

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 376 540**

51 Int. Cl.:
H04N 7/26

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **00977131 .2**

96 Fecha de presentación: **08.11.2000**

97 Número de publicación de la solicitud: **1230804**

97 Fecha de publicación de la solicitud: **14.08.2002**

54 Título: **COMPRESIÓN DE IMAGEN POR DCT DE TAMAÑO DE BLOQUE ADAPTATIVO EN BASE A LA VARIANZA.**

30 Prioridad:
08.11.1999 US 436085

45 Fecha de publicación de la mención BOPI:
14.03.2012

45 Fecha de la publicación del folleto de la patente:
14.03.2012

73 Titular/es:
**QUALCOMM INCORPORATED
5775 MOREHOUSE DRIVE
SAN DIEGO, CA 92121-1714, US**

72 Inventor/es:
**THYAGARAJAN, Kadayam y
MERRITT, Michael, J.**

74 Agente/Representante:
Carpintero López, Mario

ES 2 376 540 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

Compresión de imagen por DCT de tamaño de bloque adaptativo en base a la varianza

Antecedentes de la invención**I. Campo de la invención**

- 5 La presente invención se refiere al campo del procesamiento de imagen. Más concretamente, la presente invención se refiere a un esquema de compresión para señales de imagen que utiliza unos bloques y subbloques dimensionados de manera adaptativa de los datos de coeficientes codificados de la transformada de coseno discreta.

II. Descripción de la técnica relacionada

- 10 En el campo de la transmisión y recepción de señales de vídeo tal y como se utilizan para la proyección de "filmes" o "películas", se han realizado diversas mejoras respecto de las técnicas de compresión de imagen. Muchos de los sistemas de vídeo actuales y propuestos hacen uso de técnicas de codificación digital. La codificación digital proporciona una robustez en el enlace de comunicaciones la cual ofrece resistencia a problemas tales como el desvanecimiento por multitrayectoria y la perturbación o la transferencia de la señal, cada una de las cuales podría
- 15 en otro caso provocar una seria degradación de la calidad de la imagen. Así mismo, las técnicas digitales facilitan el uso de técnicas de encriptación de la señal, las cuales han resultado útiles o incluso necesarias en aplicaciones gubernamentales y en otras muchas de difusión comercial recientemente desarrolladas.

- El vídeo de alta definición es un área que aprovecha las técnicas mejoradas de compresión de imágenes. Cuando se propuso por primera vez, la transmisión por las ondas del vídeo de alta definición (o incluso la transmisión cableada o por fibra óptica pareció escasamente práctica debido a los excesivos condicionamientos del ancho de banda. Los sistemas de transmisión inalámbricos típicos, o de otro tipo que se estaban diseñando no se ajustaban con facilidad en la medida suficiente al ancho de banda. Sin embargo, se ha llegado al convencimiento de que la compresión de señales de vídeo digitales puede conseguirse hasta un nivel que permita la transmisión utilizando anchos de banda razonables. Dichos niveles de compresión de la señal, acoplados con la transmisión digital de la
- 20 señal, permitirá un sistema de vídeo que transmita con menor potencia y con una mayor inmunidad respecto de las perturbaciones del canal ocupando al tiempo un ancho de banda más conveniente y útil.

- Una técnica de compresión capaz de ofrecer niveles significativos de compresión preservando al tiempo el nivel de calidad deseado de las señales de vídeo utiliza un bloque y unos subbloques dimensionados de forma adaptativa de los datos de coeficientes codificados de la Transformada de Coseno Discreta (DCT). Esta técnica se designará en lo sucesivo como el procedimiento de la Transformada de Coseno Diferencial del Tamaño de Bloque Adaptativo (ABSDCT). Esta técnica se divulga en la Patente estadounidense No. 5,021,891, titulada "*Procedimiento y Sistema de Compresión de Imágenes del Tamaño de Bloque Adaptativo*" ["*Adaptive Block Size Image Compression Method And System*"] transferida al cesionario de la presente invención. Las técnicas de DCT se divulgan, así mismo, en la Patente estadounidense No. 5,107,345, titulada "*Procedimiento Y Sistema de Compresión de Imágenes del tamaño de Bloque Adaptativo*" ["*Adaptive Block Size Image Compression Method And System*"] transferida al cesionario de la presente invención. Así mismo, el uso de la técnica ABSDCT en combinación con la técnica de la Transformada de Árbol Cuaternario Diferencial se analiza en la Patente estadounidense No. 5,452,104, titulada "*Procedimiento Y Sistema de Compresión de Imágenes del tamaño de Bloque Adaptativo*" ["*Adaptive Block Size Image Compression Image Method And System*"], transferido, también al cesionario de la presente invención. Los sistemas divulgados en estas patentes utilizan lo que se designa como una codificación "intertramas", en la que cada trama de los datos de imagen es codificada sin consideración al contenido de cualquier otra trama. Al utilizar la técnica ABSDCT, puede ser reducida la tasa de transmisión de datos obtenible desde alrededor de 1,5 miles de millones de bits por segundo hasta aproximadamente 50 millones de bits por segundo sin una degradación apreciable de la calidad de la imagen.
- 30
- 35
- 40

- La técnica ABSDCT puede ser utilizada para comprimir una imagen o señal ya sea en blanco o negro o en color que represente la imagen. La señal entrante de color puede concretarse en un formato YIQ, siendo Y la luminancia, o el brillo, la muestra, y siendo I y Q la crominancia o el color, las muestras para cada bloque de píxeles de 4 x 4. Pueden, así mismo, ser utilizados otros formatos conocidos como por ejemplo los formatos YUV o RGB. Debido a la baja sensibilidad espacial del ojo al color, la mayor parte de la investigación ha mostrado que una submuestra de los componentes de color mediante un factor de cuatro en las direcciones horizontal y vertical es razonable. De acuerdo con ello, una señal de vídeo puede ser representada por cuatro componentes de luminancia y dos componentes de crominancia.
- 45
- 50

- Utilizando la ABSDCT, una señal de vídeo resultará segmentada, en términos generales, en bloques de píxeles para su procesamiento. Para cada bloque, los componentes de luminancia y crominancia son transferidos a un intercalador de bloques. Por ejemplo, un bloque de 16 x 16 (píxeles) puede ser presentado al intercalador de bloques, el cual ordena y organiza las muestras de imágenes dentro de cada bloque de 16 x 16 para producir bloques y subbloques compuestos de datos para el análisis de la transformada de coseno discreta (DCT). El operador de la DCT es un procedimiento de conversión de una señal muestreada en función del tiempo en una representación frecuencial de la misma señal. Mediante la conversión en una representación frecuencial, las
- 55

técnicas de la DCT se ha demostrado que permiten muy elevados niveles de compresión, en cuanto unos cuantificadores pueden ser diseñados para aprovechar las características de distribución de frecuencias de una imagen. En una forma de realización, DCT de 16 x 16 es aplicada en una primera ordenación, cuatro DCTs de 8 x 8 son aplicadas en una segunda ordenación. 16 DCTs de 4 x 4 son aplicadas en una tercera ordenación y 64 DCTs de 2 x 2 son aplicadas en una cuarta ordenación.

