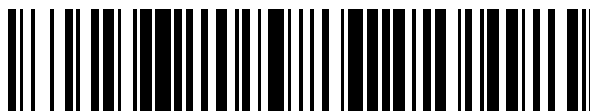


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 584 035**

51 Int. Cl.:

H04N 19/176 (2014.01)

H04N 19/129 (2014.01)

H04N 19/96 (2014.01)

H04N 19/14 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.06.2002 E 11007591 (8)**

97 Fecha y número de publicación de la concesión europea: **01.06.2016 EP 2405657**

54 Título: **Optimizador configurable de patrones**

30 Prioridad:

15.06.2001 US 882753

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

23.09.2016

73 Titular/es:

**QUALCOMM INCORPORATED (100.0%)
5775 Morehouse Drive
San Diego, CA 92121-1714, US**

72 Inventor/es:

**GOVINDASWAMY, SENTHIL;
THYAGARAJAN, KADAYAM y
LEVIN, JEFFREY A.**

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 584 035 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

Optimizador configurable de patrones

5 ANTECEDENTES DE LA INVENCION**I. Campo de la invención**

10 La presente invención se refiere al procesamiento y a la compresión de imágenes. Más específicamente, la presente invención se refiere a un optimizador configurable de patrones para imágenes comprimidas.

II. Descripción de la técnica relacionada

15 El procesamiento de imágenes digitales tiene una posición prominente en la disciplina general del procesamiento de señales digitales. La importancia de la percepción visual humana ha alentado un tremendo interés y adelantos en el arte y la ciencia del procesamiento de imágenes digitales. En el campo de la transmisión y recepción de señales de vídeo, tales como las usadas para proyectar filmes o películas, se están haciendo diversas mejoras en las técnicas de compresión de imágenes. Muchos de los sistemas de vídeo actuales y propuestos hacen uso de técnicas de codificación digital. Los aspectos de este campo incluyen la codificación de imágenes, la restauración de imágenes y
20 la selección de rasgos de imágenes. La codificación de imágenes representa los intentos de transmitir imágenes de canales de comunicación digital de forma eficaz, haciendo uso de tan pocos bits como sea posible para minimizar el ancho de banda requerido, mientras se mantienen a la vez las distorsiones dentro de ciertos límites. La restauración de imágenes representa los esfuerzos para recuperar la verdadera imagen del objeto. La imagen codificada, transmitida por un canal de comunicación, puede haber sido distorsionada por diversos factores. El origen de la degradación puede haber surgido originalmente en la creación de la imagen a partir del objeto. La selección de rasgos se refiere a la selección de ciertos atributos de la imagen. Tales atributos pueden ser requeridos en el reconocimiento, la clasificación y la decisión en un contexto más amplio.

30 La codificación digital del vídeo, tal como la del cine digital, es un área que se beneficia de las técnicas mejoradas de compresión de imágenes. La compresión de imágenes digitales puede ser clasificada, generalmente, en dos categorías: procedimientos sin pérdidas y con pérdidas. Una imagen sin pérdidas es recuperada sin ninguna pérdida de información. Un procedimiento con pérdidas implica una pérdida irrecuperable de alguna información, según la razón de compresión, la calidad del algoritmo de compresión y la implementación del algoritmo. Generalmente, los enfoques de compresión con pérdidas se consideran para obtener las razones de compresión deseadas para un enfoque del cine digital efectivo en términos de coste. Para lograr niveles de calidad del cine digital, el enfoque de compresión debería proporcionar un nivel de prestaciones sin pérdidas visualmente. De tal modo, aunque hay una pérdida matemática de información como resultado del proceso de compresión, la distorsión de imágenes provocada por esta pérdida debería ser imperceptible para un espectador en condiciones normales de visualización.

40 Las tecnologías existentes de compresión de imágenes digitales han sido desarrolladas para otras aplicaciones, esto es, para sistemas de televisión. Tales tecnologías han asumido compromisos de diseño adecuados para la aplicación pretendida, pero no satisfacen los requisitos de calidad necesarios para la presentación cinematográfica.

45 La tecnología de compresión del cine digital debería proporcionar la calidad visual que un cinéfilo ha experimentado anteriormente. Idealmente, la calidad visual del cine digital debería intentar superar la de una película grabada en versión de alta calidad. Al mismo tiempo, la técnica de compresión debería tener una alta eficacia de codificación para ser práctica. Según lo definido en el presente documento, la eficacia en la codificación se refiere a la velocidad de bits necesaria para que la calidad de la imagen comprimida satisfaga un cierto nivel cualitativo. Además, el sistema y la técnica de codificación deberían tener flexibilidad intrínseca para asimilar distintos formatos y deberían ser efectivos en términos de coste; es decir, un proceso descodificador o codificador de tamaño reducido y eficaz.

50 Muchas técnicas de compresión disponibles ofrecen niveles significativos de compresión, pero dan como resultado una degradación de la calidad de la señal de vídeo. Habitualmente, las técnicas para la transferencia de información comprimida requieren que la información comprimida sea transferida a una velocidad constante de bits.

55 Una técnica de compresión capaz de ofrecer niveles significativos de compresión, preservando a la vez el nivel deseado de calidad para las señales de vídeo, utiliza bloques y sub-bloques, de tamaño adaptativo, de datos codificados de coeficientes de la Transformación Discreta de Coseno (DCT). Esta técnica será mencionada a continuación en el presente documento como el procedimiento de Transformación Discreta de Coseno de Tamaño Adaptativo de Bloque (ABSDCT). Esta técnica está divulgada en la patente de los EE.UU. n.º 5.021.891, titulada "Procedimiento y sistema de compresión de imágenes de tamaño adaptativo de bloque", cedida al cesionario de la presente invención.

65 Las técnicas de DCT también están divulgadas en la patente de los EE.UU. N.º 5.107.345, titulada "Procedimiento y sistema de compresión de imágenes de tamaño adaptativo de bloque", cedida al cesionario de la presente invención.

Además, el uso de la técnica de ABSDCT, en combinación con una técnica de Transformación de Árbol Cuadrático Diferencial, se expone en la patente estadounidense N° 5.452.104, titulada "Procedimiento y sistema de compresión de imágenes de tamaño adaptativo de bloque", también cedida al cesionario de la presente invención.

5 Los sistemas divulgados en estas patentes utilizan lo que se menciona como una codificación "intra-trama", donde cada trama de datos de imágenes es codificada sin considerar el contenido de ninguna otra trama. Usando la técnica de ABSDCT, la velocidad de datos alcanzable puede ser reducida desde alrededor de 1,5 miles de millones de bits por segundo hasta aproximadamente 50 millones de bits por segundo, sin una degradación perceptible de la calidad de la imagen.

10 La técnica de ABSDCT puede ser usada para comprimir una imagen tanto en blanco y negro como en color, o una señal que represente la imagen. La señal de entrada en color puede estar en un formato YIQ, siendo Y la muestra de luminancia, o brillo, y siendo I y Q las muestras de crominancia, o color, para cada bloque de píxeles de tamaño 4x4. También pueden usarse otros formatos conocidos, tales como los formatos YUV, YCbCr o RGT. Debido a la
15 baja sensibilidad espacial del ojo al color, la mayor parte de las investigaciones han mostrado que es razonable una sub-muestra de los componentes de color, en un factor de cuatro, en las direcciones horizontal y vertical. En consecuencia, una señal de vídeo puede ser representada por cuatro componentes de luminancia y dos componentes de crominancia.

20 Usando la ABSDCT, una señal de vídeo será generalmente segmentada en bloques de píxeles para su procesamiento. Para cada bloque, los componentes de luminancia y crominancia son pasados a un entrelazador de bloques. Por ejemplo, un bloque de tamaño 16x16 (píxeles) puede ser presentado al entrelazador de bloques, que ordena u organiza las muestras de imágenes dentro de cada bloque de tamaño 16x16, para producir bloques y sub-
25 bloques compuestos de datos para el análisis de la transformación discreta de coseno (DCT). El operador de la DCT es un procedimiento de conversión de una señal, muestreada en el tiempo y el espacio, en una representación de frecuencia de la misma señal. Por convertir a una representación de frecuencia, se ha mostrado que las técnicas de DCT admiten muy altos niveles de compresión, ya que los cuantizadores pueden ser diseñados para aprovechar las características de distribución de frecuencia de una imagen. En un modo de realización preferente, una DCT de tamaño 16x16 es aplicada a un primer ordenamiento, cuatro DCT de tamaño 8x8 son aplicadas a un segundo
30 ordenamiento, 16 DCT de tamaño 4x4 son aplicadas a un tercer ordenamiento y 64 DCT de tamaño 2x2 son aplicadas a un cuarto ordenamiento.

La operación de la DCT reduce la redundancia espacial inherente en el origen de vídeo. Después de que la DCT es llevada a cabo, la mayor parte de la energía de la señal de vídeo tiende a concentrarse en unos pocos coeficientes de la DCT. Una transformación adicional, la Transformación de Árbol Cuadrático Diferencial (DQT), puede ser usada para reducir la redundancia entre los coeficientes de la DCT.

35 Para el bloque de tamaño 16x16 y cada sub-bloque, los valores de coeficientes de la DCT y el valor de la DQT (si se usa la DQT) son analizados para determinar el número de bits requeridos para codificar el bloque o sub-bloque. Luego, se escoge el bloque, o la combinación de sub-bloques, que requiere el mínimo número de bits para codificar, para representar el segmento de imagen. Por ejemplo, dos sub-bloques de tamaño 8x8, seis sub-bloques de tamaño 4x4 y ocho sub-bloques de tamaño 2x2 pueden ser escogidos para representar el segmento de imagen.

40 El bloque, o la combinación de sub-bloques, escogido es luego debidamente dispuesto en orden en un bloque de tamaño 16x16. Los valores de coeficientes de la DCT / DQT pueden luego someterse a ponderación de frecuencia, cuantización y codificación (tal como la codificación de longitud variable), en preparación para su transmisión. Aunque la técnica de ABSDCT descrita anteriormente rinde notablemente bien, es intensiva en términos de cálculo. Por tanto, la implementación compacta en hardware de la técnica puede ser difícil.

50 La codificación de longitud variable ha sido lograda en forma de longitud de racha y tamaño. Otros procedimientos de compresión, tales como los del Grupo Conjunto de Expertos Fotográficos (JPEG) o del Grupo de Expertos en Imágenes en Movimiento (MPEG-2), usan un procedimiento estándar de recorrido en zig-zag sobre todo el tamaño del bloque procesado. Usando la ABSDCT, sin embargo, se generan distintos tamaños de bloque, en base a la varianza dentro de los bloques de datos. En consecuencia, un procedimiento estándar de recorrido en zig-zag no siempre es óptimo sobre todo el tamaño del bloque procesado. Además, el procedimiento estándar de recorrido en zig-zag sobre cada tamaño de bloque puede ser difícil de implementar en hardware. Además, un patrón en zig-zag no siempre es el patrón óptimo para un bloque o trama dados. En consecuencia, se necesita un procedimiento y un aparato para determinar el patrón óptimo.

60 Se reclama atención al documento WO 01 / 35673 A, que se refiere a un sistema y a un procedimiento para la compresión de imágenes, utilizando bloques y sub-bloques de tamaño adaptativo, de datos de coeficientes de transformaciones discretas de coseno. Un elemento de asignación de tamaño de bloque en el codificador selecciona el bloque o sub-bloque de un bloque de entrada de píxeles a procesar. La selección se basa en la varianza de los valores de píxeles. Los bloques con varianzas mayores que un umbral son subdivididos, mientras que los bloques con varianzas menores que un umbral no son subdivididos. Un elemento de transformación transforma los valores de píxeles de los bloques seleccionados al dominio de la frecuencia.

Los valores del dominio de la frecuencia pueden luego ser cuantizados, serializados y codificados con longitud variable, en preparación para su transmisión.

