque para cada uno o más pero no toda la pluralidad de macrobloques de una segunda trama de vídeo entrelazado.

Las diversas técnicas y herramientas se pueden utilizar en combinación o de forma independiente.

20 Las características y ventajas adicionales se evidenciarán a partir de la siguiente descripción detallada de diferentes realizaciones que sigue con referencia a los dibujos adjuntos.

Breve descripción de los dibujos

25 La figura 1A es un diagrama que muestra la compresión intratrama basada en bloques de un bloque de píxeles de 8x8 según la técnica anterior. La figura 1B muestra candidatos de predicción AC para un bloque de píxeles de 8x8 según la técnica anterior.

La figura 2 es un diagrama que muestra la estimación de movimiento en un codificador de vídeo de acuerdo con la técnica anterior.

La figura 3 es un diagrama que muestra la compresión basada en bloques para un bloque de 8x8 de residuales de predicción en un codificador de vídeo de acuerdo con la técnica anterior.

30 La Figura 4 es un diagrama que muestra la descompresión basada en bloques para un bloque 8x8 de residuales de predicción en un codificador de vídeo de acuerdo con la técnica anterior.

La figura 5 es un diagrama que muestra una trama entrelazada de acuerdo con la técnica anterior.

La figura 6 es un diagrama que muestra la codificación de campo de macrobloques entrelazados según la técnica anterior.

35 La figura 7 es un diagrama de bloques de un entorno informático adecuado junto con el cual se pueden implementar varias realizaciones descritas.

La figura 8 es un diagrama de bloques de un sistema codificador de vídeo generalizado junto con el cual se pueden implementar varias realizaciones descritas.

40 La figura 9 es un diagrama de bloques de un sistema decodificador de vídeo generalizado junto con el cual se pueden implementar varias realizaciones descritas.

La figura 10 es un diagrama de un formato de macrobloque utilizado en varias realizaciones descritas.

La Figura 11A es un diagrama de parte de una trama de vídeo entrelazado, mostrando líneas alternas de un campo superior y un campo inferior. La figura 11B es un diagrama de la trama de vídeo entrelazado organizada para codificar/decodificar como una trama, y la figura 11C es un diagrama de la trama de vídeo entrelazado organizada para codificar/decodificar como campos.

45 La Figura 12 es un diagrama de flujo que muestra una técnica para señalar información de tipo de campo/trama para macrobloques utilizando codificación de plano de bits a nivel de trama o codificación a nivel de macrobloque.

La Figura 13 es un diagrama de flujo que muestra una técnica para la codificación de plano de bits de la información de predicción AC.

50 La Figura 14 es un diagrama que muestra una sintaxis de secuencia de bits de capa de trama para tramas I progresivas en una implementación combinada. La Figura 15 es un diagrama que muestra una sintaxis de secuencia de bits de capa de macrobloque para tramas I progresivas en una implementación combinada.

La Figura 16 es un diagrama que muestra una sintaxis de secuencia de bits de capa de trama para campos I entrelazados o campos BI en una implementación combinada.

55 La Figura 17 es un diagrama que muestra una sintaxis de secuencia de bits de capa de campo para campos I entrelazados en una implementación combinada.

La Figura 18 es un diagrama que muestra una sintaxis de secuencia de bits de capa de trama para tramas I entrelazadas en una implementación combinada.

60 La Figura 19 es un diagrama que muestra una sintaxis de secuencia de bits de capa de trama para tramas P entrelazadas en una implementación combinada.

La Figura 20 es un diagrama que muestra una sintaxis de secuencia de bits de capa de trama para tramas B

entrelazadas en una implementación combinada.

La Figura 21 es un diagrama que muestra una sintaxis de secuencia de bits de capa de macrobloque para macrobloques de tramas I entrelazadas en una implementación combinada.

5 La Figura 22 es un diagrama que muestra una sintaxis de secuencia de bits de capa de macrobloque para macrobloques de tramas P entrelazadas en una implementación combinada.

La Figura 23 es un diagrama que muestra una sintaxis de secuencia de bits de capa de macrobloque para macrobloques de tramas B entrelazadas en una implementación combinada.

Las Figuras 24A - 24C son diagramas que muestran mosaicos para modos de codificación de plano de bits normal-6 y diff-6 de acuerdo con la técnica anterior.

10 **Descripción detallada**

La presente solicitud se refiere a técnicas y herramientas para la compresión y descompresión eficiente de vídeo progresivo y entrelazado. En diversas realizaciones descritas, un codificador y decodificador de vídeo incorporan técnicas para codificar y decodificar vídeo progresivo y entrelazado, y las técnicas de señalización correspondientes para su uso con un formato de secuencia de bits o sintaxis que comprende diferentes capas o niveles (por ejemplo, nivel de secuencia, nivel de trama, nivel de campo, nivel de macrobloque y/o nivel de bloque).

15 Son posibles varias alternativas a las implementaciones descritas en el presente documento. Por ejemplo, las técnicas descritas con referencia a los diagramas de flujo pueden alterarse cambiando el orden de las etapas que se muestran en los diagramas de flujo, repitiendo u omitiendo ciertas etapas, etc. Como otro ejemplo, aunque algunas implementaciones se describen con referencia a formatos específicos de macrobloques, también se pueden utilizar otros formatos. Adicionalmente, las técnicas y herramientas descritas con referencia a la predicción directa también pueden ser aplicables a otros tipos de predicción.

20 Las diversas técnicas y herramientas se pueden utilizar en combinación o de forma independiente. Diferentes realizaciones implementan una o más de las técnicas y herramientas descritas. Algunas técnicas y herramientas descritas en el presente documento pueden usarse en un codificador o decodificador de vídeo, o en algún otro sistema no limitado específicamente a la codificación o decodificación de vídeo.

25 **I. Entorno informático**

La figura 7 ilustra un ejemplo generalizado de un entorno 700 informático adecuado en el que pueden implementarse varias de las realizaciones descritas. El entorno 700 informático no pretende sugerir ninguna limitación en cuanto al ámbito de uso o funcionalidad, ya que las técnicas y herramientas pueden implementarse en diversos entornos informáticos de fin general o especial.

30 Con referencia a la figura 7, el entorno 700 informático incluye al menos una unidad 710 de procesamiento y memoria 720. En la figura 7, esta configuración 730 más básica se incluye dentro de una línea discontinua. La unidad 710 de procesamiento ejecuta instrucciones ejecutables por computadora y puede ser un procesador real o virtual. En un sistema de procesamiento múltiple, múltiples unidades de procesamiento ejecutan instrucciones ejecutables por computadora para aumentar la potencia de procesamiento. La memoria 720 puede ser memoria volátil (por ejemplo, registros, caché, RAM), memoria no volátil (p. ej., ROM, EEPROM, memoria flash, etc.), o alguna combinación de los dos. La memoria 720 almacena software 780 que implementa un codificador o decodificador de vídeo con una o más de las técnicas y herramientas descritas.

40 Un entorno informático puede tener características adicionales. Por ejemplo, el entorno 700 informático incluye almacenamiento 740, uno o más dispositivos 750 de entrada, uno o más dispositivos 760 de salida y una o más conexiones 770 de comunicación. Un mecanismo de interconexión (no mostrado) como un bus, controlador o red interconecta los componentes del entorno 700 informático. Típicamente, el software del sistema operativo (no mostrado) proporciona un entorno operativo para otro software que se ejecuta en el entorno 700 informático y coordina las actividades de los componentes del entorno 700 informático.

45 El almacenamiento 740 puede ser extraíble o no extraíble e incluye discos magnéticos, cintas magnéticas o casetes, CD-ROM, DVD, o cualquier otro medio que se pueda utilizar para almacenar información y al que se pueda acceder dentro del entorno 700 informático. El almacenamiento 740 almacena instrucciones para el software 780 que implementa el codificador o decodificador de vídeo.

50 Los dispositivos 750 de entrada pueden ser un dispositivo de entrada táctil, como un teclado, ratón, bolígrafo o bola de seguimiento, un dispositivo de entrada de voz, un dispositivo de escaneo u otro dispositivo que proporciona información al entorno 700 informático. Para codificación de audio o vídeo, los dispositivos 750 de entrada pueden ser una tarjeta de sonido, tarjeta de vídeo, tarjeta sintonizadora de TV, o dispositivo similar que acepte entradas de audio o vídeo en forma analógica o digital, o un CD-ROM o CD-RW que lea muestras de audio o vídeo en el entorno 700 informático. Los dispositivos 760 de salida pueden ser un sistema de representación, impresora, altavoz, grabadora de CD u otro dispositivo que proporcione salida del entorno 700 informático.

55 Las conexiones 770 de comunicación permiten la comunicación a través de un medio de comunicación a otra entidad informática. El medio de comunicación transmite información como instrucciones ejecutables por computadora,

entrada o salida de audio o vídeo, u otros datos en una señal de datos modulada. Una señal de datos modulada es una señal que tiene una o más de sus características establecidas o cambiadas de un modo tal que se codifique la información de la señal. A modo de ejemplo, y sin limitación, los medios de comunicación incluyen técnicas cableadas o inalámbricas implementadas con un soporte eléctrico, óptico, RF, infrarrojo, acústico u otro.

5 Las técnicas y herramientas se pueden describir en el contexto general de los medios legibles por computadora. Los medios legibles por computadora son cualquier medio disponible al que se pueda acceder dentro de un entorno informático. A modo de ejemplo, y sin limitación, con el entorno 700 informático, los medios legibles por computadora incluyen memoria 720, almacenamiento 740, medios de comunicación y combinaciones de cualquiera de los anteriores.

10 Las técnicas y herramientas se pueden describir en el contexto general de instrucciones ejecutables por computadora, como las incluidas en los módulos del programa, ejecutándose en un entorno informático en un procesador real o virtual objetivo. Generalmente, los módulos de programa incluyen rutinas, programas, bibliotecas, objetos, clases, componentes, estructuras de datos, etc., que llevan a cabo tareas particulares o implementan tipos de datos abstractos particulares. La funcionalidad de los módulos de programa se puede combinar o dividir entre módulos de programa como se desee en diversas realizaciones. Las instrucciones ejecutables por computadora para los módulos de programa pueden ejecutarse dentro de un entorno informático local o distribuido.

15 Por el bien de la presentación, la descripción detallada utiliza términos como "estimar" "compensar", "predecir", y "aplicar" para describir las operaciones informáticas en un entorno informático. Estos términos son abstracciones de alto nivel para las operaciones realizadas por una computadora, y no deben confundirse con los actos realizados por un ser humano. Las operaciones informáticas reales que corresponden a estos términos varían según la implementación.

II. Codificador y decodificador de vídeo generalizado

25 La figura 8 es un diagrama de bloques de un codificador 800 de vídeo generalizado junto con el cual se pueden implementar algunas realizaciones descritas. La figura 9 es un diagrama de bloques de un decodificador 900 de vídeo generalizado junto con el cual se pueden implementar algunas realizaciones descritas.

30 Las relaciones mostradas entre los módulos dentro del codificador 800 y el decodificador 900 indican flujos generales de información en el codificador y el decodificador; otras relaciones no se muestran por simplicidad. En particular, las Figuras 8 y 9 generalmente no muestran información secundaria que indique la configuración del codificador, modos, tablas, etc. utilizados para una secuencia de vídeo, imagen, macrobloque, bloque, etc. Dicha información secundaria se envía en la secuencia de bits de salida, típicamente después de la codificación por entropía de la información secundaria. El formato de la secuencia de bits de salida puede ser un formato de Windows Media Vídeo versión 9 u otro formato.

35 El codificador 800 y el decodificador 900 procesan imágenes de vídeo, que pueden ser tramas de vídeo, campos de vídeo o combinaciones de tramas y campos. La sintaxis y la semántica de la secuencia de bits en los niveles de imagen y macrobloque pueden depender de si se usan tramas o campos. Puede haber cambios en la organización de macrobloques y el tiempo general también. El codificador 800 y el decodificador 900 están basados en bloques y utilizan un formato de macrobloque 4:2:0 para tramas, con cada macrobloque incluyendo cuatro bloques de luminancia de 8x8 (a veces tratados como un macrobloque de 16x16) y dos bloques de crominancia de 8x8. Para los campos, se puede utilizar la misma organización y formato de macrobloque o una diferente. Los bloques de 8x8 pueden subdividirse además en diferentes etapas, por ejemplo, en las etapas de transformación de frecuencia y codificación por entropía. Las organizaciones de tramas de vídeo de ejemplo se describen con más detalle a continuación.

40 Dependiendo de la implementación y el tipo de compresión deseada, se pueden agregar módulos del codificador o decodificador, omitirse, dividirse en múltiples módulos, combinarse con otros módulos, y/o reemplazarse por módulos similares. En realizaciones alternativas, los codificadores o decodificadores con diferentes módulos y/u otras configuraciones de módulos realizan una o más de las técnicas descritas.

A. Organizaciones de tramas de vídeo

45 En algunas implementaciones, el codificador 800 y el decodificador 900 procesan tramas de vídeo organizadas de la siguiente manera. Una trama contiene líneas de información espacial de una señal de vídeo. Para vídeo progresivo, estas líneas contienen muestras que comienzan desde un instante y continúan a través de líneas sucesivas hasta la parte inferior de la trama. Una trama de vídeo progresivo se divide en macrobloques como el macrobloque 1000 que se muestra en la Figura 10. El macrobloque 1000 incluye cuatro bloques de luminancia de 8x8 (Y1 a Y4) y dos bloques de crominancia de 8x8 que se ubican conjuntamente con los cuatro bloques de luminancia, pero a media resolución horizontal y verticalmente, siguiendo el formato convencional de macrobloque 4:2:0. Los bloques de 8x8 pueden subdividirse además en diferentes etapas, por ejemplo, en la transformación de frecuencia (por ejemplo, 8x4, 4x8 o 4x4 DCT) y etapas de codificación por entropía. Una trama I progresiva es una trama de vídeo progresivo intracodificado. Una trama P progresiva es una trama de vídeo progresivo codificado usando predicción directa, y una trama B progresiva es una trama de vídeo progresivo codificado usando predicción bidireccional. Las tramas P y B progresivas pueden incluir macrobloques intracodificados, así como diferentes tipos de macrobloques previstos.

Una trama de vídeo entrelazado consiste en dos escaneos de una trama: uno comprende las líneas pares de la trama (el campo superior) y el otro comprende las líneas impares de la trama (el campo inferior). Los dos campos pueden representar dos períodos de tiempo diferentes o pueden ser del mismo período de tiempo. La figura 11A muestra parte de una trama 1100 de vídeo entrelazado, incluyendo las líneas alternas del campo superior y el campo inferior en la parte superior izquierda de la trama 1100 de vídeo entrelazado.

La Figura 11B muestra la trama 1100 de vídeo entrelazado de la Figura 11A organizada para codificar/decodificar como una trama 1130. La trama 1100 de vídeo entrelazado se ha dividido en macrobloques como los macrobloques 1131 y 1132, que usan un formato 4:2:0 como se muestra en la Figura 10. En el plano de luminancia, cada macrobloque 1131, 1132 incluye 8 líneas desde el campo superior alternando con 8 líneas desde el campo inferior para un total de 16 líneas, y cada línea tiene 16 píxeles de largo. (La organización y ubicación reales de los bloques de luminancia y los bloques de crominancia dentro de los macrobloques 1131, 1132 no se muestran y, de hecho, pueden variar para diferentes decisiones de codificación.) Dentro de un macrobloque dado, la información de campo superior y la información de campo inferior pueden codificarse conjunta o separadamente en cualquiera de las diversas fases. Una trama I entrelazada son dos campos intracodificados de una trama de vídeo entrelazado, en que un macrobloque incluye información para los dos campos. Una trama P entrelazada son dos campos de una trama de vídeo entrelazado codificada usando predicción directa, y una trama B entrelazada son dos campos de una trama de vídeo entrelazado codificada usando predicción bidireccional, en que un macrobloque incluye información para los dos campos. Las tramas P y B entrelazadas pueden incluir macrobloques intracodificados, así como diferentes tipos de macrobloques previstos. Las tramas BI entrelazadas son un híbrido de tramas I entrelazadas y tramas B entrelazadas; están intracodificadas, pero no se usan como anclajes para otras tramas.

La Figura 11C muestra la trama 1100 de vídeo entrelazado de la Figura 11A organizada para codificar/decodificar como campos 1160. Cada uno de los dos campos de la trama 1100 de vídeo entrelazado se divide en macrobloques. El campo superior está dividido en macrobloques como el macrobloque 1161, y el campo inferior está dividido en macrobloques como el macrobloque 1162. (Otra vez, los macrobloques usan un formato 4:2:0 como se muestra en la Figura 10, y la organización y colocación de bloques de luminancia y bloques de crominancia dentro de los macrobloques no se muestran.) En el plano de luminancia, el macrobloque 1161 incluye 16 líneas desde el campo superior y el macrobloque 1162 incluye 16 líneas desde el campo inferior, y cada línea tiene 16 píxeles de largo. Un campo I entrelazado es un solo campo, representado por separado, de una trama de vídeo entrelazado. Un campo P entrelazado es un solo campo, representado por separado, de una trama de vídeo entrelazado codificada usando predicción directa, y un campo B entrelazado es un solo campo, representado por separado, de una trama de vídeo entrelazado codificada utilizando predicción bidireccional. Los campos P y B entrelazados pueden incluir macrobloques intracodificados, así como diferentes tipos de macrobloques previstos. Los campos BI entrelazados son un híbrido de campos I entrelazados y campos B entrelazados; están intracodificados, pero no se usan como anclajes para otros campos.

Las tramas de vídeo entrelazado organizadas para codificar/decodificar como campos pueden incluir varias combinaciones de diferentes tipos de campos. Por ejemplo, dicha trama puede tener el mismo tipo de campo en los campos superior e inferior o diferentes tipos de campo en cada campo. En una implementación, las posibles combinaciones de tipos de campo incluyen I/I, I/P, P/I, P/P, B/B, B/BI, BI/B y BI/BI.

El término imagen generalmente se refiere a datos de imagen fuente, codificados o reconstruidos. Para vídeo progresivo, una imagen es una trama de vídeo progresivo. Para vídeo entrelazado, una imagen puede referirse a una trama de vídeo entrelazado, el campo superior de la trama, o el campo inferior de la trama, dependiendo del contexto.

Como alternativa, el codificador 800 y el decodificador 900 están basados en objetos, usan un macrobloque o formato de bloque diferente, o realizan operaciones en conjuntos de píxeles de diferente tamaño o configuración que los bloques de 8x8 y macrobloques de 16x16.