La operación de la DCT reduce la redundancia espacial inherente en la fuente de vídeo. Después de que se ha llevado a cabo la DCT, la mayoría de la energía de las señales de vídeo tiende a concentrarse en unos pocos coeficientes de la DCT. Una transformada adicional, la Transformada de Árbol Cuaternario Diferencial (DQT), puede ser utilizada para reducir la redundancia entre dos coeficientes de la DCT.

Para el bloque de 16 x 16 y para cada subbloque, los valores de los coeficientes de la DCT y el valor de la DQT (si se utiliza la DQT) son analizados para determinar el número de bits requerido para codificar el bloque o el subbloque. A continuación, el bloque o la combinación de subbloques que requiera la codificación del menor número posible de bits, es elegido para representar el segmento de imagen. Por ejemplo, dos subbloques de 8 x 8, seis subbloques de 4 x 4 y ocho subbloques de 2 x 2 pueden ser elegidos para representar el segmento de imagen.

El bloque o combinación de subbloques elegido es, a continuación, adecuadamente dispuesto por orden en un bloque de 16 x 16. Los valores de los coeficientes de las DCT / DQT pueden a continuación ser sometidas a una ponderación de frecuencias, a una codificación y a una cuantificación (como por ejemplo una codificación de longitud variable) como preparación para la transmisión.

Aunque la técnica ABSDCT descrita con anterioridad funciona sorprendentemente bien, es intensiva desde el punto de vista computacional. Así, puede resultar difícil la implementación de la técnica con un hardware compacto. Se desea una técnica alternativa que consiga una implementación del hardware más eficiente. Un procedimiento y un sistema de compresión de imagen que es más eficiente desde el punto de vista computacional se proporciona mediante la presente invención de la manera descrita en las líneas que siguen.

El trabajo "Compresión de Imagen Segmentación de Tamaños de Bloque Variables" ["Image Compression with Variable Block Size Segmentation"], de Vaisey J. et al., IEEE Transactions on Signal Processing, vol. 40, No. 8 agosto de 1992, páginas 2040 - 2060 describe un procedimiento de compresión de imagen mediante la segmentación de una imagen en zonas de tamaños diferentes.

Divulga un procedimiento para la determinación de una asignación de tamaño de bloque para un bloque de entrada de los píxeles de imagen que van a ser utilizados en la compresión del bloque entrante, comprendiendo el procedimiento las etapas de lectura de un bloque de datos de píxeles, la generación de una asignación de tamaño de bloque para los bloques de los datos de píxeles basados en las varianzas de los valores de los píxeles de dichos bloques de datos de píxeles y de los bloques subdivididos de dicho bloque de datos de píxeles, en el que la etapa de generación adicional comprende las etapas de la determinación de una varianza de valores de píxeles para dicho bloque de datos de píxeles en el que la varianza se calcula sobre todos los píxeles del bloque; la comparación de dicha varianza con un umbral; la toma de una decisión para subdividir dicho bloque si dicha varianza es mayor que dicho umbral; si dicha decisión es la subdividir dicho bloque, la repetición entonces de las etapas de la determinación, la comparación y la toma de decisión para cada bloque subdividido hasta que se satisfaga un criterio predeterminado; y la designación como asignación del tamaño del bloque de cada bloque que no se subdivide más; y la provisión de una información que contiene una estructura de datos sobre dicha asignación del tamaño del bloque.

Sumario de la invención

La presente invención se refiere a un procedimiento para la determinación de la asignación del tamaño del bloque para un bloque entrante de píxeles de imagen de acuerdo con lo definido en la reivindicación 1, y a un sistema de compresión de imagen de acuerdo con lo definido en la reivindicación 12.

La presente invención es un sistema y un procedimiento de compresión de imagen que utiliza unos bloques y unos subbloques dimensionados de manera adaptativa de unos datos de coeficientes de la Transformada de Coseno Discreta. En una forma de realización, el bloque de píxeles de 16 x 16 es introducido en un codificador. El codificador comprende un elemento de asignación del tamaño del bloque, el cual segmenta el bloque de píxeles entrante para su procesamiento. La asignación del tamaño del bloque se basa en las varianzas del bloque entrante y de los bloques subdivididos. En general, las áreas con mayores varianzas serán subdivididas en bloques más pequeños, mientras que las áreas con varianzas menores no serán subdivididas, con tal de que los valores medios de los bloques y de los subbloques se incluyan en unos márgenes diferentes predeterminados.

De esta manera, en primera lugar, el umbral de la varianza de un bloque es modificado a partir de su valor nominal, dependiendo de su valor medio y, a continuación, la varianza del bloque es comparada con un umbral, y si la varianza es mayor que el umbral, entonces el bloque se subdivide.

La asignación del tamaño del bloque se proporciona a un elemento de la transformada, el cual transforma los datos de píxeles en datos en el dominio secuencial. La transformada se lleva a cabo solo sobre el bloque y los subbloques

5 seleccionados mediante la asignación del tamaño del bloque. Los datos de la transformada, a continuación, son sometidos a una cuantificación y a una serialización. Por ejemplo, puede ser utilizado un escapeo en zigzag para serializar los datos para producir un flujo de datos. El flujo de datos puede, a continuación, ser codificado mediante un codificador de longitud variable como preparación para la transmisión. Los datos codificados son enviados a través de un canal de transmisión hasta un descodificador, donde los datos de píxeles son reconstruidos como preparación para su representación.

Breve descripción de los dibujos

10 Las características distintivas, los objetivos y las ventajas de la presente invención se pondrán de manifiesto de forma más acabada a partir de la descripción detallada expuesta en las líneas que siguen tomada en combinación con los dibujos, en los cuales los mismos caracteres de referencia identifican los correspondientes elementos a lo largo de todos los dibujos, y en los que:

La FIG. 1 es un diagrama de bloques de un sistema de procesamiento de imagen que incorpora el sistema y el procedimiento de la presente invención de asignación del tamaño del bloque en base a la varianza;

15 la FIG. 2 es un diagrama de flujo que ilustra las etapas de procesamiento implicadas en la asignación del tamaño del bloque en base a la varianza;

las FIGS. 3a, 3b y 3c ilustran una asignación del tamaño del bloque ejemplar, la correspondiente descomposición de árbol cuaternario, y los correspondientes datos de PQR.

Descripción detallada de las formas de realización preferentes

20 Con el fin de facilitar la transmisión digital de señales digitales y disfrutar de los correspondientes beneficios, es necesario, en términos generales, emplear alguna forma de compresión de la señal. Para conseguir una alta definición en la imagen resultante es, así mismo, de importancia que se mantenga la alta calidad de la imagen. Así mismo, se desea una eficiencia computacional para una implementación con un hardware compacto, lo que es importante en muchas aplicaciones.