5 Se reclama atención adicional al documento EP-A-0 862 339, que se refiere a un aparato de conversión de señales en el cual las señales de imágenes digitales, comprimidas en un formato DVC, son descodificadas en señales de imágenes no comprimidas de un tipo de componente cromático 4:1:1, que están compuestas por señales de diferencia cromática y señales de luminancia. Las señales de diferencia cromática del tipo de componente cromático 4:1:1, dispuestas en una matriz, son eliminadas en una fila de cada dos, para reducir a la mitad el número de señales de diferencia cromática en una dirección y, y se añade una señal de diferencia cromática a una posición
10 entre cada par de señales de diferencia cromática adyacentes entre sí en una dirección x, para duplicar el número de señales de diferencia cromática en la dirección x. Por lo tanto, las señales convertidas de diferencia cromática de un tipo de componente cromático 4:2:0 son producidas a partir de la señales de diferencia cromática. Los valores de las señales convertidas de diferencia cromática son determinados de acuerdo a los valores de las señales de diferencia cromática. A continuación, las señales convertidas de diferencia cromática y las señales de luminancia de las señales de imágenes no comprimidas son combinadas para producir señales convertidas de imágenes digitales del tipo de componente cromático 4:2:0.

El documento US-A-5 452 104 se refiere a un procedimiento y a un sistema para comprimir los datos de imágenes, para su transmisión. Cada bloque y los correspondientes sub-bloques de datos de píxeles son sometidos a una
20 operación de transformación discreta de coseno (DCT). Los niveles variables de sub-bloques de los coeficientes resultantes de la transformación correspondiente son seleccionados para su estructuración en un bloque compuesto de coeficientes de transformación, correspondiente a cada bloque de entrada de datos de píxeles. La selección del tamaño de bloque de coeficientes de transformación para el bloque compuesto está determinada por un proceso de comparación entre la eficacia de la codificación de bloques y sub-bloques. El bloque compuesto es codificado con
25 longitud variable para reducir adicionalmente el número de bits en los datos comprimidos. Una operación de transformación de árbol cuadrático discreto (DQT) suplementa la selección de bloques de tamaño adaptativo de bloque, con reemplazo de coeficientes, para optimizar la compresión de datos de imágenes.

También se reclama atención al documento US 5 812 788 A, que se refiere a la codificación y descodificación de
30 señales de vídeo usando un conjunto de tablas de cuantización, que es generado a partir de una matriz base y una matriz de escala, ambas explícitamente codificadas en el flujo de bits de vídeo codificado. Los conjuntos de tablas de cuantización pueden ser cambiados tan a menudo como sea necesario, mientras las nuevas matrices base y de escala estén codificando explícitamente en el flujo de bits.

35 También se reclama atención al documento US 4 821 119 A, que se refiere a un codificador de vídeo inter-trama de baja velocidad de bits, del tipo conocido como codificador híbrido. Un codificador híbrido logra la compresión de imágenes usando una transformación bidimensional de señales sobre bloques de datos diferenciales de píxeles en el bucle directo de un codificador DPCM. Los coeficientes de transformación de cada bloque son luego cuantizados y codificados por entropía para su transmisión. La eficacia de la codificación está determinada, en parte, por el
40 tamaño del bloque de transformación. Los bloques más grandes son más eficaces en términos de bits, debido a la menor cantidad de datos de sobregasto requeridos, pero requieren una compleja implementación de hardware transformador. Además, los bloques más grandes producen una molesta distorsión de bloques. El codificador divulgado tiene las ventajas de la división en tamaños de bloque, tanto grandes como pequeños, de la imagen de vídeo. En el codificador, después de la combinación diferencial con un bloque correspondiente procedente de la
45 trama anterior, cada bloque de tamaño $m \times m$ de datos de píxeles es subdividido en bloques más pequeños de tamaño $n \times n$ ($n < m$) de datos, que son transformados individualmente por un transformador bidimensional discreto de coseno. Después de que los coeficientes de cada sub-bloque están cuantizados, se reconstruye el bloque principal. Un codificador por entropía recorre los sub-bloques y codifica la cadena resultante de coeficientes recorridos.

50 Finalmente, se reclama atención al documento de LEE J H ET AL: "UNA CODIFICACIÓN EFICAZ DE BLOQUES DE DCT CON RECORRIDO ADAPTATIVO DE BLOQUES", TRANSACCIONES IEICE DE COMUNICACIONES, SOCIEDAD DE COMUNICACIONES; TOKIO, JP, vol. E77-B, n.º 12, 1 de diciembre de 1994, páginas 1.489 a 1.494, XP000498064, ISSN: 0916-8516.

55 RESUMEN

De acuerdo a la presente invención, se proporcionan un procedimiento y un aparato de serialización de datos de imágenes en base a la frecuencia, según lo enunciado, respectivamente, en las reivindicaciones 1 y 18. Modos de
60 realización adicionales son reivindicados en las reivindicaciones dependientes.

Los modos de realización de la invención proporcionan un aparato y un procedimiento para un determinador óptimo de patrones. En un modo de realización, el patrón óptimo es configurable trama por trama. En otro modo de realización, se usa un patrón por omisión de un tamaño de bloque predeterminado, independientemente del tamaño efectivo de bloque, según lo determinado por la técnica de la transformación discreta de coseno de tamaño adaptativo de bloque (ABSDCT).
65

La presente invención es un sistema y un procedimiento, basados en la calidad, de compresión de imágenes que utiliza bloques y sub-bloques de tamaño adaptativo, de datos de coeficientes de la Transformación Discreta de Coseno, y un factor de escala de cuantización basado en la calidad. Un bloque de datos de píxeles es ingresado en un codificador. El codificador comprende un elemento de asignación de tamaño de bloque (BSA), que segmenta el bloque de entrada de píxeles para su procesamiento. La asignación del tamaño de bloque está basada en las varianzas del bloque de entrada y los bloques adicionalmente subdivididos. En general, las áreas con mayores varianzas son subdivididas en bloques más pequeños, y las áreas con varianzas más pequeñas no han de subdividirse, siempre que los valores medios de bloques y sub-bloques caigan en distintas gamas predeterminadas. Por tanto, en primer lugar, el umbral de varianza de un bloque es modificado a partir de su valor nominal, según su valor medio, y luego la varianza del bloque es comparada con un umbral y, si la varianza es mayor que el umbral, entonces el bloque es subdividido.

La asignación del tamaño de bloque es proporcionada a un elemento de transformación, que transforma los datos de píxeles en datos del dominio de la frecuencia. La transformación es realizada solamente en el bloque y los sub-bloques seleccionados mediante la asignación del tamaño de bloque. Los datos de transformación se someten luego al ajuste a escala mediante la cuantización y la serialización. La cuantización de los datos de transformación se cuantiza en base a una métrica de calidad de imagen, tal como un factor de escala que ajusta con respecto al contraste, el número de coeficientes, la distorsión de velocidad, la densidad de las asignaciones de tamaños de bloque y / o los factores de escala pasados. La serialización se basa en crear las más largas duraciones de racha posibles del mismo valor. En un modo de realización, se utiliza el recorrido en zigzag con un tamaño fijo de bloque para serializar los datos, para producir un flujo de datos, independientemente de la asignación de tamaño de bloque. En otro modo de realización, el tamaño del bloque es de 8x8. El flujo de datos puede ser codificado por un codificador de longitud variable, en preparación para la transmisión. Los datos codificados son enviados a través de un canal de transmisión a un decodificador, donde los datos de píxeles son reconstruidos en preparación para su exhibición.

En otro modo de realización, se describe un procedimiento de serialización de datos de imágenes en base a la frecuencia, en un sistema de cine digital. Se compila al menos un grupo de datos que pueden ser representados como un bloque de datos de tamaño 16x16. Alternativamente, se compila una trama de datos. El grupo de datos es dividido en cuatro grupos, cada uno de los cuales puede ser representado como un bloque de tamaño 8x8. Cada uno de los cuatro bloques de datos de dimensiones 8x8 es serializado usando recorrido en zig-zag, recorrido vertical y / o recorrido horizontal.

En consecuencia, es un aspecto de un modo de realización procesar bloques de datos usando un patrón fijo de recorrido en bloques de tamaño 8x8, independientemente de la asignación efectiva del tamaño de bloque.

Es otro aspecto de un modo de realización determinar e implementar una técnica óptima de recorrido trama por trama.

Es otro aspecto de un modo de realización proporcionar al usuario patrones configurables de recorrido.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

Las características y ventajas de la presente invención devendrán más evidentes a partir de la descripción detallada expuesta a continuación, cuando se considere conjuntamente con los dibujos, en los cuales los caracteres iguales de referencia identifican de forma correspondiente en toda su extensión, y en los que:

la FIG. 1 es un diagrama de bloques de una parte codificadora de un sistema de procesamiento de imágenes basado en la calidad, que incorpora el sistema y el procedimiento de asignación de tamaño de bloque basado en la varianza de la presente invención;

la FIG. 2 es un diagrama de bloques de una parte decodificadora de un sistema de procesamiento de imágenes basado en la calidad, que incorpora el sistema y el procedimiento de asignación de tamaño de bloque basado en la varianza de la presente invención;

la FIG. 3 es un diagrama de flujo que ilustra las etapas de procesamiento implicadas en la asignación de tamaño de bloque basada en la varianza;

la FIG. 4a ilustra una asignación ejemplar de tamaño de bloque;

la FIG. 4b ilustra un patrón de recorrido en zig-zag para un tamaño de bloque de 16x16;

la FIG. 4c ilustra un patrón de recorrido en zig-zag dentro de cada tamaño de bloque variable;

la FIG. 5a ilustra un patrón de recorrido en zig-zag de bloques de tamaño 8x8, independientes de los tamaños efectivos de bloque;

la FIG. 5b ilustra distintos patrones de recorrido implementados en bloques de tamaño 8x8, independientes de los tamaños efectivos de bloque;

5 la FIG. 6a ilustra un modo de realización de un proceso de serialización; y

la FIG. 6b ilustra un modo de realización alternativo de un proceso de serialización.

DESCRIPCIÓN DETALLADA DE LOS MODOS DE REALIZACIÓN PREFERENTES

10 A fin de facilitar la transmisión digital de señales digitales, y de disfrutar de las correspondientes ventajas, es generalmente necesario emplear alguna forma de compresión de señales. Para lograr una alta compresión en una imagen resultante, también es importante que se mantenga la alta calidad de la imagen. Además, se desea la eficacia de cálculo para la implementación compacta en hardware, que es importante en muchas aplicaciones.

15 Antes de que se explique en detalle un modo de realización de la invención, ha de entenderse que la invención no está limitada en su aplicación a los detalles de la construcción y la disposición de los componentes, enunciados en la siguiente descripción o ilustrados en los dibujos. La invención es capaz de otros modos de realización, y se llevan a cabo de diversas maneras. Además, se entiende que la fraseología y la terminología usadas en el presente documento son con fines de descripción y no deberían ser consideradas como limitadoras.

20 La compresión de imágenes empleada en un aspecto de un modo de realización se basa en técnicas de transformación discreta de coseno (DCT), tales como la divulgada en la Solicitud de Patente de los EE.UU., presentada a la vez, "Compresión de imágenes por DCT de tamaño de bloque adaptativo, basada en la varianza sensible al contraste", con n.º de serie 09 / 436.085, presentada el 8 de noviembre de 1999, cedida al cesionario de la presente invención. Generalmente, una imagen a procesar en el dominio digital está compuesta por datos de píxeles divididos en una formación de bloques no solapados, de tamaño NxN. Puede realizarse una DCT bidimensional en cada bloque. La DCT bidimensional está definida por la siguiente relación:

$$X(k,l) = \frac{\alpha(k)\beta(l)}{\sqrt{N * M}} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x(m,n) \cos\left[\frac{(2m+1)\pi k}{2N}\right] \cos\left[\frac{(2n+1)\pi l}{2N}\right], 0 \leq k,l \leq N-1$$

30 donde $\alpha(k), \beta(k) = \begin{cases} 1, & \text{si } k = 0 \\ \sqrt{2} & \text{si } k \neq 0 \end{cases}$,

y $x(m,n)$ es el píxel en la ubicación (m,n) dentro de un bloque de tamaño NxM, y $X(k,l)$ es el correspondiente coeficiente de la DCT.

