

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 379 072**

51 Int. Cl.:
G06K 9/36 (2006.01)
G06K 9/00 (2006.01)
H04B 1/66 (2006.01)
H04N 5/225 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **03772435 .8**
96 Fecha de presentación: **07.11.2003**
97 Número de publicación de la solicitud: **1570413**
97 Fecha de publicación de la solicitud: **07.09.2005**

54 Título: **Procedimiento de seguimiento de la región de interés y dispositivo para la codificación de video basada en ondículas**

30 Prioridad:
12.11.2002 US 293976

45 Fecha de publicación de la mención BOPI:
20.04.2012

45 Fecha de la publicación del folleto de la patente:
20.04.2012

73 Titular/es:
**2011 Intellectual Property Asset Trust
Wells Fargo Delaware Trust Company National
Association 919 North Market Street, Suite 1600
Wilmington DE 19801, US**

72 Inventor/es:
**MILED, Mohamed Khames Ben Hadj y
CHEBIL, Fehmi**

74 Agente/Representante:
Carpintero López, Mario

ES 2 379 072 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

Procedimiento de seguimiento de la región de interés y dispositivo para la codificación de video basada en ondículas

Campo de la invención

5 La presente invención se refiere, de forma general, a codificadores/descodificadores (codecs) de imagen basados en la descomposición en sub-bandas y, más particularmente, a la codificación de objetos en movimiento sobre una serie de imágenes.

Antecedentes de la invención

10 Es generalmente conocido que la compresión de imágenes es eficaz en la reducción de la cantidad de datos de la imagen para su transmisión o almacenamiento. En particular la introducción de formatos de codificación de imágenes escalables como el JPEG 2000, ha hecho posible enviar y recibir sólo una fracción de un fichero de imagen y seguir reconstruyendo una imagen de alta calidad en el extremo de recepción. La parte que se reduce de la imagen usualmente contiene información que describe los componentes de alta frecuencia presentes en la imagen, correspondientes a los detalles a los que el sistema visual humano (HVS) no es muy sensible.

15 JPEG significa el Grupo de Expertos Fotográficos Unidos para la compresión de imágenes. En 1988 este comité adoptó su primera normativa, conocida como la línea base de JPEG, que se basa en la transformada del coseno discreto (DCT) y la codificación Huffman. En 2001, el comité del JPEG desarrolló una nueva normativa de compresión, llamada JPEG 2000. Esta nueva normativa proporciona un funcionamiento de baja tasa de bits, con distorsión de tasa y características de calidad de la imagen subjetiva superiores a las normativas existentes, sin sacrificar el funcionamiento en otros puntos en el espectro de distorsión de tasa. De forma más importante, JPEG 2000 permite la extracción de diferentes resoluciones y fidelidades de píxel de una imagen comprimida a partir de la misma representación de la secuencia de código. También ofrece características tales como la codificación de la región de interés (ROI) y el acceso aleatorio a áreas de imagen. Esto permite a un usuario manipular, almacenar o transmitir sólo la información esencial de una imagen para cualquier dispositivo objetivo a partir de su representación de la secuencia de bits del JPEG 2000.

25 JPEG 2000 es un codificador del plano de bits basado en la descomposición en sub-bandas. Usa ondículas en la etapa de transformación. La imagen se descompone en resoluciones múltiples, Cada una de las resoluciones está compuesta de sub-bandas que representan las componentes de baja o/y alta frecuencia. Las muestras en las sub-bandas se codifican a continuación en planos de bits que comienzan desde el plano de bits más significativo. El uso de la transformada de ondículas y el esquema de codificación del plano de bits proporcionan la característica de escalabilidad de JPEG 2000.

30 El JPEG 2000 de Movimiento es un procedimiento de compresión de video, basado en la codificación sobre intra-tramas usando JPEG 2000. En el JPEG 2000 de Movimiento, las tramas en la secuencia de video se codifican como imágenes independientes, es decir no hay ninguna predicción de movimiento entre las imágenes. Este esquema de codificación ofrece importantes funcionalidades tales como la escalabilidad en la calidad y en la resolución, robustez a los errores de bits, y edición de tramas. Sin embargo, es ineficiente en la realización de la compresión en comparación con otras normativas, tales como el MPEG-4, donde se usan las estimaciones de los vectores de movimiento para codificar las inter-tramas. Por otra parte, la realización de la compresión del JPEG 2000 de Movimiento puede mejorarse usando las características de JPEG 2000 tales como la codificación de la ROI.

35 La codificación de la ROI es una funcionalidad útil en JPEG 2000. Permite la asignación de más bits en una ROI que en otras regiones en la misma imagen mientras se codifican, Codificando de forma desigual las partes de las imágenes de modo que pueden asignarse más bits por píxel a los objetos importantes que a los objetos menos importantes, se obtiene una mejor percepción visual haciendo esta característica muy útil especialmente en las aplicaciones de baja tasa de datos. Para codificar objetos importantes en una secuencia de video como ROI, el seguimiento de estos objetos se hace esencial para producir un flujo de video de alta calidad.

40 El seguimiento de las ROI es una característica importante en muchas aplicaciones relacionadas con lo visual, tales como el control basado en la visión, interfaces hombre-ordenador, vigilancia, automatización agrícola, captura de imágenes médicas y reconstrucción visual. La principal dificultad en el seguimiento de ROI en secuencias de video se debe a las variaciones potenciales en la región objetivo dentro de la secuencia de trama. Estas variaciones se deben usualmente a cambios en la postura, deformaciones de la forma, iluminación y oclusión parcial o total del objeto. Las fuentes de estas variaciones deberían tenerse en cuenta cuando se diseña un algoritmo de seguimiento para garantizar la robustez y la estabilidad.

45 En la técnica anterior, la mayor parte de los procedimientos de seguimiento de las ROI descansan en una estrategia de procesamiento multi-etapa, que consiste de tres etapas consecutivas. En la primera etapa, se realiza la segmentación del segundo plano basado en el color y se enmascaran las áreas de color similar que la región objetivo. La segunda etapa involucra la localización de la ROI a través de la minimización de vectores estimados, usando las proyecciones de resumen horizontal y vertical de la imagen segmentada, o con heurísticas basadas en el tamaño, la forma, la posición, la relación de aspecto y la consistencia de color en la región objetivo. La etapa final es

refinar y suavizar las fronteras de la región.

5 Hager y otros ("Seguimiento Eficaz de la Región con Modelos Paramétricos de Geometría e Iluminación", IEEE Transactions sobre el Análisis de Patrones y la Inteligencia de Máquinas, Vol. 20, Nº 10, Octubre de 1998) hace uso de modelos lineales para simplificar la ecuación del flujo óptico, y un conjunto de vectores base para modelar la variación en la iluminación realizando la descomposición de un valor singular sobre una secuencia de entrenamiento del objetivo. Kruger y otros ("Procedimiento del Subespacio de Ondículas para el Seguimiento de Caras en Tiempo Real". Procedimiento de Reconocimiento de Patrones, 23º Simposio DAGM, Munich, Alemania 2002) usa un número mayor de funciones de ondícula para construir una red de ondículas que representan las características del objetivo, y a continuación sigue la ROI en base a la distancia Euclídea entre los coeficientes de la ondícula.

10 Kundu (Patente de los Estados Unidos Nº 5.974.192) desvela un procedimiento para emparejar bloques en una secuencia de imágenes basado en las texturas, los píxel dominantes de borde así como la intensidad y el color. Ponticos (Patente de los Estados Unidos Nº 6.035.067) desvela un procedimiento donde se usa un "hueco potencial" para impedir los movimientos grandes del objeto del segundo plano usando criterios predefinidos para la clasificación de las regiones de la imagen. Poggio y otros (Patente de los Estados Unidos Nº 6.421.463) desvela un procedimiento para la búsqueda de objetos en imágenes donde el sistema de búsqueda debe entrenarse usando una plantilla de coeficientes de ondículas.