B. Codificador de vídeo

La figura 8 es un diagrama de bloques de un sistema 800 codificador de vídeo generalizado. El sistema 800 codificador recibe una secuencia de imágenes de vídeo que incluye una imagen 805 actual (por ejemplo, trama de vídeo progresivo, trama de vídeo entrelazado, o campo de una trama de vídeo entrelazado), y produce información 895 de vídeo comprimido como salida. Las realizaciones particulares de los codificadores de vídeo usan típicamente una versión variada o complementada del codificador 800 generalizado.

El sistema 800 codificador comprime las imágenes previstas y las imágenes clave. Por el bien de la presentación, la Figura 8 muestra una ruta para imágenes clave a través del sistema 800 codificador y una ruta para imágenes previstas. Muchos de los componentes del sistema 800 codificador se utilizan para comprimir tanto imágenes clave como imágenes previstas. Las operaciones exactas realizadas por esos componentes pueden variar según el tipo de información que se esté comprimiendo.

Una imagen prevista (por ejemplo, trama P o trama B progresiva, campo P o campo B entrelazado, o trama P o trama B entrelazada) se representa en términos de predicción (o diferencia) de una o más imágenes (que generalmente se denominan imágenes de referencia o anclas). Una predicción residual es la diferencia entre lo que se predijo y la

imagen original. Por el contrario, una imagen clave (por ejemplo, trama I progresiva, campo I entrelazado, o trama I entrelazada) se comprime sin referencia a otras imágenes.

Si la imagen 805 actual es una imagen prevista directa, un estimador 810 de movimiento estima el movimiento de macrobloques u otros conjuntos de píxeles de la imagen 805 actual con respecto a una o más imágenes de referencia, por ejemplo, la imagen 825 anterior reconstruida almacenada en el almacén 820 de imágenes. Si la imagen 805 actual es una imagen prevista bidireccionalmente, un estimador 810 de movimiento estima el movimiento en la imagen 805 actual con respecto a hasta cuatro imágenes de referencia reconstruidas (para un campo B entrelazado, por ejemplo). Típicamente, un estimador de movimiento estima el movimiento en una imagen B con respecto a una o más imágenes de referencia temporalmente anteriores y una o más imágenes de referencia temporalmente futuras. Por consiguiente, el sistema 800 codificador puede usar los almacenes 820 y 822 separados para múltiples imágenes de referencia. Para obtener más información sobre tramas B progresivas y tramas B y campos B entrelazados, véase la Solicitud de Patente de los Estados Unidos N.º de serie 10/622.378, titulada, "Codificación predictiva bidireccional avanzada de tramas de vídeo", presentada el 18 de julio de 2003 y la Solicitud de Patente de los Estados Unidos N.º de serie 10/882.135, titulada, "Codificación predictiva bidireccional avanzada de vídeo entrelazado", presentada el 29 de junio de 2004, que se incorpora en el presente documento como referencia.

El estimador 810 de movimiento puede estimar el movimiento por píxel, $\frac{1}{2}$ píxel, $\frac{1}{4}$ píxel u otros incrementos, y puede cambiar la precisión de la estimación de movimiento basado en imagen por imagen o de otra manera. El estimador 810 de movimiento (y el compensador 830) también pueden cambiar entre los tipos de interpolación de píxeles de la imagen de referencia (por ejemplo, entre bicúbico y bilineal) por trama o con otra base. La precisión de la estimación de movimiento puede ser igual o diferente horizontal y verticalmente. El estimador 810 de movimiento sale como información 815 de movimiento de información secundaria tal como información de vector de movimiento diferencial.

El compensador 830 de movimiento aplica los vectores de movimiento a las imágenes 825 reconstruidas para formar una imagen 835 actual con compensación de movimiento. La predicción rara vez es perfecta, sin embargo, y las diferencias entre la imagen 835 actual con compensación de movimiento y la imagen 805 actual original son los residuales 845 de predicción. Durante la reconstrucción posterior de la imagen, los residuales 845 de predicción se añaden a la imagen 835 actual con compensación de movimiento para obtener una imagen reconstruida que está más cerca de la imagen 805 actual original. En compresión con pérdida, sin embargo, todavía se pierde parte de la información de la imagen 805 actual original. Como alternativa, un estimador de movimiento y un compensador de movimiento aplican otro tipo de estimación/compensación de movimiento.

Un transformador 860 de frecuencia convierte la información de vídeo del dominio espacial en dominio de frecuencia (es decir, datos espectrales). Para imágenes de vídeo basadas en bloques, el transformador 860 de frecuencia aplica una DCT, variante de DCT u otra transformación de bloque a bloques de datos de píxeles o datos residuales de predicción, produciendo bloques de coeficientes de transformación de frecuencia. Como alternativa, el transformador 860 de frecuencia aplica otra transformación de frecuencia convencional tal como una transformada de Fourier o utiliza análisis de ondículas o de sub-banda. El transformador 860 de frecuencia puede aplicar una transformación de frecuencia de 8x8, 8x4, 4x8, 4x4 u otro tamaño.

Un cuantificador 870 cuantifica luego los bloques de coeficientes de datos espectrales. El cuantificador aplica una cuantificación uniforme, escalar, a los datos espectrales con un tamaño de paso que varía con base de una imagen a otra o de otra manera. Como alternativa, el cuantificador aplica otro tipo de cuantificación a los coeficientes de datos espectrales, por ejemplo, un vector no uniforme, o cuantificación no adaptativa, o cuantifica directamente datos de dominio espacial en un sistema codificador que no utiliza transformaciones de frecuencia. Además de la cuantificación adaptativa, el codificador 800 puede usar caída de trama, filtrado adaptativo u otras técnicas para el control de velocidad.

El codificador 800 puede usar señalización especial para un macrobloque omitido, que es un macrobloque que no tiene información de ciertos tipos.

Cuando se necesita una imagen actual reconstruida para la estimación/compensación de movimiento posterior, un cuantificador 876 inverso realiza una cuantificación inversa en los coeficientes de datos espectrales cuantificados. Un transformador 866 de frecuencia inverso realiza entonces la inversa de las operaciones del transformador 860 de frecuencia, produciendo una predicción residual reconstruida (para una imagen prevista) o una imagen clave reconstruida. Si la imagen 805 actual era una imagen clave, la imagen clave reconstruida se toma como la imagen actual reconstruida (no se muestra). Si la imagen 805 actual era una imagen prevista, el residual de predicción reconstruido se agrega a la imagen 835 actual con compensación de movimiento para formar la imagen actual reconstruida. Uno o ambos de los almacenes 820, 822 de imágenes almacenan la imagen actual reconstruida para su uso en la predicción con compensación de movimiento. En algunas realizaciones, el codificador aplica un filtro de desbloqueo a la trama reconstruida para suavizar de forma adaptativa las discontinuidades y otros artefactos en la imagen.

El codificador 880 de entropía comprime la salida del cuantificador 870, así como cierta información secundaria (por ejemplo, información 815 de movimiento, tamaño del paso de cuantificación). Las técnicas típicas de codificación por entropía incluyen codificación aritmética, codificación diferencial, codificación Huffman, codificación de longitud de

ejecución, codificación LZ, codificación de diccionario y combinaciones de lo anterior. El codificador 880 de entropía generalmente usa diferentes técnicas de codificación para diferentes tipos de información (por ejemplo, coeficientes DC, coeficientes AC, diferentes tipos de información secundaria), y puede elegir entre múltiples tablas de códigos dentro de una técnica de codificación particular.

- 5 El codificador 880 de entropía proporciona información 895 de vídeo comprimido al multiplexor ["MUX"] 890. El MUX 890 puede incluir una memoria intermedia, y un indicador de nivel de memoria intermedia puede retroalimentarse a los módulos adaptativos de velocidad de bits para el control de velocidad. Antes o después del MUX 890, la información 895 de vídeo comprimido puede codificarse por canal para su transmisión a través de la red. La codificación de canal puede aplicar datos de detección y corrección de errores a la información 895 de vídeo comprimido.

10 C. Decodificador de vídeo

La figura 9 es un diagrama de bloques de un sistema 900 decodificador de vídeo general. El sistema 900 decodificador recibe información 995 para una secuencia comprimida de imágenes de vídeo y produce una salida que incluye una imagen 905 reconstruida (por ejemplo, trama de vídeo progresivo, trama de vídeo entrelazado, o campo de una trama de vídeo entrelazado). Las realizaciones particulares de los decodificadores de vídeo usan típicamente una versión
15 variada o complementada del decodificador 900 generalizado.

El sistema 900 decodificador descomprime las imágenes previstas y las imágenes clave. Por el bien de la presentación, la Figura 9 muestra una ruta para imágenes clave a través del sistema 900 decodificador y una ruta para imágenes previstas directas. Muchos de los componentes del sistema 900 decodificador se utilizan para descomprimir tanto las imágenes clave como las imágenes previstas. Las operaciones exactas realizadas por esos componentes pueden
20 variar según el tipo de información que se descomprime.

Un DEMUX 990 recibe la información 995 para la secuencia de vídeo comprimido y pone la información recibida a disposición del decodificador 980 de entropía. El DEMUX 990 puede incluir una memoria intermedia de fluctuación y otras memorias intermedias también. Antes o después del DEMUX 990, la información de vídeo comprimido puede ser decodificada y procesada por canales para la detección y corrección de errores.

- 25 El decodificador 980 de entropía decodifica por entropía datos cuantificados codificados por entropía, así como información secundaria codificada por entropía (por ejemplo, información 915 de movimiento, tamaño del paso de cuantificación), aplicando típicamente la inversa de la codificación de entropía realizada en el codificador. Las técnicas de decodificación por entropía incluyen decodificación aritmética, decodificación diferencial, decodificación Huffman, decodificación de longitud de ejecución, decodificación LZ, decodificación de diccionario y combinaciones de lo anterior. El decodificador 980 de entropía típicamente usa diferentes técnicas de decodificación para diferentes tipos
30 de información (por ejemplo, coeficientes DC, coeficientes AC, diferentes tipos de información secundaria), y puede elegir entre múltiples tablas de códigos dentro de una técnica de decodificación particular.

El decodificador 900 decodifica la información 915 de movimiento al, por ejemplo, calcular uno o más predictores para vectores de movimiento, vectores de movimiento diferencial de decodificación de entropía y combinación de vectores de movimiento diferencial decodificados con predictores para reconstruir vectores de movimiento.
35

Un compensador 930 de movimiento aplica información 915 de movimiento a una o más imágenes 925 de referencia para formar una predicción 935 de la imagen 905 que se está reconstruyendo. Por ejemplo, el compensador 930 de movimiento usa uno o más vectores de movimiento de macrobloques para encontrar macrobloques en las imágenes 925 de referencia. Uno o más almacenes de imágenes (por ejemplo, almacén 920, 922 de imágenes) almacenan imágenes reconstruidas anteriores para su uso como imágenes de referencia. Típicamente, las imágenes B tienen más de una imagen de referencia (p. ej., al menos una imagen de referencia temporal anterior y al menos una imagen de referencia temporal futura). Por consiguiente, el sistema 900 decodificador puede usar almacenes 920 y 922 de imágenes separados para múltiples imágenes de referencia. El compensador 930 de movimiento puede compensar el movimiento en píxeles, $\frac{1}{2}$ píxel, $\frac{1}{4}$ píxel u otros incrementos, y puede cambiar la precisión de la compensación de movimiento imagen por imagen o de otra manera. El compensador 930 de movimiento también puede cambiar entre tipos de interpolación de píxeles de imagen de referencia (por ejemplo, entre bicúbico y bilineal) por trama o con otra base. La precisión de la compensación de movimiento puede ser igual o diferente horizontal y verticalmente. Como alternativa, un compensador de movimiento aplica otro tipo de compensación de movimiento. La predicción del compensador de movimiento rara vez es perfecta, por lo que el decodificador 900 también reconstruye los residuales de predicción.
40
45
50

Un cuantificador 970 inverso cuantifica inversamente datos decodificados por entropía. En general, el cuantificador inverso aplica una cuantificación uniforme, inversa y escalar a los datos decodificados por entropía con un tamaño de paso que varía de una imagen a otra o de otra manera. Como alternativa, el cuantificador inverso aplica otro tipo de cuantificación inversa a los datos, por ejemplo, para reconstruir después de un vector no uniforme, o cuantificación no adaptativa, o directamente cuantifica a la inversa datos de dominio espacial en un sistema decodificador que no utiliza transformaciones de frecuencia inversa.
55

Un transformador 960 de frecuencia inversa convierte los datos cuantificados, de dominio de frecuencia en información de vídeo de dominio espacial. Para imágenes de vídeo basadas en bloques, el transformador 960 de frecuencia inversa

5 aplica una DCT inversa ["IDCT"], variante de IDCT, u otra transformación de bloques inversa a bloques de los coeficientes de transformación de frecuencia, produciendo datos de píxeles o datos residuales de predicción para imágenes clave o imágenes previstas, respectivamente. Como alternativa, el transformador 960 de frecuencia inversa aplica otra transformada de frecuencia inversa convencional tal como una transformada inversa de Fourier o usa síntesis de ondícula o sub-banda. El transformador 960 de frecuencia inversa puede aplicar una transformada de frecuencia inversa de 8x8, 8x4, 4x8, 4x4 u otro tamaño.

10 Para una imagen prevista, el decodificador 900 combina el residual 945 de predicción reconstruida con la predicción 935 con compensación de movimiento para formar la imagen 905 reconstruida. Cuando el decodificador necesita una imagen 905 reconstruida para la compensación de movimiento posterior, uno o ambos almacenes de imágenes (por ejemplo, almacén 920 de imágenes) almacena la imagen 905 reconstruida para su uso en la predicción de la siguiente imagen. En algunas realizaciones, el decodificador 900 aplica un filtro de desbloqueo a la imagen reconstruida para suavizar de forma adaptativa las discontinuidades y otros artefactos en la imagen.

III. Codificación de plano de bits

15 En algunas implementaciones, la información binaria específica de macrobloque puede codificarse como un plano de bits y transmitirse a un nivel superior al nivel de macrobloque en una secuencia de bits (por ejemplo, nivel de trama, nivel de campo, o algún otro nivel). Un codificador puede seleccionar entre varios modos de codificación, basado en, por ejemplo, la eficiencia relativa de los diferentes modos de codificación para codificar ciertos tipos de datos. Las técnicas de codificación de plano de bits aprovechan la correlación espacial de ciertos tipos de información binaria entre macrobloques en una trama o campo. Cuando se combina con la señalización a nivel de macrobloque, las técnicas y herramientas descritas proporcionan señalización flexible y eficiente de información específica de macrobloques.

20 En algunas implementaciones, un codificador usa los elementos de sintaxis INVERT, IMODE y DATABITS para codificación de plano de bits, como se describe a continuación.

25 El elemento INVERT es un código de un bit que indica si el plano de bits tiene más bits iguales a 0 o más bits iguales a 1. Dependiendo de INVERT y el modo de codificación de plano de bits, el decodificador puede invertir el plano de bits decodificado para recrear el original. El elemento IMODE es un VLC que representa el modo de codificación de plano de bits. En general, los códigos más cortos se utilizan para codificar modos de codificación más frecuentes. El elemento DATABITS es una secuencia de símbolos codificada por entropía basada en el modo de codificación indicado en el elemento IMODE. Por ejemplo, la información binaria para macrobloques en una trama o campo puede codificarse como matrices bidimensionales y transmitirse en un encabezado de trama o campo. El tamaño de cada matriz es $rowMB \times colMB$, en que $rowMB$ y $colMB$ son el número de filas y columnas de macrobloques, respectivamente, en la trama o campo.

35 En algunas realizaciones, cada matriz se codifica como un conjunto de bits consecutivos (por ejemplo, representando macrobloques en orden de escaneo de ráster) en uno de los siete modos de codificación de plano de bits. Los siete modos de codificación de plano de bits se describen a continuación.

Modo saltar fila

En modo Saltar fila, el elemento ROWSKIP indica si el elemento ROWBITS está presente para cada fila en el plano de bits. Si una fila completa de valores en el plano de bits es cero, ROWSKIP = 0 y ROWBITS se omite. Si al menos un valor en la fila no es cero, ROWSKIP = 1 y ROWBITS contiene un bit para cada valor en la fila.

40 *Modo saltar columna*

En modo saltar columna, el elemento COLUMNSKIP indica si el elemento COLUMNBITS está presente para cada columna en el plano de bits. Si una columna completa de valores en el plano de bits es cero, COLUMNSKIP = 0 y COLUMNBITS se omite. Si al menos un valor en la columna no es cero, COLUMNSKIP = 1 y COLUMNBITS contiene un bit para cada valor en la columna.

45 *Modo normal-2*

En modo Normal-2, Si $rowMB \times colMB$ es impar, el primer símbolo se representa con un bit que coincide con el valor del primer símbolo, y los símbolos posteriores se codifican en pares usando una tabla VLC.

Modo normal-6

50 En modo Normal-6, el plano de bits está codificado en grupos de seis píxeles en mosaicos de 2x3 o 3x2. El plano de bits se agrupa en mosaicos de 2x3 y 3x2 utilizando un conjunto de reglas de mosaico, y los píxeles restantes (si los hay) se codifican utilizando una variante de los modos Saltar fila y Saltar columna. Si el tamaño de la matriz es un múltiplo de 3x2 o de 2x3, el plano de bits está en mosaico con solo mosaicos rectangulares de seis elementos. Se utilizan mosaicos "verticales" de 3x2 si y solo si $rowMB$ es un múltiplo de tres y $colMB$ no lo es. En caso contrario, se utilizan mosaicos "horizontales" 2x3. Los mosaicos rectangulares de seis elementos se codifican mediante una

combinación de códigos de longitud variable y de longitud fija.

Modos Diff-2 y Diff-6

5 Si se utiliza cualquiera de los modos diferenciales (Diff-2 o Diff-6), se decodifica un plano de bits de "bits diferenciales" utilizando el modo normal correspondiente (Normal-2 o Normal-6, respectivamente). Los bits diferenciales se utilizan para regenerar el plano de bits original.

El procedimiento de regeneración es un DPCM 2-D en un alfabeto binario. Para regenerar el bit en la ubicación (i, j), el predictor $b_p(i,j)$ se genera de la siguiente manera (a partir de los bits $b(i, j)$ en las posiciones (i, j)):

$$b_p(i,j) = \begin{cases} A & i = j = 0, \text{ o } b(i,j-1) \neq b(i-1,j) \\ b(0,j-1) & i == 0 \\ b(i-1,j) & \text{en caso contrario} \end{cases}$$

10 Para el modo de codificación diferencial, no se realiza el procedimiento de inversión bit a bit basado en INVERT. Sin embargo, el indicador INVERT se usa en una capacidad diferente para indicar el valor del símbolo A para la derivación del predictor que se muestra arriba. Más específicamente, A = 0 si INVERT = 0 y A = 1 si INVERT = 1. El valor real del plano de bits se obtiene haciendo un proceso xor del predictor con el valor de bit diferencial decodificado. En la ecuación anterior, $b(i,j)$ es el bit en la posición i, j-ésima después de la decodificación final (es decir, después de hacer Norm-2/Norm-6, seguido de diferencial xor con su predictor).