25 La presente invención proporciona un sistema o un aparato y un procedimiento de compresión de imagen que toma en cuenta tanto la calidad de la imagen como la eficiencia computacional en la ejecución de la compresión de la imagen. La compresión de la imagen de la presente invención se basa en técnicas de la Transformada de Coseno Discreta (DCT). En general, una imagen que va a ser procesada en el dominio digital estaría compuesta por unos datos de píxeles en un conjunto de bloques no superpuestos, de tamaño $N \times N$. Una DCT de dos dimensiones puede ejecutarse sobre cada bloque. La DCT de dos dimensiones se define mediante la siguiente relación:

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

35 donde

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

40 $x(m, n)$ es la localización de los píxeles (m, n) dentro de un bloque $N \times M$, y

$X(k, l)$ es el correspondiente coeficiente de la DCT.

45 Dado que los valores de los píxeles son no negativos, el componente $x(0, 0)$ de la DCT es siempre positivo, y generalmente tiene la mayor energía. De hecho, para imágenes típicas, la mayoría de la energía de la transformada se concentra alrededor del componente $X(0, 0)$. Esta propiedad de compactación de la energía hace que la técnica de la DCT sea un procedimiento de compresión tan atractivo.

La técnica de compresión de la imagen de la presente invención utiliza una codificación adaptativa de contraste para conseguir una reducción adicional de la tasa de transmisión de bits. Se ha observado que las imágenes más naturales están compuestas por unas áreas planas de variación relativamente lenta, y unas áreas activas como por

ejemplo en los límites del objeto y con una textura de gran contraste. Los esquemas de codificación adaptativos de contraste utilizan este factor mediante la asignación de más bits a las áreas activas y menos bits a las áreas menos activas.

5 La codificación adaptativa de contraste, así mismo, es útil para la reducción del efecto de bloqueo. En la implementación de otras técnicas de codificación de la DCT, el efecto de bloqueo es quizás la perturbación más importante para la calidad de la imagen. Así mismo, el efecto de bloqueo tiende a ser más perceptible en áreas activas de la imagen. Sin embargo, se ha llegado a comprender que el efecto de bloqueo se reduce cuando se utiliza una DCT de menor tamaño. El efecto de bloqueo resulta virtualmente invisible cuando se utiliza una DCT de 2×2 , aunque pueden resentirse las prestaciones de bit por píxel. De esta manera, la codificación adaptativa de contraste
10 puede reducir el efecto de bloqueo mediante la asignación de tamaños de bloque más pequeños (y con ello de más bits) hacia las áreas activas y tamaños de bloque mayores hacia las áreas relativamente vacías.

Otra característica distintiva de la presente invención es que utiliza una codificación intertramas (procesamiento espacial) en lugar de la codificación intertramas (procesamiento espacio temporal). Una razón para la adopción de la codificación intertramas es la gran complejidad del receptor requerida para procesar las señales de codificación intertramas. La codificación intertramas requiere de forma inherente múltiples memorias intermedias de trama además de unos circuitos de procesamiento más complejos. En muchas aplicaciones, se necesita una reducción de la complejidad para la implementación efectiva.
15

Una segunda razón para la utilización de la codificación intertramas es que puede existir una situación, o un material de programa, que pueda provocar que el esquema de codificación espacio - temporal se estropee y lleve a cabo un rendimiento deficiente. Por ejemplo, las películas de 24 tramas por segundo pueden incluirse dentro de esta categoría dado que el tiempo de integración, debido al obturador mecánico es relativamente corto. El corto periodo de tiempo de integración hace posible un grado más alto de solapamiento temporal. La asunción de la correlación de trama a trama se rompe en el caso de un movimiento rápido en cuanto resulta entrecortado.
20

Una razón adicional para la utilización de la codificación intertramas es un esquema de codificación espacio - temporal es más difícil de estandarizar cuando están implicadas frecuencias tanto a 50 Hz como a 60 Hz de la línea de transporte de energía. La televisión en la actualidad transmite señales ya sea a 50 Hz o 60 Hz. El uso de un esquema de intertramas, al ser un sistema digital, puede adaptar tanto el funcionamiento a 50 Hz como a 60 Hz, o incluso para películas de 24 tramas por segundo mediante la transacción entre la tasa de tramas y la resolución espacial.
25

A los fines del procesamiento de imagen, la operación de la DCT se lleva a cabo sobre unos datos de píxeles que son divididos en un conjunto de bloques no superpuestos. Nótese que, aunque los tamaños de los bloques son analizados en la presente memoria teniendo un tamaño de $N \times N$, se contempla que pueden ser utilizados tamaños de bloques diferentes. Por ejemplo, un tamaño de bloque $N \times M$ puede ser utilizado cuando tanto N como M sean números enteros, siendo M o bien mayor o menor que N . Otro importante aspecto es que el bloque es divisible en al menos un nivel de subbloques, como por ejemplo $N/i \times N/j$, $N/i \times N/j$, $N/i \times M/j$, etc., cuando i y j sean números enteros. Así mismo, el tamaño de bloque ejemplar, tal y como se analiza en la presente memoria es un bloque de píxeles de 16×16 con unos correspondientes bloques y subbloques de coeficientes de la DCT. Se contempla, así mismo, que pueden ser utilizados otros distintos números enteros, como por ejemplo tanto valores de números pares como impares, por ejemplo, de 9×9 .
30
35

Con referencia ahora a la FIG. 1, en ella se muestra un sistema 100 de procesamiento de imagen, el cual incorpora el sistema de compresión de la presente invención. El sistema 100 de procesamiento de imagen comprende un codificador 102 que comprime una señal de vídeo recibida. La señal comprimida es transmitida por medio de un canal de transmisión 104, y es recibida por un descodificador 106. El descodificador 106 descodifica la señal recibida en muestras de la imagen, las cuales pueden a continuación ser representadas.
40

En general, una imagen es dividida en bloques de píxeles para su procesamiento. Una señal de color puede ser convertida desde el espacio RGB hasta el espacio $Y C_1 C_2$, siendo Y el componente de luminancia, o de brillo, y siendo C_1 y C_2 los componentes de la crominancia o el color. Debido a la baja sensibilidad espacial del ojo al color, muchos sistemas submuestran los componentes C_1 y C_2 mediante un factor de cuatro en las direcciones horizontal y vertical. Sin embargo, el submuestreo no es necesario. Una imagen de resolución total, conocida como el formato 4:4:4, puede ser o bien muy útil o bien necesario en algunas aplicaciones, como por ejemplo las designadas como englobadoras del "cine digital". Dos posibles representaciones de $Y C_1 C_2$ son la representación $Y I Q$ y la representación $Y U V$, las dos bien conocidas en la técnica. Así mismo, es posible emplear una variante de la representación $Y U V$ conocida como $Y C b C r$.
45
50

