



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 346 160**

51 Int. Cl.:

G06T 7/20 (2006.01)

G06K 9/64 (2006.01)

G08B 13/194 (2006.01)

G06K 9/00 (2006.01)

G06K 9/62 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **07824097 .5**

96 Fecha de presentación : **10.10.2007**

97 Número de publicación de la solicitud: **2076884**

97 Fecha de publicación de la solicitud: **08.07.2009**

54 Título: **Mejoras relacionadas con la detección de sucesos.**

30 Prioridad: **16.10.2006 GB 0620353**
16.10.2006 EP 06270091

45 Fecha de publicación de la mención BOPI:
11.10.2010

45 Fecha de la publicación del folleto de la patente:
11.10.2010

73 Titular/es: **BAE Systems plc.**
6 Carlton Gardens
London SW1Y 5AD, GB

72 Inventor/es: **Williams, Mark Lawrence**

74 Agente: **González Palmero, Fe**

ES 2 346 160 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

Mejoras relacionadas con la detección de sucesos.

Esta invención se refiere a un procedimiento y aparato para detectar el acaecimiento de sucesos en multitudes, sucesos capturados en grabaciones de vídeo (como grabaciones de circuito cerrado de televisión (CCTV)). La presente invención se puede usar para monitorizar una grabación de vídeo en tiempo real, y alertar a un operador cuando se produzca un suceso, o bien se puede usar para analizar una grabación de vídeo grabada previamente con el fin de buscar sucesos.

Hay muchísimos “sucesos” que merecen la detección mediante sistemas de seguridad dirigidos a hacer frente al terror y al comportamiento antisocial. Un subconjunto importante es el de los sucesos que se producen en o cerca de multitudes, ya que las multitudes son a menudo el objetivo, o bien la razón, de los sucesos de interés. Puede haber una amplia variedad de diversos sucesos de interés que sería deseable detectar. Un área de particular interés y preocupación es el problema de detectar sucesos imprevistos, ya que, *a priori*, puede no saberse ni siquiera qué tipo de suceso es deseable detectar. Los sucesos que son entendidos, o esperados, se pueden detectar usando un conjunto de reglas heurísticas, pero esto no es posible si el tipo de suceso detectar es desconocido, y, además, pasa a ser poco práctico cuando hay un deseo de detectar muchos tipos diferentes de sucesos. Los sucesos de interés pueden incluir ataques terroristas; incendios, terremotos, u otros desastres; los comienzos de malestar social; la presencia de paquetes inusuales; o incluso una persona que actúe de forma inusual.

En Boghossian B.A. y col., Motion-based machine vision techniques for the management of large crowds, Proceedings of the 6th IEEE International Conference on Electronics, Circuits and Systems, ICECS'99, 5. Septiembre de 1999, Vol. 2, 961-964 se presenta un procedimiento para la vigilancia de multitudes. El procedimiento estima las trayectorias y direcciones de los flujos de multitud y usando algoritmos basados en el movimiento detecta emergencias relacionadas con la multitud. Por ejemplo, usa la transformada de Hough de vectores de movimiento para detectar patrones circulares que indican salidas de embotellamiento o puntos de divergencia que indican un lugar de peligro (por ejemplo pelea, incendio).

Hay un número de medidas de complejidad bien conocidas que se refieren al número efectivo de grados de libertad en un sistema, o dimensionalidad de un conjunto de datos. Una medida de complejidad conocida se describe por W. A. Wright, R. E. Smith, M. Danek, y P. Greenway en su ponencia “A Generalisable Measure of Self-Organisation and Emergence”, publicada en las páginas 857 a 864 de “Artificial Neural Networks - ICANN 2001” por Springer-Verlag (2001) y editado por G. Dorffner, H. Bischof, y K. Homik. Wright y col. asocian una habilidad del sistema para mostrar un comportamiento emergente con transiciones repentinas en su medida de complejidad, con relación a cambios suaves en los parámetros del sistema. En este contexto, el comportamiento emergente puede incluir, por ejemplo, la dispersión repentina de un banco de peces en respuesta a la llegada de un depredador. Wright y col. sugieren que su medida se podría usar para evaluar y fomentar comportamientos deseados en una variedad de sistemas complejos, incluyendo sistemas neurales artificiales, sistemas de computación evolutiva, y sistemas de aprendizaje por refuerzo. Sin embargo, desafortunadamente no comprenden cómo se podría aplicar su medida de complejidad en otros campos.

Es un objeto de la presente invención proporcionar un procedimiento y aparato mejorados para la detección de sucesos. Es un objeto adicional de la presente invención proporcionar un procedimiento y sistema para la detección de sucesos que supere, o al menos mitigue, al menos algunos de los problemas mencionados anteriormente asociados con sistemas y procedimientos de detección de sucesos previamente conocidos.

En términos generales, la presente invención reside en la comprensión de que el análisis de la dinámica de multitudes se puede aplicar a la detección de sucesos que se producen en multitudes. Esto hace posible que un suceso de interés sea detectado por la reacción de una multitud a ese suceso, más que por la observación directa del suceso. Los sistemas y procedimientos de detección de sucesos de acuerdo con la presente invención son por tanto sensibles a una amplia variedad de sucesos: la invención proporciona un detector de sucesos que es sensible al modo en el que reacciona la multitud, como se refleja por los cambios en el movimiento de la multitud, más que a un tipo específico de suceso.

Por lo tanto, de acuerdo con un primer aspecto de la presente invención, se proporciona un procedimiento de detección del acaecimiento de un suceso en una multitud, comprendiendo el procedimiento las etapas de:

(a) provisión de una grabación de vídeo de la multitud, comprendiendo la grabación de vídeo una serie de fotogramas de imágenes;

(b) cálculo, para cada uno de la serie de fotogramas de imágenes, de un valor representativo de la dinámica de la multitud, siendo calculado el valor en dependencia de dicho cada uno de la serie de fotogramas de imágenes, calculándose de ese modo una serie de valores;

(c) análisis de la serie de valores para identificar cualquier discrepancia que se produzca en la serie de valores; y

(d) en respuesta a identificar una discrepancia, provisión de una señal de salida para alertar a un operador que un suceso ha sido detectado.