El documento EP 1.061 748 A2 desvela un sistema basado en plantillas para el seguimiento de un objeto de imagen y la segmentación. El sistema entre otros incluye un procedimiento de búsqueda auto-predictivo de correlación de dos dimensiones para el seguimiento de objetos. El objeto de una trama anterior sirve como una plantilla para la localización del objeto en una trama de la imagen actual. La trama de la imagen actual sirve como un área de búsqueda. La plantilla se superpone sobre una ventana dentro del área de búsqueda. Los puntos de los datos de la plantilla se comparan con los puntos de datos de la ventana para determinar si los puntos de datos correlacionan en el grado deseado. Si lo hacen, entonces se ha encontrado una coincidencia para la plantilla. Dependiendo del tamaño de la plantilla y el tamaño del área de búsqueda, la plantilla puede situarse sobre un cierto número de ventanas dentro del área de búsqueda. Para reducir el número de ventanas al que se compara la plantilla, se deducen un tamaño de paso a lo largo de un primer eje y un tamaño de paso a lo largo de un segundo eje. En lugar de comparar la plantilla con cada ventana posible del área de búsqueda, la plantilla se mueve a lo largo de cualquiera o ambos ejes, el primer eje y el segundo eje por el correspondiente tamaño de paso del primer eje o el tamaño del paso del segundo eje. La plantilla se compara con las diversas ventanas del área de búsqueda. La comparación puede ser una función de correlación de la plantilla y la ventana respectiva y resulta un coeficiente de correlación. Cualquier ventana en la cual el coeficiente de correlación con la plantilla se encuentre que excede un valor específico es una coincidencia local para la plantilla. A continuación se realiza una búsqueda total en la vecindad de cualquier localización que es una coincidencia local, una búsqueda total de tal vecindad abarca realizar una correlación entre la plantilla y cada ventana del área de búsqueda potencial entre la ventana de localización de coincidencia local y las ventanas en las etapas anterior y siguiente en cada uno de los ejes horizontal y vertical. Cualesquiera localizaciones entre las localizaciones de coincidencia local y las localizaciones comprobadas durante la búsqueda total de la vecindad que exceden el valor umbral se consideran coincidencias de plantillas. En algunas realizaciones, sólo la localización que tiene la más alta correlación se considera una coincidencia. En otras realizaciones puede haber múltiples coincidencias. De este modo, las coincidencias superiores o todas las coincidencias por encima del umbral se seleccionan como coincidencias resultantes.

La principal desventaja de los procedimientos de la técnica anterior es una complejidad de cálculo. Además, algunos de ellos requieren un algoritmo de seguimiento para entrenarse para una ROI del mismo tipo, color o forma.

De este modo, es ventajoso y deseable proporcionar un procedimiento y un dispositivo para una forma arbitraria y el seguimiento del tipo de ROI con una baja complejidad de cálculo y un coste de la memoria adecuado para los codec JPEG 2000 de Movimiento.

Sumario de la invención

Es un objetivo principal de la presente invención proporcionar un procedimiento y un dispositivo para el seguimiento de al menos una porción de un objeto en una secuencia de imágenes, en el que las imágenes se codifican como imágenes individuales y no se requieren plantillas y objetos pre-entrenados para el seguimiento. El objetivo puede conseguirse combinando la detección de fronteras en la banda de baja frecuencia por píxel coincidente en el espacio de crominancia y el refinamiento de la frontera detectada en la banda de alta frecuencia por el análisis de bordes usando la componente de luminancia.

De este modo, de acuerdo con un primer aspecto de la presente invención, se proporciona un procedimiento para el seguimiento de una región objetivo en una trama de imagen de acuerdo con la reivindicación 1.

55 Por consiguiente, al menos una parte de la región objetivo en la trama anterior tiene un contorno para la definición de un contorno correspondiente en dicha trama de imagen, y los primeros píxel incluyen los píxel adyacentes al contorno correspondiente en dicha trama de imagen.

Los valores del primer y segundo píxel son indicativos de al menos una de las componentes de crominancia de los coeficientes de la ondícula en una sub-banda baja.

5 La región objetivo de dicha trama de imagen incluye una frontera y dicha pluralidad de píxel en dicha trama de imagen incluye una pluralidad de terceros píxel adyacentes a la frontera, teniendo cada uno de los terceros píxel al menos un valor correspondiente de tercer píxel. El procedimiento está caracterizado además por

la determinación del tipo de borde de los píxel terceros de cómo modificar la región objetivo en dicha trama de imagen en base al tipo de borde de los píxel terceros.

Los valores de los píxel terceros son indicativos de la componente de luminancia de los coeficientes de la ondícula en una sub-banda alta.

10 De acuerdo con un segundo aspecto de la presente invención se proporciona un programa de ordenador para su uso en un codificador de imagen de acuerdo con la reivindicación 11.

Por consiguiente, los valores de los píxel primero y segundo son indicativos de al menos una de las componentes de la crominancia de los coeficientes de la ondícula en una sub-banda baja.

15 Ventajosamente, la región objetivo de la primera trama de imagen incluye una frontera y dicha pluralidad de píxel en la primera trama de imagen incluye una pluralidad de terceros píxel adyacentes a la frontera, teniendo cada uno de los terceros píxel al menos un valor de tercer píxel correspondiente. El programa de ordenador está caracterizado además por

un código para la determinación del tipo de borde de los terceros píxel de cómo modificar la región objetivo en la primera trama de imagen en base al tipo de borde de los terceros píxel.

20 Por consiguiente, los valores de los terceros píxel son indicativos de la componente de luminancia de los coeficientes de la ondícula en una sub-banda alta.

De acuerdo con un tercer aspecto de la invención, se proporciona un codificador de imágenes para la codificación de una secuencia de imágenes de acuerdo con la reivindicación 16.

25 El codificador de imágenes está adaptado para codificar dicha región objetivo en la primera trama de imagen con una calidad visual más alta que otra región en dicha primera trama de imagen.

De acuerdo con el cuarto aspecto de la presente invención, se proporciona un sistema de codificación de imágenes de acuerdo con la reivindicación 21.

La presente invención será evidente con la lectura de la descripción tomada conjuntamente con las Figuras 1 a 5.

Breve descripción de los dibujos

30 La Figura 1 es una representación esquemática que ilustra la proyección de la frontera de una ROI en una trama de imagen anterior a la trama de imagen actual.

La Figura 2 es una representación esquemática que ilustra los píxel en un área que encierra la frontera de la ROI proyectada en la trama de imagen actual que se usa para propósitos de emparejamiento de píxel.

35 La Figura 3a es una representación esquemática que ilustra una ROI estimada en la trama de imagen actual.

La Figura 3b es una representación esquemática que ilustra la detección de bordes horizontal y vertical.

La Figura 4 es un diagrama de flujo que ilustra el procedimiento de búsqueda de la ROI, de acuerdo con la presente invención.

La Figura 5 es un diagrama de bloques que ilustra un sistema de codificación de imágenes capaz del seguimiento de la ROI, de acuerdo con la presente invención.

Mejor modo de realización de la invención

El procedimiento de seguimiento de ROI, de acuerdo con la presente invención, está diseñado para secuencias a comprimir por el JPEG 2000 de Movimiento. El procedimiento consiste en dos etapas, realizadas en las sub-bandas de baja resolución en el dominio de las ondículas. En la primera etapa, se realiza una estimación de la región objetivo en una trama actual emparejando los píxel hacia atrás en base a la correspondiente región objetivo en la trama anterior. El uso de las sub-bandas de baja frecuencia reduce la complejidad de cálculo y también introduce estabilidad al seguimiento. El emparejamiento de píxel hacia atrás elimina la necesidad de asumir un modelo para el movimiento. Preferiblemente, la primera etapa se realiza usando las componentes de crominancia o cromática (U, V) para el seguimiento. A diferencia de la luminancia, las componentes de crominancia son usualmente estables sobre la secuencia de video y tienen variaciones pequeñas. Esto hace más robusto al algoritmo de seguimiento frente a cambios en la iluminación, eliminando por lo tanto la necesidad de secuencias de entrenamiento de la estimación de la iluminación incidente. En la segunda etapa, las sub-bandas de alta frecuencia (es decir, HL y LH) de la componente de luminancia Y se usan para realizar una detección de bordes para refinar la región objetivo estimada en la trama actual. El uso de la componente Y posibilita la detección robusta de los bordes incluso cuando la región objetivo se pone en contacto con otra región del mismo tipo de color.