En una forma de realización preferente, cada uno de los componentes Y , C_b y C_r es procesado sin submuestreo. De esta manera, se suministra una entrada de un bloque de píxeles de 16×16 al codificador 102. El codificador 102 comprende un elemento 108 de asignación de tamaño de bloque, el cual efectúa la asignación del tamaño del bloque para la preparación para la compresión del vídeo. El elemento 108 de asignación del tamaño del bloque determina la descomposición del bloque de 16×16 en base a las características perceptuales de la imagen del bloque. La asignación del tamaño del bloque subdivide cada bloque de 16×16 en bloques más
55

pequeños en una forma de árbol cuaternario dependiendo de la actividad existente dentro de un bloque de 16 x 16. El elemento 108 de asignación del tamaño del bloque genera unos datos de árbol cuaternario llamados los datos de PQR, cuya longitud puede oscilar entre 1 y 21 bits. De esta manera, si la asignación del tamaño del bloque determina que un bloque de 16 x 16 va a ser dividido, se establece el bit R de los datos de PQR, seguidos por cuatro bits adicionales de datos P correspondientes a los cuatro bloques divididos de 8 x 8. Si la asignación del tamaño del bloque determina que ninguno de los bloques de 8 x 8 va a ser subdividido, entonces se añaden cuatro bits adicionales de los datos Q para cada bloque de 8 x 8.

Con referencia ahora a la FIG. 2, se ofrece un diagrama de flujo que muestra detalles de la interacción del elemento 108 de asignación del tamaño del bloque. El algoritmo utiliza la varianza de un bloque como una métrica en la decisión de subdividir un bloque. Comenzando en la etapa 202, se lee un bloque de píxeles de 16 x 16. En la etapa 204, se computa la varianza, v_{16} , del bloque de 16 x 16. La varianza se computa como sigue:

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

donde $N = 16$, y x_{ij} es el píxel de la fila i ésima, la columna j ésima dentro del bloque $N \times N$. En la etapa 206, en primer lugar, se modifica el umbral de varianza T_{16} para proporcionar un nuevo umbral T'_{16} si el valor medio del bloque se sitúa entre dos valores predeterminados, a continuación la varianza del bloque es comparada con el nuevo umbral, T'_{16} .

Si la varianza v_{16} no es mayor que el umbral T_{16} , entonces, en la etapa 208, se escribe la dirección de comienzo del bloque de 16 x 16, y el bit R de los datos de PQR se establece en 0 para indicar que el bloque de 16 x 16 no está subdividido. El algoritmo a continuación lee el siguiente bloque de píxeles de 16 x 16. Si la varianza v_{16} es mayor que el umbral T_{16} , entonces, en la etapa **210**, el bit R de los datos de PQR se establece en 1, para indicar que el bloque de 16 x 16 va a ser subdividido en cuatro bloques de 8 x 8.

Los cuatro bloques de 8 x 8, $i = 1 : 4$, son considerados de forma secuencial para una división ulterior, tal y como se muestra en la etapa **212**. Para cada bloque de 8 x 8, la varianza, v_{8i} , es computada, en la etapa **214**. En la etapa **216**, el primer umbral de varianza T_8 es modificado para proporcionar un nuevo umbral T'_8 si el valor medio del bloque se sitúa entre dos valores predeterminados, entonces la varianza del bloque se compara con este nuevo umbral.

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 8 x 8, y el correspondiente bit Q, Q_i se fija en 0. El siguiente bloque de 8 x 8 es entonces procesado. Si la varianza v_{8i} es mayor que el umbral T_8 , entonces, en la etapa **220**, el correspondiente bit Q, Q_i , se establece en 1 para indicar que el bloque de 8 x 8 va a ser subdividido en cuatro bloques de 4 x 4.

Los cuatro bloques 4 x 4, $j_i = 1 : 4$, son considerados de forma secuencial para una ulterior división, tal y como se muestra en la etapa **222**. Para cada bloque de 4 x 4, la varianza, v_{4j_i} , es computada, en la etapa **224**. En la etapa **226**, en primer lugar el umbral de varianza T_4 es modificado para proporcionar un nuevo umbral T'_4 si el valor medio del bloque se sitúa entre dos valores predeterminados, a continuación la varianza del bloque es comparada con este nuevo umbral.

Si la varianza v_{4j_i} no es mayor que el umbral T_4 , entonces, en la etapa **228**, se escribe la dirección del bloque de 4 x 4, y el correspondiente bit P, P_{ij} , se establece en 0. El siguiente bloque de 4 x 4 es a continuación procesado. Si la varianza v_{4j_i} es mayor que el umbral T_4 , entonces, en la etapa **230**, el correspondiente bit P, P_{ij} , se establece en 0 para indicar que el bloque de 4 x 4 va a ser subdividido en 4 bloques de 2 x 2. Así mismo, se escribe la dirección de los 4 bloques de 2 x 2.

Los umbrales T_{16} , T_8 y T_4 pueden ser unas constantes predeterminadas. Esto es conocido como decisión dura. Como alternativa, puede ser implementada una decisión adaptativa o suave. La decisión suave varía los umbrales de las varianzas dependiendo del valor medio de los píxeles de los bloques $2N \times 2N$, donde N puede ser 8, 4 o 2. De esta manera, las funciones de los valores medios de los píxeles pueden ser utilizadas como umbrales.

Con fines ilustrativos, considérese el siguiente ejemplo. Supóngase que los umbrales de varianza predeterminados del componente Y sean 50, 1100 y 880, respectivamente, para los bloques de 16 x 16, 8 x 8 y 4 x 4. En otras palabras, $T_{16} = 50$, $T_8 = 1100$ y $T_4 = 880$. Supóngase que el rango de los valores medios sea de 80 y 100. Supóngase que la varianza computada para el bloque de 16 x 16 es 60. Dado que 60 y su valor medio 90 es mayor que T_{16} , el bloque de 16 x 16 es subdividido en cuatro subbloques de 8 x 8. Supóngase que las varianzas computadas para los bloques de 8 x 8 son 1180, 935, 980, y 1210. Dado que dos de los bloques de 8 x 8 presentan unas varianzas que exceden de T_8 , estos dos bloques se dividen a su vez para obtener un total de ocho subbloques de 4 x 4. Finalmente, supóngase que las varianzas de los ocho bloques de 4 x 4 son 620, 630, 670, 610, 590, 525,

930 y 690, con unos correspondientes valores medios de 90, 120, 110, 115. Dado que el valor medio del primer bloque de 4 x 4 cae en el rango (80, 100), su umbral se reducirá hasta $T^{*4} = 200$ que es menor de 880. De esta manera, este bloque de 4 x 4 será subdividido así como el séptimo bloque de 4 x 4. La asignación de tamaños de bloques resultantes se muestra en la FIG. 3a. La correspondiente descomposición de árbol cuaternario se muestra en la FIG. 3b. Así mismo, los datos de PQR generados por esta asignación de tamaños de bloques se muestra en la FIG. 3c.