35 Dado que los valores de píxel son no negativos, el componente de la DCT $X(0,0)$ es siempre positivo y generalmente tiene la mayor parte de la energía. De hecho, para imágenes típicas, la mayoría de la energía de transformación está concentrada alrededor del componente $X(0,0)$. Esta propiedad de compactación de la energía es lo que hace de la técnica de la DCT un procedimiento de compresión tan atractivo.

40 La técnica de compresión de imágenes utiliza la codificación adaptativa por contraste para lograr una reducción adicional de la tasa de bits. Se ha observado que las imágenes más naturales están compuestas por áreas planas de variación relativamente lenta, y áreas activas tales como fronteras de objetos y textura de alto contraste. Los esquemas de codificación adaptativa por contraste aprovechan este factor asignando más bits a las áreas activas y menos bits a las áreas menos activas.

45 Los procedimientos adaptativos por contraste utilizan la codificación intra-trama (procesamiento espacial) en lugar de la codificación inter-trama (procesamiento espacial-temporal). La codificación inter-trama requiere inherentemente múltiples almacenes temporales de tramas, además de circuitos de procesamiento más complejos. En muchas aplicaciones se necesita una complejidad reducida para la implementación efectiva. La codificación intra-trama también es útil en una situación que puede hacer que un esquema de codificación espacial-temporal se desmorone y rinda malamente. Por ejemplo, las películas de 24 tramas por segundo pueden caer en esta categoría, dado que el tiempo de integración, debido al obturador mecánico, es relativamente corto. El corto tiempo de integración permite un mayor grado de irregularidad temporal de trazos. La hipótesis de correlación de trama a trama se desmorona para el movimiento rápido, según se torna espasmódico. La codificación intra-trama también es más fácil de estandarizar cuando están implicadas frecuencias de línea de energía tanto de 50 Hz como de 60 Hz. La televisión transmite actualmente señales, bien a 50 Hz o bien a 60 Hz. El uso de un esquema intra-trama, siendo un enfoque

digital, puede adaptarse al funcionamiento tanto de 50 Hz como de 60 Hz, o incluso a películas de 24 tramas por segundo, equilibrando la velocidad de tramas con respecto a la resolución espacial.

5 Con fines de procesamiento de imágenes, la operación de la DCT se realiza en datos de píxeles que están divididos en una formación de bloques no solapados. Obsérvese que, aunque los tamaños de bloque se exponen en el presente documento como de tamaño NxN, se prevé que pueden ser usados diversos tamaños de bloque. Por ejemplo, puede utilizarse un tamaño de bloque de NxM, donde tanto N como M son enteros, siendo M mayor o menor que N. Otro aspecto importante es que el bloque es divisible en al menos un nivel de sub-bloques, tales como N/ixN/i, N/ixN/j, N/ixM/j, etc., donde i y j son enteros. Además, el tamaño de bloque ejemplar, según lo expuesto en el
10 presente documento, es un bloque de píxeles de tamaño 16x16, con bloque y sub-bloques correspondientes de coeficientes de la DCT. Se prevé además que pueden ser usados otros enteros diversos, tales como valores enteros, tanto pares como impares, p. ej., 9x9.

15 Las FIGs. 1 y 2 ilustran un sistema de procesamiento de imágenes 100 que incorpora el concepto de serializador configurable. El sistema de procesamiento de imágenes 100 comprende un codificador 104 que comprime una señal de vídeo recibida. La señal comprimida es transmitida usando un canal de transmisión o un medio físico 108, y es recibida por un descodificador 112. El descodificador 112 descodifica los datos codificados recibidos en muestras de imágenes, que pueden ser exhibidas luego.

20 En general, una imagen es dividida en bloques de píxeles para su procesamiento. Una señal cromática puede ser convertida desde el espacio RGB al espacio YC₁C₂, usando un convertidor de RGB a YC₁C₂ 116, donde Y es el componente de luminancia, o brillo, y C₁ y C₂ son los componentes de crominancia, o color. Debido a la baja sensibilidad espacial del ojo al color, muchos sistemas sub-muestran los componentes C₁ y C₂ en un factor de cuatro, en las direcciones horizontal y vertical. Sin embargo, el sub-muestreo no es necesario. Una imagen de
25 resolución completa, conocida como de formato 4:4:4, puede ser muy útil o necesaria en algunas aplicaciones, tales como las mencionadas como capaces de abarcar el "cine digital". Dos posibles representaciones de YC₁C₂ son la representación YIQ y la representación YUV, ambas muy bien conocidas en la técnica. También es posible emplear una variación de la representación YUV, conocida como YCbCr. Esto puede ser descompuesto adicionalmente en componentes impares y pares. En consecuencia, en un modo de realización se usa la representación Y-par, Y-impar, Cb-par, Cb-impar, Cr-par, Cr-impar.
30

En un modo de realización preferente, cada uno de los componentes Y, Cb y Cr, par e impar, es procesado sin sub-muestreo. Por tanto, una entrada de cada uno de los seis componentes de un bloque de píxeles de tamaño 16x16 es proporcionada al codificador 104. Con fines de ilustración, se ilustra el codificador 104 para el componente Y-par.
35 Se usan codificadores similares para el componente Y-impar, y los componentes Cb y Cr par e impar. El codificador 104 comprende un elemento de asignación de tamaño de bloque 120, que realiza la asignación del tamaño de bloque, en preparación para la compresión de vídeo. El elemento de asignación del tamaño de bloque 120 determina la descomposición en bloques del bloque de tamaño 16x16, en base a las características perceptivas de la imagen en el bloque. La asignación del tamaño de bloque subdivide cada bloque de tamaño 16x16 en bloques más pequeños, tales como de tamaño 8x8, 4x4 y 2x2, en forma de árbol cuadrático, según la actividad dentro de un
40 bloque de tamaño 16x16. El elemento de asignación del tamaño de bloque 120 genera datos de árbol cuadrático, llamados los datos PQR, cuya longitud puede estar entre 1 y 21 bits. Por tanto, si la asignación del tamaño de bloque determina que un bloque de tamaño 16x16 ha de ser dividido, el bit R de los datos PQR es activado y está seguido por cuatro bits adicionales de datos Q, correspondientes a los cuatro bloques divididos de tamaño 8x8. Si la
45 asignación del tamaño de bloque determina que cualquiera de los bloques de tamaño 8x8 ha de ser subdividido, entonces se añaden cuatro bits adicionales de datos P para cada bloque subdividido de tamaño 8x8.

Con referencia ahora a la FIG. 3, se proporciona un diagrama de flujo que muestra detalles del funcionamiento del elemento de asignación del tamaño de bloque 120. La varianza de un bloque es usada como una métrica en la
50 decisión de subdividir un bloque. Comenzando en la etapa 202, se lee un bloque de píxeles de tamaño 16x16. En la etapa 204, se calcula la varianza, v16, del bloque de tamaño 16x16. La varianza se calcula de la siguiente manera:

$$\text{var} = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} x^2_{i,j} - \left(\frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} x_{i,j} \right)^2$$

55 donde N = 16 y x_{ij} es el píxel en la i-ésima fila, j-ésima columna dentro del bloque de tamaño NxN. En la etapa 206, primero se modifica el umbral de varianza T16 para proporcionar un nuevo umbral T'16 si el valor medio del bloque está entre dos valores predeterminados, y luego la varianza del bloque se compara con el nuevo umbral, T'16.

60 Si la varianza v16 no es mayor que el umbral T16, entonces en la etapa 208 se escribe la dirección de inicio del bloque de tamaño 16x16 en un almacenamiento temporal, y el bit R de los datos PQR se fija en 0 para indicar que el bloque de tamaño 16x16 no está subdividido. El algoritmo lee luego el siguiente bloque de píxeles de tamaño 16x16.

Si la varianza v_{16} es mayor que el umbral T_{16} , entonces en la etapa **210** el bit R de los datos PQR se fija en 1 para indicar que el bloque de tamaño 16x16 ha de ser subdividido en cuatro bloques de tamaño 8x8.

5 Los cuatro bloques de tamaño 8x8, $i=1:4$, son considerados secuencialmente para una subdivisión adicional, según se muestra en la etapa **212**. Para cada bloque de tamaño 8x8, se calcula la varianza, v_{8i} , en la etapa **214**. En la etapa **216**, primero se modifica el umbral de varianza T_8 para proporcionar un nuevo umbral T'_8 si el valor medio del bloque está entre dos valores predeterminados, y luego la varianza del bloque se compara con este nuevo umbral.

10 Si la varianza v_{8i} no es mayor que el umbral T_8 , entonces en la etapa **218** se escribe la dirección de inicio del bloque de tamaño 8x8 en un almacenamiento temporal, y el bit Q correspondiente, Q_i , se fija en 0. Luego se procesa el próximo bloque de tamaño 8x8. Si la varianza v_{8i} es mayor que el umbral T_8 , entonces en la etapa **220** se fija el correspondiente bit Q, Q_i , en 1 para indicar que el bloque de tamaño 8x8 ha de ser subdividido en cuatro bloques de tamaño 4x4.

15 Los cuatro bloques de tamaño 4x4, $j=1:4$, son considerados secuencialmente para su subdivisión adicional, según se muestra en la etapa **222**. Para cada bloque de tamaño 4x4, se calcula la varianza, v_{4j} , en la etapa **224**. En la etapa **226** se modifica primero el umbral de varianza T_4 para proporcionar un nuevo umbral T'_4 si el valor medio del bloque está entre dos valores predeterminados, y luego se compara la varianza del bloque con este nuevo umbral.

20 Si la varianza v_{4j} no es mayor que el umbral T_4 , entonces en la etapa **228** se escribe la dirección del bloque de tamaño 4x4, y se fija en 0 el correspondiente bit P, P_j . Se procesa luego el próximo bloque de tamaño 4x4. Si la varianza v_{4j} es mayor que el umbral T_4 , entonces en la etapa **230** se fija en 1 el correspondiente bit P, P_j , para indicar que el bloque de tamaño 4x4 ha de ser subdividido en cuatro bloques de tamaño 2x2. Además, la dirección de los 4 bloques de tamaño 2x2 se escribe en un almacenamiento temporal.

25 Los umbrales T_{16} , T_8 y T_4 pueden ser constantes predeterminadas. Esto se conoce como la decisión dura. Alternativamente, puede ser implementada una decisión adaptativa o suave. Por ejemplo, la decisión suave varía los umbrales para las varianzas, según el valor medio de píxeles de los bloques de tamaño $2N \times 2N$, donde N puede ser 8, 4 o 2. Por tanto, las funciones de los valores medios de píxeles pueden ser usadas como los umbrales.

30 Con fines de ilustración, considérese el siguiente ejemplo. Sean los umbrales de varianza predeterminados para el componente Y, 50, 1.100 y 880, respectivamente, para los bloques de tamaño 16x16, 8x8 y 4x4. En otras palabras, $T_{16} = 50$, $T_8 = 1.100$ y $T_4 = 880$. Sea la gama de valores medios 80 y 100. Supongamos que la varianza calculada para el bloque de tamaño 16x16 es 60. Dado que 60 es mayor que T_{16} , y el valor medio 90 está entre 80 y 100, el bloque de tamaño 16x16 es subdividido en cuatro sub-bloques de tamaño 8x8. Supongamos que las varianzas calculadas para los bloques de tamaño 8x8 son 1.180, 935, 980 y 1.210. Dado que dos de los bloques de tamaño 8x8 tienen varianzas que superan T_8 , estos dos bloques son adicionalmente subdivididos para producir un total de ocho sub-bloques de tamaño 4x4. Finalmente, supongamos que las varianzas de los ocho bloques de tamaño 4x4 son 620, 630, 670, 610, 590, 525, 930 y 690, con los correspondientes valores medios 90, 120, 110, 115. Dado que el valor medio del primer bloque de tamaño 4x4 cae en la gama (80, 100), su umbral será reducido a $T'_4 = 200$, que es menor que 880. Luego, este bloque de tamaño 4x4 será subdividido, así como el séptimo bloque de tamaño 4x4. La asignación resultante del tamaño de bloque se muestra en la FIG. **4a**. La correspondiente descomposición en árbol cuadrático se muestra en la FIG. **4b**. Adicionalmente, los datos PQR generados por esta asignación del tamaño de bloque se muestran en la FIG. **4c**.