55 La Figura 1 ilustra la primera etapa en la búsqueda de ROI, de acuerdo con la presente invención, para obtener una frontera de ROI estimada en una trama de imagen actual k . Como se muestra en la Figura 1, el contorno de una ROI

20 en una trama de imagen ($k - 1$) se proyecta sobre la trama de imagen k . La trama ($k - 1$) es la trama anterior. El contorno proyectado 30 en la trama k se muestra como un círculo discontinuo. Una banda de píxel limitada por una frontera exterior 32 y una frontera interior 34 se usa como un área de búsqueda 130 en la trama actual k . El ancho de la banda de píxel 130 es $2d$. Como se muestra en la Figura 2, el área de búsqueda 130 representa el espacio de movimiento del contorno de la ROI y d es una frontera superior sobre cualquier vector de desplazamiento. Sólo los píxel dentro del área de búsqueda 130, o los píxel "relevantes" se emparejan hacia atrás en base a los coeficientes de la ondícula de paso bajo de las componentes de crominancia. Debería observarse que la frontera de la ROI 20, representada por un círculo, es sólo para propósitos de ilustración. En general, la frontera o contorno de un ROI es irregular en la forma, y el área de búsqueda 130 se ocupa por un grupo de píxel "relevante" que encierra el contorno de ROI proyectado.

El emparejamiento de píxel hacia atrás consiste en el mapeo de cada píxel "relevante" con otro píxel en la trama anterior ($k - 1$). El criterio de mapeo es la distancia Euclídea entre los coeficientes de la ondícula de paso bajo de las componentes de crominancia (U, V). Esta métrica se minimiza a través de una búsqueda local dentro de una pequeña ventana. Como se muestra en la Figura 2, un píxel 40 en la trama k se usa para encontrar una coincidencia dentro de la ventana de búsqueda 50 que rodea un píxel correspondiente en la trama ($k - 1$). El área de la ventana de búsqueda está dada por $(2r + 1)^2$, donde r es un número entero mayor o igual que d . Si el punto de imagen 40 se expresa como $x_k(i, j)$, entonces la ventana de búsqueda 50 incluye un conjunto de puntos de píxel $x_{k-1}(p, q)$ donde $(i - r) \leq p \leq (i + r)$, $(j - r) \leq q \leq (j + r)$. Para cada píxel en el área de búsqueda 130 en la trama k , se realiza una búsqueda hacia atrás para encontrar un píxel óptimo correspondiente x_{opt} en la trama ($k - 1$) dentro de la ventana de búsqueda 50. En particular, si w_k es una banda de paso bajo de la trama k y w_{k-1} es la misma banda de paso bajo de la trama ($k - 1$), entonces el píxel óptimo x_{opt} se define como

$$X_{opt} = \min_{p, q} | w_k(i, j) - w_{k-1}(p, q) | \quad (1)$$

Como tal, la ventana de búsqueda 50 es un conjunto de coincidencias posibles, y el emparejamiento se consigue a través de la minimización de las distancias Euclídeas entre $w_k(i, j)$ y $w_{k-1}(p, q)$ en términos de componentes de crominancia. Ignorar la componente de luminancia Y hace el algoritmo más robusto frente a los cambios en la iluminación. Debería observarse que las componentes $w_k(i, j)$ y $w_{k-1}(p, q)$ están relacionados con los coeficientes de ondículas en el espacio de la crominancia. Si indicamos a los coeficientes de la ondícula en el espacio U y en el espacio V como u y v , entonces w puede ser u o v , o preferiblemente, un vector de 2D que tiene como valores (u, v) .

La siguiente etapa en la búsqueda de ROI es averiguar el estatus de la región de cada uno de los píxel $x_k(i, j)$ - si pertenece al segundo plano (no ROI) de la imagen o a la ROI. El estatus de la región del píxel $x_k(i, j)$ depende del estatus de la región del píxel correspondiente x_{opt} que es conocido. De este modo, si el píxel óptimo en la trama ($k - 1$) está en el ROI, el píxel correspondiente $x_k(i, j)$ también se establece que está en la ROI. De forma similar, si el píxel óptimo en la trama ($k - 1$) está en el segundo plano, el píxel correspondiente $x_k(i, j)$ se establece que también es un píxel del segundo plano. Debería observarse que la ROI de la primera trama en una secuencia debe alimentarse al algoritmo de búsqueda, y el estatus de la región de cada uno de los píxel en el interior y el exterior de la ROI de la primera trama puede marcarse consecuentemente.

Los píxel $x_k(i, j)$ dentro del área de búsqueda 130 que tienen el estatus del segundo plano se usarán para formar una máscara para definir una ROI estimada en la trama k . La Figura 3a es una representación esquemática de una ROI estimada en la trama k después del enmascaramiento.

Preferiblemente, un límite superior especificado para la distancia Euclídea mínima se usa para determinar si el emparejamiento es satisfactorio. El límite superior especificado puede fijarse en base al promedio de la intensidad del área, por ejemplo, si x_{opt} en la ecuación 1 es mayor que el límite superior especificado, entonces el estado de la región del píxel $x_k(i, j)$ correspondiente a ese x_{opt} se fijará como segundo plano.

La siguiente etapa es un procedimiento de sintonización, que se usa para refinar la ROI en la trama k en base a los píxel de la frontera de la ROI estimada. El propósito del procedimiento de sintonización es detectar cualesquiera bordes a través de una búsqueda local alrededor de estos puntos píxel de frontera para determinar la ROI final en la trama k . El dominio de búsqueda para el procedimiento de sintonización, de acuerdo con la presente invención, consiste de los coeficientes de las sub-bandas HL y LH en el espacio de la luminancia, lo que permite la extracción de los bordes vertical y horizontal, respectivamente (Figura 3b). El uso de la componente de luminancia Y posibilita la detección robusta de los bordes incluso cuando la región objetivo se pone en contacto con otra región del mismo tipo de color. El píxel con una mayor magnitud es probable que sea un píxel de borde. Si ocurre que el píxel de borde está en el segundo plano, entonces sus coeficientes de color se comprobarán antes de incluirlos en la ROI. Debe tenerse cuidado de no incluir un píxel de un color diferente en la región de contorno de la ROI. Tal píxel podría conducir a una propagación de errores cuando se realiza el seguimiento para toda la secuencia de imágenes.

55 Algoritmo

Después de proyectar el contorno de la ROI de la trama ($k - 1$) para la trama k , la búsqueda de la ROI en la trama k se realiza de acuerdo con el siguiente algoritmo:

Bucle sobre todos los píxel "relevantes" $x_k(i, j)$ dentro del área de búsqueda en la trama k ;

Definir una ventana de búsqueda en la trama ($k - 1$) para cada $x_k(i, j)$;

Bucle sobre todos los píxel de la ventana $x_{k-1}(p, q)$;

Calcular las distancias Euclídeas $|w_k(i, j) - w_{k-1}(p, q)|$;

Elegir la distancia Euclídea mínima para identificar x_{opt} ;

Determinar el estado de la región de $x_k(i, j)$ de acuerdo con x_{opt} ;

5 **Establecer la frontera de la ROI estimada en base al estado de la región de $x_k(i, j)$;**

Bucle sobre los píxel de la frontera de la ROI estimada;

Clasificar el tipo de frontera;

Comparar los coeficientes de alta frecuencia y comprobar los coeficientes de color para

decidir sobre el píxel de borde.