Nótese que un procedimiento similar es utilizado para asignar tamaños de bloques para los componentes de color C_1 y C_2 . Los componentes de color pueden ser decimados horizontal, verticalmente, o en ambos sentidos.

Así mismo, nótese que, aunque la asignación de tamaños de bloques ha sido descrita como un sistema de arriba abajo, en el cual el bloque de mayor tamaño (de 16 x 16 en el presente ejemplo), es utilizado en primer término, puede utilizarse en su lugar un sistema de abajo arriba. El sistema de abajo arriba evaluará en primer término los bloques más pequeños, (de 2 x 2 en el presente ejemplo).

Con referencia de nuevo a la FIG. 1, se describirá el resto del sistema 110 de procesamiento de imagen. Los datos de PQR, junto con las direcciones de los bloques seleccionados, se suministran a un elemento 110 de DCT de la DCT. El elemento 110 de la DCT utiliza los datos de PQR para llevar a cabo las transformadas de coseno discretas de los tamaños apropiados sobre los bloques seleccionados. Solo los bloques seleccionados necesitan ser sometidos al procesamiento de la DCT.

El sistema 100 de procesamiento de imagen puede de manera opcional comprender el elemento DCT 112 para la reducción de la redundancia entre los coeficientes de DC de las DCTs. Un coeficiente de DC se encuentra en la esquina superior izquierda de cada bloque de la DCT. Los coeficientes de la DCT son, en general, de gran tamaño en comparación con los coeficientes de AC. La discrepancia en cuanto a los tamaños hace difícil diseñar un codificador de longitud variable eficiente. De acuerdo con ello, es ventajoso reducir la redundancia entre los coeficientes DC.

El elemento 112 de la DQT lleva a cabo unas DCTs de 2 - D sobre los coeficientes de DC, tomados 2 x 2 a un tiempo. Empezando por los bloques de 2 x 2 dentro de los bloques de 4 x 4 se lleva a cabo una DCT de 2 - D sobre cuatro coeficientes de DC. Esta DCT de 2 x 2 se denomina la transformada de árbol cuaternario diferencial, o DQT, de cuatro coeficientes de DC. A continuación, el coeficiente de DC de la DQT junto con los tres coeficientes de DC con un bloqueo de 8 x 8 son utilizados para compactar la DQT del nivel siguiente. Finalmente, los coeficientes de DC de los cuatro bloques de 8 x 8 dentro de un bloque de 16 x 16 son utilizados para computar la DQT. De esta manera, en un bloque de 16 x 16, hay un auténtico coeficiente de DC y el resto son coeficientes de AC correspondientes a la DCT y a la DQT.

Los coeficientes de la transformada (tanto de la DCT como de la DQT), son suministrados a un cuantificador para su cuantificación. En una forma de realización preferente, los coeficientes de la DCT son cuantificados utilizando unas máscaras de ponderación de frecuencias (FWMs) y un factor de cuantificación de escala. Una FWM es una tabla de ponderaciones de frecuencias de las mismas dimensiones que el bloque de los coeficientes de entrada de la DCT. Las ponderaciones de frecuencias aplican diferentes ponderaciones a los diferentes coeficientes de la DCT. Las ponderaciones están diseñadas para otorgar importancia a las muestras entrantes que presentan un contenido de frecuencias respecto de las cuales el sistema visual humano es más sensible y para quitar importancia a las muestras que presentan un contenido de frecuencias a las que el sistema visual es menos sensible. Las ponderaciones pueden, así mismo, ser diseñadas en base a factores tales como las distancias de visualización, etc.

Las ponderaciones son seleccionadas en base a datos empíricos. Un procedimiento para diseñar las máscaras de ponderación para los coeficientes de la DCT de 8 x 8 se divulga en el documento ISO / IEC J TCA CD 1 0918 "Compresión y codificación Digital de imágenes fijas de tono continuo - parte 1: Condiciones y directrices" ["Digital Compression and Encoding of Continuous-tone still images - part 1: Requirements and guidelines"] International Standards Organization, 1994. En general. Se diseñan dos FWMs, una para el componente de luminancia y otra para los componentes de crominancia. Las tablas de FWM para los tamaños de bloques de 2 x 2, 4 x 4 se obtienen por decimación y el de 16 x 16 por interpolación para el del bloque de 8 x 8. El factor de escala controla la calidad y la tasa de transmisión de datos de los coeficientes cuantificados.

De esta manera, cada coeficiente de la DCT es cuantificado de acuerdo con la relación:

50

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

5 donde DTC (i, j) es el coeficiente de la DTC entrante, $f_{wm}(i, j)$ es la máscara de ponderación de frecuencias, q es el factor de escala y $DCTq(i, j)$ es el coeficiente cuantificado. Nótese que, dependiendo del signo del coeficiente de la DCT, el primer término situado dentro de los corchetes es redondeado arriba o abajo. Los coeficientes de la DQT son, así mismo, cuantificados utilizando una máscara de ponderación apropiada. Sin embargo, pueden ser utilizadas múltiples tablas o máscaras, y aplicadas a cada uno de los componentes Y, Cb y Cr.

10 Los coeficientes cuantificados son suministrados a un serializador **116** de escaneo en zigzag. El serializador **116** escanea los bloques de los coeficientes cuantificados de una forma en zigzag para producir un flujo serializado de coeficientes cuantificados. Así mismo puede ser escogida una pluralidad de pautas de escaneo en zigzag diferentes, así como de pautas en zigzag distintas. Una técnica preferente, emplea unos tamaños de bloque de 8 x 8 para el escaneo en zigzag aunque pueden ser empleados otros tamaños.

Nótese que el serializador **116** de escaneo en zigzag puede ser situado bien antes o bien después del cuantificador **114**. Los resultados netos son los mismos.

15 En cualquier caso, el flujo de coeficientes cuantificados es suministrado a un codificador **118** de longitud variable. El codificador **118** de longitud variable puede hacer uso de una codificación de ceros de recorrido seguida por una codificación de Huffman. Esta técnica se analiza con detalle en las Patentes estadounidenses anteriormente mencionadas 5,021,891, 5,107,345 y 5,452,104, y ahora se expone de forma resumida en la presente memoria. Un codificador de longitud de recorrido tomaría los coeficientes cuantificados y separaría los coeficientes nulos de los no nulos. Los valores nulos son designados como valores de longitud de recorrido y son codificados mediante el sistema de Huffman. Los valores no nulos son codificados por el sistema Huffman de forma separada.