45 Obsérvese que se usa un procedimiento similar para asignar tamaños de bloque para los componentes cromáticos Y-impar, C_{b-par} , $C_{b-impar}$, C_{r-par} y $C_{r-impar}$. Los componentes cromáticos pueden ser diezmados horizontalmente, verticalmente, o en ambos sentidos.

50 Adicionalmente, obsérvese que, aunque la asignación del tamaño de bloque ha sido descrita como un enfoque de arriba hacia abajo, en el cual el bloque más grande (de tamaño 16x16 en el presente ejemplo) se evalúa primero, puede usarse en cambio un enfoque de abajo hacia arriba. El enfoque de abajo hacia arriba evaluará primero los bloques más pequeños (de tamaño 2x2 en el presente ejemplo).

55 Con referencia de nuevo a la FIG. **1**, los datos PQR, junto con las direcciones de los bloques seleccionados, son proporcionados a un elemento de DCT **124**. El elemento de DCT **124** usa los datos PQR para realizar transformaciones discretas de coseno de los tamaños adecuados sobre los bloques seleccionados. Solamente los bloques seleccionados necesitan someterse al procesamiento de la DCT.

60 El sistema de procesamiento de imágenes **100** también comprende el elemento de DQT **128** para reducir la redundancia entre los coeficientes DC de las DCT. Un coeficiente DC se encuentra en la esquina superior izquierda de cada bloque de la DCT. Los coeficientes DC son, en general, grandes en comparación con los coeficientes AC. La discrepancia en los tamaños dificulta diseñar un codificador eficaz de longitud variable. En consecuencia, es ventajoso reducir la redundancia entre los coeficientes DC.

65

- El elemento de DQT **128** realiza DCT bidimensionales sobre los coeficientes DC, tomados 2x2 a la vez. A partir de bloques de tamaño 2x2 dentro de bloques de tamaño 4x4 se realiza una DCT bidimensional sobre los cuatro coeficientes DC. Esta DCT de tamaño 2x2 se llama la transformación de árbol cuadrático diferencial, o DQT, de los cuatro coeficientes DC. Luego, el coeficiente DC de la DQT, junto con los tres coeficientes DC vecinos dentro de un
- 5 bloque de tamaño 8x8, se usan para calcular la DQT del próximo nivel. Finalmente, los coeficientes DC de los cuatro bloques de tamaño 8x8 dentro de un bloque de tamaño 16x16 se usan para calcular la DQT. Por tanto, en un bloque de tamaño 16x16, hay un coeficiente DC verdadero y el resto son coeficientes AC correspondientes a la DCT y la DQT.
- 10 Los coeficientes de transformación (tanto de DCT como de DQT) son proporcionados a un cuantizador para su cuantización. En un modo de realización preferente, los coeficientes de DCT son cuantizados usando máscaras de ponderación de frecuencia (FWM) y un factor de escala de cuantización. Una FWM es una tabla de ponderaciones de frecuencia de las mismas dimensiones que el bloque de coeficientes de DCT de entrada. Las ponderaciones de frecuencia aplican distintas ponderaciones a los distintos coeficientes de DCT. Las ponderaciones están diseñadas
- 15 para destacar las muestras de entrada que tengan contenido de frecuencia al que el sistema visual u óptico humano es más sensible y para depreciar las muestras que tengan contenido de frecuencia al que el sistema visual u óptico es menos sensible. Las ponderaciones también pueden ser diseñadas en base a factores tales como distancias de visualización, etc.
- 20 Las ponderaciones son seleccionadas en base a datos empíricos. Un procedimiento para diseñar las máscaras de ponderación para coeficientes de DCT de tamaño 8x8 se divulga en el documento ISO / IEC JTC1 CD 10918, "Compresión y codificación digital de imágenes fijas de tono continuo – parte 1: Requisitos y directrices", Organización Internacional de Normas, 1994, que está incorporado al presente documento por referencia. En general, se diseñan dos FWM, una para el componente de luminancia y uno para los componentes de crominancia.
- 25 Las tablas de FWM, para tamaños de bloques 2x2 y 4x4, se obtienen por diezmado, y los de 16x16 por interpolación de la del bloque de tamaño 8x8. El factor de escala controla la calidad y la tasa de bits de los coeficientes cuantizados.

Por tanto, cada coeficiente de DCT es cuantizado de acuerdo a la relación:

30

$$DCT_q(i, j) = \left\lfloor \frac{8 * DCT(i, j)}{fwm(i, j) * q} \pm \frac{1}{2} \right\rfloor$$

donde DCT(i,j) es el coeficiente de DCT de entrada, fwm(i,j) es la máscara de ponderación de frecuencia, q es el factor de escala y DCT_q(i,j) es el coeficiente cuantizado. Obsérvese que, según el signo del coeficiente de DCT, el primer término dentro de los corchetes es redondeado hacia arriba o hacia abajo. Los coeficientes de DQT también son cuantizados usando una máscara de ponderación adecuada. Sin embargo, pueden usarse múltiples tablas o máscaras y aplicarse a cada uno de los componentes Y, Cb y Cr.

35

El bloque de datos de píxeles y las máscaras de ponderación de frecuencia son luego ajustados a escala por un cuantizador **130**, o un elemento de factor de escala. En un modo de realización preferente, hay 32 factores de escala correspondientes a tasas medias de bits. A diferencia de otros procedimientos de compresión, tales como MPEG2, la tasa media de bits se controla en base a la calidad de la imagen procesada, en lugar de la tasa de bits deseada y el estado del almacén temporal.

40

Los coeficientes cuantizados son proporcionados a un serializador de recorrido **152**. El serializador **152** recorre los bloques de coeficientes cuantizados para producir un flujo serializado de coeficientes cuantizados. Pueden ser empleados los recorridos en zig-zag, el recorrido por columnas o el recorrido por filas. También puede ser escogido un cierto número de distintos patrones de recorrido en zigzag, así como patrones distintos al zigzag. Una técnica preferente emplea tamaños de bloque de 8x8 para el recorrido en zigzag, aunque pueden ser empleados otros tamaños.

50

Se describen en el presente documento distintas técnicas de recorrido, con respecto a las FIGs. **4** y **5**. La FIG. **4b** ilustra un recorrido en zig-zag por todo un bloque **400** de tamaño 16x16. En un bloque basado en la frecuencia, tal como la DCT, los valores son codificados y representados de modo que el valor de DC esté en la esquina superior izquierda, y los valores de AC disminuyen en valor al aproximarse a la esquina inferior derecha. Por lo tanto, una técnica de recorrido en zig-zag por todo el bloque de tamaño 16x16, independientemente de la asignación del tamaño de bloque dentro del bloque de tamaño 16x16, lleva a ineficacia en la codificación. En otras palabras, el recorrido en zig-zag de tal forma lleva a longitudes más cortas de racha del mismo valor.

55

La FIG. **4c** ilustra una técnica más eficaz de recorrido, aprovechando el orden en el que los coeficientes están ordenados en un bloque dado. Cada bloque **404, 406, 408, 410, 412, 414, 416, 418, 420, 422, 424, 426** y **428** emplea un recorrido distinto en zig-zag. En un modo de realización, cada bloque puede emplear distintos patrones de recorrido, tales como verticales u horizontales, o en zig-zag inverso. Aunque este modo de realización está

60

sumamente optimizado en la preservación de las máximas longitudes de racha, calcular recorridos en zig-zag por separado para cada bloque es más intensivo en términos de cálculo y, en hardware, puede ser más difícil de implementar.

5 En consecuencia, se ha determinado que implementaciones del recorrido, tales como las descritas en las FIGs. **5a** y **5b**, pueden equilibrar la optimización de la maximización de longitudes de racha, equilibradas con la facilidad en la implementación de hardware. La FIG. **5a** ilustra un bloque **500** de tamaño 16x16, subdividido por la asignación del tamaño de bloque en los bloques **504, 506, 508, 510, 512, 514, 516, 518, 520, 522, 524, 526** y **528**. En un modo de realización, independientemente de la descomposición de la BSA, se emplea el recorrido en zig-zag en cada cuadrante de tamaño 8x8 del bloque de tamaño 16x16. Por tanto, los bloques **504, 506, 508** y **510** son serializados por un recorrido en zig-zag, el bloque **512** es serializado por un recorrido en zig-zag, el bloque **514** es serializado por un recorrido en zig-zag y los bloques **516, 518, 520, 522, 524, 526** y **528** son serializados por un recorrido en zig-zag.

15 La FIG. **5b** ilustra un bloque **550** de tamaño 16x16, subdividido por la asignación del tamaño de bloque en los bloques **554, 556, 558, 560, 562, 564, 568, 570, 572, 574, 576** y **578**. En este modo de realización, se emplean distintos tipos de recorrido en cada cuadrante de tamaño 8x8 del bloque de tamaño 16x16. El tipo de recorrido empleado está determinado por la evaluación de los valores dentro del bloque de tamaño 8x8 y la determinación de un procedimiento de recorrido que sea el más eficaz. Por ejemplo, en la FIG. **5b**, se emplea el recorrido horizontal para los bloques **554, 556, 558, 560**, el bloque **562** es serializado por un recorrido en zig-zag, el bloque **564** es serializado por un recorrido vertical, y los bloques **566, 568, 570, 572, 574, 576** y **578** son serializados por un recorrido en zig-zag. En un modo de realización alternativo, el procedimiento óptimo de recorrido es determinado trama por trama, a diferencia de bloque por bloque. La determinación del procedimiento óptimo de recorrido trama por trama es menos intensivo en términos de cálculo, a diferencia del procedimiento bloque por bloque.

25 La FIG. **6a** ilustra un proceso **600** por el cual ocurre la serialización. Se lee **604** un grupo de datos. Dado que los datos leídos están basados en tamaños variables de bloque, los datos leídos no son de tamaño o longitud uniforme. Los datos son compilados **608** o estructurados en una forma que pueda ser representada como un bloque de tamaño 16x16. Los datos son luego divididos **612** en cuatro tamaños de bloque de 8x8. Se realiza **616** luego un recorrido en zig-zag en cada bloque de tamaño 8x8. Los datos son luego encaminados **620** a un almacén temporal.

30 La FIG. **6b** ilustra un modo de realización alternativo **650** de la serialización. Se lee una trama de datos **654**. La trama de datos es evaluada **658** para determinar la técnica óptima de serialización. En base a la evaluación, se emplea un recorrido en zig-zag **662**, un recorrido vertical **664** o un recorrido horizontal **668**. Tras serializar en base a uno de los procedimientos de recorrido, los datos son encaminados **672** a un almacén temporal.

35 Con referencia de nuevo a la FIG. **1**, se proporciona el flujo de coeficientes serializados y cuantizados a un codificador de longitud variable **156**. El codificador de longitud variable **156** puede hacer uso de la codificación por longitud de racha de los ceros, seguida por la codificación de Huffman. Esta técnica está expuesta en detalle en las precitadas Patentes de los EE.UU. con n.º 5.021.891, 5.107.345 y 5.452.104, que están incorporadas por referencia y resumidas en el presente documento. Un codificador por longitud de racha tomaría los coeficientes cuantizados y separaría los coeficientes nulos de los no nulos. Los valores nulos son mencionados como valores de longitud de racha, y son codificados según Huffman. Los valores no nulos son codificados según Huffman por separado.

45 También es posible una codificación de Huffman modificada de los coeficientes cuantizados y se usa en el modo de realización preferente. Aquí, después del recorrido en zigzag, un codificador por longitud de racha determinará los pares de longitud de racha y tamaño dentro de cada bloque de tamaño 8x8. Estos pares de longitud de racha y tamaño son luego codificados según Huffman.