Modo en bruto

15 El modo en bruto codifica la información binaria específica del macrobloque con un bit por símbolo binario. Aunque los bits en modo en bruto se pueden codificar a nivel de trama o campo en un plano de bits sin comprimir, los bits de modo en bruto también se pueden codificar de uno en uno para cada macrobloque en la misma capa en la secuencia de bits que el resto de la información de nivel de macrobloque para el macrobloque.

20 Otros modos codifican planos de bits a nivel de trama o campo y utilizan una segunda pasada a través de la trama o campo durante la codificación. Por ejemplo, el codificador recopila la información binaria para cada macrobloque (por ejemplo, estado de predicción AC) a nivel de macrobloque en la primera pasada, y codifica la información específica de macrobloque a nivel de trama o campo en la segunda pasada. Sin embargo, en situaciones de baja latencia, evitar la segunda pasada puede ser beneficioso. Por lo tanto, la capacidad de cambiar la codificación de la información binaria del nivel de trama o campo al nivel de macrobloque proporciona flexibilidad adicional.

25 En algunas implementaciones, un codificador usa el modo en bruto para tramas codificadas en segmentos. Un segmento representa una o más filas contiguas de macrobloques. Con respecto a la predicción espacial, la primera fila de macrobloques en el segmento generalmente se trata como la primera fila de macrobloques en una imagen. Cuando se usan segmentos, la información que generalmente se representa en un plano de bits en la capa de imagen se señala en modo en bruto en la capa de macrobloque, para que cada macrobloque lleve su propia información local.

30 Además de los modos de codificación de plano de bits descritos anteriormente, se pueden utilizar otros modos de codificación de plano de bits. Por ejemplo, un codificador puede codificar un plano de bits en grupos de píxeles de tamaño arbitrario. En una variación del modo Normal-6, un codificador puede emplear un modo de "rectángulos más grandes", como cuando la información binaria en el plano de bits se organiza en un patrón predecible. El codificador puede agrupar el plano de bits en mosaicos de tamaño arbitrario utilizando un conjunto diferente de reglas de mosaico. Los mosaicos se pueden codificar así utilizando una combinación de códigos de longitud variable y de longitud fija, o alguna otra disposición de codificación.

35 Para más información sobre la codificación de plano de bits en algunas implementaciones, véanse las Secciones IV, V y VI, a continuación.

40 **IV. Innovaciones en la señalización de tipo campo/trama para imágenes codificadas con trama entrelazada**

Las realizaciones descritas incluyen técnicas y herramientas para señalar información de tipo de campo/trama en imágenes codificadas con trama entrelazada (por ejemplo, tramas I entrelazadas, tramas P entrelazadas, tramas B entrelazadas, etc.). Por ejemplo, las técnicas y herramientas descritas incluyen un codificador capaz de utilizar la codificación de plano de bits a nivel de trama o la codificación a nivel de macrobloque para codificar información de tipo de campo/trama. Un decodificador realiza la decodificación correspondiente. Las técnicas y herramientas descritas se pueden usar en combinación entre sí o con otras técnicas y herramientas, o se pueden usar de forma independiente.

45 En general, los macrobloques codificados en campo son más eficientes cuando hay un alto movimiento entre campos, y los macrobloques codificados por trama son más eficientes cuando hay un movimiento más bajo. Un codificador puede elegir un tipo de campo/trama para un macrobloque determinado e indicar si está codificado por trama o

codificado por campo en la secuencia de bits. El tipo de campo/trama de un macrobloque indica la organización interna de un macrobloque. Refiriéndose nuevamente a la Figura 6, en codificación de campo, un macrobloque 610 entrelazado está permutado de modo que todas las líneas de campo superior (por ejemplo, líneas pares 0, 2,... 14) en los bloques de luminancia se colocan en la mitad superior del macrobloque 620 codificado por campo, y todas las líneas de campo inferior (por ejemplo, líneas impares 1, 3,... 15) en los bloques de luminancia se colocan en la mitad inferior del macrobloque codificado por campo. Para un macrobloque codificado por trama, las líneas de campo superior y las líneas de campo inferior se alternan en todo el macrobloque, como en el macrobloque 610 entrelazado. Los bloques de crominancia permanecen entrelazados tanto para macrobloques codificados por campo como para macrobloques codificados por trama.

Por consiguiente, en algunas implementaciones, un codificador selecciona un bit por modo de macrobloque entre varios modos de codificación y envía un solo bit (por ejemplo, el indicador FIELDTX) a nivel de macrobloque para cada macrobloque en una trama para el cual se toma la decisión de modo de informar explícitamente al decodificador si el macrobloque está codificado por campo o por trama. La capacidad de seleccionar un nivel de macrobloque, un bit por modo de macrobloque o un modo a nivel de trama, comprimido de plano de bits, proporciona flexibilidad adicional para codificar dicha información.

En particular, la señalización a nivel de macrobloque puede ser eficiente cuando un bit de decisión de modo de codificación de campo/trama se señala para algunos, pero no todos, los macrobloques de una trama. Por ejemplo, en una trama P entrelazada, el bit de decisión de campo/trama se señala para intra macrobloques pero no entre macrobloques (para los cuales el bit de decisión de campo/trama se señala de otro modo). La codificación de plano de bits en dicho contexto da como resultado la señalización de información extraña. Por otra parte, en algunas tramas, cada macrobloque tiene un bit de decisión de campo/trama. Agrupar los bits de tipo campo/trama juntos y codificarlos en un plano de bits a nivel de trama puede aprovechar la correlación espacial en el tipo de campo/trama que puede existir entre macrobloques en la trama.

Por ejemplo, la Figura 12 muestra una técnica 1200 para señalar información de tipo de campo/trama para macrobloques utilizando codificación de plano de bits o codificación a nivel de macrobloque. Un codificador selecciona un modo de codificación en 1210. En 1220, si el modo de codificación es un modo de codificación de plano de bits, el codificador señala la información del tipo de campo/trama como un plano de bits en 1230. En caso contrario, el codificador determina en 1240 si enviar un bit de señal (en 1250) para indicar una decisión de campo/trama para el macrobloque. En 1260, si hay más macrobloques para los que se debe tomar una decisión de señalización, el codificador determina si enviar bits de señal para esos macrobloques. Un decodificador realiza la decodificación correspondiente.

Por ejemplo, en una implementación, FIELDTX es un elemento de sintaxis de 1 bit presente en todos los macrobloques de trama I entrelazados y macrobloques de trama BI entrelazados, y en macrobloques intracodificados en tramas P y B entrelazadas. FIELDTX = 1 indica que el macrobloque está codificado por campo, y FIELDTX = 0 indica que el macrobloque está codificado por trama. Los bits FIELDTX pueden codificarse en plano de bits a nivel de trama en la secuencia de bits, o señalizarse a nivel de macrobloque en un bit por macrobloque para algunos o todos los macrobloques. Por ejemplo, en tramas I entrelazadas y tramas BI entrelazadas, los bits FIELDTX están codificados en plano de bits a nivel de trama en la secuencia de bits, o a nivel de macrobloque con base en un bit por macrobloque. En tramas P entrelazadas y tramas B entrelazadas, los bits FIELDTX se envían para intra macrobloques a nivel de macrobloque en un bit por macrobloque.

Como alternativa, el tipo de campo/trama se señala de manera diferente. Por ejemplo, el tipo de campo/trama puede señalizarse (por ejemplo, a nivel de macrobloque) con elementos de secuencia de bits de diferentes tamaños o disposiciones. Como otra alternativa, el tipo de campo/trama se puede señalar a un nivel inferior al nivel de trama que no sea el nivel de macrobloque (por ejemplo, nivel de segmento o algún otro nivel). Como otra alternativa, el tipo de campo/trama puede señalizarse selectivamente a nivel de macrobloque (o algún otro nivel) por menos de todos los macrobloques en una trama intracodificada.

V. Innovaciones en la señalización de predicción AC para imágenes progresivas y entrelazadas

Las técnicas descritas incluyen técnicas y herramientas para señalar información de predicción AC en imágenes progresivas y entrelazadas (por ejemplo, tramas I progresivas, tramas I entrelazadas, etc.). Por ejemplo, las técnicas y herramientas descritas incluyen técnicas y herramientas para señalar el uso de la predicción AC en macrobloques en una trama o campo utilizando codificación de plano de bits. Las técnicas y herramientas descritas se pueden usar en combinación entre sí o con otras técnicas y herramientas, o se pueden usar de forma independiente.

Por ejemplo, un codificador/decodificador predice coeficientes AC antes de la codificación por entropía. El codificador resta de los valores del coeficiente AC en el bloque actual los valores del coeficiente AC correspondientes en un bloque predictor vecino para obtener diferenciales del coeficiente AC. Los diferenciales se codifican por entropía. Dependiendo de la dirección de predicción, el bloque predictor es el bloque inmediatamente superior o el bloque a la izquierda del bloque actual. Refiriéndose nuevamente a la Figura 1B, para la predicción superior, la fila 177 superior de coeficientes AC en el bloque 175 inmediatamente superior al bloque actual se usa como el predictor para la fila 129 superior de coeficientes AC en el bloque 125 actual. Para la predicción izquierda, la columna 137 más a la izquierda

de los coeficientes AC en el bloque 135 a la izquierda inmediata del bloque actual se usa como el predictor para la columna más a la izquierda de los coeficientes AC en el bloque 125 actual. En una implementación, una dirección de predicción elegida para la predicción del coeficiente DC se usa como la dirección de predicción para la predicción del coeficiente AC. Como alternativa, también se señala una dirección de predicción AC.

- 5 La mecánica precisa de la predicción AC puede ser diferente para diferentes tipos de imágenes. En su forma más simple, la predicción AC implica simplemente usar los coeficientes AC de la fila superior o de la columna izquierda de un bloque vecino como predictores de coeficientes. En predicciones AC más complejas, los predictores de coeficientes se escalan por un factor relacionado con los niveles de cuantificación en el bloque actual y el bloque vecino. Como alternativa, el codificador y el decodificador usan predicción AC con otras mecánicas.
- 10 Si no existe un bloque en la dirección de predicción especificada, los valores pronosticados para los siete coeficientes AC en la columna más a la izquierda o en la fila superior se establecen en cero. Por ejemplo, si la dirección de predicción está hacia arriba y el bloque actual está en la fila superior, cada uno de los coeficientes AC previstos en la fila superior del bloque actual se establece en cero porque no hay un bloque adyacente en la dirección hacia arriba. Los coeficientes AC en la fila o columna prevista se agregan a los coeficientes AC decodificados correspondientes (que son diferenciales) en el bloque actual para producir el bloque de coeficiente de transformación cuantificado completamente reconstruido.

- Realizar predicciones de coeficientes AC para todos los macrobloques en una trama o campo puede ser ineficiente; la predicción será ineficaz en algunos macrobloques. Por lo tanto, el codificador/decodificador usa señalización para indicar si la predicción AC está habilitada o deshabilitada para macrobloques individuales en una trama o campo. Se puede enviar un solo bit de información para cada macrobloque a nivel de macrobloque para informar explícitamente al decodificador si se usa predicción AC en el macrobloque. Sin embargo, un codificador/decodificador que señala la predicción AC solo enviando un bit por macrobloque no puede aprovechar la posible correlación espacial en el estado de predicción AC para macrobloques en el campo o trama.
- 20

- Por consiguiente, en algunas implementaciones, un codificador/decodificador explota dicha correlación espacial en una trama o campo al agrupar la información de predicción AC para macrobloques individuales y codificar la información como un plano de bits (por ejemplo, a nivel de trama, nivel de campo, o algún otro nivel por encima del nivel de macrobloque). En una implementación, el plano de bits es el plano de bits a nivel de trama/nivel de campo ACPRED. El plano de bits ACPRED se describe con más detalle en la Sección VII, a continuación.
- 25

- La Figura 13 muestra una técnica 1300 para la codificación o decodificación de plano de bits de la información de predicción AC. En 1310, un codificador/decodificador selecciona un modo de codificación de un grupo de múltiples modos de codificación disponibles. En 1320, el codificador/decodificador procesa un plano de bits de acuerdo con el modo de codificación seleccionado, comprendiendo el plano de bits información binaria que significa si los macrobloques en una trama de video están codificados mediante predicción AC.
- 30

- Las técnicas de codificación de plano de bits descritas también se pueden usar en combinación con codificación a nivel de macrobloque, de un bit por macrobloque. Por ejemplo, en una implementación, el elemento de sintaxis ACPRED cuando se señala en una capa de macrobloque indica si la predicción AC se usa para los bloques en ese macrobloque. Cuando ACPRED = 1, los coeficientes AC en bloques decodificados del macrobloque se tratan como valores diferenciales basados en los coeficientes en un bloque predictor. El elemento de secuencia de bits ACPRED a nivel de macrobloque se describe con más detalle en la Sección VII, a continuación.
- 35

- Como alternativa, la información de predicción AC se señala de manera diferente. Por ejemplo, la información de predicción AC puede señalizarse con elementos de secuencia de bits de diferentes tamaños o disposiciones. Como otra alternativa, la información de predicción AC puede codificarse en un plano de bits a un nivel superior al nivel de macrobloque que no sea el nivel de trama o el nivel de campo (por ejemplo, nivel de segmento o algún otro nivel). Como otra alternativa, la información de predicción AC puede señalizarse selectivamente a nivel de macrobloque (o algún otro nivel) para menos de todos los macrobloques en una trama intracodificada, con codificación de plano de bits (o alguna otra técnica de codificación) potencialmente utilizada a nivel de trama (o algún otro nivel) para señalar el tipo de campo/trama para macrobloques no señalizados a nivel de macrobloque.
- 40
- 45

VI. Implementaciones combinadas

- Una implementación combinada detallada para una sintaxis de secuencia de bits, semántica y decodificador se describe ahora, además de una implementación combinada alternativa con pequeñas diferencias con respecto a la implementación combinada principal.
- 50

A. Sintaxis de secuencia de bits

- En varias implementaciones combinadas, los datos se presentan en forma de una secuencia de bits que tiene varias capas (por ejemplo, secuencia, punto de entrada, trama, campo, segmento, macrobloque, capas de bloques y/o subbloques).
- 55

En los diagramas de sintaxis, las rutas de flecha muestran los posibles flujos de elementos de sintaxis. Los elementos

de sintaxis que se muestran con límites de bordes cuadrados indican elementos de sintaxis de longitud fija; aquellos con límites redondeados indican elementos de sintaxis de longitud variable y aquellos con un límite redondeado dentro de un límite redondeado externo indican un elemento de sintaxis (por ejemplo, un plano de bits) compuesto por elementos de sintaxis más simples. Un elemento de sintaxis de longitud fija se define como un elemento de sintaxis para el cual la longitud del elemento de sintaxis no depende de los datos en el elemento de sintaxis en sí; la longitud de un elemento de sintaxis de longitud fija es constante o está determinada por datos anteriores en el flujo de sintaxis. Una capa inferior en un diagrama de capas (por ejemplo, una capa de macrobloque en un diagrama de capa de trama) se indica mediante un rectángulo dentro de un rectángulo.

Los elementos de nivel de secuencia se utilizan para decodificar una secuencia de imágenes de vídeo comprimido. Los datos de nivel de secuencia pueden afectar a la interpretación o la presencia de elementos en otros niveles (por ejemplo, nivel de punto de entrada, nivel de trama, nivel de macrobloque, etc.) En general, un punto de entrada marca una posición en una secuencia de bits (por ejemplo, una trama I u otra trama clave) en la que un decodificador puede comenzar a decodificar. En otras palabras, no se necesitan imágenes antes del punto de entrada en la secuencia de bits para decodificar imágenes después del punto de entrada. Se puede usar un encabezado de punto de entrada para señalar cambios en los parámetros de control de codificación (por ejemplo, habilitar o deshabilitar herramientas de compresión para tramas que siguen a un punto de entrada).

Para tramas I progresivas, los elementos de secuencia de bits a nivel de trama se muestran en la Figura 14. (Los elementos de secuencia de bits a nivel de trama para las tramas BI progresivas son idénticos a los de las tramas I progresivas.) Los datos para cada trama consisten en un encabezado de trama seguido de datos para la capa de macrobloques (para intra macrobloques). Los elementos de secuencia de bits a nivel de macrobloque para tramas I progresivas se muestran en la Figura 15.

Para tramas de vídeo entrelazado con campos I entrelazados y/o campos BI, los elementos de secuencia de bits a nivel de trama se muestran en la Figura 16. Los datos para cada trama consisten en un encabezado de trama seguido de datos para las capas de campo (se muestra como el elemento repetido "FieldPicLayer" por campo). Los elementos de secuencia de bits que componen los encabezados de campo para campos I entrelazados se muestran en la Figura 17. (Los elementos de secuencia de bits a nivel de campo para campos BI entrelazados son idénticos a los de los campos I entrelazados.) Los elementos de secuencia de bits que forman la capa de macrobloques para campos I entrelazados y campos BI entrelazados son idénticos a los de macrobloques en tramas I progresivas.)

Para tramas I, tramas P y tramas B entrelazadas, los elementos de secuencia de bits a nivel de trama se muestran en las Figuras 18, 19 y 20, respectivamente. (Los elementos de secuencia de bits a nivel de trama para tramas BI entrelazadas son idénticos a los de las tramas I entrelazadas.) En tramas I, tramas P y tramas B entrelazadas, los datos para cada trama consisten en un encabezado de trama seguido de datos para la capa de macrobloques (ya sea para intra macrobloques o varios tipos de inter macrobloques). Los elementos de secuencia de bits que forman la capa de macrobloque para tramas I, tramas P y tramas B entrelazadas se muestran en las Figuras 21, 22 y 23, respectivamente. (Los elementos de secuencia de bits para macrobloques en tramas BI entrelazadas son idénticos a los de macrobloques en tramas I entrelazadas.)

Las siguientes secciones describen elementos de secuencia de bits seleccionados en capas de trama, campo y macrobloques. Aunque los elementos de secuencia de bits seleccionados se describen en el contexto de una capa particular, algunos elementos de secuencia de bits se pueden utilizar en más de una capa.