Así las reacciones de la multitud a un suceso, analizadas en términos de la dinámica de la multitud, se usan para detectar el acaecimiento de ese suceso. La multitud tiene numerosos puntos de vista proximales con los que monitorizar sus alrededores, y además la multitud podría reaccionar a olores, como humo, y a indicios audibles, como avisos o explosiones. Tales olores e indicios audibles no se pueden detectar directamente por un sistema de vigilancia de vídeo. Además, la multitud puede identificar un comportamiento peculiar o inusual usando pistas sutiles, como el lenguaje corporal, que no se pueden leer de modo fiable desde una visión más remota. El ataque de gas neurotóxico de 1995 en el metro de Tokio es un ejemplo de un suceso - la liberación de un gas tóxico - que se podría haber detectado indirectamente por la reacción de las personas expuestas al gas, pero no directamente por la grabación de vídeo. Analizando la dinámica de la multitud, y monitorizando la dinámica en tiempo real mediante el uso de algún valor representativo de la dinámica, la invención permite que una amplia variedad de sucesos sean detectados como discrepancias en el valor representativo.

La etapa de análisis de la serie de valores puede comprender la etapa de provisión de un intervalo normal de valores, y entonces se puede identificar una discrepancia cuando el valor está fuera del intervalo normal. El intervalo normal puede ser un intervalo predeterminado, o se puede determinar en dependencia del promedio y la desviación estándar de la serie de valores. De forma ventajosa, esto simplifica el proceso de identificación de discrepancias, y se implementa fácilmente de forma automática por medios de procesamiento apropiados como un ordenador.

En una forma de realización, el procedimiento comprende además las etapas de rastreo de un número de objetivos que se mueven dentro de la multitud para determinar un número de rastros de objetivos, y la actualización de los rastros de objetivos en respuesta a cada fotograma de imagen sucesivo; en el que la etapa de cálculo de una serie de valores comprende la determinación de un valor en dependencia de los rastros de objetivos actualizados. Los rastros de objetivos actualizados son una medida directa de la dinámica de la multitud, y por lo tanto se pierde poca información útil en el procesamiento de la serie de imágenes. Por conveniencia, la etapa de cálculo de una serie de valores comprende la etapa de formación de una matriz de rastros que comprende filas, representando cada fila un rastro de objetivos, y evaluación de la métrica de complejidad de Wright para la matriz de rastros. En este documento, el término "métrica de complejidad de Wright" se usa para referirse a la métrica de complejidad deducida por Wright y col. en su ponencia citada anteriormente.

De forma alternativa, la etapa de cálculo de una serie de valores puede comprender el cálculo de una serie de imágenes de diferencias en dependencia de la serie de fotogramas de imágenes, estando relacionada la imagen de diferencias para un fotograma de imagen actual con las diferencias entre el fotograma de imagen actual y un fotograma de imagen precedente. Por conveniencia, el fotograma de imagen precedente puede ser el fotograma de imagen inmediatamente anterior al primer fotograma de imagen en la serie de fotogramas de imágenes. El cálculo de la serie de valores representativos de la dinámica de la multitud a partir de una serie de imágenes de diferencias, más que a partir de una serie de rastros de objetivos, evita el problema de requerir técnicas de rastreo complejas para rastrear objetivos individuales dentro de la multitud, simplificándose así el procedimiento de detección de sucesos.

Por conveniencia, la etapa de cálculo de la serie de valores en dependencia de las imágenes de diferencias puede comprender, para cada uno de la serie de valores, las etapas de: formación de una matriz de diferencias, comprendiendo la matriz de diferencias un número de filas y representando cada fila una de la serie de imágenes de diferencias; y evaluación de la métrica de complejidad de Wright para la matriz de diferencias para proporcionar dicho cada uno de la serie de valores. El procedimiento puede comprender además la etapa de disminución de la resolución de al menos uno de la serie de fotogramas de imágenes antes de calcular la serie de valores, mejorándose de forma ventajosa la eficiencia del procedimiento de la matriz de diferencias.

Opcionalmente, el procedimiento se realiza en tiempo real.

Por conveniencia, la serie de valores comprende una serie de valores de complejidad. Los valores de complejidad pueden estar relacionados con el número de grados de libertad efectivos mostrados en la dinámica de la multitud. El valor de complejidad es una cantidad escalar que se puede monitorizar fácilmente.

La invención se extiende a un producto de programa informático que comprende instrucciones legibles por ordenador, cuyo producto es operable, cuando se inserta en un ordenador en asociación operable con una cámara de vídeo, para realizar un procedimiento según cualquier reivindicación precedente.

Según un segundo aspecto de la invención, se proporciona un sistema de detección de sucesos que comprende una cámara de vídeo, un procesador, y un medio de alerta; siendo operable la cámara de vídeo para grabar una grabación de vídeo, y para transmitir dicha grabación de vídeo al procesador, comprendiendo dicha grabación de vídeo una serie de fotogramas de imágenes; y comprendiendo el procesador medios para calcular, para cada uno de la serie de fotogramas de imágenes, un valor representativo de la dinámica de la multitud, siendo calculado el valor en dependencia de dicho cada uno de la serie de fotogramas de imágenes, de tal manera que se calcula una serie de valores; y medios para analizar la serie de valores para identificar cualquier discrepancia que se produzca en la serie de valores de complejidad; y en el que el medio de alerta está configurado para alertar a un operador en respuesta al acaecimiento de cualquier discrepancia identificada por el proceso.

ES 2 346 160 T3

Las características anteriores y adicionales de la invención se exponen en las reivindicaciones adjuntas, y se explicarán a continuación, sólo a modo de ejemplo, con referencia a diversas formas de realización específicas que se ilustran en los dibujos anexos, en los que:

la fig. 1 es una vista esquemática de un aparato según una primera forma de realización de la invención;

la fig. 2 es un diagrama de flujos que ilustra esquemáticamente las etapas realizadas en la operación de la forma de realización de la figura 1;

la fig. 3 es una gráfica que ilustra datos grabados durante el ensayo del procedimiento ilustrado en el diagrama de flujos de la figura 2; y

la fig. 4 es un diagrama de flujos que ilustra esquemáticamente las etapas realizadas como parte de un procedimiento de acuerdo con una forma de realización adicional de la invención.