50 Los códigos de Huffman son diseñados a partir de las estadísticas, bien medidas o bien teóricas, de una imagen. Se ha observado que la mayoría de las imágenes naturales están compuestas por áreas planas, o de variación relativamente lenta, y áreas activas, tales como fronteras de objetos y textura de alto contraste. Los codificadores de Huffman con transformaciones del dominio de la frecuencia, tales como la DCT, explotan estas características asignando más bits a las áreas activas y menos bits a las áreas planas. En general, los codificadores de Huffman hacen uso de tablas de consulta para codificar la longitud de racha y los valores no nulos. Se usan generalmente tablas múltiples, siendo preferentes 3 tablas en la presente invención, aunque pueden emplearse 1 o 2, según se desee.

55 La señal de imagen comprimida, generada por el codificador **104**, puede ser almacenada temporalmente usando un almacén temporal **160**, y luego transmitida al descodificador **112** usando el canal de transmisión **108**. Los datos PQR, que contienen la información de asignación de tamaño de bloque, también se proporcionan al descodificador **112**. El descodificador **112** comprende un almacén temporal **164** y un descodificador de longitud variable **168**, que descodifica los valores de longitud de racha y los valores no nulos.

65 La salida del descodificador de longitud variable **168** es proporcionada a un serializador inverso **172**, que ordena los coeficientes de acuerdo al esquema de recorrido empleado. Por ejemplo, si se usaron una mezcla de recorrido en

zig-zag, recorrido vertical y recorrido horizontal, el serializador inverso **172** reordenaría adecuadamente los coeficientes con el conocimiento del tipo de recorrido empleado. El serializador inverso **172** recibe los datos PQR para asistir en el ordenamiento adecuado de los coeficientes, en un bloque compuesto de coeficientes.

5 El bloque compuesto es proporcionado, usando un selector **174**, a un cuantizador inverso **176**, para deshacer el procesamiento debido al uso del factor de escala del cuantizador y las máscaras de ponderación de frecuencia.

10 El bloque de coeficientes es luego proporcionado a un elemento de IDQT **186**, seguido por un elemento de IDCT **186**, si se había aplicado la transformación del Árbol Cuadrático Diferencial. En otro caso, el bloque de coeficientes es proporcionado directamente al elemento de IDCT **190**. El elemento de IDQT **186** y el elemento de IDCT **190** transforman inversamente los coeficientes para producir un bloque de datos de píxeles. Los datos de píxeles luego pueden tener que ser interpolados, convertidos al formato RGB y almacenados luego para su exhibición futura.

15 En consecuencia, se presenta un sistema y un procedimiento para la compresión de imágenes que realiza la asignación del tamaño de bloque en base a la varianza de píxeles. La asignación del tamaño de bloque basada en la varianza ofrece varias ventajas. Debido a que la Transformación Discreta de Coseno es realizada después de que son determinados los tamaños de bloque, se logra un cálculo eficaz. La transformación intensiva en términos de cálculo solamente necesita ser realizada en los bloques seleccionados. Además, el proceso de selección de bloques es eficaz, ya que la varianza de valores de píxeles es matemáticamente sencilla de calcular. Otra ventaja más de la asignación del tamaño de bloque en base a la varianza es que está basada perceptivamente. La varianza de píxeles es una medida de la actividad en un bloque y proporciona una indicación de la presencia de bordes, texturas, etc.

20 Tiende a capturar los detalles de un bloque mucho mejor que medidas tales como el promedio de valores de píxel. Por tanto, el esquema basado en la varianza de la presente invención asigna bloques más pequeños a regiones con más bordes y bloques más grandes a las regiones más planas. Como resultado, puede lograrse una calidad sobresaliente en las imágenes reconstruidas.

25 Como ejemplos, los diversos bloques lógicos ilustrativos, diagramas de flujo y etapas descritas con relación a los modos de realización divulgados en el presente documento pueden ser implementados o realizados en hardware o software, con un circuito integrado específico de la aplicación (ASIC), un dispositivo lógico programable, compuertas discretas o lógica de transistores, componentes discretos de hardware, tales como, p. ej., registros y FIFO, un procesador ejecutando un conjunto de instrucciones de firmware, cualquier software programable convencional y un procesador, o cualquier combinación de los mismos. El procesador, ventajosamente, puede ser un microprocesador pero, como alternativa, el procesador puede ser cualquier procesador convencional, controlador, micro-controlador o máquina de estados. El software podría residir en memoria RAM, memoria flash, memoria ROM, registros, un disco rígido, un disco extraíble, un CD-ROM, un DVD-ROM o cualquier otra forma de medio de almacenamiento conocido en la técnica.

30

35

40 La descripción anterior de los modos de realización preferentes se proporciona para permitir a cualquier persona experta en la técnica hacer o usar la presente invención. Las diversas modificaciones para estos modos de realización serán inmediatamente evidentes para los expertos en la técnica y los principios genéricos definidos en el presente documento pueden ser aplicados a otros modos de realización sin el uso de la facultad inventiva. Por tanto, la presente invención no está concebida para limitarse a los modos de realización mostrados en el presente documento, sino que ha de acordársele el más amplio ámbito congruente con los principios y características novedosas divulgados en el presente documento.

45 Otras características y ventajas de la invención se enuncian en las siguientes reivindicaciones.

50

REIVINDICACIONES

1. Un procedimiento de serialización de datos de vídeo, basado en representaciones del dominio de la frecuencia de los datos de vídeo, comprendiendo el procedimiento:
 - 5 compilar al menos un grupo de los datos de vídeo que puedan ser representados como un bloque de datos de vídeo;
 - 10 dividir el bloque de datos de vídeo en grupos que puedan ser representados como una pluralidad de primeros bloques de datos subdivididos;
 - dividir al menos uno, entre la pluralidad de primeros bloques subdivididos, en dos o más segundos bloques subdivididos;
 - 15 evaluar uno o más valores dentro de cada uno entre la pluralidad de primeros bloques subdivididos, para determinar un esquema de recorrido de serialización para cada uno entre la pluralidad de primeros bloques subdivididos; y
 - 20 serializar independientemente los coeficientes de transformación de cada uno, entre la pluralidad de primeros bloques subdivididos, de acuerdo al esquema determinado, sin serializar independientemente los coeficientes de transformación de cada uno de los dos o más segundos bloques subdivididos.
2. El procedimiento enunciado en la Reivindicación 1, en el que el bloque de datos de vídeo es de tamaño $M \times M$, donde M es un entero mayor que 1.
- 25 3. El procedimiento enunciado en la Reivindicación 2, en el que cada uno entre la pluralidad de primeros bloques de datos subdivididos es de tamaño $N \times N$, donde N es un entero mayor que 1, siendo M mayor que N .
- 30 4. El procedimiento enunciado en la Reivindicación 3, en el que la serialización de los coeficientes de transformación de cada uno entre la pluralidad de primeros bloques subdivididos comprende al menos el recorrido en zig-zag de cada uno entre la pluralidad de primeros bloques de datos subdivididos, el recorrido vertical de cada uno entre la pluralidad de primeros bloques de datos subdivididos o el recorrido horizontal de cada uno entre la pluralidad de primeros bloques de datos subdivididos.
- 35 5. El procedimiento enunciado en la Reivindicación 2, en el que la compilación de dicho al menos un grupo comprende compilar una trama de datos que pueda ser representada como una pluralidad de bloques.
- 40 6. El procedimiento enunciado en la Reivindicación 1, en donde los datos de vídeo están separados en componentes cromáticos Y , Cb y Cr .
7. El procedimiento enunciado en la Reivindicación 6, en el que los componentes cromáticos Y , Cb y Cr son adicionalmente separados en componentes cromáticos pares e impares.
- 45 8. El procedimiento enunciado en la Reivindicación 3, que comprende además:
 - dividir dicho al menos un grupo de datos en grupos que puedan ser representados como bloques subdivididos, de tamaño menor que 8×8 ; y
 - 50 serializar los bloques divididos de tamaño 8×8 que incluyen bloques subdivididos de tamaño 4×4 o 2×2 , de acuerdo al esquema de recorrido determinado.
9. El procedimiento enunciado en la Reivindicación 3, en el que la división de dicho al menos un grupo de datos comprende dividir un bloque de tamaño 16×16 en cuatro bloques subdivididos de tamaño 8×8 .
- 55 10. El procedimiento enunciado en la Reivindicación 5, en el que cada uno entre la pluralidad de bloques es de tamaño 16×16 .
- 60 11. El procedimiento enunciado en la Reivindicación 1, que comprende además:
 - leer un grupo de componentes cromáticos de datos de píxeles;
 - generar una asignación de tamaño de bloque para dividir el grupo de componentes cromáticos en sub-bloques de datos de píxeles;
 - 65 transformar los sub-bloques en correspondientes representaciones del dominio de la frecuencia; y

ajustar a escala las correspondientes representaciones del dominio de la frecuencia, en un flujo de datos, en donde el ajuste a escala de las correspondientes representaciones del dominio de la frecuencia está basado en una métrica de calidad correlacionada con una calidad de la imagen digital.

- 5
12. El procedimiento de la reivindicación 11, en el que el ajuste a escala de las correspondientes representaciones del dominio de la frecuencia comprende además:
- 10 proporcionar una máscara ponderada de frecuencia a los sub-bloques de los datos de píxeles, de modo que la máscara de ponderación de frecuencia proporcione mayor énfasis a las partes de la imagen digital donde un sistema visual humano es más sensible y proporcione menor énfasis a las partes de la imagen digital donde el sistema visual humano es menos sensible; y
- 15 cuantizar los sub-bloques de los datos de píxeles en base a la calidad de la imagen digital.
13. El procedimiento enunciado en la reivindicación 11, en el que la métrica de calidad es una razón entre señal y ruido.
14. El procedimiento enunciado en la reivindicación 11, en el que la transformación de los sub-bloques comprende realizar una técnica de compresión definida.
- 20 15. El procedimiento enunciado en la Reivindicación 14, en el que la técnica de compresión definida es la Transformación Discreta de Coseno.
- 25 16. El procedimiento enunciado en la Reivindicación 11, en el que la transformación de los sub-bloques comprende realizar una primera técnica de compresión definida, seguida por una segunda técnica de compresión definida.
- 30 17. El procedimiento enunciado en la Reivindicación 16, en el que la primera técnica de compresión definida es la Transformación Discreta de Coseno y la segunda técnica de compresión definida es la Transformación de Árbol Cuadrático Diferencial.
- 35 18. Un aparato para serializar datos de vídeo en base a representaciones del dominio de la frecuencia de los datos de vídeo, comprendiendo el aparato:
- 40 medios para compilar al menos un grupo de los datos de vídeo que puedan ser representados como un bloque de datos de vídeo;
- medios para dividir el bloque de datos de vídeo en grupos que puedan ser representados como una pluralidad de primeros bloques de datos subdivididos;
- 45 medios para dividir al menos uno, entre la pluralidad de primeros bloques subdivididos, en dos o más segundos bloques subdivididos;
- medios para evaluar uno o más valores dentro de cada uno entre la pluralidad de primeros bloques subdivididos, para determinar un esquema de recorrido de serialización para el primer bloque subdividido; y
- 50 medios para serializar independientemente los coeficientes de transformación de cada uno, entre la pluralidad de primeros bloques de datos subdivididos, de acuerdo al esquema de recorrido determinado, sin serializar independientemente los coeficientes de transformación de cada uno de los dos o más segundos bloques subdivididos.