1. Elementos seleccionados de capa de trama y campo

Las Figuras 14, 16, 18, 19 y 20 son diagramas que muestran sintaxis de secuencia de bits a nivel de trama para tramas I progresivas/tramas BI progresivas, tramas que tienen campos I o campos BI entrelazados, tramas I entrelazadas/tramas BI entrelazadas, tramas P entrelazadas y tramas B entrelazadas, respectivamente. La Figura 17 es un diagrama que muestra una sintaxis de secuencia de bits a nivel de campo para campos I entrelazados. Algunos elementos (p. ej., el plano de bits ACPRED) también están presentes en una sintaxis de secuencia de bits a nivel de campo para campos BI entrelazados, y potencialmente en sintaxis para otros tipos de imágenes. Los elementos de secuencia de bits específicos se describen a continuación.

Modo de codificación de trama (FCM) (tamaño variable)

FCM es una palabra de código de longitud variable ["VLC"] utilizada para indicar el tipo de codificación de imagen. FCM adquiere valores para los modos de codificación de trama como se muestra en la Tabla 1 a continuación:

Tabla 1: Modo de codificación de trama VLC

Valor de FCM	Modo de codificación de trama
0	Progresivo
10	Trama-entrelazada
11	Campo-entrelazado

Tipo de imagen de campo (FPTYPE) (3 bits)

FPTYPE es un elemento de sintaxis de tres bits presente en el encabezado de la trama para una trama que incluye campos I entrelazados y/o campos BI entrelazados, y potencialmente otros tipos de campos. FPTYPE toma valores para diferentes combinaciones de tipos de campo en la trama de vídeo entrelazado, de acuerdo con la Tabla 2 a continuación.

5

Tabla 2: Tipo de imagen de campo FLC

FPTYPE FLC	Primer tipo de campo	Segundo tipo de campo
000	I	I
001	I	P
010	P	I
011	P	P
100	B	B
101	B	BI
110	BI	B
111	BI	BI

Tipo de imagen (PTYPE) (tamaño variable)

PTYPE es un elemento de sintaxis de tamaño variable presente en el encabezado de trama para tramas P entrelazadas y tramas B entrelazadas (u otros tipos de tramas entrelazadas tales como tramas I entrelazadas). PTYPE toma valores para diferentes tipos de trama de acuerdo con la Tabla 3 a continuación.

10

Tabla 3: Tipo de imagen VLC

PTYPE VLC	Tipo de imagen
110	I
0	P
10	B
1110	BI
1111	Omitido

Si PTYPE indica que se omite la trama, entonces la trama se trata como una trama P que es idéntica a su trama de referencia. La reconstrucción de la trama omitida es conceptualmente equivalente a copiar la trama de referencia. Una trama omitida significa que no se transmiten más datos para esta trama.

Plano de bits de transformación de campo (FIELDTX) (tamaño variable)

15 A nivel de trama o de campo, FIELDTX es un plano de bits que indica si los macrobloques en una trama I entrelazada tienen codificación por trama o codificación por campo. FIELDTX se explica con más detalle a continuación y en la Sección IV, anteriormente.

Plano de bits de predicción AC (ACPRED) (tamaño variable)

20 Para tramas I y tramas BI progresivas y para tramas I y tramas BI entrelazadas, los elementos de sintaxis ACPRED de 1 bit que estarían presentes en todos los macrobloques se codifican conjuntamente usando un elemento de sintaxis codificado en un plano de bits que indica el estado de predicción AC para cada macrobloque en la trama. El plano de bits decodificado representa el estado de predicción AC para cada macrobloque como valores de 1 bit. El plano de bits ACPRED se describe con más detalle a continuación y en la Sección V, anteriormente.

3. Elementos seleccionados de la capa de macrobloque

25 Las Figuras 15, 21, 22 y 23 son diagramas que muestran sintaxis de secuencia de bits a nivel de macrobloque para macrobloques en tramas I progresivas/campos I entrelazados/campos BI entrelazados, tramas I entrelazadas/tramas BI entrelazadas, tramas P entrelazadas y tramas B entrelazadas, respectivamente, en la implementación combinada. Los elementos de secuencia de bits específicos se describen a continuación. Los datos para un macrobloque consisten en un encabezado de macrobloque seguido de datos de capa de bloque. Los elementos de secuencia de bits en la
30 capa de macrobloque (por ejemplo, ACPRED, FIELDTX, etc.) también pueden estar presentes para macrobloques para otros tipos de imágenes.

Indicador de predicción AC (ACPRED) (1 bit)

35 En la implementación combinada, el elemento de sintaxis ACPRED está presente en todas las tramas I progresivas, trama I entrelazada, trama BI entrelazada, campo I entrelazado, macrobloques de campo BI entrelazado e intra macrobloques en campos P entrelazados, campos B entrelazados, tramas P entrelazadas y tramas B entrelazadas. ACPRED también está presente en macrobloques en tramas P y tramas B progresivas. A nivel de macrobloque, ACPRED es un elemento de sintaxis de 1 bit que especifica si los bloques en el macrobloque se codificaron mediante predicción AC. ACPRED = 0 indica que no se utiliza la predicción AC. ACPRED = 1 indica que se usa la predicción

AC. ACPRED también se puede codificar como un plano de bits a nivel de trama, como se explica con más detalle en la Sección V, anteriormente.

Indicador de transformación de campo (FIELDTX) (1 bit)

5 FIELDTX es una sintaxis de 1 bit presente en macrobloques intracodificados en tramas I entrelazadas, tramas BI entrelazadas, tramas P entrelazadas y tramas B entrelazadas. Este elemento de sintaxis indica si un macrobloque está codificado por trama o campo (básicamente, la organización interna del macrobloque). FIELDTX = 1 indica que el macrobloque está codificado por campo. En caso contrario, el macrobloque está codificado por trama. En macrobloques intercodificados, este elemento de sintaxis se puede inferir del elemento de secuencia de bits a nivel de macrobloque MBMODE. FIELDTX también se puede codificar como un plano de bits a nivel de trama, como se explica con más detalle en la Sección IV, anteriormente.

B. Codificación de plano de bits

15 La información binaria específica del macrobloque, como los bits de salto, puede codificarse en un símbolo binario por macrobloque. Por ejemplo, si se omite o no un macrobloque se puede indicar con un bit. En estos casos, el estado para todos los macrobloques en un campo o trama puede codificarse como un plano de bits y transmitirse en el encabezado del campo o trama. Una excepción para esta regla es si el modo de codificación de plano de bits se establece en Modo en bruto, en cuyo caso, el estado de cada macrobloque se codifica como un bit por símbolo y se transmite junto con otros elementos de sintaxis de nivel de macrobloque a nivel de macrobloque.

20 La codificación de plano de bits a nivel de trama/campo se utiliza para codificar matrices binarias bidimensionales. El tamaño de cada matriz es $rowMB \times colMB$, en que $rowMB$ y $colMB$ son el número de filas y columnas de macrobloques, respectivamente, en el campo o trama en cuestión. Dentro de la secuencia de bits, cada matriz se codifica como un conjunto de bits consecutivos. Se utiliza uno de los siete modos para codificar cada matriz. Los siete modos son:

- 1. modo en bruto: información codificada como un bit por símbolo y transmitida como parte de la sintaxis de nivel MB;
- 2. modo normal-2: dos símbolos codificados conjuntamente;
- 25 3. modo diferencial-2: codificación diferencial del plano de bits, seguido de la codificación de dos símbolos residuales conjuntamente;
- 4. modo normal-6: seis símbolos codificados conjuntamente;
- 5. modo diferencial-6: codificación diferencial del plano de bits, seguido de codificar seis símbolos residuales conjuntamente;
- 30 6. modo saltar fila: salto de un bit para señalar filas sin bits establecidos; y
- 7. modo saltar columna: salto de un bit para señalar columnas sin bits establecidos.

Los elementos de sintaxis para un plano de bits en el nivel de trama o campo son en la siguiente secuencia: INVERT, IMODE y DATABITS.

Invertir indicador (INVERTIR)

35 El elemento de sintaxis INVERT es un valor de 1 bit, que si se establece indica que el plano de bits tiene más bits establecidos que bits cero. Dependiendo de INVERT y el modo, el decodificador invertirá el plano de bits interpretado para recrear el original. Debe apreciarse que el valor de este bit se ignorará cuando se utilice el modo en bruto. A continuación se proporciona una descripción de cómo se usa el valor INVERT para decodificar el plano de bits.

Modo de codificación (IMODE)

40 El elemento de sintaxis IMODE es un valor de longitud variable que indica el modo de codificación utilizado para codificar el plano de bits. La Tabla 4 muestra la tabla de códigos utilizada para codificar el elemento de sintaxis IMODE. A continuación se proporciona una descripción de cómo se usa el valor IMODE para decodificar el plano de bits.

Tabla 4: Tabla de códigos IMODE VLC

IMODE VLC	Modo de codificación
10	Norm-2
11	Norm-6
010	Saltar fila
011	Saltar columna
001	Diff-2
0001	Diff-6
0000	En bruto

Bits de codificación de plano de bits (DATABITS)

El elemento de sintaxis DATABITS es un elemento de sintaxis de tamaño variable que codifica la secuencia de símbolos para el plano de bits. El procedimiento utilizado para codificar el plano de bits está determinado por el valor de IMODE. Los siete modos de codificación se describen en las siguientes secciones.

Modo en bruto

- 5 En este modo, el plano de bits se codifica como un bit por símbolo (por ejemplo, en orden de escaneo de ráster) y se envía como parte de la capa de macrobloque. Como alternativa, la información se codifica en modo en bruto en el nivel de campo o trama y DATABITS tiene una longitud de rowMB × colMB.

Modo normal-2

- 10 Si rowMB × colMB es impar, el primer símbolo se codifica en bruto. Los símbolos posteriores se codifican por pares, en orden de escaneo natural. La tabla binaria VLC en la Tabla 5 se usa para codificar pares de símbolos.

Tabla 5: Tabla de códigos Norm-2/Diff-2

Símbolo 2n	Símbolo 2n + 1	Palabra de código
0	0	0
1	0	100
0	1	101
1	1	11

Modo Diff-2

El procedimiento Normal-2 se usa para producir el plano de bits como se describió anteriormente, y luego la operación Diff¹ se aplica al plano de bits como se describe a continuación.

15 *Modo normal-6*

- 20 En los modos Norm-6 y Diff-6, el plano de bits está codificado en grupos de seis píxeles. Estos píxeles se agrupan en mosaicos de 2x3 o 3x2. El plano de bits se coloca en mosaico al máximo usando un conjunto de reglas, y los píxeles restantes se codifican usando una variante de los modos saltar fila y saltar columna. Los mosaicos "verticales" de 2x3 se usan si y solo si rowMB es un múltiplo de 3 y colMB no lo es. En caso contrario, se utilizan mosaicos "horizontales" 3x2. La Figura 24A muestra un ejemplo simplificado de mosaicos "verticales" de 2x3. Las Figuras 24B y 24C muestran ejemplos simplificados de mosaicos "horizontales" de 3x2 para los cuales los rectángulos oscuros alargados tienen 1 píxel de ancho y están codificados mediante la codificación de saltar fila y saltar columna. Para un plano en mosaico como se muestra en la Figura 24C, con mosaicos lineales a lo largo de los bordes superior e izquierdo de la imagen, el orden de codificación de los mosaicos sigue el siguiente patrón. Los mosaicos de 6 elementos se codifican primero, seguido por los mosaicos lineales codificados de saltar columna y saltar fila. Si el tamaño de la matriz es un múltiplo de 2x3 o de 3x2, los últimos mosaicos lineales no existen y el plano de bits está perfectamente en mosaico.

- 30 Los mosaicos rectangulares de 6 elementos se codifican usando una combinación de códigos de longitud variable y de longitud fija. Se asume que N es el número de bits establecidos en el mosaico, es decir, 0 ≤ N ≤ 6. Para N < 3, se utiliza una VLC para codificar el mosaico. Para N = 3, un código de escape de longitud fija va seguido por un código de longitud fija de 5 bits. Para N > 3, otro código de escape de longitud fija va seguido por una VLC. Para N > 3, la VLC que sigue al código de escape es idéntica a la VLC utilizada para el código del complemento de este mosaico para el caso N < 3. El escape de longitud fija utilizado para el caso de N > 3 es diferente del código de escape de longitud fija para el caso N = 3. El mosaico rectangular contiene seis bits de información. Se asume que k es el código asociado con el mosaico, en que $k = \sum_i b_i 2^i$, b_i es el valor binario del bit ⁱésimo en orden de escaneo natural dentro del mosaico. De este modo, 0 ≤ k < 64. Se utiliza una combinación de VLC y códigos de escape más códigos de longitud fija para señalar k.

Modo Diff-6

El procedimiento Normal-6 se usa para producir el plano de bits como se describió anteriormente, y luego la operación Diff¹ se aplica al plano de bits como se describe a continuación.

40 *Modo saltar fila*

- 45 En el modo de codificación saltar fila, las filas que son completamente cero se omiten con una sobrecarga de un bit. La sintaxis es la siguiente: para cada fila, un solo bit ROWSKIP indica si se omite la fila; si se omite la fila, el bit ROWSKIP para la siguiente fila es el siguiente; de lo contrario (la fila no se omite), los bits ROWBITS (un bit para cada macrobloque en la fila) son los siguientes. Por lo tanto, si toda la fila es cero, se envía un bit cero como el símbolo ROWSKIP y se omite ROWBITS. Si hay un bit establecido en la fila, ROWSKIP se establece en 1, y toda la fila se envía en bruto (ROWBITS). Las filas se escanean de arriba a abajo del campo o trama.

Modo saltar columna

Saltar columna es la transposición de saltar fila. Las columnas se escanean de izquierda a derecha del campo o trama.

Diff¹:: Decodificación diferencial inversa

5 Si se utiliza cualquiera de los modos diferenciales (Diff-2 o Diff-6), primero se decodifica un plano de bits de "bits diferenciales" utilizando los modos normales correspondientes (Norm-2 o Norm-6 respectivamente). Los bits diferenciales se utilizan para regenerar el plano de bits original. El procedimiento de regeneración es un DPCM 2-D en un alfabeto binario. Para regenerar el bit en la ubicación (i, j), el predictor bp(i, j) se genera de la siguiente manera (a partir de los bits b(i, j) en las posiciones (i, j)):

$$b_p(i,j) = \begin{cases} A & i = j = 0, \text{ o } b(i,j-1) \neq b(i-1,j) \\ b(0,j-1) & i == 0 \\ b(i-1,j) & \text{en caso contrario} \end{cases}$$

10 Para el modo de codificación diferencial, no se realiza el procedimiento de inversión bit a bit basado en INVERT. Sin embargo, el indicador INVERT se usa en una capacidad diferente para indicar el valor del símbolo A para la derivación del predictor que se muestra arriba. Más específicamente, A = 0 si INVERT = 0 y A = 1 si INVERT = 1. El valor real del plano de bits se obtiene haciendo un proceso xor del predictor con el valor de bit diferencial decodificado. En la ecuación anterior, b(i,j) es el bit en la posición i, j-ésima después de la decodificación final (es decir, después de hacer Norm-2/Norm-6, seguido de diferencial xor con su predictor).

15 Habiendo descrito e ilustrado los principios de la invención con referencia a diversas realizaciones, se reconocerá que las diversas realizaciones pueden modificarse en disposición y detalle sin apartarse de tales principios. Debe entenderse que los programas, procedimientos o procesamientos descritos en el presente documento no están relacionados o limitados a ningún tipo particular de entorno informático, a menos que se indique lo contrario. Se pueden usar varios tipos de entornos informáticos de fin general o especializado o realizar operaciones de acuerdo con las enseñanzas descritas en el presente documento. Los elementos de las realizaciones mostradas en software pueden
20 implementarse en hardware y viceversa.

En vista de las muchas realizaciones posibles a las que se pueden aplicar los principios de la invención, se reclaman como la invención todas esas realizaciones que puedan entrar dentro del ámbito de las siguientes reivindicaciones y equivalentes a las mismas.

REIVINDICACIONES

1. Un procedimiento que comprende:

5 para una trama (1100; 1130) I entrelazada en una secuencia de vídeo, decodificar un plano de bits señalado (1230) en la capa (1800) de trama de una secuencia de bits para la trama I entrelazada, en el que el plano de bits representa los tipos de transformación de campo/trama para una pluralidad de macrobloques (1000; 1131, 1132; 1161, 1162) de la trama I entrelazada; y
 10 para una trama P entrelazada en la secuencia de vídeo, para cada uno de al menos uno pero no todos de la pluralidad de macrobloques de la trama P entrelazada, procesar un bit de tipo de transformación de campo/trama por macrobloque señalado (1250) en la capa de macrobloque de la secuencia de bits, en el que la trama P entrelazada incluye uno o más intra macrobloques y uno o más inter macrobloques que están codificados en la secuencia de bits, y en el que el bit de tipo de transformación de campo/trama por macrobloque se señala para cada uno de los uno o más intra macrobloques pero ninguno de los uno o más inter macrobloques que están codificados en la secuencia de bits.

15 2. El procedimiento de la reivindicación 1, en el que un elemento de secuencia de bits diferente representa el tipo de transformación de campo/trama para cada uno de los uno o más inter macrobloques que están codificados en la secuencia de bits.

3. El procedimiento de la reivindicación 1 en el que la decodificación del plano de bits incluye seleccionar un modo de decodificación entre una pluralidad de modos de decodificación disponibles.

20 4. El procedimiento de la reivindicación 3, en el que la pluralidad de los modos de decodificación disponibles incluyen saltar fila, saltar columna, VLC por pares, VLC de grupos de seis y uno o más modos diferenciales.

5. El procedimiento de la reivindicación 1 que además comprende:
 para la trama I entrelazada, después de decodificar el plano de bits, para cada uno de la pluralidad de macrobloques de la trama I entrelazada, procesar información de capa de macrobloque para el macrobloque basado, al menos en parte, en el tipo de transformación de campo/trama señalado para el macrobloque en el plano de bits.

25 6. Un medio legible por computadora que tiene almacenado en el mismo instrucciones ejecutables por computadora para hacer que un sistema informático programado de ese modo realice el procedimiento de la reivindicación 1.

7. Un procedimiento que comprende:

30 para una trama (1100; 1130) I entrelazada en una secuencia de vídeo, codificar un plano de bits y señalar (1230) el plano de bits en la capa (1800) de trama de una secuencia de bits para la trama I entrelazada, en el que el plano de bits representa los tipos de transformación de campo/trama para la pluralidad de macrobloques (1000; 1131, 1132; 1161, 1162) de la trama I entrelazada; y
 35 para una trama P entrelazada en la secuencia de vídeo, para cada uno de al menos uno pero no todos de la pluralidad de macrobloques de la trama P entrelazada, señalar (1250) un bit de tipo de transformación de trama/campo por macrobloque en la capa (2200) de macrobloque de la secuencia de bits, en el que la trama P entrelazada incluye uno o más intra macrobloques y uno o más inter macrobloques que están codificados en la secuencia de bits, y en el que el bit de tipo de transformación de campo/trama por macrobloque se señala para cada uno de los uno o más intra macrobloques pero ninguno de los uno o más inter macrobloques que están codificados en la secuencia de bits.