Con referencia en primer lugar a la figura 1, se muestra un aparato 100 de acuerdo con una primera forma de realización de la invención. El aparato 100 comprende una cámara de vídeo 110, un procesador 120, y un medio de alerta 130. La cámara de vídeo 110 se dispone para monitorizar un área de interés en la que haya una multitud. El área de interés puede ser una calle principal abarrotada, una terminal de aeropuerto, o incluso una carretera en la que esté fluyendo tráfico. Así, se entenderá que el término “multitud”, según se usa en este documento, deberá ser interpretado ampliamente de forma que incluya cualquier grupo de objetivos de interés, y también que incluya grupos de objetivos de interés que se muevan por un área, tanto si los objetivos están o no presentes simultáneamente en esa área. La cámara 110 puede ser cualquier tipo de dispositivo operable para grabar una serie de imágenes de su campo de visión, como se entenderá por los expertos en la materia, pero en esta forma de realización particular, la cámara 110 es una cámara de CCTV.

La grabación de vídeo, en la forma de una serie de fotogramas de imágenes, se emite desde la cámara 110 al procesador 120. En la presente forma de realización, el procesador 120 monitoriza las imágenes que recibe en tiempo real con el fin de detectar el acaecimiento de sucesos. Para cada imagen recibida, se calcula un valor de complejidad, para que el procesador 120 calcule una serie de valores de complejidad. La complejidad es una cantidad útil para medir puesto que es representativa de la dinámica de la multitud. Más especialmente, se refiere al número efectivo de grados de libertad presentes en el movimiento de los miembros de la multitud. Considérese una situación simple en la que el sistema 110 está monitorizando una multitud de gente en una gran sala. Bajo circunstancias normales, la gente en la multitud se moverá en gran medida al azar, con poca correlación entre las trayectorias tomadas por los individuos. Sin embargo, en caso de que hubiera, por ejemplo, un incendio en un extremo de la habitación, la gente en la multitud empezará a moverse hacia las salidas, y habrá una correlación mucho mayor entre el movimiento de los individuos dentro de la multitud. En la última situación, el número efectivo de grados de libertad de movimiento de la multitud se disminuye, y así el incendio sería detectado como un suceso. En especial, el suceso sería detectado incluso si el incendio no estuviese en el campo de visión de la cámara 110. La manera en la que se realiza el cálculo del valor de complejidad se describirá en detalle más adelante.

Se calcula la desviación estándar y el promedio de la serie de valores de complejidad, y un intervalo normal se define como dos desviaciones estándares a cada lado del promedio. Si el valor calculado para un fotograma de imagen dado está fuera de este intervalo normal, el procesador 120 proporciona una señal, a través del medio de alerta 130, a un operador. En esta forma de realización específica, el medio de alerta 130 es una alarma audible. Como se explica anteriormente, la desviación del intervalo normal de los valores de complejidad se puede asociar con el acaecimiento de un suceso de interés, y la alerta proporcionada al operador permite que el operador investigue más atentamente. Por ejemplo, donde el operador debe monitorizar la salida no sólo de la cámara 110, sino también de varias cámaras, la alerta puede hacer que el operador mire la grabación grabada por la cámara 110 más atentamente.

La operación del procesador 120 se muestra esquemáticamente en el diagrama de flujos 200, que ilustra las etapas llevadas a cabo por el procesador 120 en la determinación de si se debería dar una alerta en respuesta a la recepción de un nuevo fotograma de imagen. El diagrama de flujos 200 toma por tanto, como su punto de partida, la recepción de un nuevo fotograma de imagen por el procesador 120, cuya etapa se ilustra como etapa 210. A continuación, este fotograma de imagen recibido nuevamente se denomina fotograma de imagen actual. En la etapa 220, se calcula una imagen de diferencias. La imagen de diferencias es indicativa de cómo ha cambiado el fotograma de imagen actual con relación a un fotograma precedente. Las imágenes se pueden representar en la forma de una matriz de los valores de píxel, y la imagen de diferencias se calcula sustrayendo, de la matriz que representa el fotograma de imagen actual, la matriz que representa el fotograma de imagen recibido inmediatamente antes del fotograma de imagen actual.

El cálculo de una imagen de diferencias en la etapa 220 simplifica considerablemente las computaciones posteriores que se llevarán a cabo en las imágenes, ya que los píxeles sin cambios entre los dos fotogramas serán todos representados, en la imagen de diferencias, por un elemento de matriz que tiene un valor de cero. Además, esto extrae la información de antecedentes redundante del fotograma de imagen actual. El procesamiento se puede simplificar aún más por la aplicación de un umbral a la matriz de imágenes de diferencias, de tal manera que a los elementos que tienen un valor por debajo de un valor umbral preestablecido se les da un valor de cero, y a los elementos que tienen un valor por encima de un valor umbral preestablecido se les da un valor de uno. La matriz de imágenes de diferencias es así mapeada a una matriz binaria, y consiste en elementos que tienen un valor de 1 o bien 0. En algunas circunstancias,

y especialmente donde sólo algunos píxeles cambian los valores entre los fotogramas, es ventajoso intercambiar los elementos, de tal manera que los elementos que tienen previamente un valor de 0 son cambiados para tener un valor de 1, y viceversa. Esto es debido a que puede haber inestabilidades numéricas en el cálculo de complejidad descrito más abajo (y en particular en la etapa de descomposición de valor singular) donde la matriz tiene una gran población de elementos con valor de cero.