10 Debería observarse que las operaciones aritméticas usadas en el algoritmo descrito anteriormente son adiciones y comparaciones. No se requiere ninguna multiplicación. Dado el hecho de que el codificador ya calculó los coeficientes de la ondícula, la complejidad más alta en el algoritmo son los bucles de búsqueda. Debido a que el seguimiento se realiza a baja resolución, y el objeto objetivo generalmente ocupa sólo una pequeña parte de la imagen (requerida para la eficacia de la compresión), el número total promedio de bucles es también pequeño.

15 El procedimiento de seguimiento de una región objetivo dentro de una trama de imagen en una secuencia de tramas de imagen, de acuerdo con la presente invención, se resume en el diagrama de flujo **500**, como se muestra en la Figura 4.

20 Como se muestra en la Figura 4, el contorno de la ROI en la trama (**k-1**) se proyecta sobre la trama **k** en la etapa **510**. Un área de búsqueda que incluye una pluralidad de píxel en la trama **k** que encierra el contorno de la ROI proyectada se establece en la etapa **512**. Para cada uno de los píxel $x_k(i, j)$ dentro del área de búsqueda, se determina un píxel óptimo $x_{k-1}(p, q)$ dentro de una ventana de búsqueda en la trama (**k-1**) en base a la distancia Euclídea en la etapa **514**. El estado de la región de $x_k(i, j)$ se determina en base al píxel óptimo correspondiente en la etapa **516**. Si se determina el estado de la región de cada $x_k(i, j)$ dentro del área de búsqueda como se decide en la etapa **518**, a continuación se establece una ROI estimada en la trama **k** en base al estatus de la región de $x_k(i, j)$ en la etapa **520**. Usando las sub-bandas de alta frecuencia en el espacio de la luminancia para detectar los píxel de borde alrededor de la frontera de la ROI estimada para refinar la ROI en la trama **k** se realiza en la etapa **522**. Si se decide en la etapa **524** que la trama **k** actual es la última trama de la secuencia, a continuación el procedimiento continúa en la siguiente secuencia. De otro modo, el procedimiento vuelve a la etapa **510** para determinar la ROI en la siguiente trama.

30 Un sistema de codificación de imágenes que es capaz de realizar el procedimiento de seguimiento de la región objetivo, de acuerdo con la presente invención, se muestra en la Figura 5. El sistema de codificación de imágenes **600** comprende un codificador **610** y un decodificador **630**. Como se muestra en la Figura 5, una secuencia de imágenes **700** consistente de una pluralidad de tramas de imagen digitales, cada una separable en tres planos de color (R, G, B) se proporciona al módulo convertidor **612**. El convertidor **612** convierte las tramas de imagen digitales en una serie de tramas de la imagen digital **710** en formato YUV. Un módulo de transformada directa de las ondículas **614** se usa para transformar los píxel en cada una de las tramas de imagen **710** en los coeficientes de ondícula en diversas sub-bandas. Los datos de la imagen transformados **720** se proporcionan a un módulo de asignación de multi-resolución **616** donde se asignan más bits por píxel a las regiones importantes de la imagen dentro de una trama. Preferiblemente, la frontera de cada una de las regiones importantes de la imagen en la primera trama se proporciona a un módulo de seguimiento de la ROI **618** de modo que la región importante de la imagen en las tramas posteriores puede seguirse automáticamente. El módulo de seguimiento de la ROI **618** usa el algoritmo de seguimiento de la ROI, de acuerdo con la presente invención, para determinar las regiones importantes de la imagen en las tramas posteriores. Los datos de imagen de multi-resolución **730** se dirigen a un módulo de codificación **620** para la codificación y compresión. La secuencia de código **740** puede transmitirse en tiempo real o almacenarse para su uso futuro. En el extremo receptor, el decodificador **630** reconstruye la secuencia de imagen de la entrada **700** en base a la secuencia de código **740**. Como se muestra, se usa un módulo de decodificación **632** para reconstruir los datos transformados en base a la secuencia de código **740**. Los datos transformados reconstruidos **750** se transforman por el módulo de la transformada inversa de ondículas **634** en una serie de tramas de imagen digital **760** en el espacio YUV. Después de convertirse por el módulo de conversión **636**, los datos de imagen se convierten en una serie de imágenes digitales **770** en tres planos de color (R, G, B).

Resultados de la simulación

El procedimiento de seguimiento de la ROI, de acuerdo con la presente invención, se ha usado sobre tres secuencias de imagen diferentes: "MOBILE", "FOREMAN" e "IRENE". Todas las secuencias tienen un tamaño CIF (Formato Intermedio Común) y una tasa de 30 tramas por segundo. Para operar a tasas más bajas, se saltan 3 tramas en cada iteración de seguimiento en las secuencias "MOBILE" y "FOREMAN", y el seguimiento se realiza a una tasa de 7,5 tramas por segundo. En la secuencia "IRENE" se mantiene la tasa de trama original.

En la secuencia "MOBILE", la ROI es una bola roja con pequeñas regiones blancas alrededor. La bola se somete a movimientos de rotación y de traslación, hacia atrás y adelante, y se pone en contacto con otras regiones de colores similares.

5 En la secuencia "FOREMAN", la cara humana que se sigue cambia las posturas y posiciones. Al final de la secuencia, se produce un cambio abrupto de escena y la región objetivo se ocluye totalmente.

En la secuencia "IRENE", se eligen dos ROI, cada una sobre una mano diferente del locutor. Las manos muestran una variación relativamente rápida de posiciones, posturas y colores. Las manos también se solapan con otras regiones de la piel, tales como la región del cuello, lo que hace el seguimiento incluso más difícil.

10 En la simulación, la ROI se alimenta al algoritmo, en la primera trama. En cada iteración, sólo se usa la transformada de ondículas de la trama anterior. La búsqueda se realiza en la segunda resolución (sub-banda LL) y la frontera superior d se elige para que sea igual a 6 píxel, mientras que el tamaño de la ventana de búsqueda es de 16 por 16 píxel ($r = 8$). La frontera superior sobre la distancia Euclídea mínima es 3 en las secuencias "MOBILE" Y "FOREMAN". En la secuencia "IRENE", la frontera superior no se fija para compensar los cambios en el color en la ROI.

15 En la secuencia "MOBILE", la dificultad principal es la presencia de muchas regiones de segundo plano con colores similares al objetivo. Algunas de estas regiones no aparecen inicialmente en la escena (inclusión total o parcial). En la primera trama, la región objetivo no incluye la bola completa, sino que sólo aparece parte de la bola. Los resultados de la simulación muestran que el algoritmo de seguimiento es capaz de mantener el seguimiento de la bola a través de la secuencia de video. Sin embargo, a medida que la iteración progresa, se incluyen algunas de las
20 regiones del segundo plano en la ROI. Las regiones incluidas están muy cerca de la región objetivo en colores y posiciones y no están inicialmente en la escena.

En la secuencia "FOREMAN", el algoritmo muestra su capacidad de seguimiento de la cara a pesar de los cambios en las posturas faciales en movimiento relativamente rápido durante algunos periodos de la secuencia. A medida que progresa la iteración, el algoritmo incluye parte del cuello en la ROI. La región incluida se conecta con la cara sin un borde de separación. Una parte de la pared también se incluye en la ROI. La región incluida se conecta a la cara sin un borde de separación. Una parte de la pared también se incluye en la ROI debido a la similitud en las componentes de color. Los resultados de la simulación también muestran la capacidad del algoritmo para detectar la oclusión completa de la región objetivo desde la escena.