20 Una codificación de Huffman modificada de los coeficientes cuantificados es, así mismo, posible, y es utilizada en la forma de realización preferente. Aquí, después del escaneo en zigzag, un codificador de longitud de recorrido determinará los pares de longitud de recorrido / tamaño dentro de cada bloque de 8 x 8. Estos pares de longitud de recorrido / tamaño son a continuación codificados por el sistema de Huffman.

25 Los códigos de Huffman están diseñados a partir de estadísticas ya sea medidas o ya sea teóricas de una imagen. Se ha observado que las imágenes más naturales están compuestas por áreas en blanco o por áreas de variación relativamente lenta, y por áreas ocupadas, como por ejemplo los límites de dos objetos y la textura de gran contraste. Los codificadores de Huffman con transformadas en el dominio frecuencial, por ejemplo, la DCT, explotan estas características mediante la asignación de más bits a las áreas más ocupadas y de menos bits a las áreas en blanco. En general, los codificadores de Huffman hacen uso de tablas de consulta para codificar los valores de longitud de recorrido y los no nulos. En general se utilizan múltiples tablas, siendo 3 tablas preferentes en la presente invención, aunque, si se desea, pueden ser empleadas 1 o 2.

30 La señal de imagen comprimida generada por el codificador **102** es transmitida al descodificador **106** por medio del canal de transmisión **104**. Los datos de PQR, los cuales contienen la información de la asignación de los tamaños de los bloques es, así mismo, suministrada al descodificador **106**. El descodificador **106** comprende un descodificador **120** de longitud variable, el cual descodifica los valores de longitud de recorrido y los valores no nulos.

35 La salida del descodificador **120** de longitud variable se suministra a un serializador inverso **122** de escaneo en zigzag que ordena los coeficientes de acuerdo con el esquema de escaneo empleado. El serializador inverso **122** de escaneo en zigzag recibe los datos de PQR para ayudar a la ordenación adecuada de los coeficientes para obtener un bloque de coeficientes compuesto.

40 El bloque compuesto es suministrado a un codificador inverso **124** para deshacer el procesamiento debido al uso de las máscaras de ponderación de frecuencias.

45 El bloque de coeficientes es a continuación suministrado a un elemento **126** de la IDQT, seguido por un elemento **128** de la IDCT, si se ha aplicado la transformada de Árbol Cuaternario Diferencial. En otro caso, el bloque de coeficientes es suministrado directamente al elemento **128** de la IDCT. El elemento **126** de la IDQT y el elemento **128** de la IDCT efectúan una transformación inversa de los coeficientes para producir un bloque de datos de píxeles. Los datos de píxeles, pueden, a continuación, tener que ser interpolados, convertidos en forma de RGB y, a continuación, almacenados para su representación futura.

50 De acuerdo con ello, se presenta un sistema y un procedimiento para la compresión de imagen que lleva a cabo una asignación del tamaño de los bloques en base a una varianza de píxeles. La asignación del tamaño de los bloques en base a la varianza ofrece diversas ventajas. Debido a que la Transformada de Coseno Discreta se lleva a cabo después de que los tamaños de los bloques son determinados, se consigue una computación eficiente. La transformada intensiva de forma computacional solo necesita llevarse a cabo sobre los bloques seleccionados. Así mismo, el procedimiento de selección de bloques es eficiente, en cuanto la varianza de los valores de los píxeles es matemáticamente sencilla de calcular. Otra ventaja adicional de la asignación del tamaño de los bloques en base a la varianza es que está basada de forma perceptual. La varianza de los píxeles es una medida de la actividad existente en un bloque y, proporciona una indicación de la presencia de los bordes, las texturas, etc. Tiende a captar los detalles de un bloque mucho mejor que mediciones tales como la media de los valores de los píxeles. De esta manera, el esquema en base a la varianza de la presente invención asigna bloques más pequeños a zonas con más

bordes y bloques de mayor tamaño a las zonas más planas. Como resultado de ello, puede conseguirse una calidad sobresaliente de las imágenes reconstruidas.

5 Otra ventaja importante adicional es que, dado que la asignación del tamaño de los bloques se efectúa antes de la cuantificación, se obtiene una flexibilidad mayor en el control de la tasa de transmisión de bits y de la calidad. Dado que el umbral de la varianza está adaptado a la media local, se asignan bloques pequeños incluso en áreas relativamente oscuras. Esto preserva los detalles en todas las áreas que están por encima exactamente del umbral de visibilidad perceptible. Así mismo, la compresión de imagen en base a la varianza proporciona una degradación vistosa de la calidad de la imagen cuando el factor de cuantificación de escala se modifica de valores bajos a valores altos, a diferencia de procedimientos tales como el MPEG. Esto es especialmente crítico en aplicaciones tales como en el área del cine digital.

10 Dada la gran demanda de vídeo digital, la piratería es una seria amenaza. El marcaje con agua digital es un requisito importante para disuadir la violación de los derechos de autor y de la pérdida de ingresos. El marcaje con agua se lleva a cabo en áreas de una imagen que son perceptiblemente significativas, la asignación de tamaños de bloques en base a la varianza es un candidato para el marcaje con agua.

REIVINDICACIONES

- 1.- Un procedimiento para la determinación de una asignación de tamaño de bloque para un bloque entrante de píxeles de imagen que va a ser utilizado en la compresión de dicho bloque entrante, que comprende las etapas de:
- la lectura de un bloque de datos de píxeles;
 - 5 la generación de una asignación de tamaño de bloque para el bloque de datos de píxeles en base a las varianzas de los valores de píxeles de dicho bloque de datos de píxeles y de bloques subdivididos de dicho bloque de datos de píxeles, en el que dicha etapa de generación comprende así mismo las etapas de:
 - la determinación de una varianza de valores de píxeles para dicho bloque de datos de píxeles, en el que la varianza se calcula sobre todos los píxeles del bloque;
 - 10 la comparación de dicha varianza con un umbral, en el que dicho umbral es una función del valor de píxel medio del bloque que está siendo evaluado;
 - la toma de una decisión para subdividir dicho bloque si dicha varianza es mayor que dicho umbral;
 - si dicha decisión es la de subdividir dicho bloque, entonces la repetición de las etapas de determinación, comparación y toma de decisión para cada bloque subdividido hasta que se satisfaga un criterio predeterminado; y
 - 15 la designación como dicha asignación de tamaño de bloque a cada bloque que no está más subdividido; y
 - la provisión de una estructura de datos que contenga la información acerca de dicha asignación de tamaño de bloque.
- 20 2.- Un procedimiento para la compresión de un bloque de datos de píxeles de una imagen, que comprende las etapas de:
- la determinación de una asignación de tamaño de bloque para el bloque de datos de píxeles de acuerdo con el procedimiento de la reivindicación 1;
 - 25 la transformación de dichos datos de píxeles de bloques seleccionados de acuerdo con lo indicado por dicha estructura de datos en una representación en el dominio frecuencial;
 - la cuantificación de dichos datos en el dominio frecuencial;
 - el escaneo de dichos datos cuantificados para obtener un flujo serializado de datos; y
 - la codificación de dicho flujo serializado de datos en preparación para una transmisión.
- 3.- El procedimiento de la reivindicación 2, en el que la Transformada de Coseno Discreta se lleva a cabo seguida por una Transformada de Árbol Cuaternario Diferencial durante dicha etapa de transformación.
- 30 4.- El procedimiento de la reivindicación 2, en el que una Transformada de Coseno Discreta se lleva a cabo seguida por una Transformada de Árbol Cuaternario Diferencial durante dicha etapa de transformación.
- 5.- El procedimiento de la reivindicación 2, en el que se lleva a cabo un escaneo en zigzag durante dicha etapa de escaneo.
- 35 6.- El procedimiento de la reivindicación 5, en el que el escaneo en zigzag se lleva a cabo utilizando un tamaño de bloque de 8 x 8.
- 7.- El procedimiento de la reivindicación 2, en el que se lleva a cabo una codificación de Huffman durante dicha etapa de codificación.
- 8.- El procedimiento de la reivindicación 7, en el que dicha codificación de Huffman emplea múltiples tablas de consulta para codificar valores de longitud de recorrido y no nulos.
- 40 9.- El procedimiento de la reivindicación 8, en el que hay tres tablas de consulta.
- 10.- El procedimiento de las reivindicaciones 1 o 2, en el que dicho umbral cambia para cada nivel de subdivisión.
- 11.- El procedimiento de las reivindicaciones 1 o 2, en el que dicho criterio predeterminado para no seguir repitiendo las etapas de determinación, comparación y toma de decisiones, se basa en un tamaño de bloque mínimo preseleccionado de datos de píxeles.
- 45