55

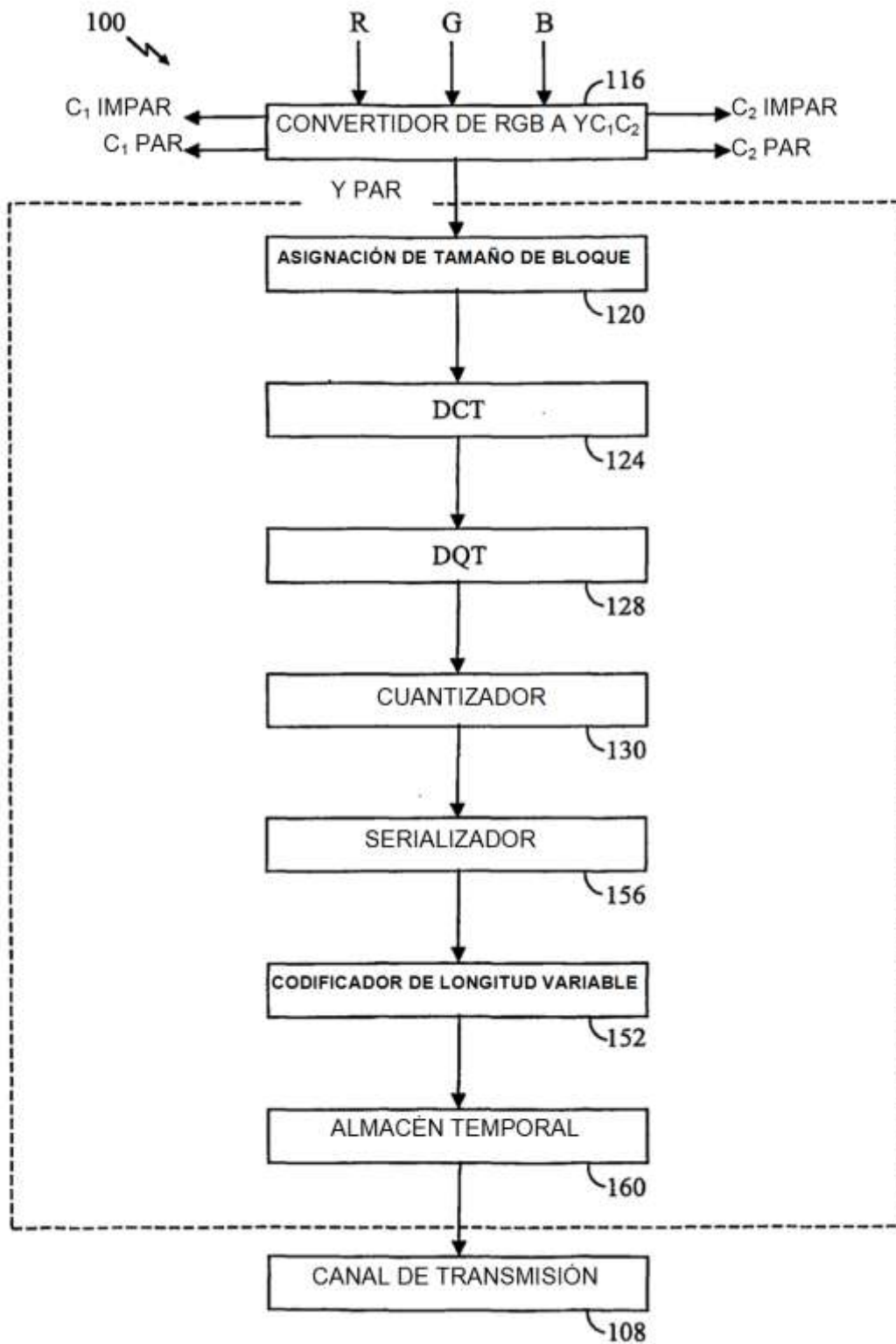


FIG. 1

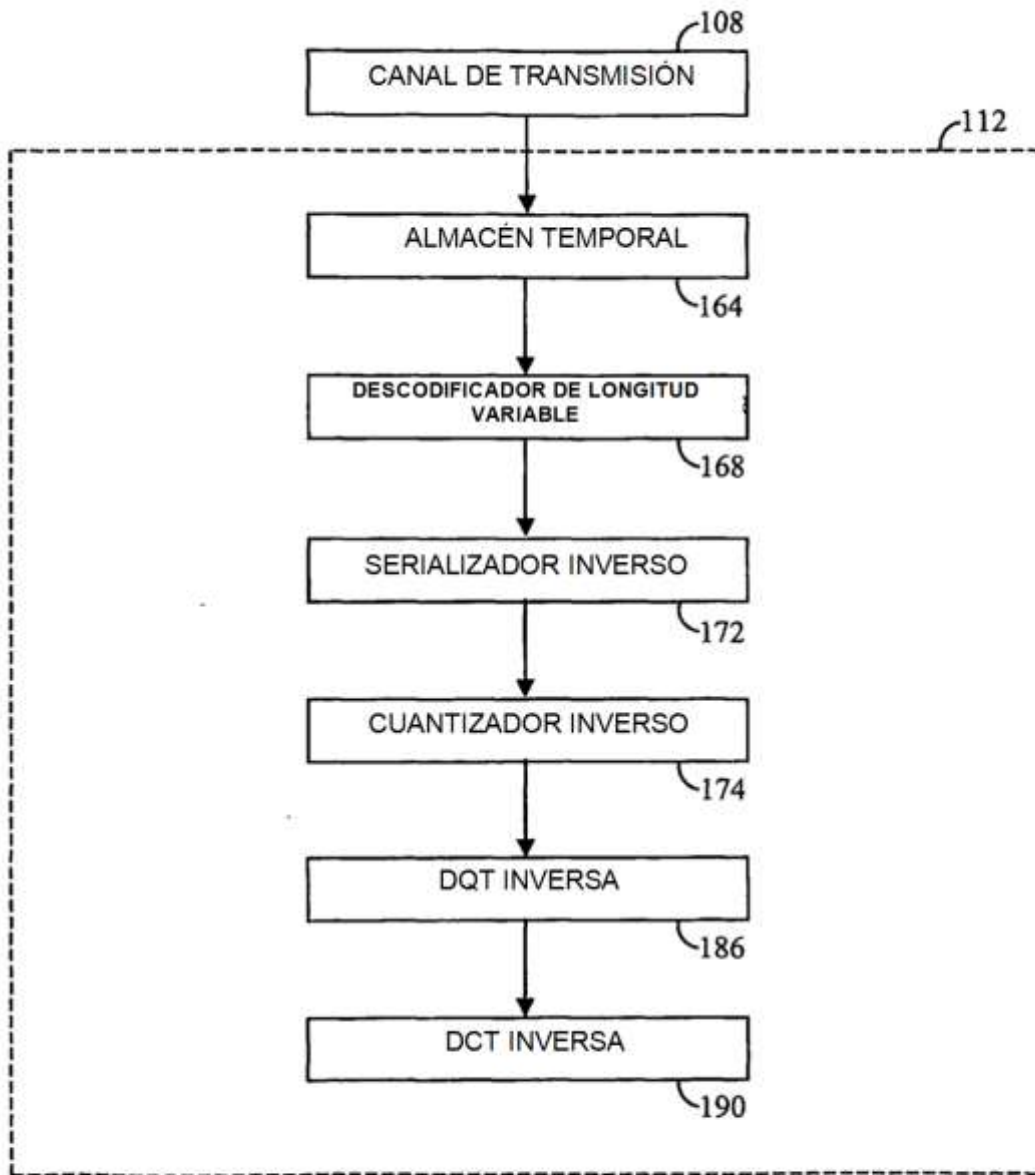


FIG. 2

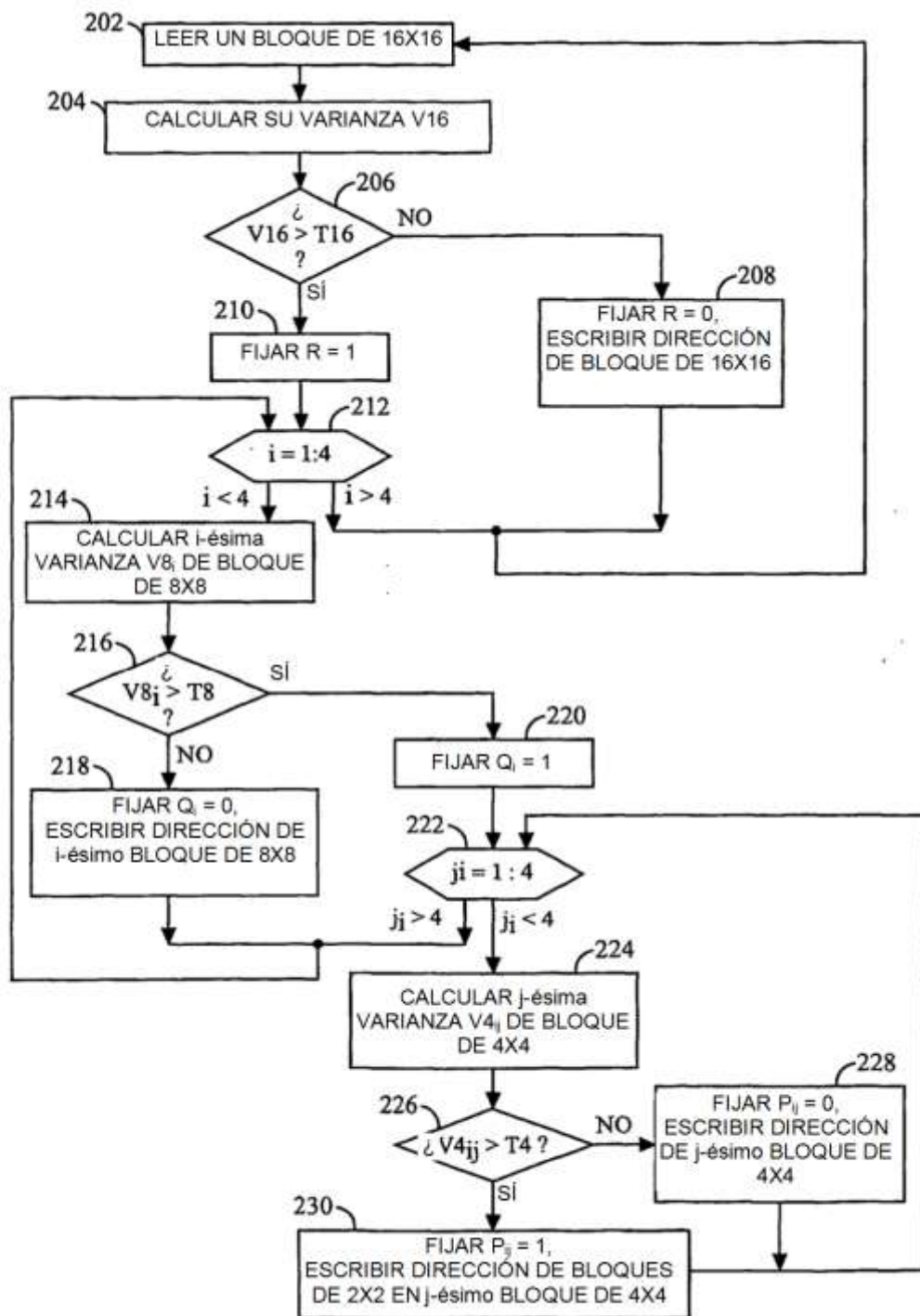


FIG. 3

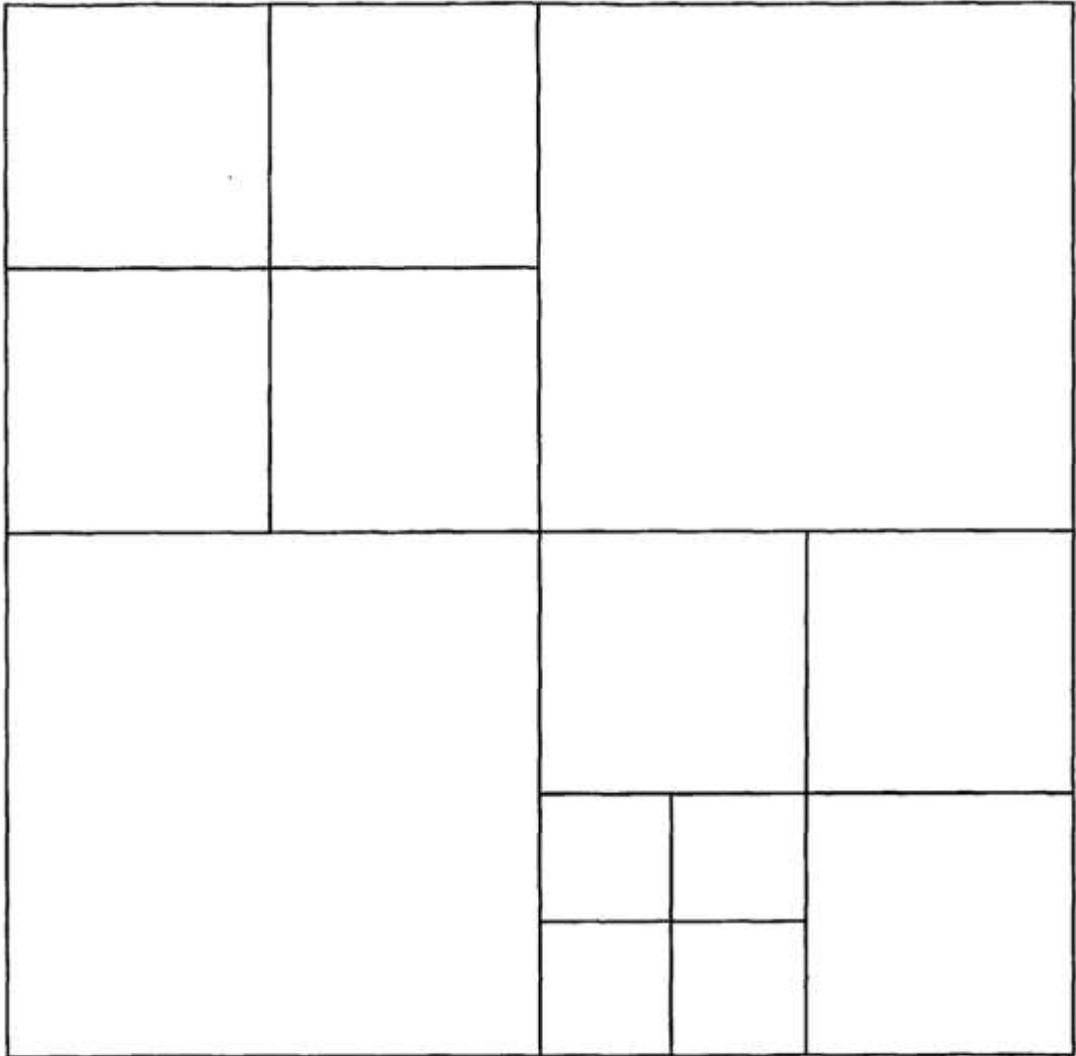
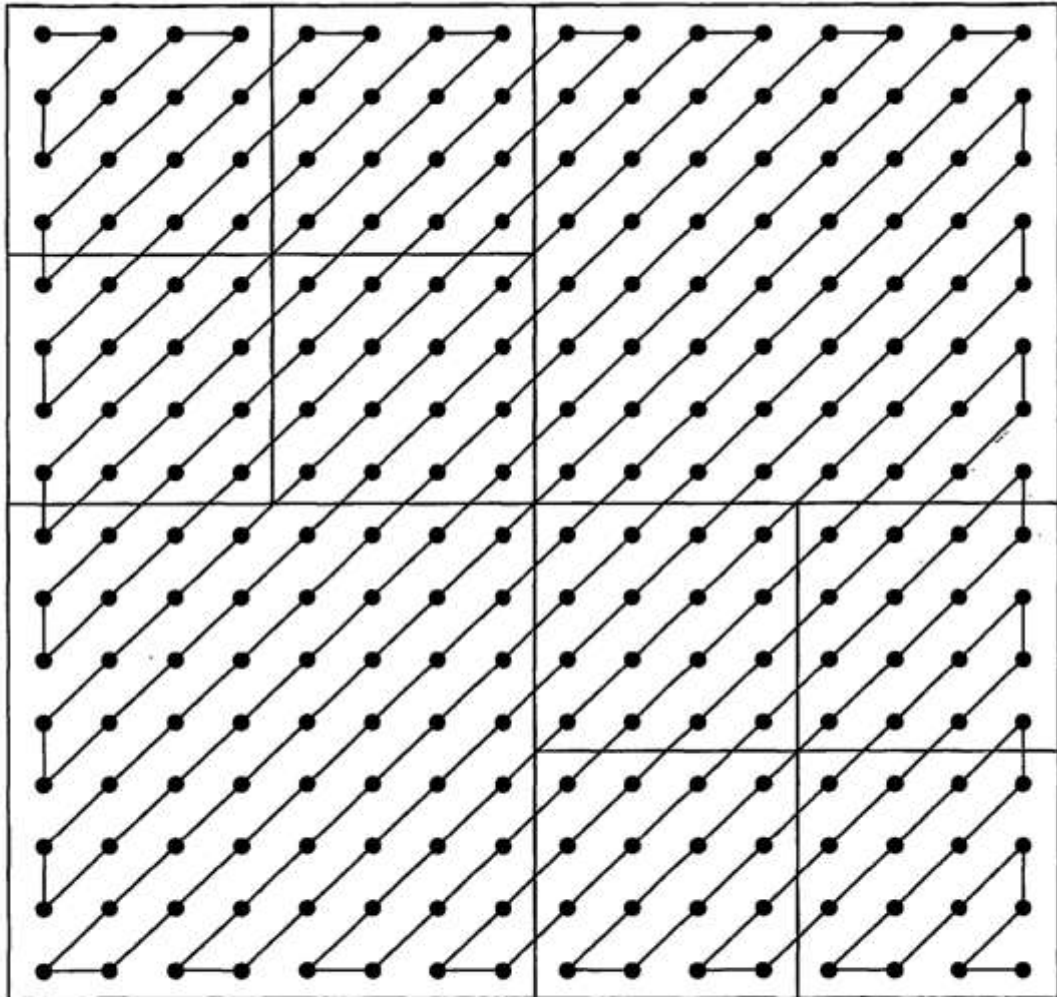