En la etapa 230, se calcula un valor de complejidad usando datos de un número de imágenes de diferencias almacenadas por el procesador 100, y datos de la imagen de diferencias calculada a partir del fotograma de imagen actual. Las imágenes de diferencias almacenadas son las que preceden inmediatamente al fotograma de imagen actual. El procesador 120 almacena un número constante de imágenes de diferencias para que, en respuesta al fotograma de imagen actual que se recibe (en la etapa 210) y su imagen de diferencias asociada que se calcula (en la etapa 220), la imagen de diferencias más antigua almacenada sea eliminada. En la presente forma de realización, se almacenan veinte imágenes de diferencias en un momento dado. El valor de complejidad se calcula entonces a partir de una matriz de diferencias Z definida por:

$$Z = \begin{pmatrix} x_{t=t_0}^1 & x_{t=t_0}^2 & x_{t=t_0}^3 & \cdots & x_{t=t_0}^N \\ x_{t=t_0+1}^1 & x_{t=t_0+1}^2 & x_{t=t_0+1}^3 & \cdots & x_{t=t_0+1}^N \\ \vdots & & & & \\ x_{t=T}^1 & x_{t=T}^2 & x_{t=T}^3 & \cdots & x_{t=T}^N \end{pmatrix}$$

en el que el elemento $x_{t=j}^n$ es el valor del píxel n -ésimo en la imagen de diferencias grabada en $t = j$. Así, la matriz que representa la imagen de diferencias en $t = j$ se convierte en una fila de valores de píxel, y la matriz de diferencias Z está compuesta de diecinueve filas que representan las imágenes de diferencias calculadas a partir de una serie de diecinueve fotogramas de imágenes consecutivos recibidos inmediatamente antes del fotograma de imagen actual, y la imagen de diferencias calculada para el fotograma de imagen actual. La imagen de diferencias calculada a partir del fotograma de imagen actual es la designada $t = T$, y la imagen de diferencias más antigua es la designada $t = t_0$.

En efecto, por tanto, cuando se calcula cada imagen de diferencias nueva, se evalúa un valor de complejidad usando una ventana de fotogramas de imágenes, cuya ventana es de un tamaño constante $(T - t_0)$ y, en esta forma de realización específica, es equivalente a veinte fotogramas de imágenes. Como será claro para los expertos en la materia, el Índice de tiempo se mide en términos del número de fotogramas de imágenes recibidos. Como también será claro para los expertos en la materia, es necesario inicializar una matriz de diferencias Z cuando comience la monitorización, pero cómo se hace esto no es crítico ya que el procedimiento de inicialización no afecta a los cálculos una vez que se hayan grabado más de $(T - t_0)$ imágenes por la cámara de vídeo 110. Una técnica de inicialización posible que se usa en la presente forma de realización es definir una matriz inicial Z_0 , de un tamaño apropiado, cuyos elementos están todos valorados en la unidad.

El valor de complejidad se evalúa usando la métrica de complejidad de Wright:

$$\Omega = 2^S$$

donde

$$S = \sum_{i=1}^M \sigma_i' \log_2 \sigma_i'$$

y σ_i' , es el valor singular normalizado i -ésimo de la matriz de diferencias Z . Los valores singulares normalizados se obtienen usando una descomposición de valores singulares, una operación bien conocida aplicable a cualquier matriz que produzca una matriz diagonal con entradas diagonales que sean los valores singulares de la matriz original. Para cualquier matriz real A , los valores singulares son las raíces cuadradas de los valores característicos de la matriz A^*A , donde A^* denota la transpuesta de la matriz A . Estos valores se normalizan usando un procedimiento de escalado simple que escala el valor singular más grande de la matriz Z a 1. La métrica de complejidad Ω es deducida en la ponencia por Wright y col. mencionada anteriormente, y se puede considerar que representa el número efectivo de grados de libertad mostrados por la multitud.

El cálculo de Ω , y especialmente la etapa de descomposición de valores singulares que lleva a los valores σ_i' , puede ser numéricamente largo si el número de píxeles en cada fotograma de imagen es extenso. El tiempo de procesamiento requerido para este cálculo se puede reducir si se disminuye la resolución de los fotogramas de imágenes, reduciéndose de esta manera el número de píxeles. Ya que cada elemento en la matriz de diferencias Z está relacionado con un píxel en una imagen de diferencias, el número de elementos en cada fila de la matriz de diferencias Z , y por

tanto el tamaño global de la matriz de diferencias Z , se reduce disminuyéndose la resolución de los fotogramas de imágenes de entrada. Esto acorta la cantidad del tiempo del procesador requerido para el cálculo de Ω , en gran medida debido a la reducción en la cantidad de tiempo necesario para la etapa de descomposición de valores singulares. Se puede determinar por tanteo cuánto se puede reducir la resolución de las imágenes, pero se observa que los requisitos encontrados de velocidad de procesamiento, y la suficiencia de resolución para capturar información relativa a la dinámica de la multitud, ambos deben ser satisfechos.

En la etapa 240, la complejidad evaluada en la etapa 230 se compara con un intervalo normal. El intervalo normal, como se describe anteriormente, se define de forma que está dentro de valores umbrales predeterminados, establecidos por un operador, que varían según la desviación estándar de la serie de valores de complejidad. En esta forma de realización particular, se elige que el intervalo normal sea un intervalo de dos desviaciones estándares a cada lado del promedio de los valores de complejidad. Si el valor de complejidad evaluado en la etapa 230 está dentro del intervalo normal, el procesador 120 retorna a la etapa 210 y espera un nuevo fotograma de imagen adicional. Si el valor de complejidad está fuera del intervalo normal, el procesador 120, en la etapa 250, proporciona una salida al medio de alerta 130 de tal manera que un operador sea alertado de la discrepancia en la serie de valores de complejidad. El operador puede entonces investigar más aún las causas de la discrepancia, por ejemplo centrando la atención en la salida de la cámara 110.

Se ha demostrado que el procedimiento ilustrado por el diagrama de flujos 200 mostrado en la figura 2 funciona con éxito usando datos de muestra de un aparcamiento. Se creó un suceso artificial colocando un “espantapájaros” en el aparcamiento aproximadamente a mitad del periodo de prueba. La figura 3 es una representación gráfica de la variación de la medida de complejidad con el tiempo, según se mide usando un procedimiento de acuerdo con la segunda forma de realización de la invención. La presencia del espantapájaros altera la trayectoria de las personas que caminan por la toma del aparcamiento, y esto se observa en la gráfica 300 de la figura 3 como una transición (designada 310) en la variación de la complejidad con el tiempo.