30 En la secuencia "IRENE" ambas manos del locutor se siguen a una tasa de 30 muestras por segundo. Tal región es muy difícil de seguir debido a las variaciones significativas en las posiciones, las posturas (volteando los lados y diferentes posiciones de los dedos) y los colores. Para superar el problema del cambio en el color, la frontera superior de la distancia mínima Euclídea se relaja. En algunas tramas, ambas manos voltearon las caras. Sin embargo, el algoritmo es aún capaz de seguir la ROI, perdiendo sólo una parte de la mano derecha. En algunas tramas, se consiguió un buen funcionamiento del seguimiento a pesar del solapamiento con el cuello y la cara.
35 Cuando las manos se alejan de otras regiones de la piel, se pierde una parte significativa de la mano derecha asociada con el segundo plano. Finalmente, toda la ROI sobre la mano derecha desaparece completamente. Esto principalmente es debido a que la mano derecha se mueve relativamente rápida y de este modo es más difícil de seguir. Debería observarse que es posible incluir una característica de retroalimentación en el algoritmo de seguimiento para impedir la desaparición de la ROI. Por ejemplo, el área de búsqueda y la ventana de búsqueda pueden aumentarse si la ROI se hace progresivamente más pequeña sobre un cierto número de tramas.

La principal ventaja del procedimiento de seguimiento de la región objetivo, de acuerdo con la presente invención, es que no descansa en información anterior respecto a la región de interés. Por ejemplo, en el seguimiento de la cara humana, el algoritmo no descansa en características específicas tales como los valores de color, y las singularidades de los ojos y la boca. Usualmente, se requiere información específica fuera de línea a través de un
45 entrenamiento del objetivo. El algoritmo de seguimiento de acuerdo con la presente invención, es adecuado para una región no específica de la región de interés a especificar por el usuario.

El seguimiento de la ROI es una funcionalidad útil en un codec de imagen, tal como el JPEG 2000. Permite la asignación de más bits en una región objetivo que en otras regiones en la trama de imagen mientras se codifica la trama de imagen. Como tal, puede conseguirse una alta calidad visual alrededor de la región objetivo.

50 Aunque la invención se ha descrito con respecto a la realización preferida de la misma, se entenderá por los especialistas en la técnica que los cambios anteriores y otros diversos cambios, omisiones y desviaciones en la forma y detalle de la misma pueden realizarse sin apartarse del alcance de las reivindicaciones.

REIVINDICACIONES

1. Un procedimiento de seguimiento de una región objetivo en una trama de imagen (k) en base a una región objetivo (20) de una trama de imagen anterior (k - 1) en una secuencia de tramas de imagen, comprendiendo cada una de dichas secuencias de tramas de imagen una pluralidad de píxeles, comprendiendo dicho procedimiento:
- 5 determinar un área de búsqueda (130) en dicha trama de imagen (k) en base a, al menos una parte de la región objetivo (20) en dicha trama anterior (k - 1), siendo los píxel de dicha área de búsqueda (130) una pluralidad de primeros píxel (40), teniendo cada uno de los primeros píxel al menos un valor de primer píxel correspondiente; estando dicho procedimiento **caracterizado por** para cada uno de los primeros píxel (40) en el área de búsqueda (130):
- 10 determinar un área de búsqueda adicional (50) en dicha trama anterior (k - 1), incluyendo dicha área de búsqueda adicional (50) una pluralidad de segundos píxel de entre la pluralidad de píxel en la trama anterior (k - 1), teniendo cada uno de los segundos píxel al menos un valor de segundo píxel correspondiente y un estatus de región, en el que el estatus de región del segundo píxel es indicativo de si dicho segundo píxel está localizado dentro de la región objetivo (20) en dicha trama anterior (k - 1);
- 15 encontrar una coincidencia entre el valor de primer píxel del primer píxel (40) de entre los valores del segundo píxel para la localización de un segundo píxel de referencia; y
- 20 determinar el estatus de la región de al menos uno de dichos primeros píxel (40) en base al estatus de la región del segundo píxel de referencia para la determinación de la región objetivo en dicha trama de imagen (k) en base al estatus de la región de dicho, al menos un primer píxel (40).
2. El procedimiento de acuerdo con la reivindicación 1, **caracterizado porque** dicha al menos una parte de la región objetivo (20) en dicha trama anterior (k -1) tiene un contorno para la definición de un contorno correspondiente (30) en dicha trama de imagen (k), y en el que los primeros píxel (40) incluyen píxel adyacentes al contorno correspondiente (30) en dicha trama de imagen (k).
- 25 3. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 2, en el que la región objetivo de dicha trama de imagen (k) incluye una frontera y dicha pluralidad de píxel en dicha trama de imagen (k) incluyen una pluralidad de terceros píxel adyacentes a la frontera, teniendo cada uno de los terceros píxel al menos un valor del tercer píxel correspondiente, **caracterizado** dicho procedimiento además por
- 30 la determinación del tipo de borde de los terceros píxel para modificar la región objetivo en dicha trama de imagen (k) en base al tipo de borde de los terceros píxel.
4. El procedimiento de acuerdo con la reivindicación 3, **caracterizado porque** el tipo de borde se determina en base al componente de la luminancia de los coeficientes de ondícula.
5. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 4, **caracterizado porque** los valores del primer y segundo píxel son indicativos de los coeficientes de ondícula.
- 35 6. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 5, **caracterizado porque** los valores del primer y segundo píxel son indicativos de los coeficientes de ondícula en una sub-banda baja.
7. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 5, **caracterizado porque** los valores del primer y segundo píxel son indicativos de al menos uno de los componentes de la crominancia de los coeficientes de ondícula en una sub-banda baja.
- 40 8. El procedimiento de acuerdo con la reivindicación 3 o la reivindicación 4, **caracterizado porque** los valores del tercer píxel son indicativos de los coeficientes de ondícula.
9. El procedimiento de acuerdo con la reivindicación 3 o la reivindicación 4, **caracterizado porque** los valores del tercer píxel son indicativos de los coeficientes de ondícula en una sub-banda alta.
- 45 10. El procedimiento de acuerdo con una cualquiera de las reivindicaciones 3, 4, 8 y 9, **caracterizado porque** los valores del tercer píxel son indicativos de la componente de luminancia de los coeficientes de ondícula en una sub-banda alta.
- 50 11. Un programa de ordenador para su uso en un codificador (610) de imagen que tiene un medio (620) para la codificación de una secuencia (700) de las tramas de imagen dentro de una secuencia de código (740), teniendo dicha secuencia (700) de tramas de imagen al menos un primera trama de imagen (k) y una segunda trama de imagen anterior (k - 1), teniendo cada una de las tramas de imagen primera (k) y segunda (k - 1) una pluralidad de píxel, en el que la trama de imagen segunda (k - 1) tiene un región objetivo (20), comprendiendo dicho programa de ordenador:
- un código para definir un área de búsqueda (130) en la primera trama de imagen (k) en base a al menos una parte de la región objetivo (20) en la segunda trama de imagen (k - 1), siendo los píxel de dicha área de

búsqueda (130) una pluralidad de primeros píxel (40), teniendo cada uno de los primeros píxel al menos uno correspondiente al valor del primer píxel; dicho programa de ordenador **caracterizado por** un código para determinar, para cada uno de los primeros píxel (40) en el área de búsqueda (130);

5 un área de búsqueda adicional (50) en la segunda trama de imagen ($k - 1$), incluyendo dicha área de búsqueda adicional (50) una pluralidad de segundos píxel de entre la pluralidad de píxel de la segunda trama de imagen ($k - 1$), teniendo cada uno de los segundos píxel al menos un valor de segundo píxel correspondiente y un estatus de la región, en el que el estatus de la región del segundo píxel es indicativo de si dicho segundo píxel está localizado dentro de la región objetivo (20) en dicha trama anterior ($k - 1$);

10 un segundo píxel de referencia en dicha área de búsqueda adicional (50) basado en una coincidencia entre el valor de primer píxel del primer píxel (40) de entre los valores del segundo píxel; y el estatus de la región de al menos uno de dichos primeros píxel (40) basado en el estatus de la región del segundo píxel de referencia para determinar una región objetivo en la primera trama de imagen (k) en base al estatus de la región de dicho al menos un primer píxel (40).

15 12. El programa de ordenador de acuerdo con la reivindicación 11, **caracterizado porque** los valores del primer y segundo píxel son indicativos de al menos una de las componentes de la crominancia de los coeficientes de ondícula en una sub-banda baja.