- 12.- Un sistema de compresión de imagen para comprimir un bloque de datos de píxeles, que comprende:
- 5 un medio de asignación de tamaño de bloque para la selección de dicho bloque o de dichos bloques subdivididos de dicho bloque que va a ser comprimido en base a las varianzas de los valores de píxeles de dicho bloque de datos de píxeles y de bloques subdivididos de dicho bloque de datos de píxeles, en el que dicho medio de asignación de tamaño de bloque:
 - determina una varianza de valores de píxeles para dicho bloque de datos de píxeles, en el que la varianza se calcula sobre todos los píxeles del bloque, compara dicha varianza con un umbral, en el que dicho umbral es una función del valor de píxel medio del bloque que está siendo evaluado;
 - 10 la toma de una decisión para subdividir dicho bloque si dicha varianza es mayor que dicho umbral, si dicha decisión es la de subdividir dicho bloque, a continuación repite la determinación de la varianza, la comparación con un umbral, y la decisión de subdividir para cada bloque subdividido hasta que satisfaga un criterio predeterminado; y
 - designa como dicha asignación de tamaño de bloque cada bloque que no se subdivide más
 - 15 un medio de transformación para la transformación del tamaño de píxeles de dicho bloque seleccionado o de dichos bloques subdivididos en datos en el dominio frecuencial;
 - un medio cuantificado para la cuantificación de dichos datos en el dominio frecuencial;
 - un medio de serializador para escanear dichos datos cuantificados para obtener un flujo serializado de datos; y
 - 20 un medio de codificación de longitud variable para la codificación de dicho flujo de datos serializado en preparación para la transmisión.
- 13.- El sistema de la reivindicación 12, en el que dicho umbral cambia para cada nivel de subdivisión.
- 14.- El sistema de la reivindicación 12, en el que dicho criterio predeterminado para no seguir subdividiendo se basa en un tamaño de bloque mínimo preseleccionado de los datos de píxeles que están siendo conseguidos.
- 25 15.- El sistema de la reivindicación 12, en el que dicho medio de transformación lleva a cabo una Transformada de Coseno Discreta.
- 16.- El sistema de la reivindicación 12, en el que dicho medio de transformación lleva a cabo una Transformada de Coseno Discreta seguida por una Transformada de Árbol Cuaternario Diferencial.
- 17.- El sistema de la reivindicación 12, en el que dicho medio serializador comprende un escaneador en zigzag.
- 30 18.- El sistema de la reivindicación 17, en el que el escaneador en zigzag emplea un tamaño de bloque de 8 x 8 para el escaneo en zigzag.
- 19.- El sistema de la reivindicación 12, en el que dicho medio de codificación de longitud variable comprende un codificador de Huffman,
- 20.- El sistema de la reivindicación 19, en el que dicho codificador de Huffman emplea múltiples tablas de consulta para codificar valores de longitud de recorrido y valores no nulos.
- 35 21.- El sistema de la reivindicación 20, en el que hay tres tablas de consulta.