La transición 310 es claramente una transición estadísticamente significativa del tipo que sería fácilmente detectable por medios automáticos. De hecho, el desplazamiento del valor de complejidad medio al colocar el espantapájaros en el aparcamiento es equivalente aproximadamente a trece desviaciones estándares. La medida de complejidad fue evaluada, en este ejemplo, usando una resolución de imagen de aproximadamente un píxel por persona, equivalente a aproximadamente 500 píxeles por fotograma. Se comprobó que tal resolución era suficientemente eficiente computacionalmente e incluso permitía la posibilidad de una detección de sucesos satisfactoria. Se comprobó que la disminución de la resolución más aún tenía el efecto de disminuir la importancia de la transición 310 en la medida de complejidad, ya que la desviación estándar de la serie de valores de complejidad aumentó. Así la detección de cualquier discrepancia en la serie de valores se hizo menos probable al disminuirse la resolución de los fotogramas de imágenes.

Como se describe anteriormente, la medida de complejidad deducida por Wright y col., está relacionada con la dinámica de los objetos a los que se aplica. Es por tanto sorprendente que la aplicación de la métrica de complejidad de Wright a una matriz Z que comprende elementos que representan valores de píxel diera como resultado cualquier valor significativo. Sin embargo, se piensa que las imágenes de diferencias retienen, de hecho, cierta información sobre la dinámica de la multitud que se puede extraer mediante la aplicación de la métrica de complejidad. La información contenida dentro de las imágenes de diferencias se refiere sólo a los píxeles cuyos valores han cambiado entre fotogramas consecutivos. Estos píxeles cambiados son claramente indicativos del movimiento dentro de la multitud, ya que cualquier movimiento semejante hará que cambien los valores de píxel. Así la matriz de diferencias Z se puede usar con éxito en combinación con la medida de complejidad a los efectos de la detección de sucesos.

Adicionalmente, mientras que el “suceso” en el ejemplo del aparcamiento citado anteriormente, de un espantapájaros que aparece junto a un coche aparcado, es completamente artificial, los expertos en la materia reconocerán inmediatamente que hay muchas circunstancias en las que una multitud de gente reaccionaria de forma similar a un suceso digno de detección. Por ejemplo, sería probable que una multitud reaccionara a la presencia de un paquete desatendido sospechoso: sería probable que los individuos en la multitud alteraran su trayectoria con el fin de alejarse más del paquete. Tal reacción sería detectada por un sistema según la primera forma de realización de la invención. En especial, es la reacción de la multitud, más que la presencia del paquete sospechoso, la que se detecta. Un sistema capaz de detectar automáticamente tales sucesos tiene una utilidad clara en el campo de la seguridad nacional.

Un procedimiento de acuerdo con una forma de realización adicional de la invención se ilustra esquemáticamente por el diagrama de flujos 400 mostrado en la figura 4. El procedimiento de esta forma de realización usa el aparato 100 ilustrado en la figura 1, y es similar al procedimiento de esta forma de realización ilustrado esquemáticamente por el diagrama de flujos 200 mostrado en la figura 2, excepto en el procesamiento de los datos de imagen usados para calcular la medida de complejidad. En el diagrama de flujos 400 de la figura 4, la etapa 410 representa un nuevo fotograma de imagen que es recibido por el procesador 120 (figura 1). Se considera que la imagen es de un área en la que hay un número de objetivos de interés moviéndose, cuyos objetivos pueden, por ejemplo, ser coches que se mueven en el tráfico en una calle concurrida. Los objetivos son rastreados por el procesador 120, y sus rastros grabados y actualizados en respuesta a la recepción de cada fotograma de imagen nuevo. La etapa de actualización de los rastros de objetivos se representa como etapa 420 en el diagrama de flujos 400. Los objetivos se pueden rastrear usando cualquier algoritmo de rastreo conocido, como un algoritmo del Filtro de Kalman (dado a conocer por R. E. Kalman en “A New approach to linear filtering and prediction problems”, publicado en Transactions of the ASME

Journal of Engineering, páginas 33 a 45 (1960)), o un algoritmo de filtro de partículas (dado a conocer por N. Gordon, D. Salmond, y A. Smith en su ponencia "Novel Approach to non-linear/non-Gaussian Bayesian State Estimation", publicada en IEE Proc. F, Volume 140, no. 2, páginas 107 a 113 (1993)), en conjunción con un procedimiento de detección de objetivos adecuado, como, por ejemplo, una técnica de emparejamiento de plantillas. La técnica de emparejamiento de plantillas requiere que un procesador intente emparejar una plantilla sabida por adelantado a una parte de una imagen recibida. Donde se encuentra un emparejamiento, se identifica un objetivo.

En la etapa 430, la complejidad de los rastros se evalúa usando la medida de complejidad de Wright ? definida anteriormente. En este caso, la matriz Z comprende filas de datos que representan los rastros de objetivos más que datos de las imágenes de diferencias, y por tanto se denomina a continuación matriz de rastros. Así, por ejemplo, cada fila puede comprender coordenadas, en el fotograma de imagen, de cada objetivo, con la evolución del tiempo de los rastros siendo grabada como nuevas filas añadidas a la matriz de rastros con cada nueva imagen recibida. El valor de complejidad se calcula entonces como se cita anteriormente, usando la métrica de complejidad de Wright. Se observa que la disminución de la resolución de las imágenes antes de calcular el valor de complejidad es innecesaria en este caso, ya que no hay ningún vínculo entre la resolución de la imagen y el tamaño de la matriz de rastros. De hecho, se espera que una reducción en la resolución de la imagen tuviera el efecto no deseado de disminuir la fiabilidad del algoritmo de rastreo usado para rastrear objetivos por la grabación de vídeo. En la etapa 440, se determina si el valor de complejidad se halla o no dentro de un intervalo normal definido, como se describe anteriormente en relación a la etapa 240 del diagrama de flujos 200 ilustrado en la figura 1. Si el valor de complejidad está dentro del intervalo normal, el proceso retorna a la etapa 410. Si está fuera del intervalo normal, un operador es alertado de la discrepancia en la etapa 450, y el proceso retorna a la etapa 410.