40 8. El procedimiento de la reivindicación 7, en el que un elemento de secuencia de bits diferente representa el tipo de transformación de campo/trama para cada uno de los uno o más inter macrobloques que están codificados en la secuencia de bits.

9. El procedimiento de la reivindicación 7, en el que la codificación del plano de bits incluye seleccionar un modo de codificación entre una pluralidad de modos de codificación disponibles.

45 10. El procedimiento de la reivindicación 9 en el que la pluralidad de modos de codificación disponibles incluyen saltar fila, saltar columna, VLC por pares, VLC de grupos de seis y uno o más modos diferenciales.

11. Un medio legible por computadora que tiene almacenado en el mismo instrucciones ejecutables por computadora para hacer que un sistema informático programado de ese modo realice el procedimiento de la reivindicación 7.

12. Un decodificador (900) que comprende:

50 medios para decodificar planos de bits señalados (1230) en la capa (1800) de trama de la secuencia de bits, en el que los planos de bits incluyen un primer plano de bits para una trama I (1100; 1130) entrelazada, y en el que el primer plano de bits representa los tipos de transformación de campo/trama para una pluralidad de macrobloques (1000; 1131, 1132; 1161, 1162) de la trama I entrelazada; y
 medios para procesar los bits de tipo de transformación de trama/campo por macrobloque señalados (1250) en la capa (2200) de macrobloque de la secuencia de bits, en el que los bits de tipo de transformación de trama/campo

- 5 por macrobloque incluyen un bit de tipo de transformación de trama/campo por macrobloque para cada uno o más pero no todos de la pluralidad de macrobloques de una trama P entrelazada, en el que la trama P entrelazada incluye uno o más intra macrobloques y uno o más inter macrobloques que están codificados en la secuencia de bits, y en el que el bit de tipo de transformación de campo/trama por macrobloque se señala para cada uno de los uno o más intra macrobloques pero ninguno de los uno o más inter macrobloques que están codificados en la secuencia de bits.
13. El decodificador de la reivindicación 12 que comprende además: medios para seleccionar un modo de decodificación de plano de bits entre la pluralidad de modos de decodificación de plano de bits disponibles.
- 10 14. El decodificador de la reivindicación 13 en el que la pluralidad de modos de codificación disponibles incluyen saltar fila, saltar columna, VLC por pares, VLC de grupos de seis y uno o más modos diferenciales.
15. El decodificador de la reivindicación 14, en el que un elemento de secuencia de bits diferente representa el tipo de transformación de campo/trama para cada uno de los uno o más inter macrobloques que están codificados en la secuencia de bits.