20 13. El programa de ordenador de acuerdo con la reivindicación 11 o la reivindicación 12, en el que la región objetivo de la primera trama de imagen (k) incluye una frontera y dicha pluralidad de píxel en la primera trama de imagen incluye una pluralidad de terceros píxel adyacentes a la frontera, teniendo cada uno de los terceros píxel al menos un valor de tercer píxel correspondiente, **caracterizado** dicho programa de ordenador además por

un código para la determinación del tipo de borde de los terceros píxel para modificar la región objetivo en la primera trama de imagen (k) en base al tipo de borde en los terceros píxel.

25 14. El programa de ordenador de acuerdo con la reivindicación 13 **caracterizado porque** los valores del tercer píxel son indicativos de la componente de luminancia de los coeficientes de ondícula en una sub-banda alta.

15. El programa de ordenador de acuerdo con una cualquiera de las reivindicaciones 11 a 14, **caracterizado porque** el estatus de la región del segundo píxel es indicativo de si dicho segundo píxel está localizado dentro de la región objetivo (20) en la segunda trama de imagen ($k - 1$).

30 16. Un codificador (610) de imagen para la codificación de una secuencia (700) de tramas de imagen que comprenden al menos una primera trama de imagen (k) y una segunda trama de imagen anterior ($k - 1$), incluyendo cada una de las tramas de imagen primera (k) y segunda ($k - 1$) una pluralidad de píxel, teniendo la segunda trama de imagen ($k - 1$) una región objetivo (20), teniendo dicho codificador de imagen (610):

35 un medio (614) para la descomposición de cada una de las tramas de imagen en una pluralidad de componentes de sub-bandas; y

un medio (620) para la codificación de las componentes de sub-bandas en una secuencia de código (740), comprendiendo dicho codificador de imágenes (610):

40 un primer algoritmo, que responde a los componentes de las sub-bandas, para definir un área de búsqueda (130) en la primera trama de imagen (k) en base a al menos una parte de la región objetivo (20) en la segunda trama de imagen ($k - 1$), siendo los píxel del área de búsqueda (130) un pluralidad de primeros píxel (40), teniendo cada uno al menos un valor del primer píxel correspondiente; dicho codificador de imagen (610) **caracterizado por**

un segundo algoritmo que responde a cada uno de los primeros píxel (40), para determinar:

45 un área de búsqueda adicional (50) en la segunda trama de imagen ($k-1$) que incluye una pluralidad de segundos píxel de entre la pluralidad de píxel en la segunda trama de imagen ($k - 1$), teniendo cada uno de los segundos píxel al menos un valor de segundo píxel correspondiente y un estatus de la región, en el que el estatus de la región del segundo píxel es indicativo de si dicho segundo píxel está localizado dentro de la región objetivo (20) en dicha trama anterior ($k - 1$);

50 un segundo píxel de referencia en el área de búsqueda adicional (50) en base a una coincidencia entre al valor de primer píxel del primer píxel (40) de entre los valores del segundo píxel; y

55 un estatus de la región de al menos uno de los primeros píxel (40) en base al estatus de la región del segundo píxel de referencia para la determinación de una región objetivo en la primera trama de imagen (k) basada en el estatus de la región de dicho al menos un primer píxel (40).

17. El codificador (610) de imagen de acuerdo con la reivindicación 16, **caracterizado porque** los valores del primer y segundo píxel son indicativos de al menos una de las componentes de crominancia de los coeficientes de ondícula en una sub-banda baja.
- 5 18. El codificador (610) de imagen de acuerdo con la reivindicación 16 o la reivindicación 17, en la que la región objetivo de la primera trama de imagen (k) incluye una frontera y dicha pluralidad de píxel en la primera trama de imagen (k) incluyen una pluralidad de terceros píxel adyacentes a la frontera, teniendo cada uno de los terceros píxel al menos un valor del tercer píxel correspondiente, **caracterizado** dicho programa de ordenador además por un código para la determinación del tipo de borde de los terceros píxel para modificar la región objetivo en dicha trama de imagen (k) en base al tipo de borde de los terceros píxel.
- 10 19. El codificador (610) de imagen de acuerdo con la reivindicación 18, **caracterizado porque** los valores del tercer píxel son indicativos de la componente de luminancia de los coeficientes de ondícula en una sub-banda alta.
20. El codificador (610) de imagen de acuerdo con una cualquiera de las reivindicaciones 16 a 19, **caracterizado porque** el codificador de imagen (610) está adaptado para codificar dicha región objetivo en la primera trama de imagen (k) con una calidad visual más alta que la otra región en dicha primera trama de imagen (k).
- 15 21. Un sistema (600) de codificación de imagen que tiene un codificador (610) de acuerdo con una cualquiera de las reivindicaciones 16 a 20 para la codificación de una secuencia (700) de tramas de imagen en una secuencia de código (740), y un decodificador (630) para la reconstrucción de la secuencia (700) de las tramas de imagen en base a la secuencia de código (740).