El procedimiento de la forma de realización ilustrada en la figura 4 también es aplicable a otros objetivos aparte de coches, y, en particular, es aplicable a multitudes de gente. El procedimiento puede por tanto ser comparado cuantitativamente al ilustrado en la figura 2, y los inventores han mostrado que el procedimiento de la figura 4 puede producir mejores resultados que el procedimiento de la figura 2. Un ensayo comparativo, que analiza los mismos datos de vídeo mostrados en la figura 2 usando el procedimiento de la forma de realización ilustrada en la figura 4 mostró un desplazamiento de 22, más que de 13, desviaciones estándares del valor de complejidad al colocar el espantapájaros en el aparcamiento. Se cree que esto se debe a que los datos de rastros están más estrechamente relacionados con la dinámica de la multitud objetivo que los datos basados en píxeles. Sin embargo, en este ejemplo, el rastreo de gente que se mueve por el aparcamiento se realizó de forma manual, más que usándose cualquier algoritmo de rastreo conocido, puesto que los algoritmos de rastreo conocidos son frecuentemente poco fiables para rastrear a la gente. La gente puede cambiar de forma cuando se mueve, puede cambiar su apariencia de forma significativa al girarse y cambiar el aspecto presentado a la cámara, y puede estar oculta parcialmente o completamente por los objetos en el fotograma. Así puede ser muy difícil rastrear con éxito a una persona por una serie de formas de imagen debido a los problemas asociados con la detección de objetivos. Se pueden usar técnicas de emparejamiento de plantillas (como se describe brevemente anteriormente), al igual que técnicas de bases de datos basadas en modelos, en las que se usa el conocimiento previo del sistema esquelético conocido, y la manera en que se mueve, para imponer una restricción sobre la detección de objetivos. Sin embargo, tales técnicas pueden requerir una capacidad de procesamiento significativa.

Un procedimiento de acuerdo con otra forma de realización más de la invención se ilustra esquemáticamente en el diagrama de flujos 500 mostrado en la figura 5. El procedimiento se puede realizar usando el aparato 100 mostrado en la figura 1 donde la cámara 110 es operable para emitir fotogramas de imágenes de color al procesador 120. Las etapas ilustradas en el diagrama de flujos 500 son similares a las etapas ilustradas en el diagrama de flujos 200 mostrado en la figura 2, excepto en que el color de la imagen es cuantizado en la etapa 520 antes de que se calculen las imágenes de diferencias y los valores de complejidad. Así, en la etapa 510, se recibe un nuevo fotograma de imagen en el procesador 120 (mostrado en la figura 1). En la etapa 520, se cuantiza el color del fotograma de imagen. En el fotograma de imagen de color, cada píxel tiene un valor de color. El proceso de cuantización de color asigna un valor de 1 ó 0 a cada píxel en dependencia de si su color se halla entre un intervalo predeterminado de valores de color. Se puede producir por tanto un número de matrices de imágenes con color cuantizado, que cubren el espectro de colores a los que es sensible la cámara 110. Las imágenes de diferencias se pueden calcular entonces, en la etapa 530, para cada una de las imágenes con color cuantizado, y así una matriz de diferencias puede ser calculada para cada intervalo de color. Así cada fotograma de imagen nuevo genera un número de matrices de diferencias, cuyos valores de complejidad se evalúan en la etapa 540 usando la métrica de complejidad de Wright definida anteriormente. Los diversos valores de complejidad para cada una de las matrices de diferencias se suman con el fin de llegar a un único valor de complejidad para el fotograma de imagen. En la etapa 550, se determina si el único valor de complejidad está o no dentro de un intervalo normal (del mismo modo que se hace en la etapa 240 en el procedimiento ilustrado en el diagrama de flujos 200). Si el valor está dentro del intervalo normal, el procesador retorna a la etapa 510 y espera un nuevo fotograma de imagen; si el valor de complejidad está fuera del intervalo normal, un operador es alertado en la etapa 560. Se ha comprobado que la inclusión de la etapa de cuantización de color da como resultado un procedimiento de detección de sucesos más efectivo en algunas pruebas que usan una grabación real.

Habiéndose descrito la invención con referencia a diversas formas de realización específicas, se observa que estas formas de realización son ejemplares en todos los sentidos. Son posibles variaciones y modificaciones sin desviarse del ámbito de la invención, la cual se define en las reivindicaciones adjuntas. Tales variaciones y modificaciones serán inmediatamente obvias para los expertos en la materia. Por ejemplo, mientras que, anteriormente, se ha descrito la definición de un intervalo normal en términos de la desviación estándar y el promedio de la serie de valores de complejidad, también sería posible por supuesto para un operador proporcionar umbrales predeterminados, o en efecto

ES 2 346 160 T3

para una discrepancia ser identificada cuando el valor cambia en más de una cantidad predefinida entre dos (o más) fotogramas de imágenes.

5 También se ha descrito anteriormente el cálculo de una serie de valores de complejidad en dependencia de la serie de fotogramas de imágenes usándose la métrica de complejidad de Wright Ω , pero se reconocerá inmediatamente por los expertos en la materia que se podrían usar otras medidas de complejidad, o, en efecto, cualquier medida representativa de la dinámica de la multitud en lugar de la medida de complejidad de Wright Ω . Tales valores pueden ser otros valores relacionados con el número de grados de libertad efectivos, o pueden ser valores estadísticos marginales. Algunos ejemplos incluyen los exponentes de Lyapunov, la velocidad media de los objetivos de interés que se extienden por la multitud, o la longitud de la trayectoria media en línea recta de los objetivos de interés que se extienden por la multitud. Además, se debe entender que la invención se puede aplicar tanto en tiempo real, como se describe anteriormente, con el fin de alertar a un operador del acaecimiento de un suceso, como en análisis posteriores de grabación de vídeo capturada en un momento anterior. La habilidad para analizar una grabación de vídeo obtenida previamente es útil para las grabaciones adquiridas de cámaras de seguridad, ya que puede reducir de forma significativa la cantidad de trabajo manual requerido para identificar una secuencia de interés en la grabación.