Figura 1A,
técnica anterior

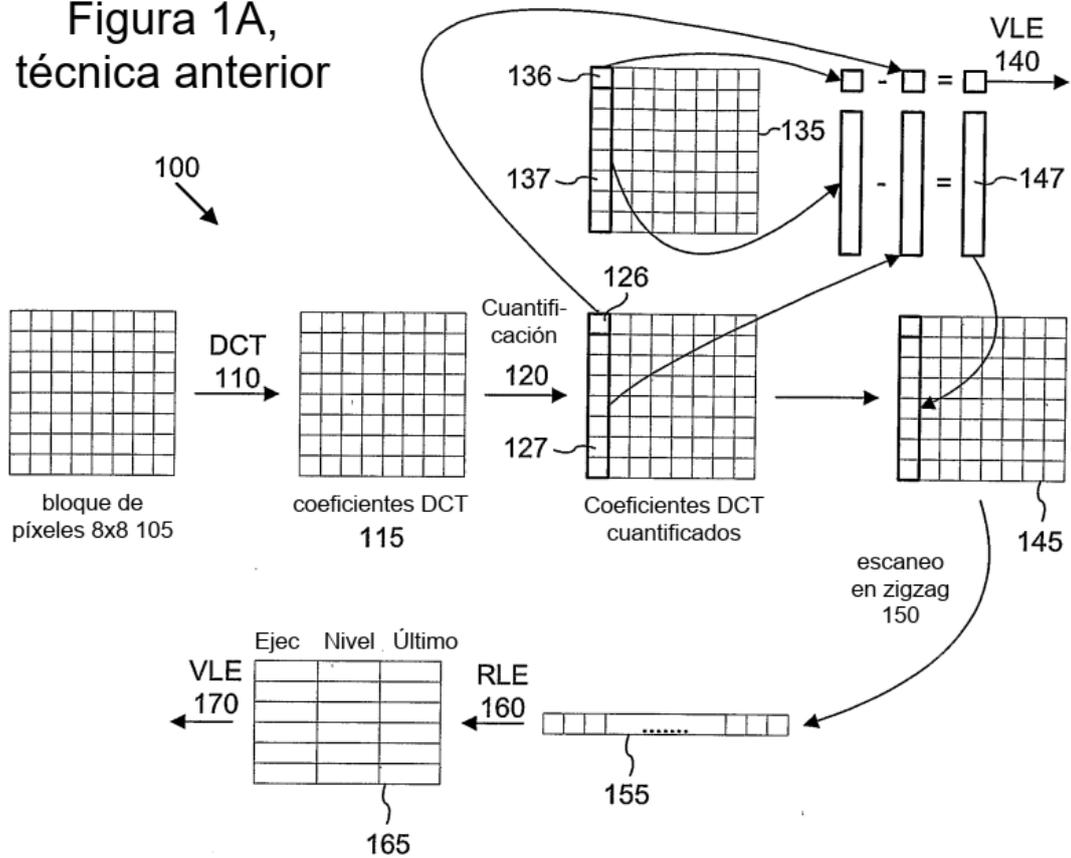


Figura 1B, técnica anterior

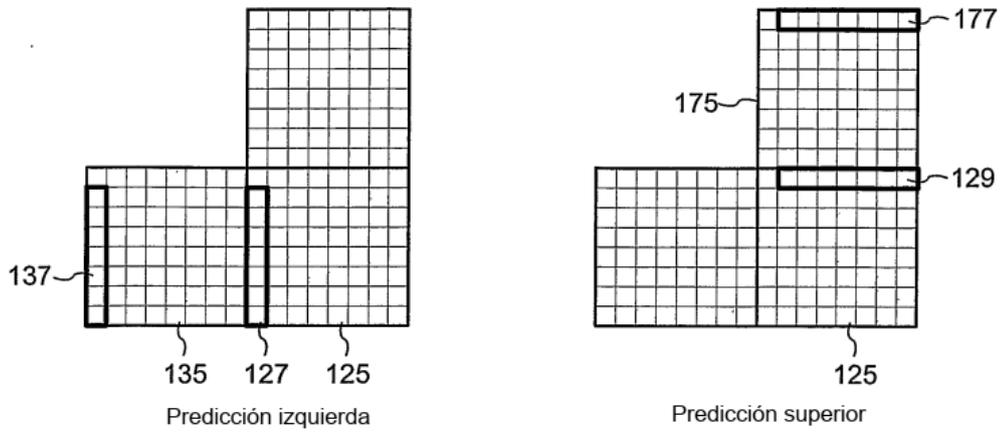


Figura 2, técnica anterior

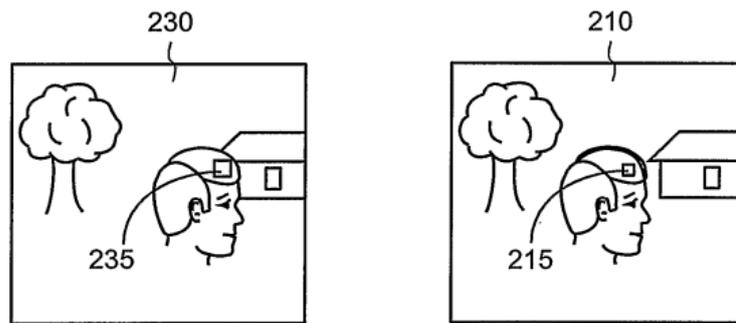


Figura 5, técnica anterior

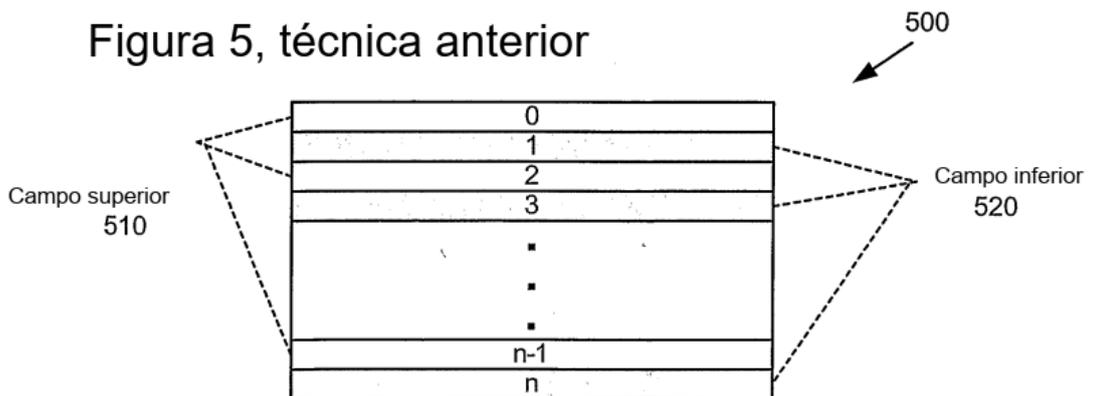


Figura 3, técnica anterior

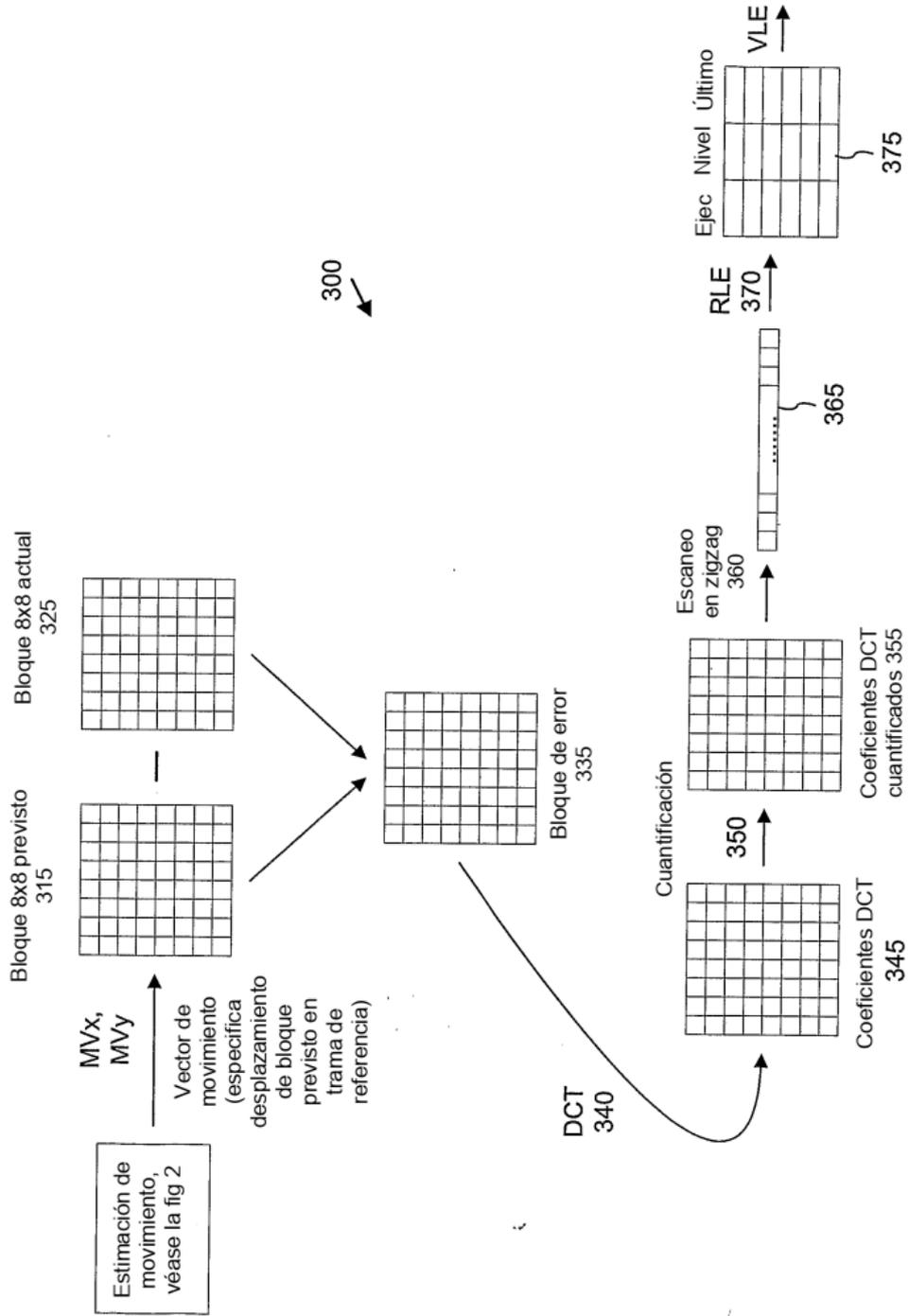


Figura 4, técnica anterior

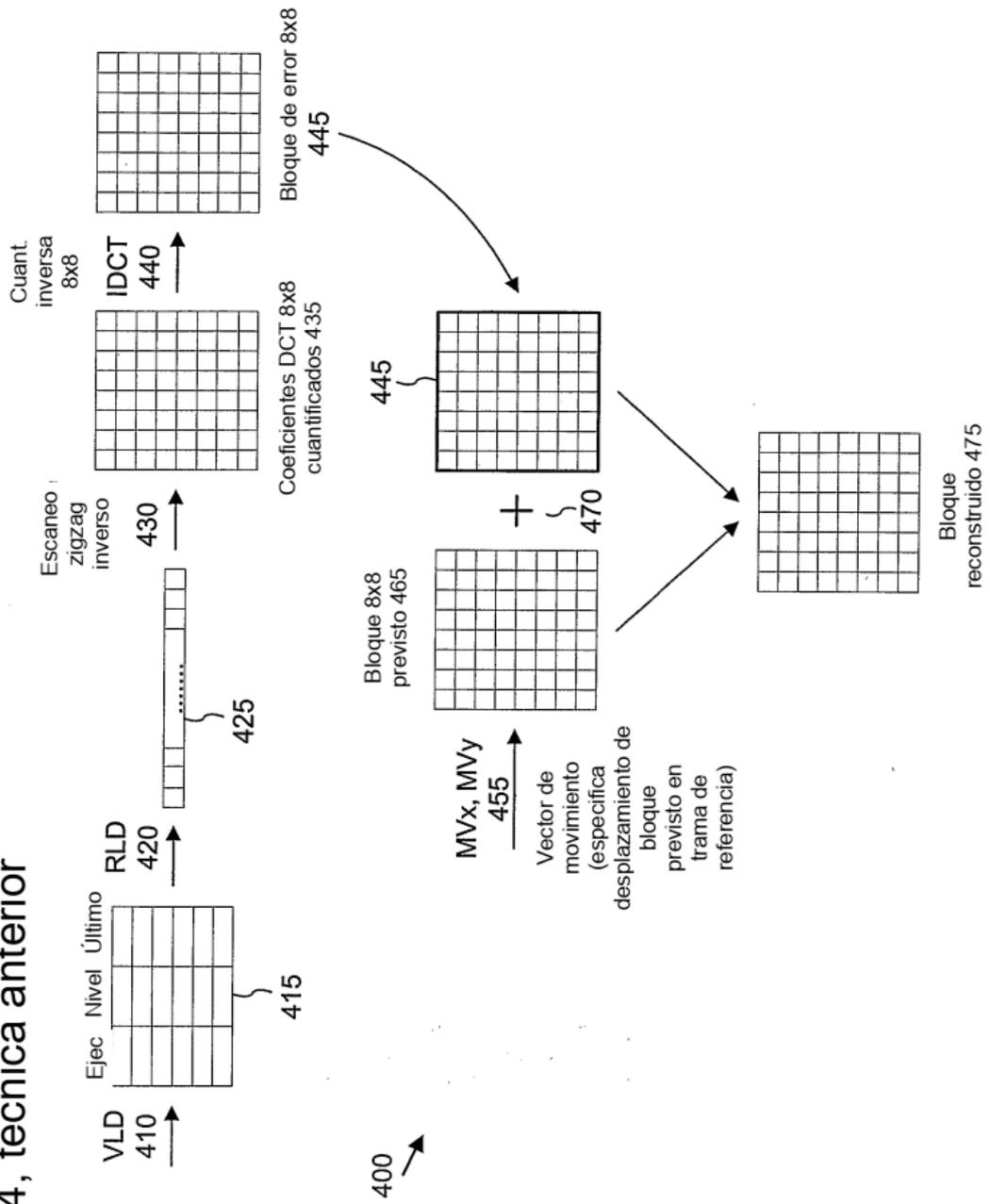


Figura 6, técnica anterior

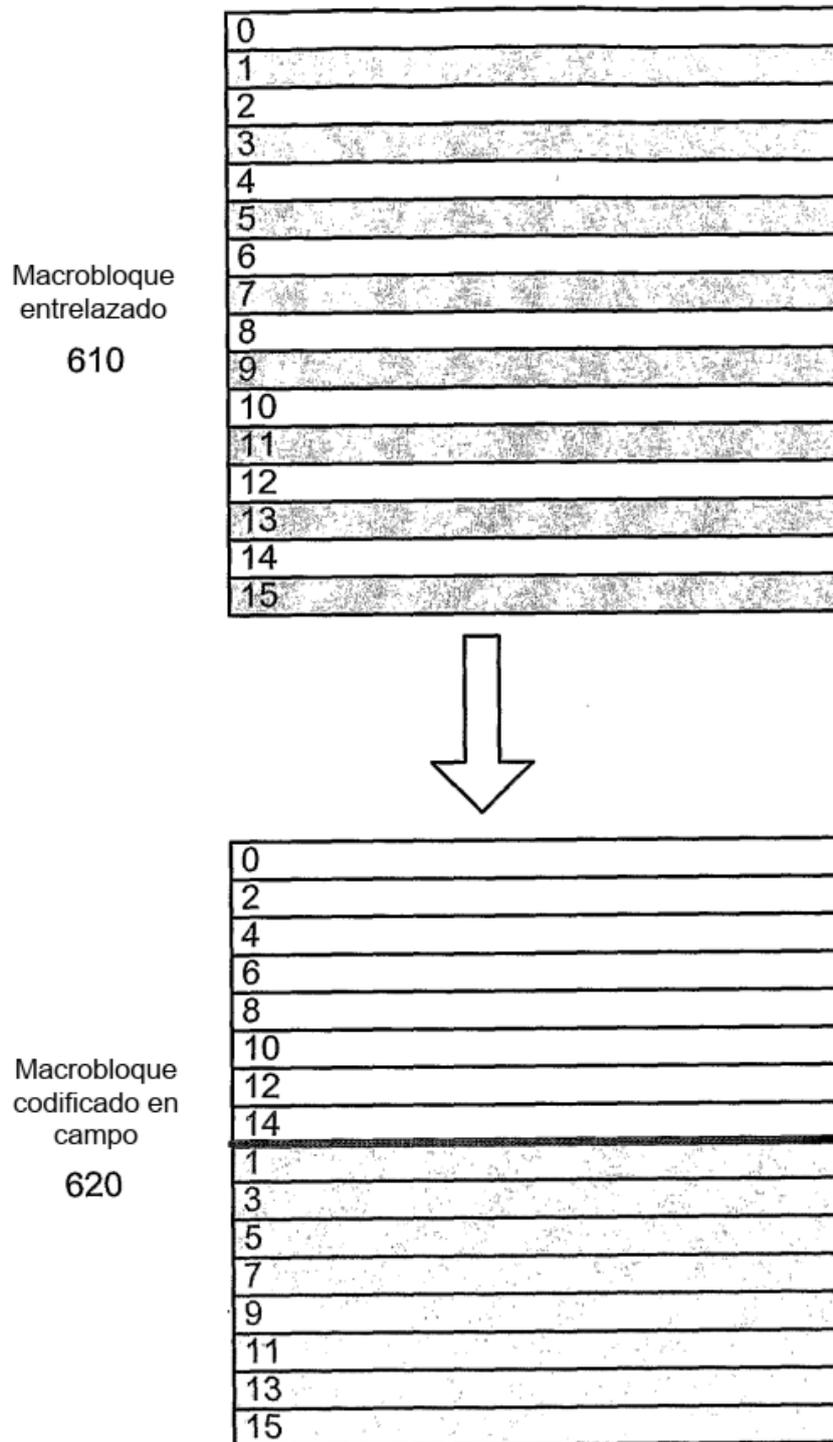


Figura 7

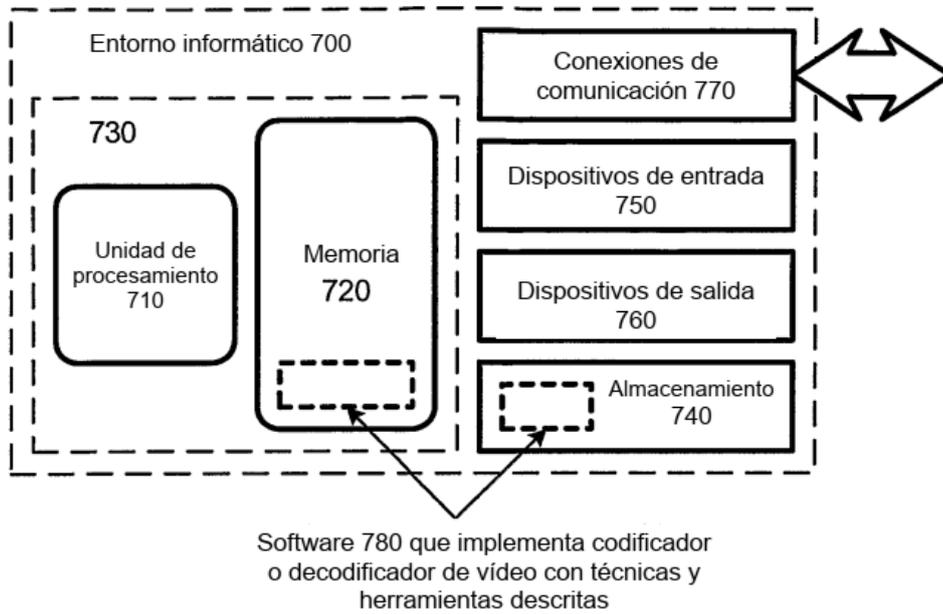


Figura 8

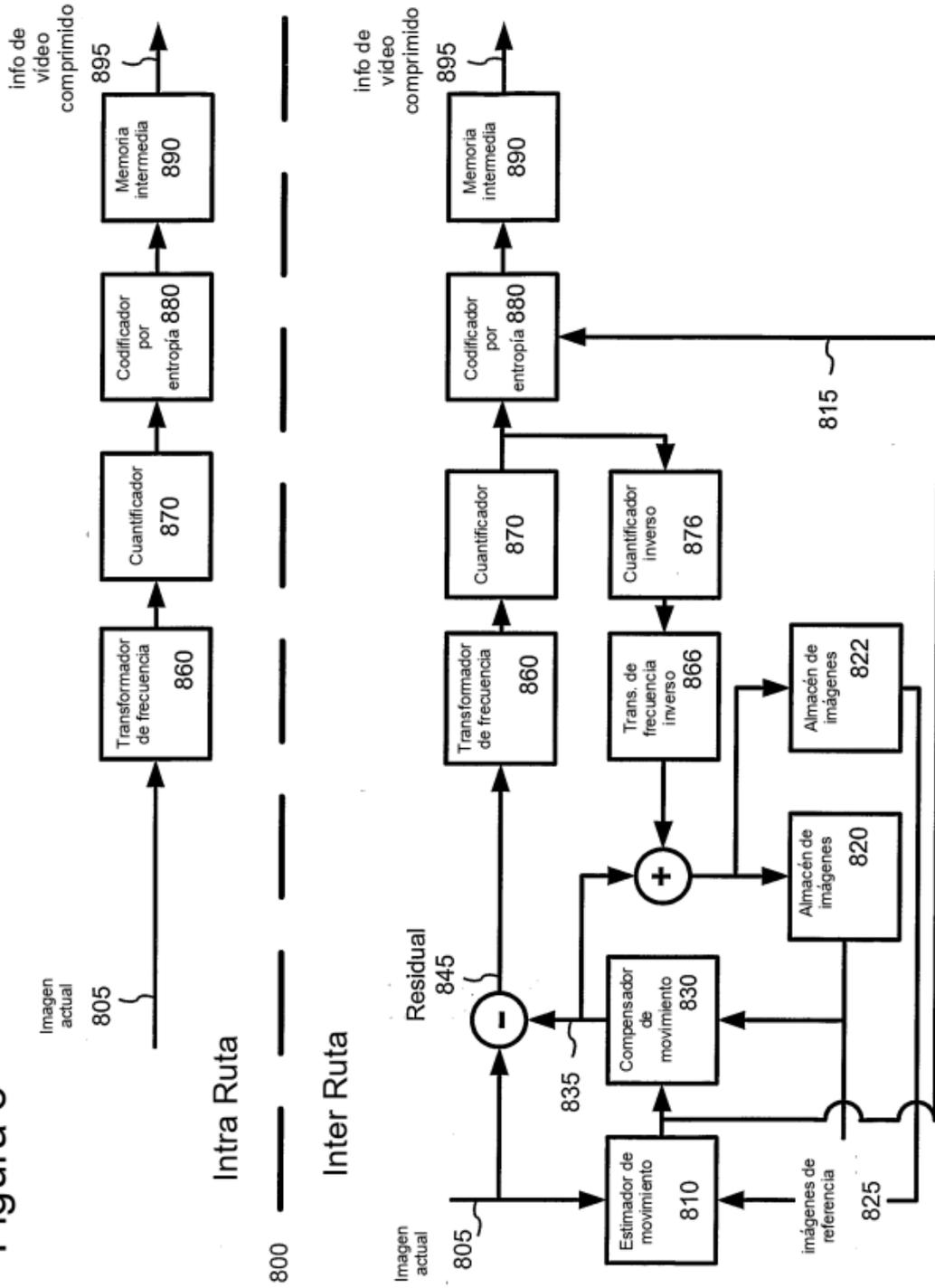


Figura 9

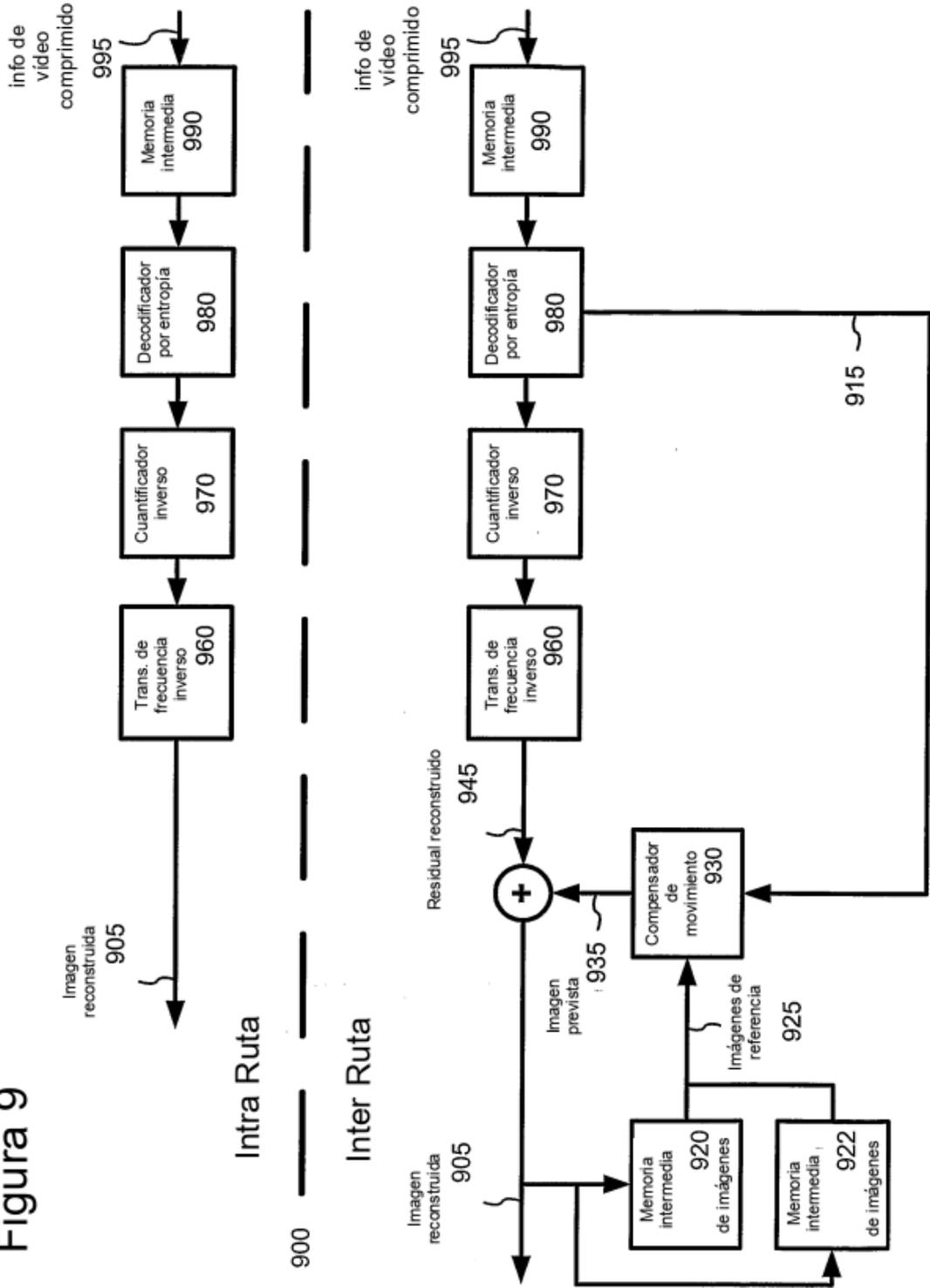


Figura 10