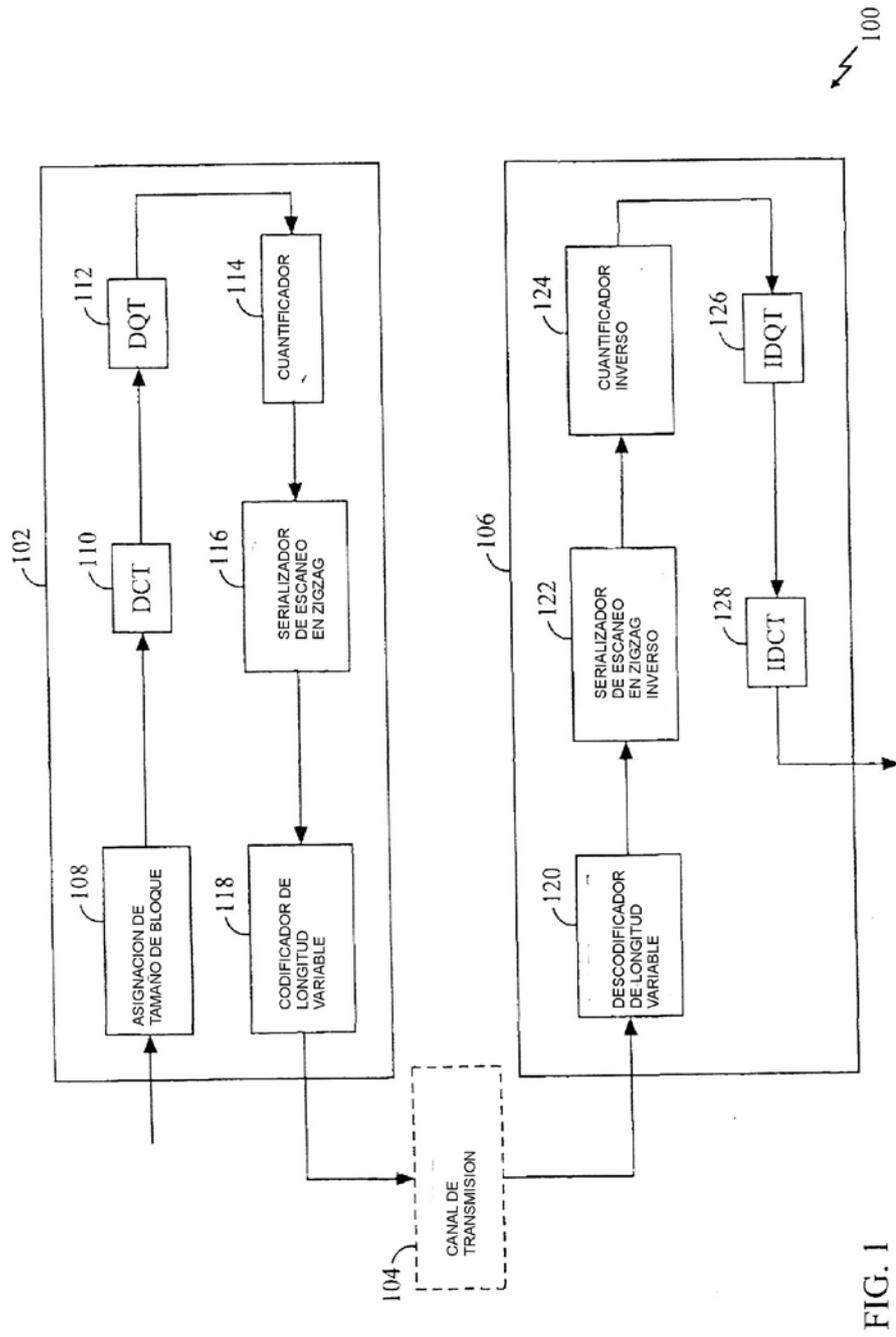


FIG. 1

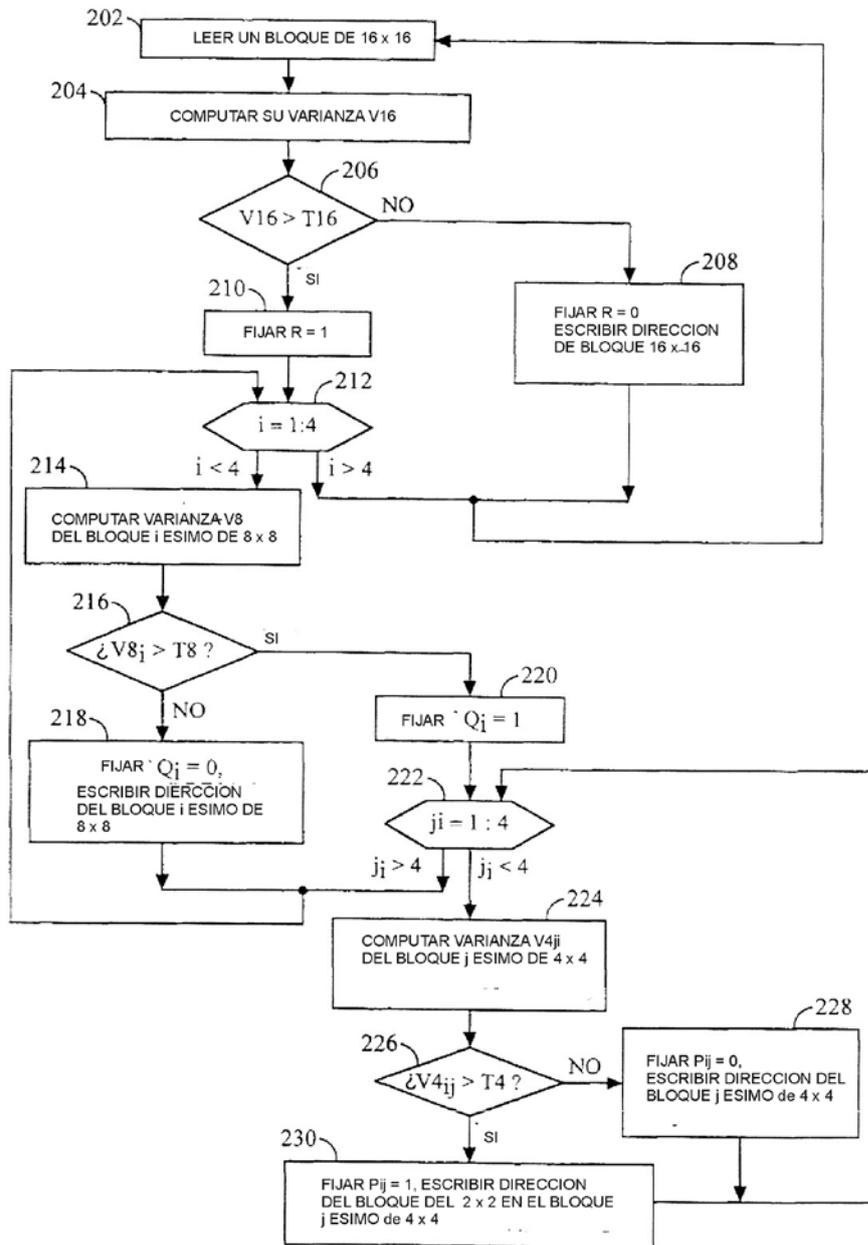


FIG. 2

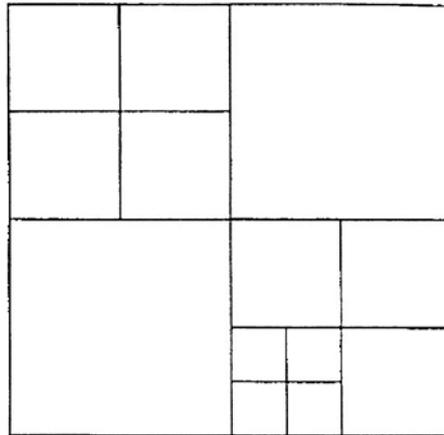


FIG. 3A

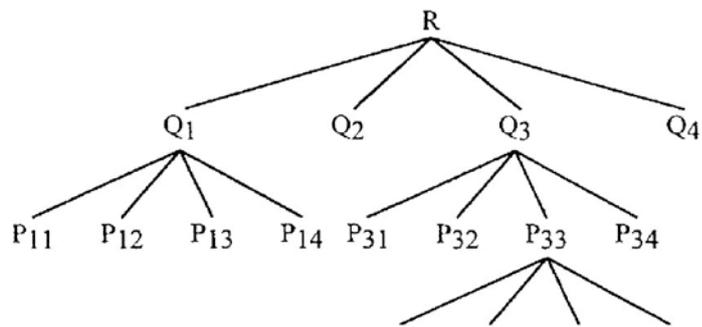
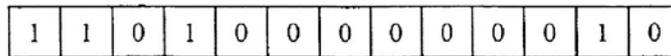


FIG. 3B



DATOS DE PQR

FIG. 3C