FIG. 4a





402 ↗

FIG. 4b

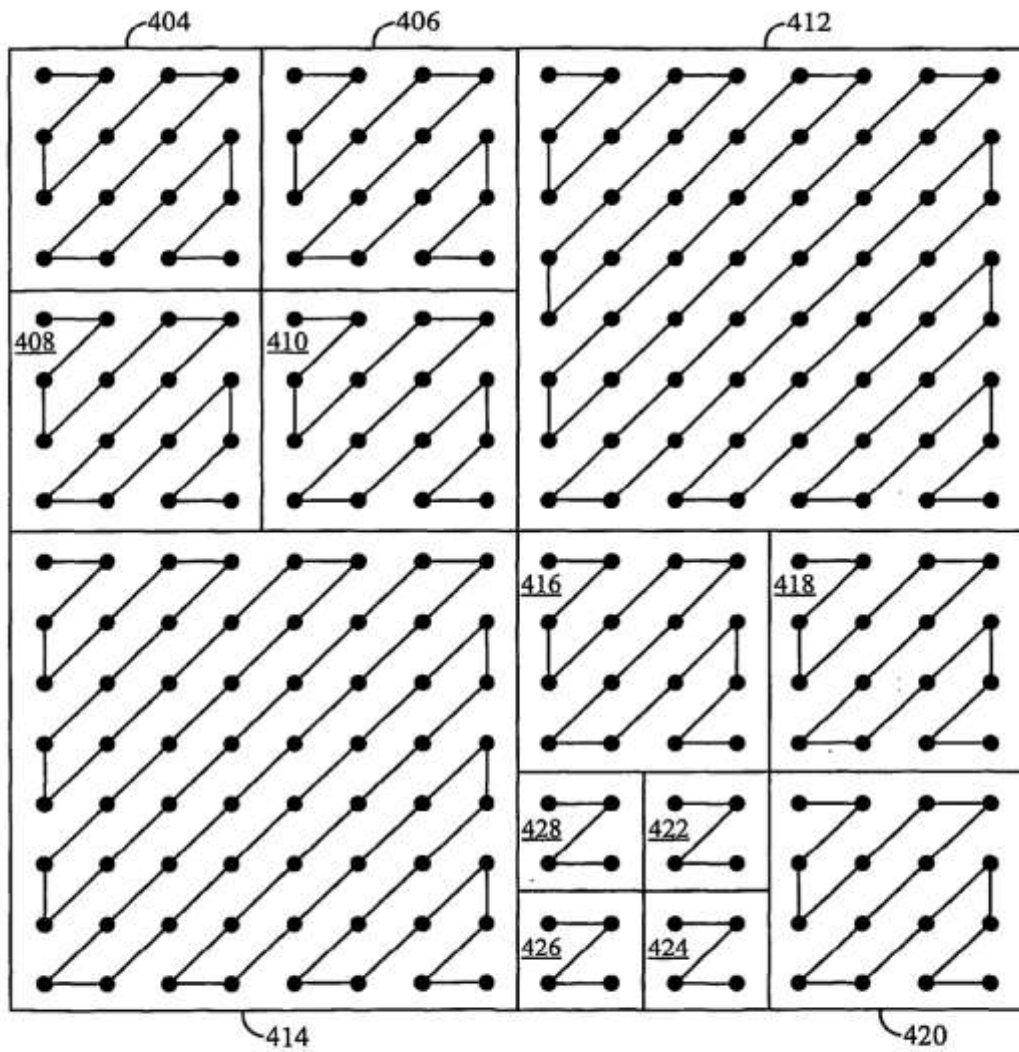


FIG. 4c

403

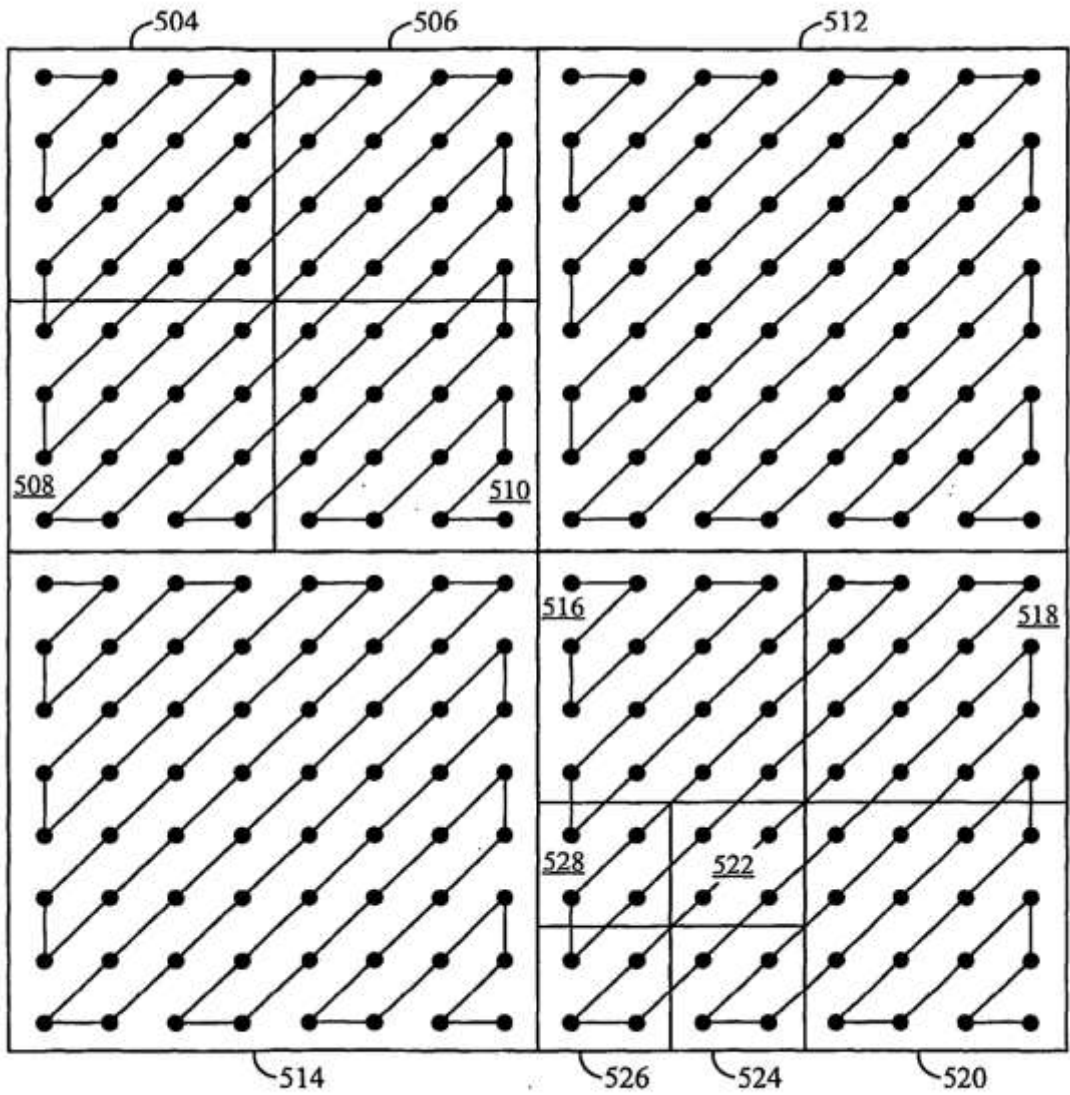


FIG. 5a