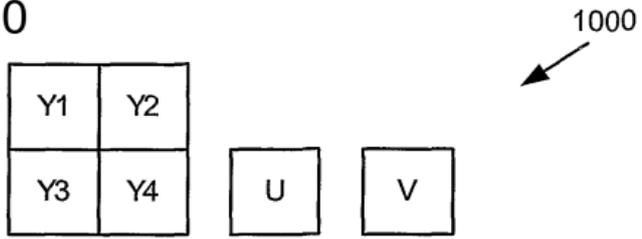


Figura 11A

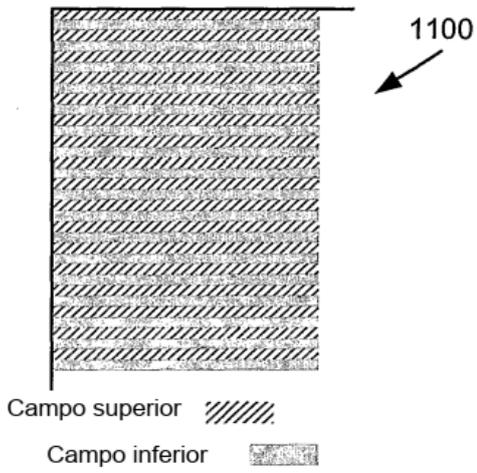


Figura 11B

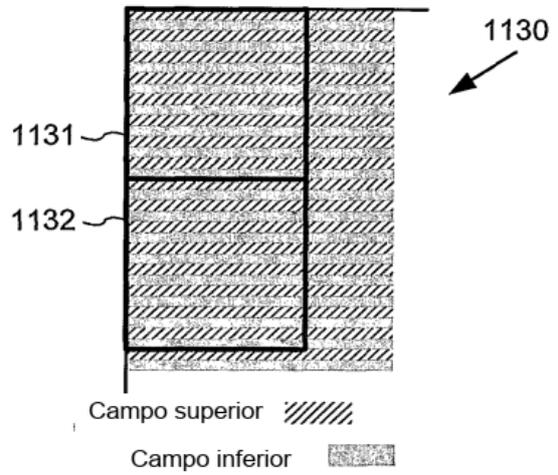


Figura 11C

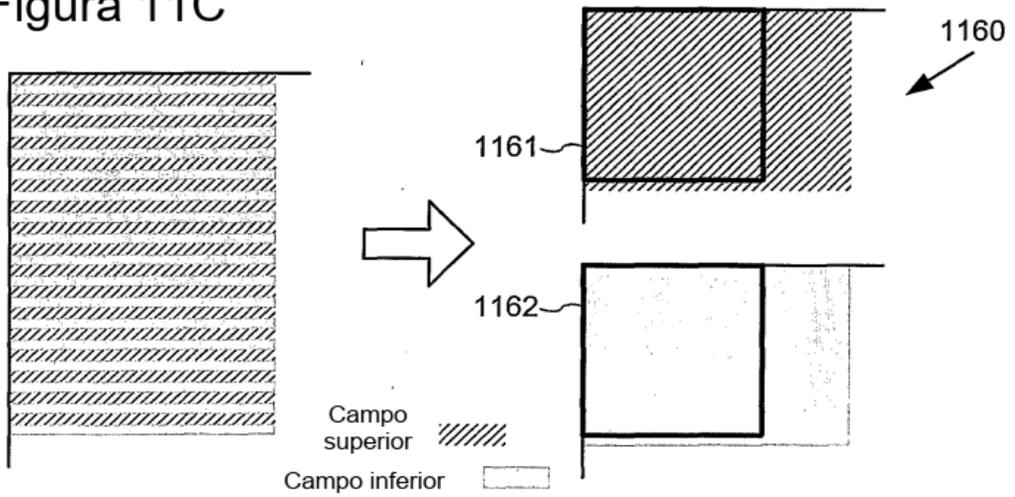


Figura 12

1200

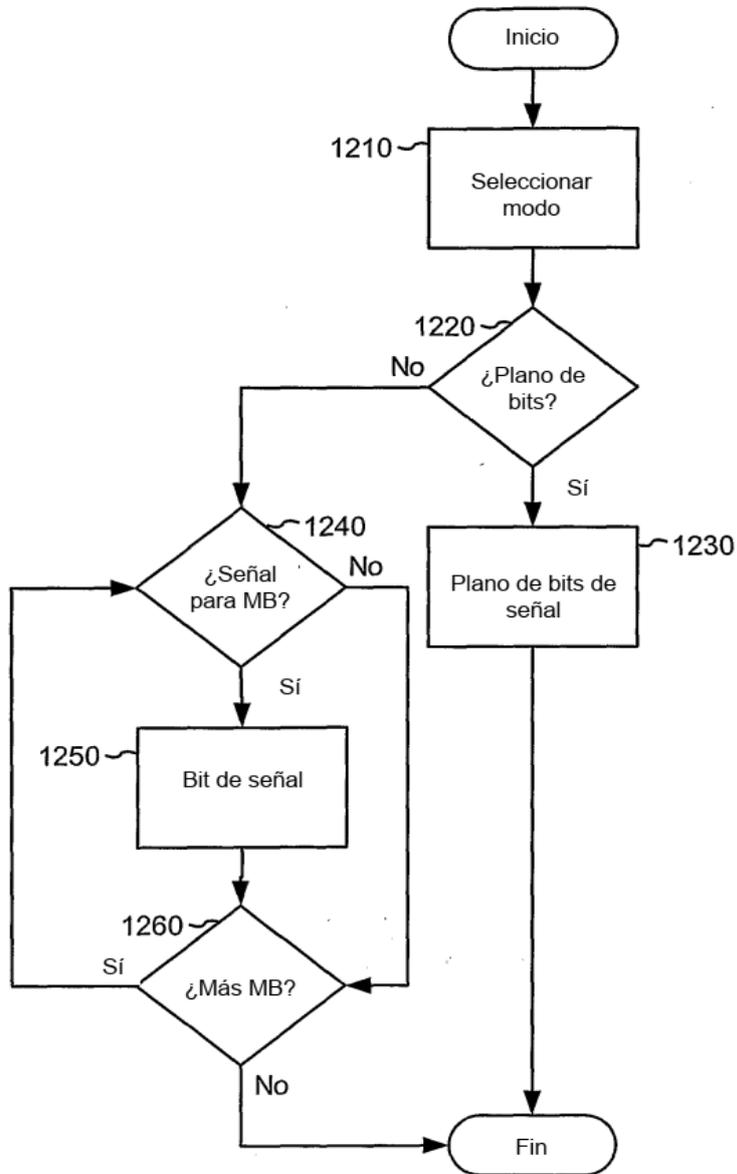


Figura 13

1300

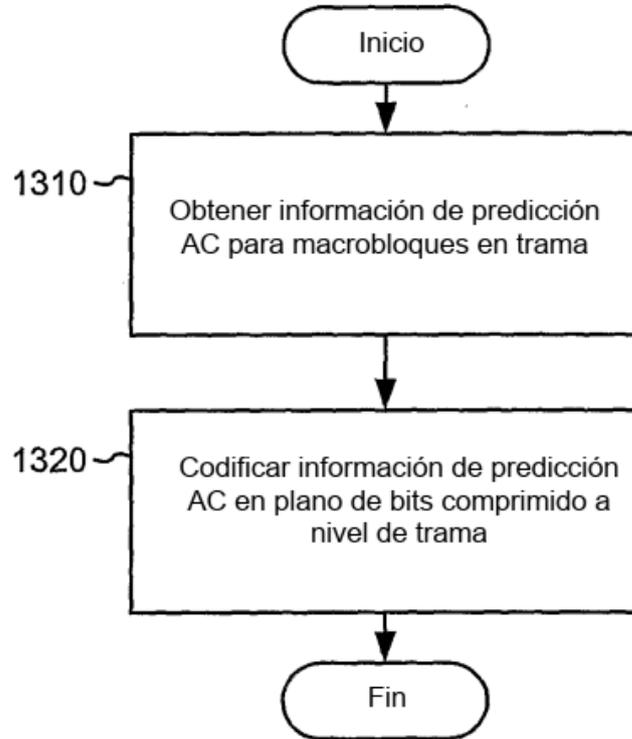


Figura 14

Sintaxis de secuencia de bits de trama I progresiva de capa de trama 1400

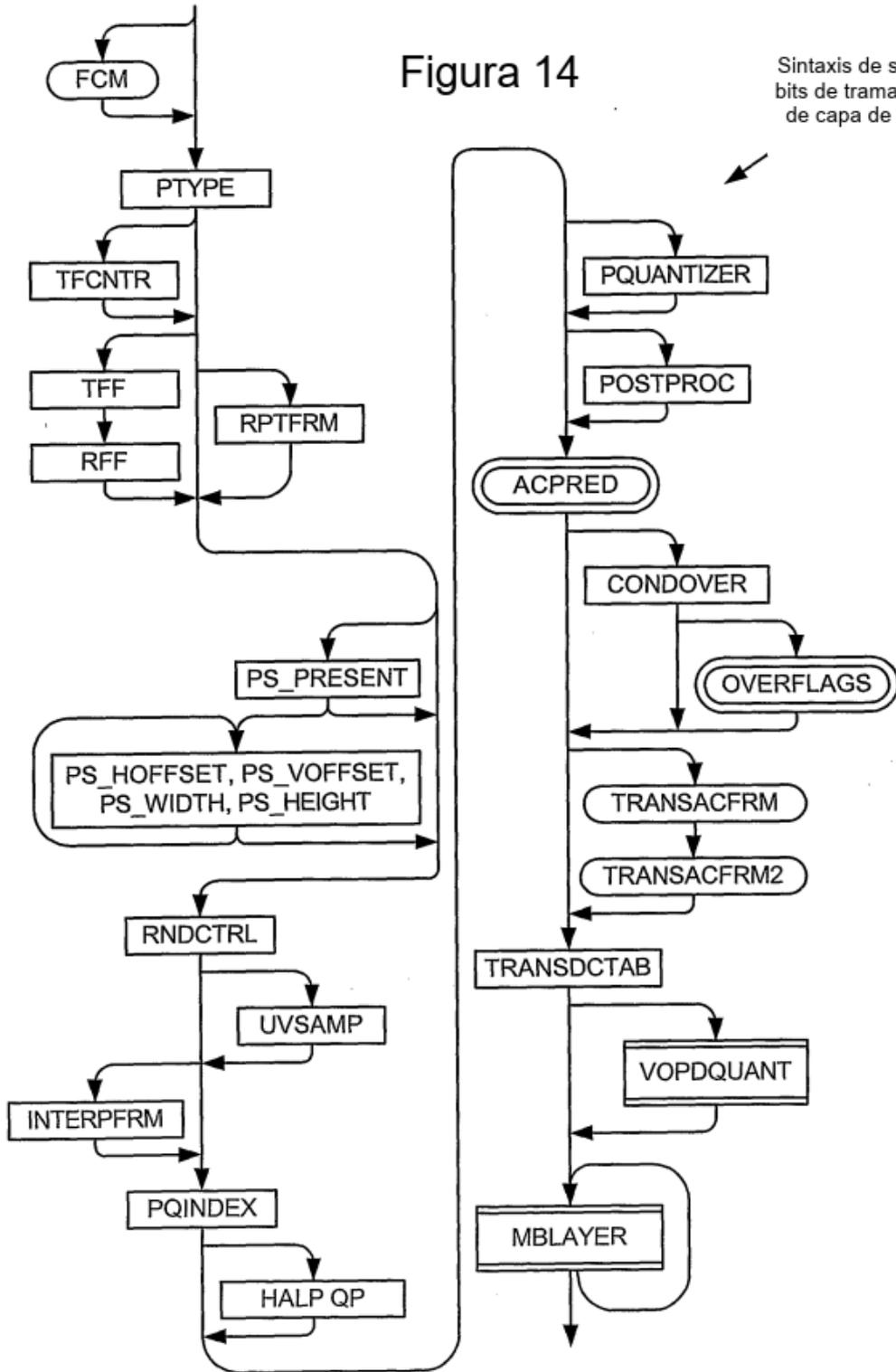


Figura 15

Sintaxis de secuencia de bits de trama I progresiva de capa de macrobloque 1500

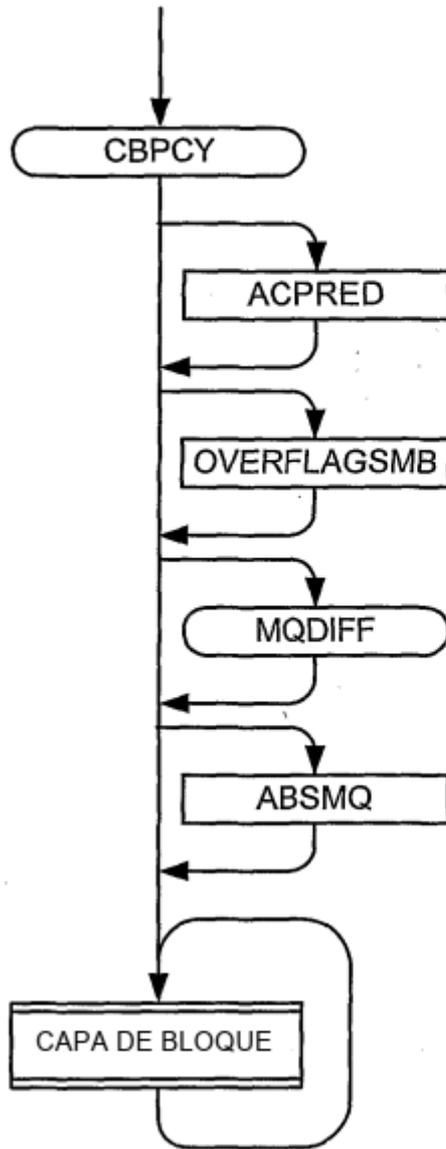


Figura 16

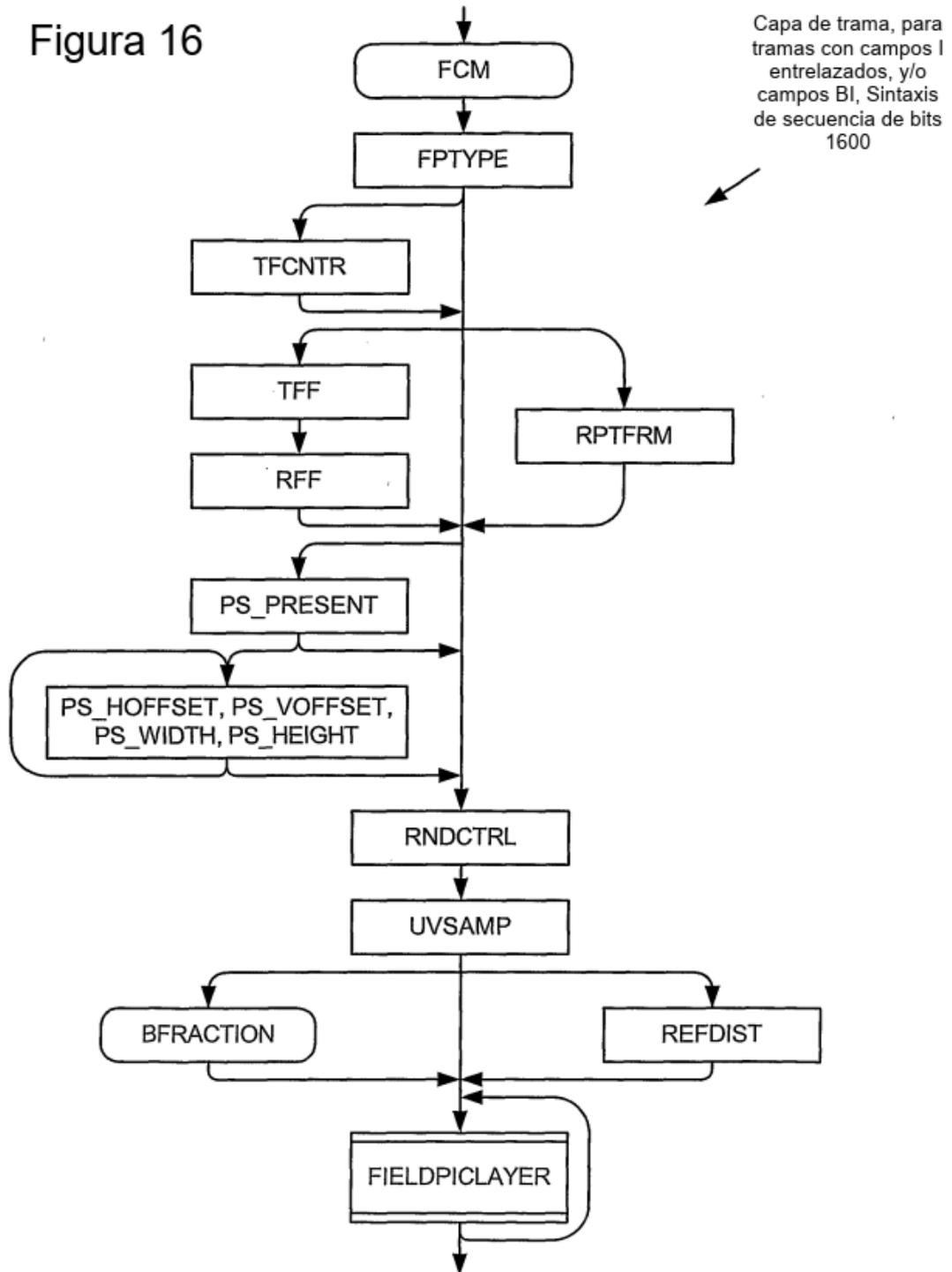
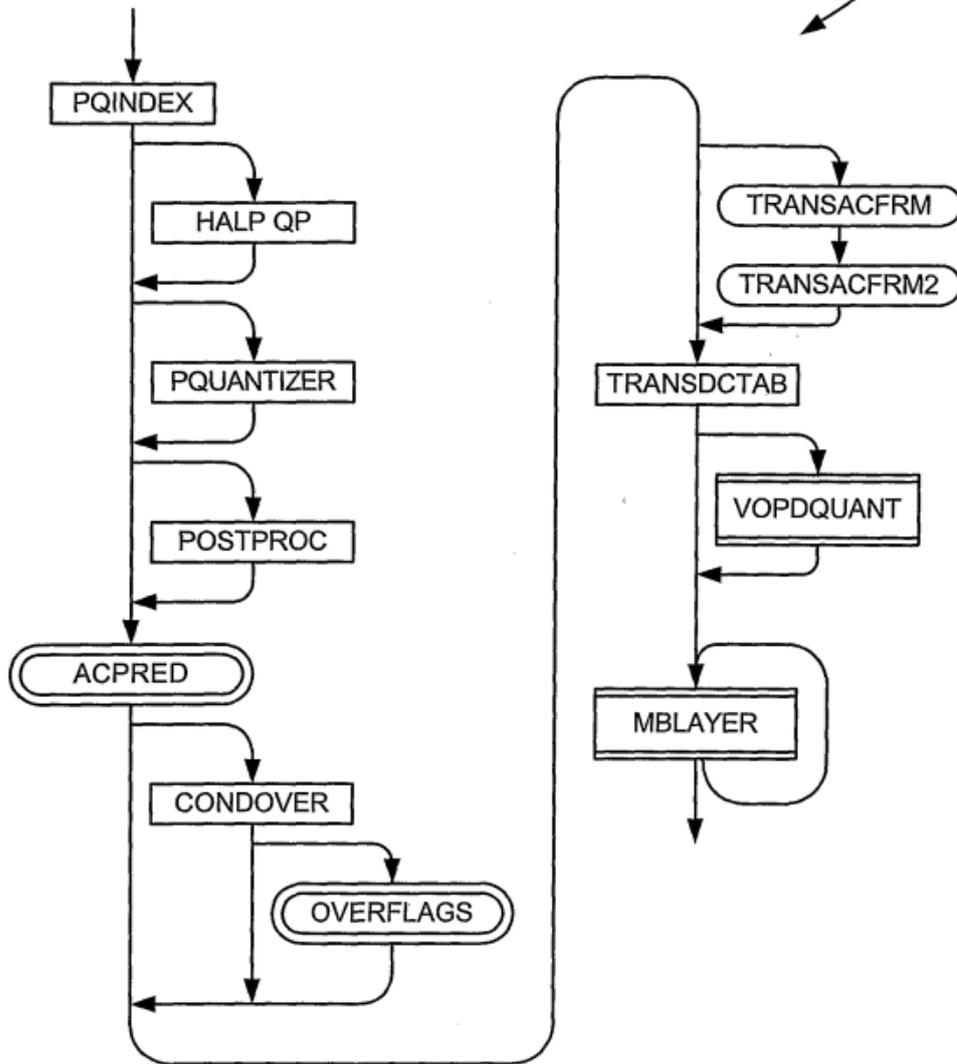