20 Finalmente, se observa que se entenderá claramente que cualquier característica descrita anteriormente con relación a cualquier forma de realización se puede usar sola, o en combinación con otras características descritas, y también se puede usar en combinación con una o más características de cualquier otra de las formas de realización, o cualquier combinación de cualquier otra de las formas de realización.

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

1. Un procedimiento para detectar el acaecimiento de un suceso en una multitud, comprendiendo el procedimiento las etapas de:
 - (a) provisión de una grabación de vídeo de la multitud, comprendiendo la grabación de vídeo una serie de fotogramas de imágenes;
 - (b) y estando **caracterizado** el procedimiento por las etapas de:
 - cálculo, para cada uno de la serie de fotogramas de imágenes, de un valor representativo de la dinámica de la multitud, siendo calculado el valor en dependencia de dicha cada serie de fotogramas de imágenes, calculándose de ese modo una serie de valores;
 - (c) análisis de la serie de valores para identificar cualquier discrepancia que se produzca en la serie de valores; y
 - (d) en respuesta a la identificación de una discrepancia, provisión de una señal de salida para alertar a un operador de que se ha detectado un suceso.
2. Un procedimiento según la reivindicación 1 en el que la etapa de análisis de la serie de valores comprende la etapa de provisión de un intervalo normal de valores, y en el que se identifica una discrepancia cuando el valor está fuera del intervalo normal.
3. Un procedimiento según la reivindicación 2 en el que el intervalo se determina en dependencia del promedio y la desviación estándar de la serie de valores.
4. Un procedimiento según cualquier reivindicación precedente que comprende además las etapas de rastreo de un número de objetivos que se mueven dentro de la multitud para determinar un número de rastros de objetivos, y actualización de los rastros de objetivos en respuesta a cada fotograma de imagen sucesivo; y en el que la etapa de cálculo de una serie de valores comprende la determinación de un valor en dependencia de los rastros de objetivos actualizados.
5. Un procedimiento según la reivindicación 4 en el que la etapa de cálculo de una serie de valores comprende la etapa de formación de una matriz de rastros que comprende filas, representando cada fila un rastro de objetivos, y evaluación de la métrica de complejidad de Wright para la matriz de rastros.
6. Un procedimiento según cualquiera de las reivindicaciones 1 a 3 en el que la etapa de cálculo de una serie de valores comprende el cálculo de una serie de imágenes de diferencias en dependencia de la serie de fotogramas de imágenes, estando relacionada la imagen de diferencias para un fotograma de imagen actual con las diferencias entre el fotograma de imagen actual y un fotograma de imagen precedente.
7. Un procedimiento según la reivindicación 6, en el que el fotograma de imagen precedente es el fotograma de imagen inmediatamente anterior al fotograma de imagen actual en la serie de fotogramas de imágenes.
8. Un procedimiento según la reivindicación 6 o la reivindicación 7 en el que la etapa de cálculo de la serie de valores comprende, para cada uno de la serie de valores, las etapas de:
 - formación de una matriz de diferencias, comprendiendo la matriz de diferencias un número de filas y representando cada fila una de la serie de imágenes de diferencias; y
 - evaluación de la métrica de complejidad de Wright para la métrica de diferencias para proporcionar dicho cada uno de la serie de valores.
9. Un procedimiento según una cualquiera de las reivindicaciones 6 a 8, que comprende además la etapa de disminución de la resolución de al menos uno de la serie de fotogramas de imágenes antes de calcular la serie de valores.
10. Un procedimiento según cualquier reivindicación precedente en el que el procedimiento se realiza en tiempo real.
11. Un procedimiento según cualquier reivindicación precedente en el que la serie de valores comprende una serie de valores de complejidad.
12. Un producto de programa informático que comprende instrucciones legibles por ordenador, cuyo producto es operable, cuando se ejecuta en un ordenador en asociación operable con una cámara de vídeo, para realizar un procedimiento según cualquier reivindicación precedente.

ES 2 346 160 T3

13. Un sistema de detección de sucesos que comprende una cámara de vídeo, un procesador, y un medio de alerta; siendo operable la cámara de vídeo para grabar una grabación de vídeo, y para transmitir dicha grabación de vídeo al procesador, comprendiendo dicha grabación de vídeo una serie de fotogramas de imágenes; y comprendiendo el procesador medios para calcular, para cada uno de la serie de fotogramas de imágenes, un valor representativo de la
5 dinámica de la multitud, siendo calculado el valor en dependencia de dicho cada uno de la serie de fotogramas de imágenes, de tal manera que se calcule una serie de valores; y medios para analizar la serie de valores para identificar cualquier discrepancia que se produzca en la serie de valores de complejidad; y en el que el medio de alerta está configurado para alertar a un operador en respuesta al acaecimiento de cualquier discrepancia identificada por el
10 procesador.