20

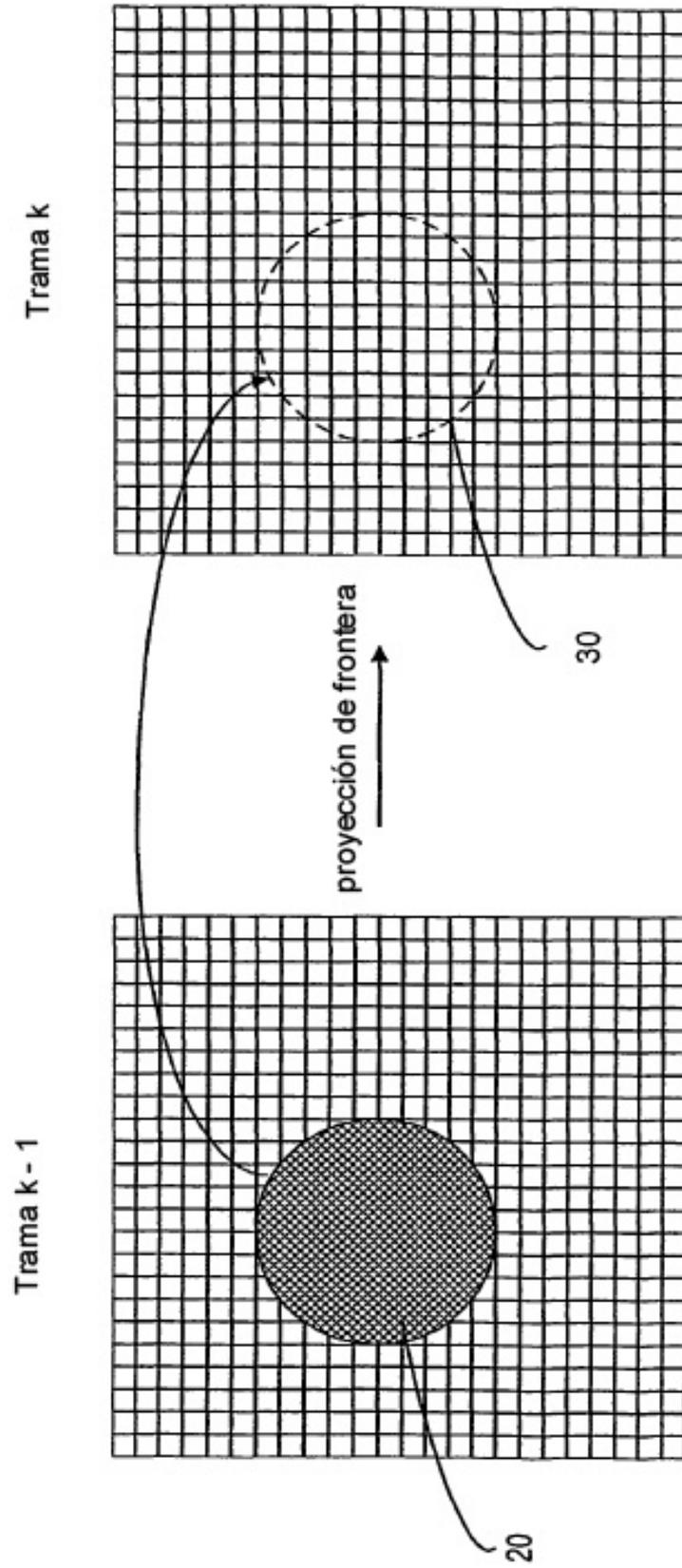


FIG. 1

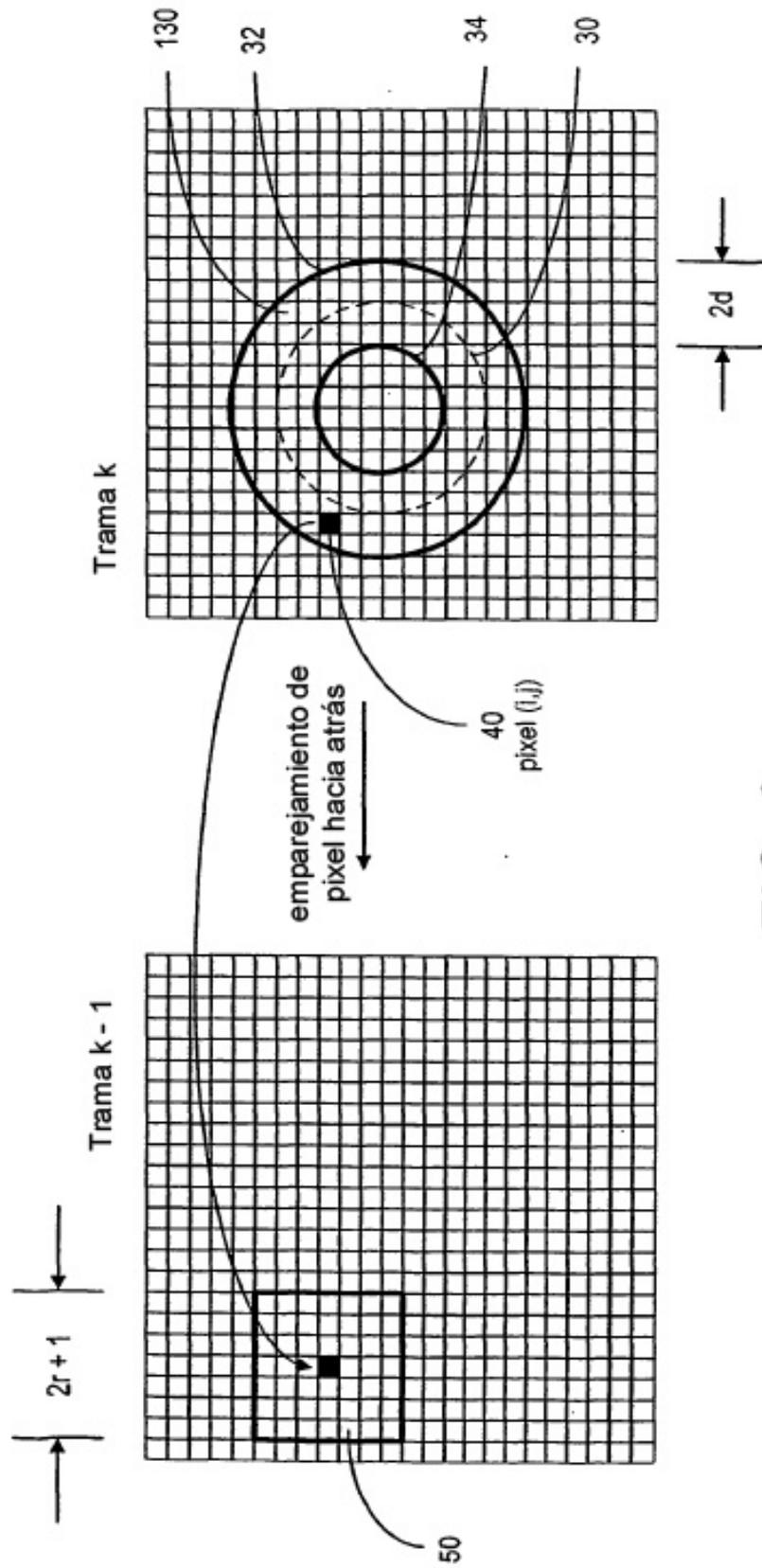


FIG. 2

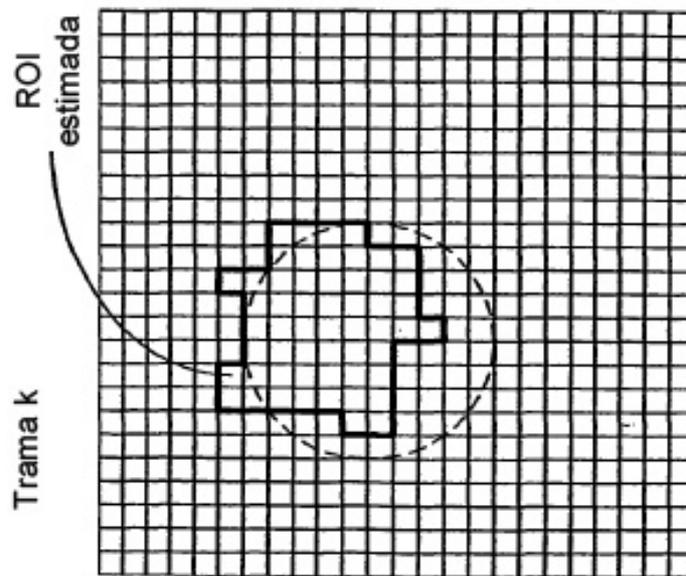
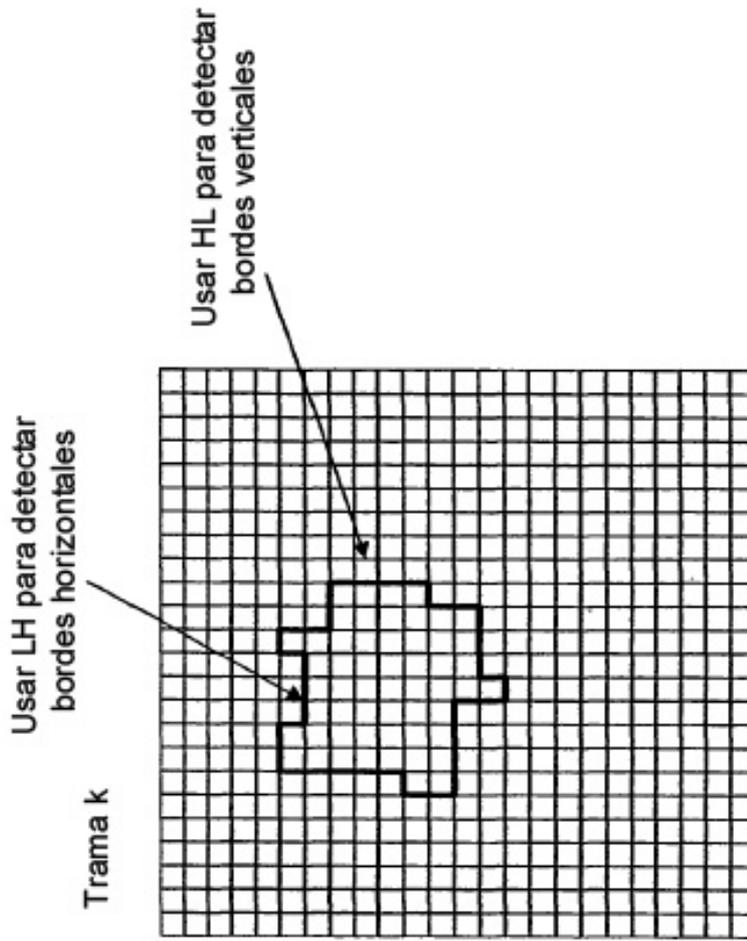