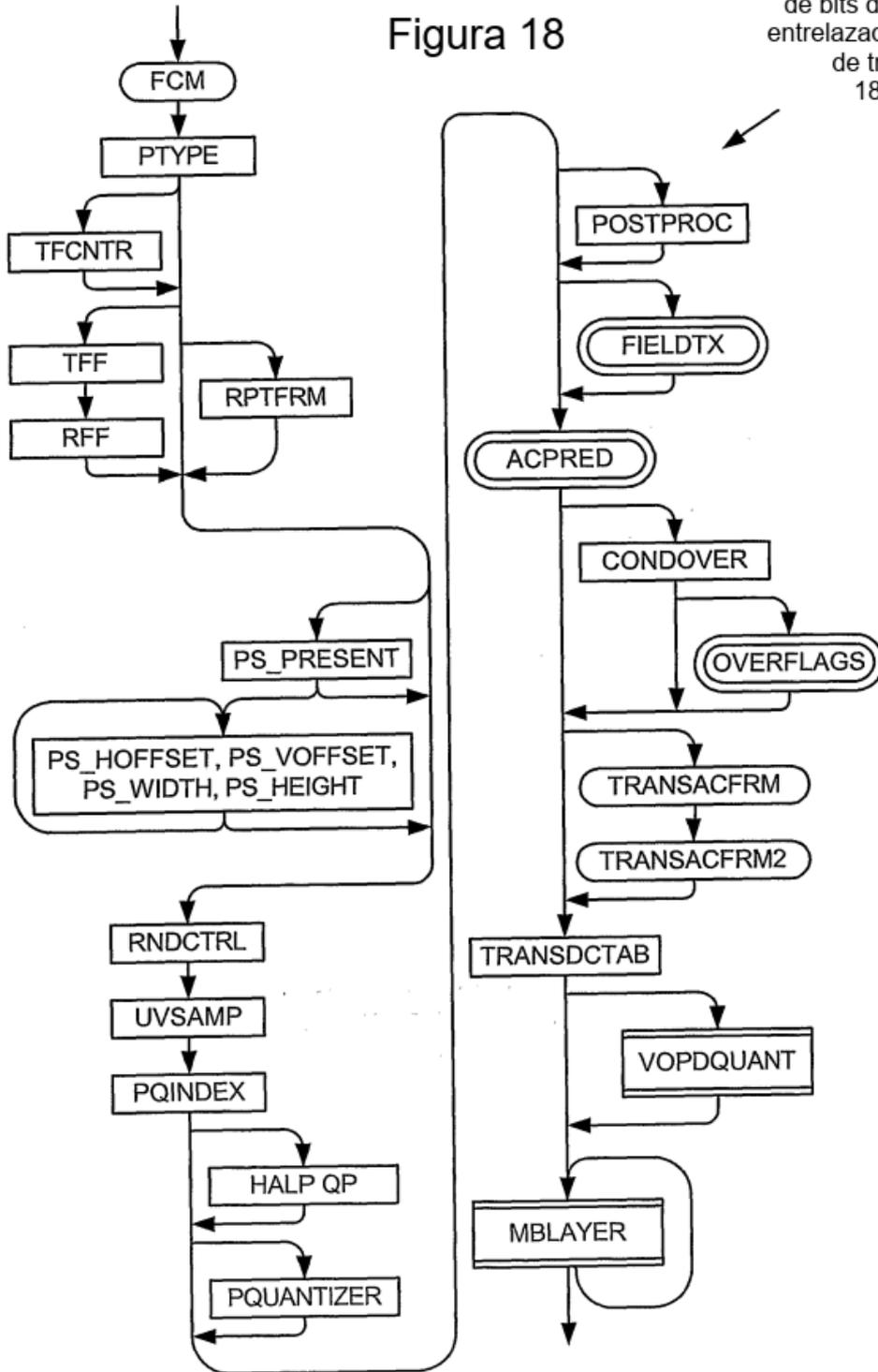
Figura 17

Sintaxis de secuencia de bits de trama I entrelazada de capa de campo 1700



Sintaxis de secuencia de bits de trama I entrelazada de capa de trama 1800

Figura 18



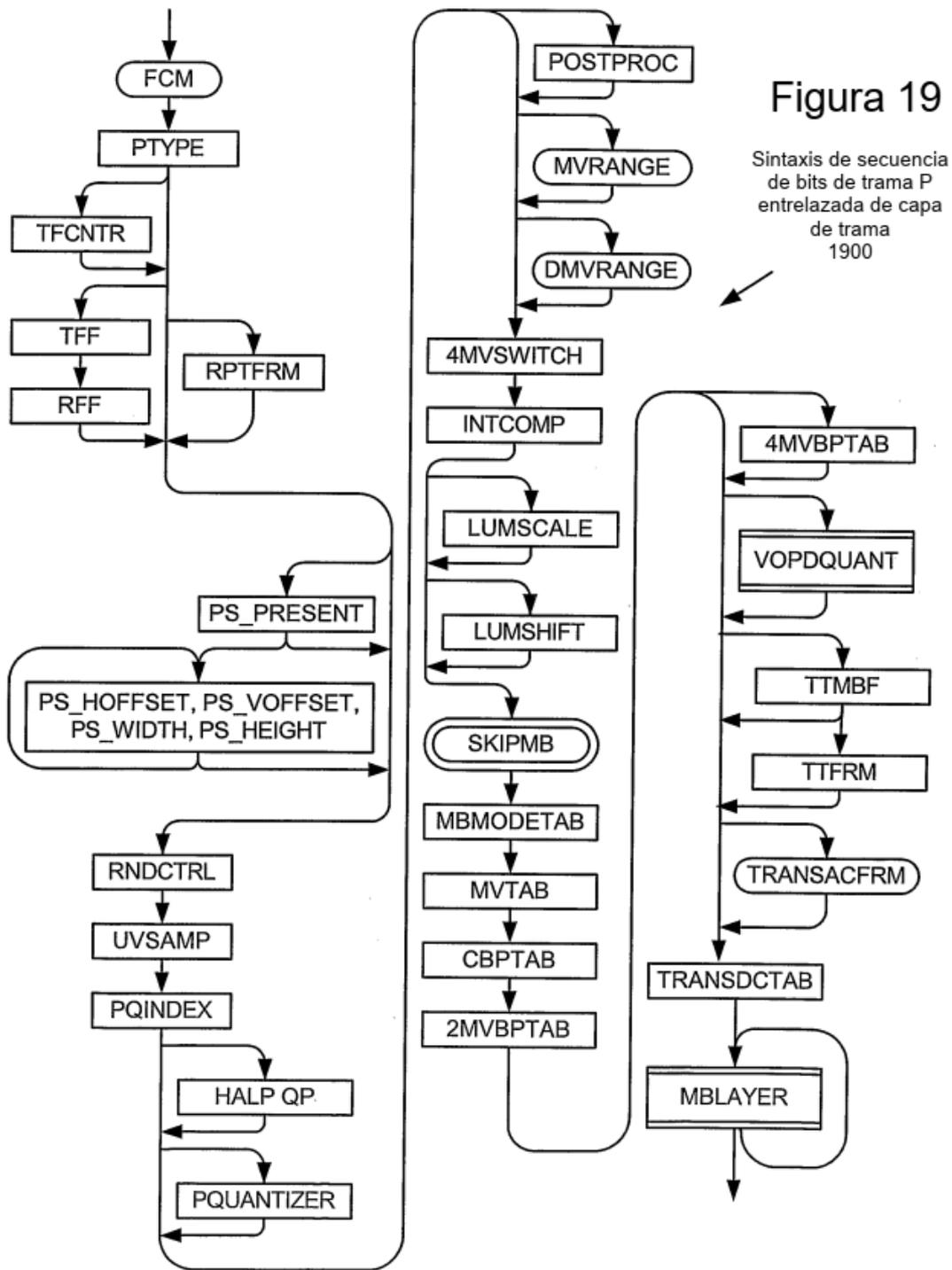


Figura 20

Sintaxis de secuencia de bits de trama B entrelazada de capa de trama 2000

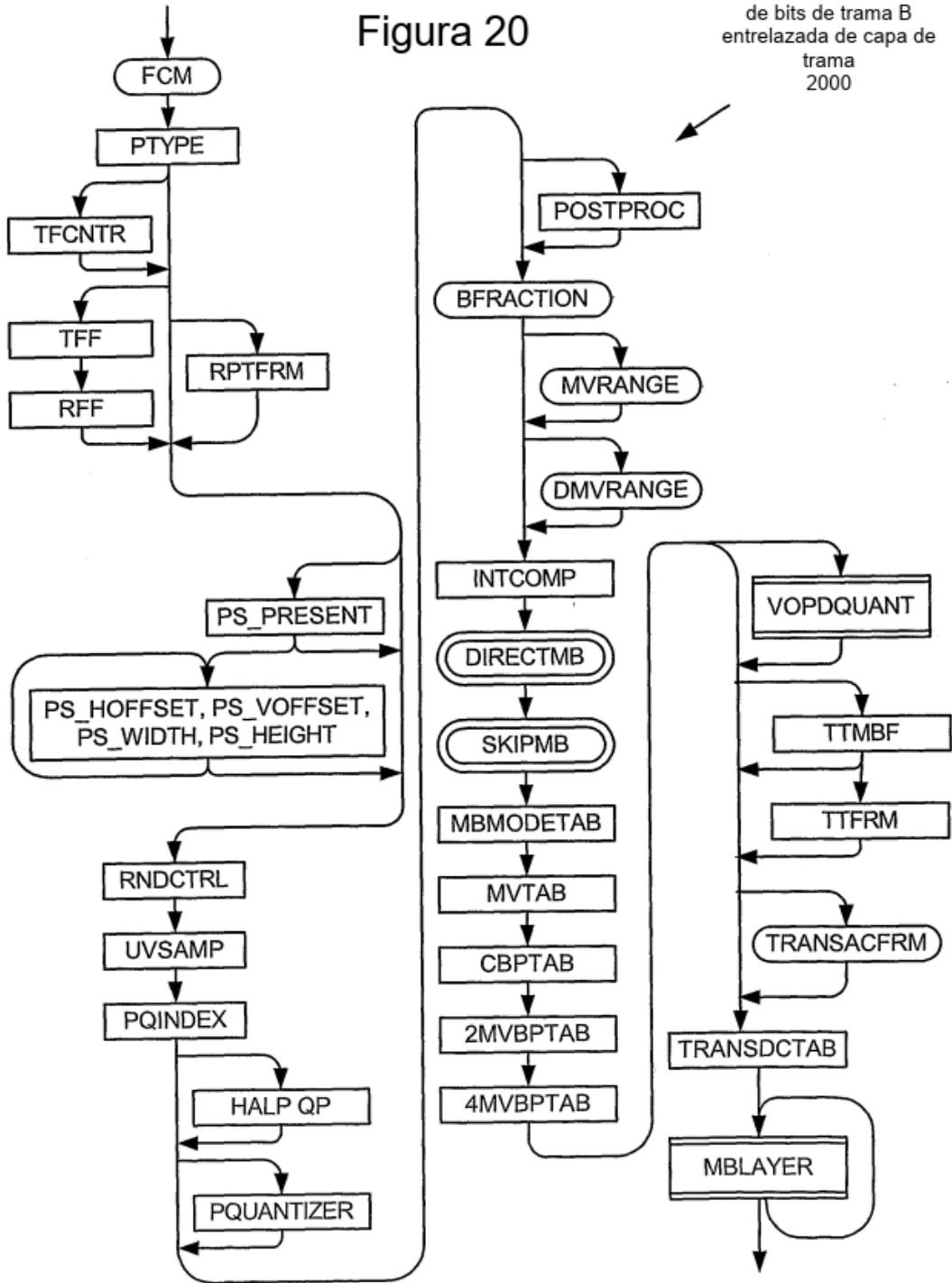


Figura 21

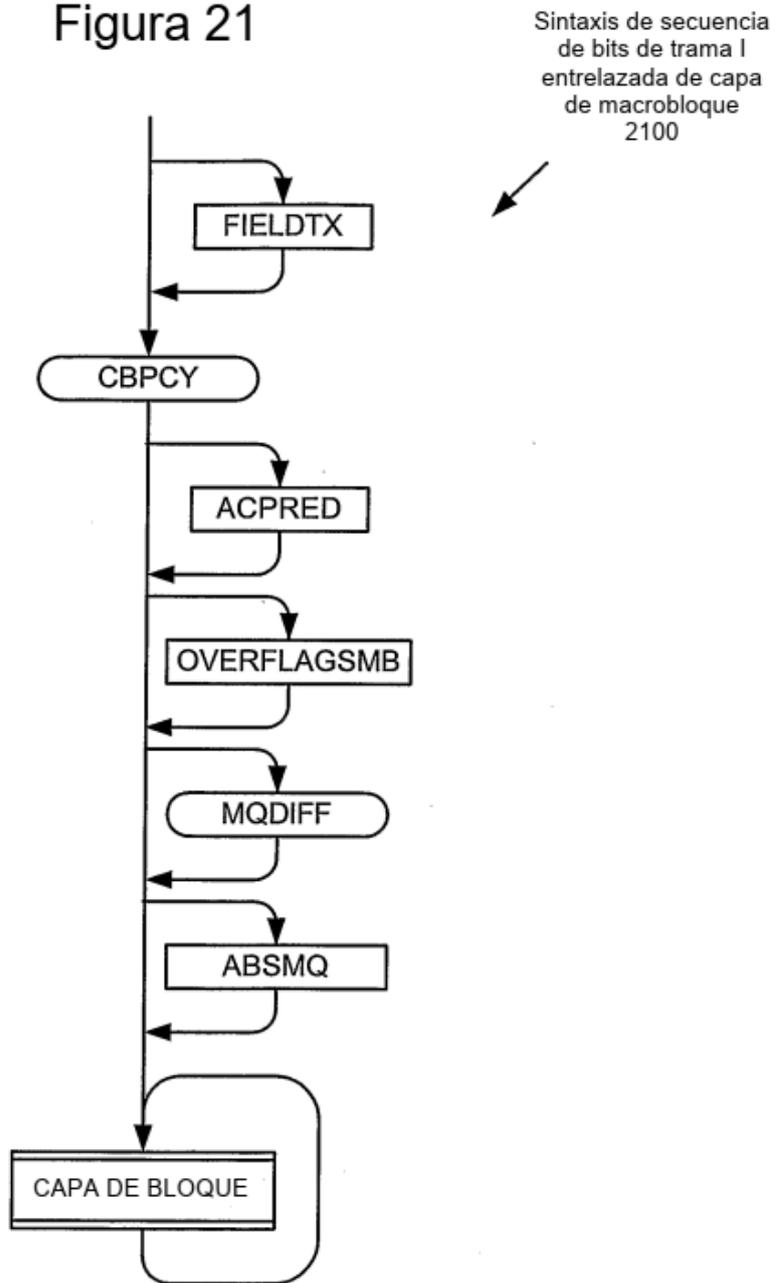


Figura 22

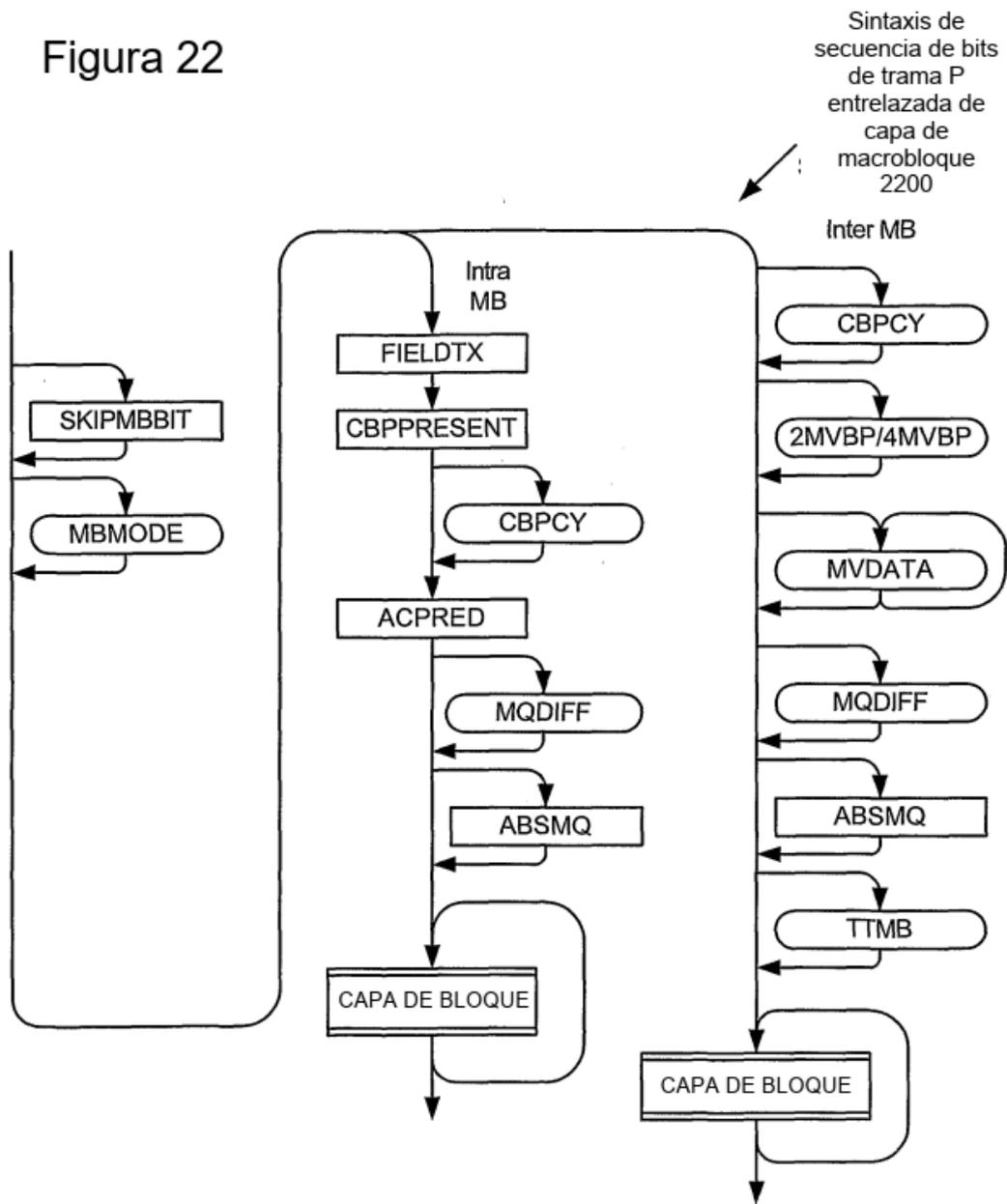


Figura 23

Sintaxis de secuencia de bits de trama B entrelazada de capa de macrobloque 2300

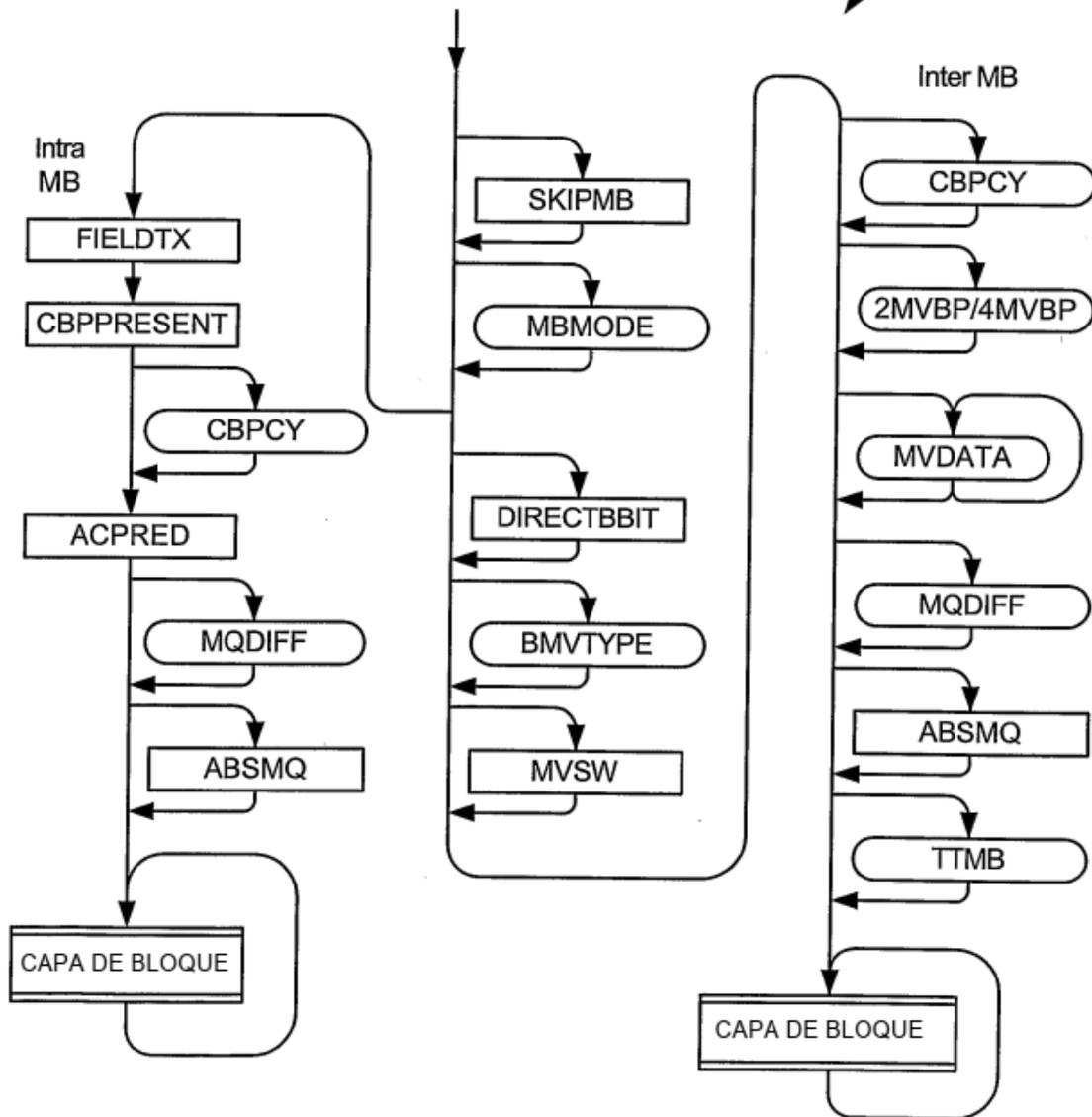


Figura 24A

2400

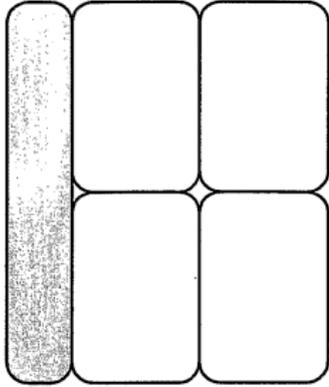


Figura 24B

2401

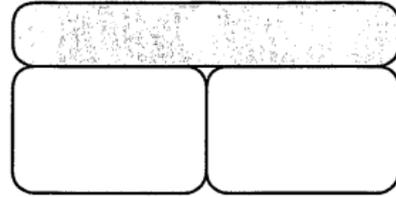


Figura 24C

2402