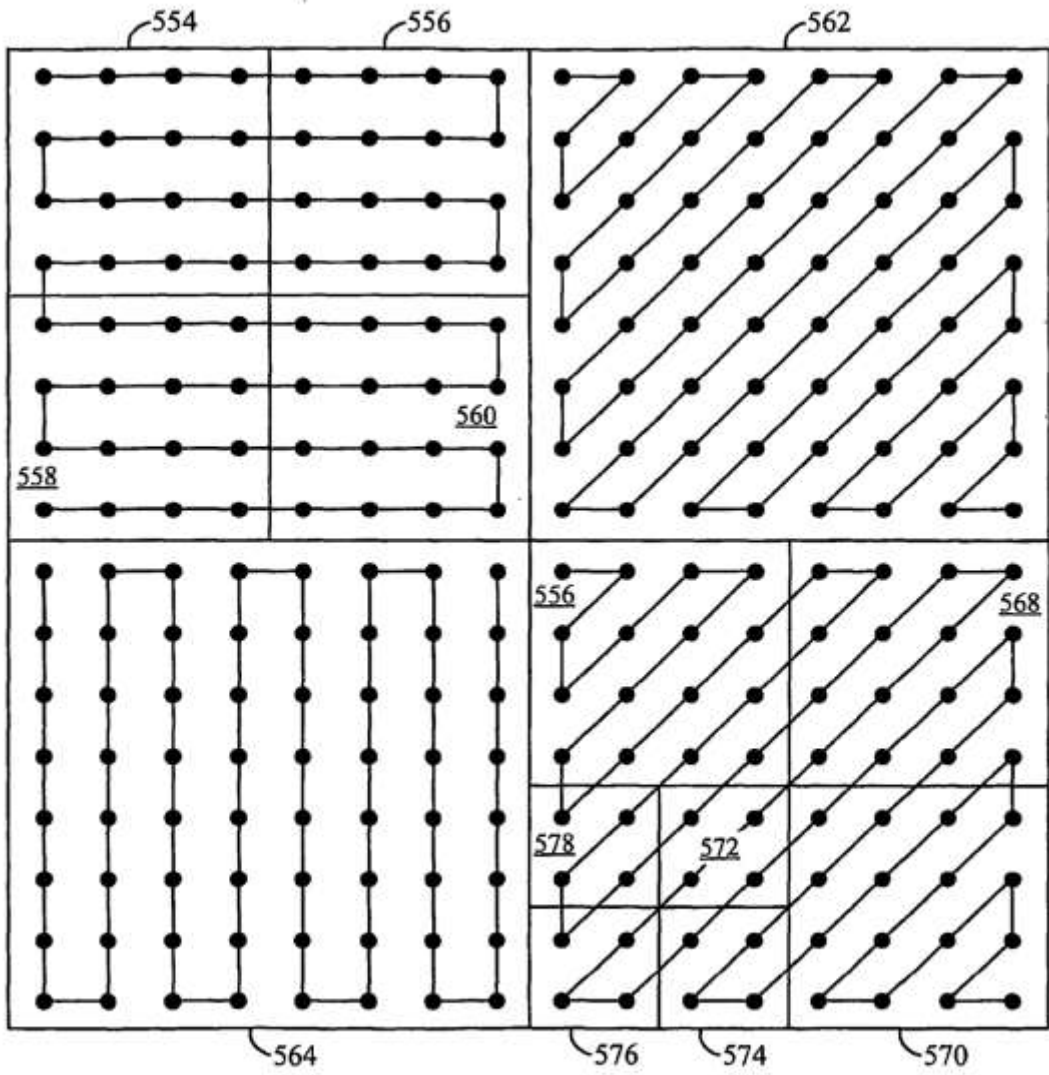


FIG. 5b

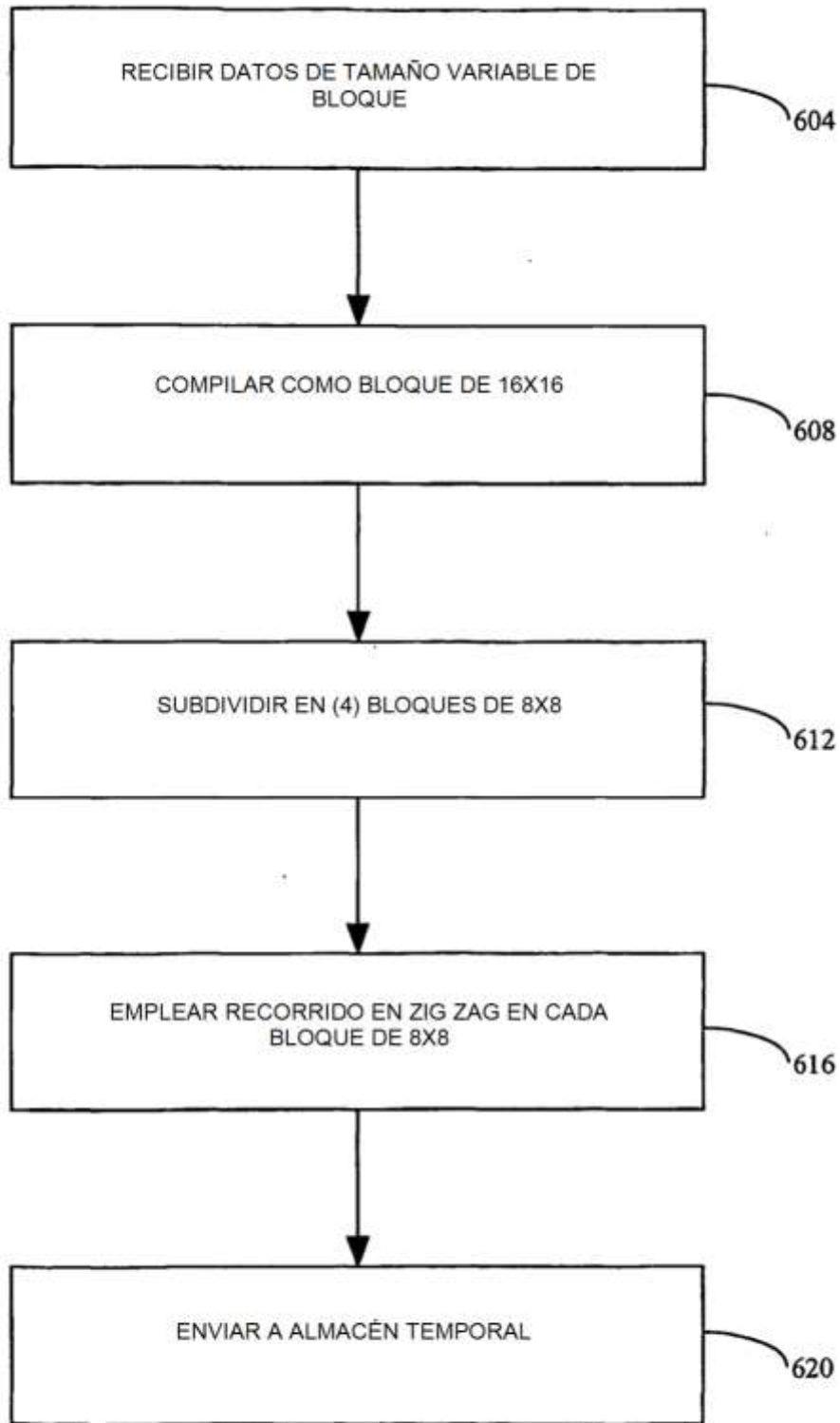


FIG. 6a

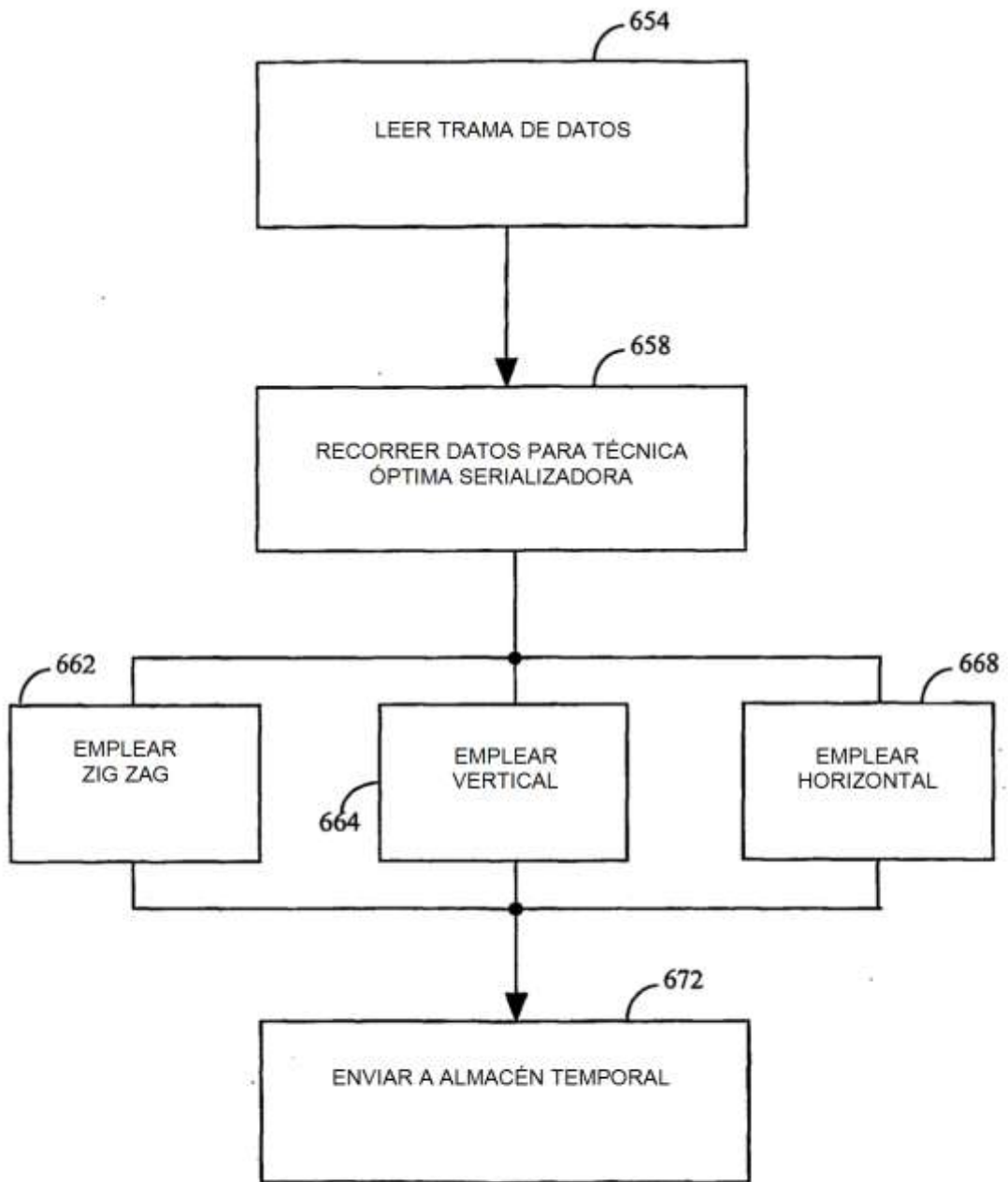


FIG. 6b