FIG. 3b

FIG. 3a

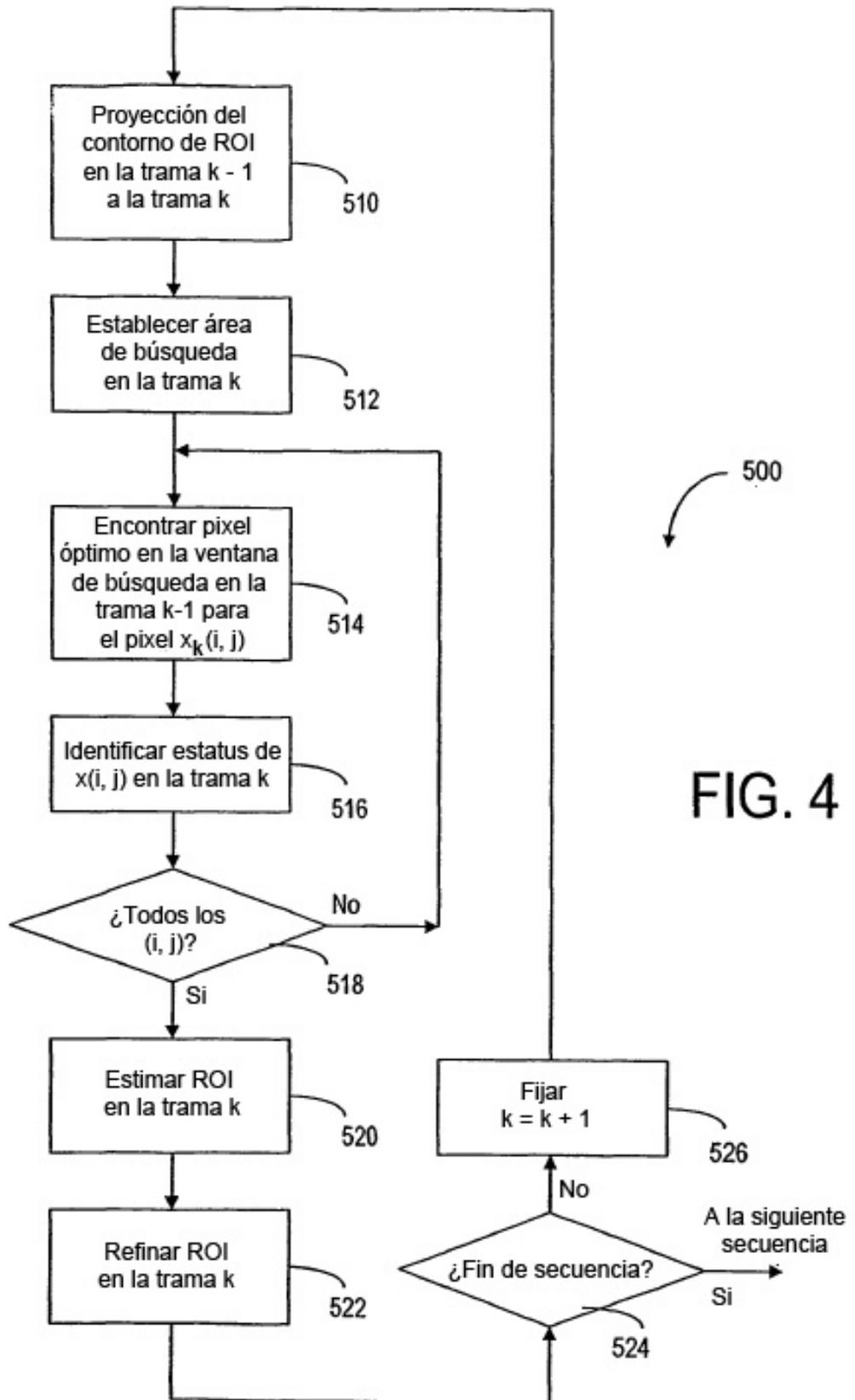


FIG. 4

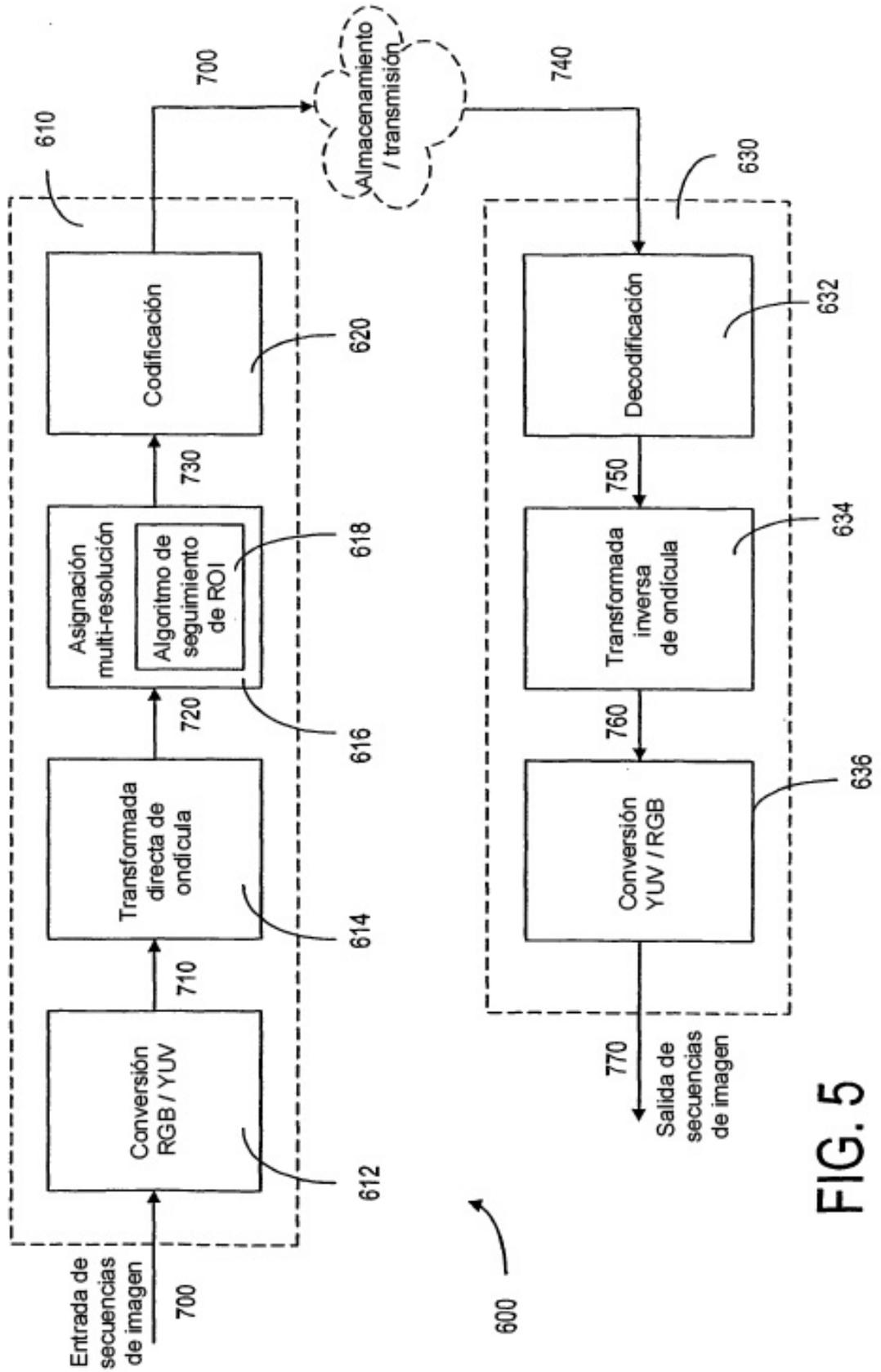


FIG. 5