14. Un sistema de detección de sucesos según la reivindicación 13 configurado para operar en tiempo real.

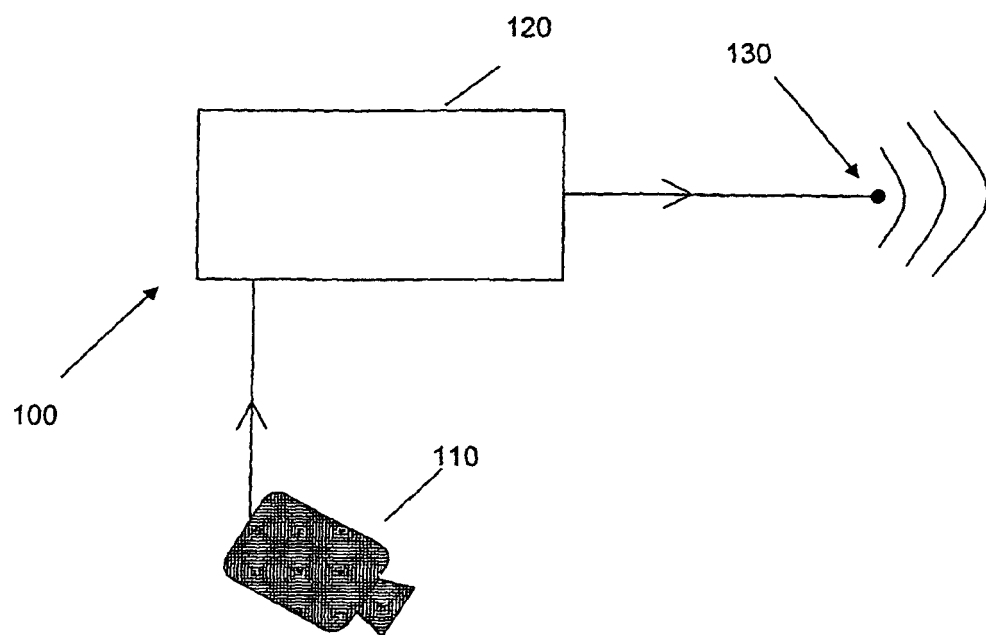


FIGURA 1

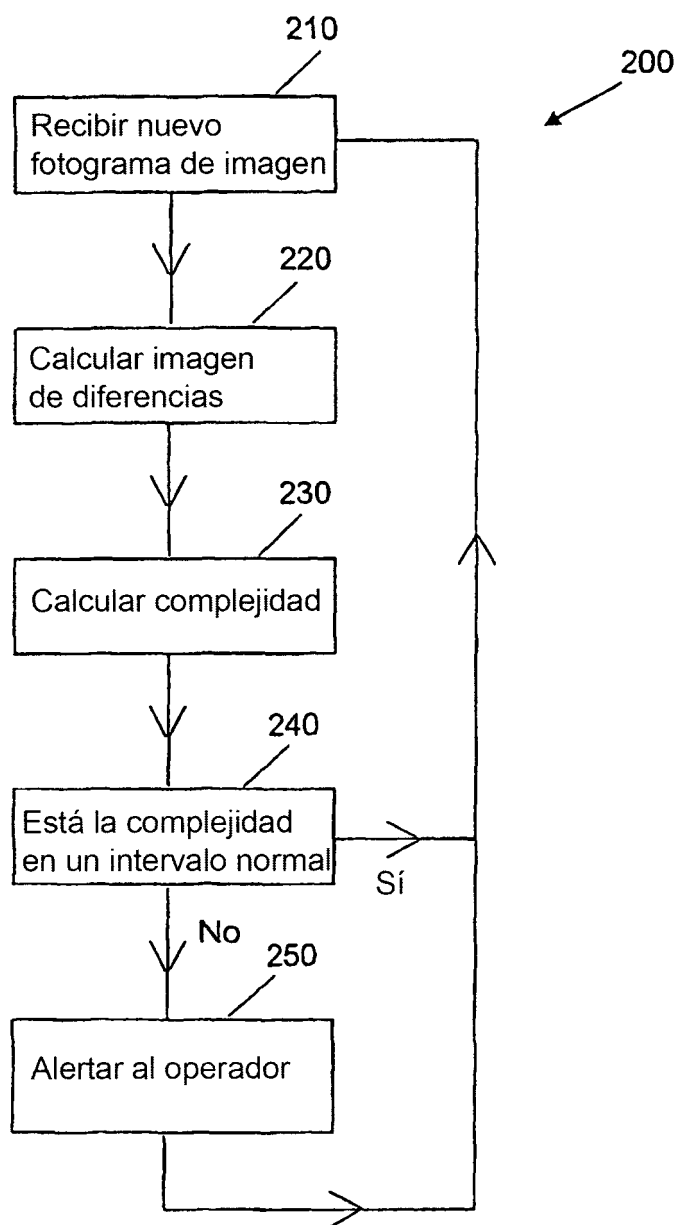


FIGURA 2

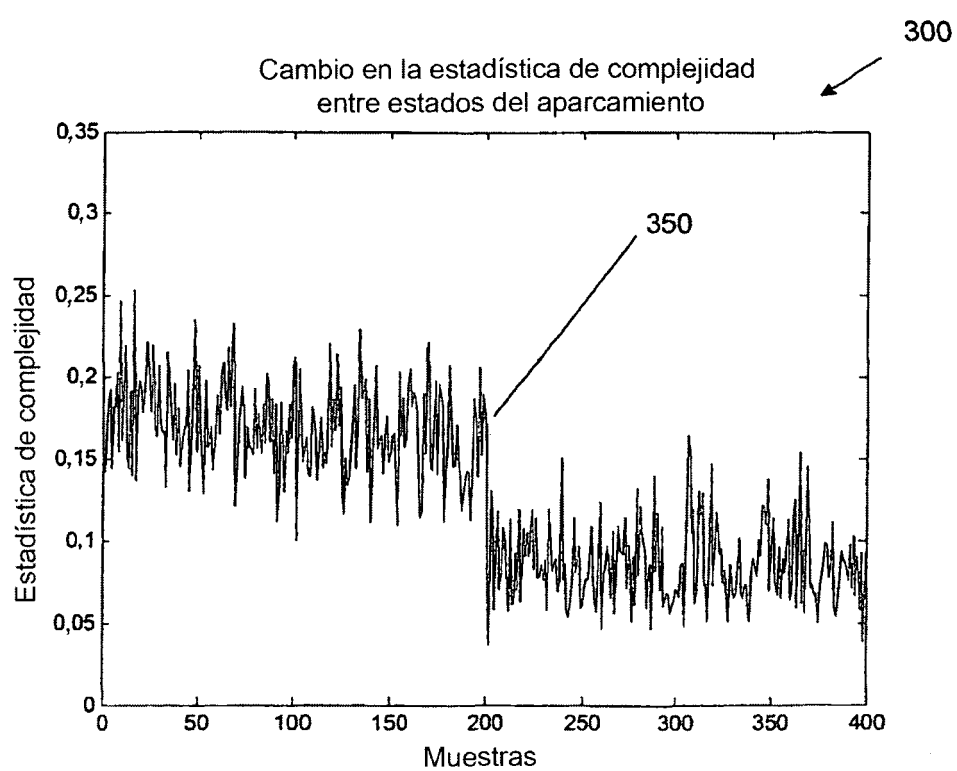


FIGURA 3

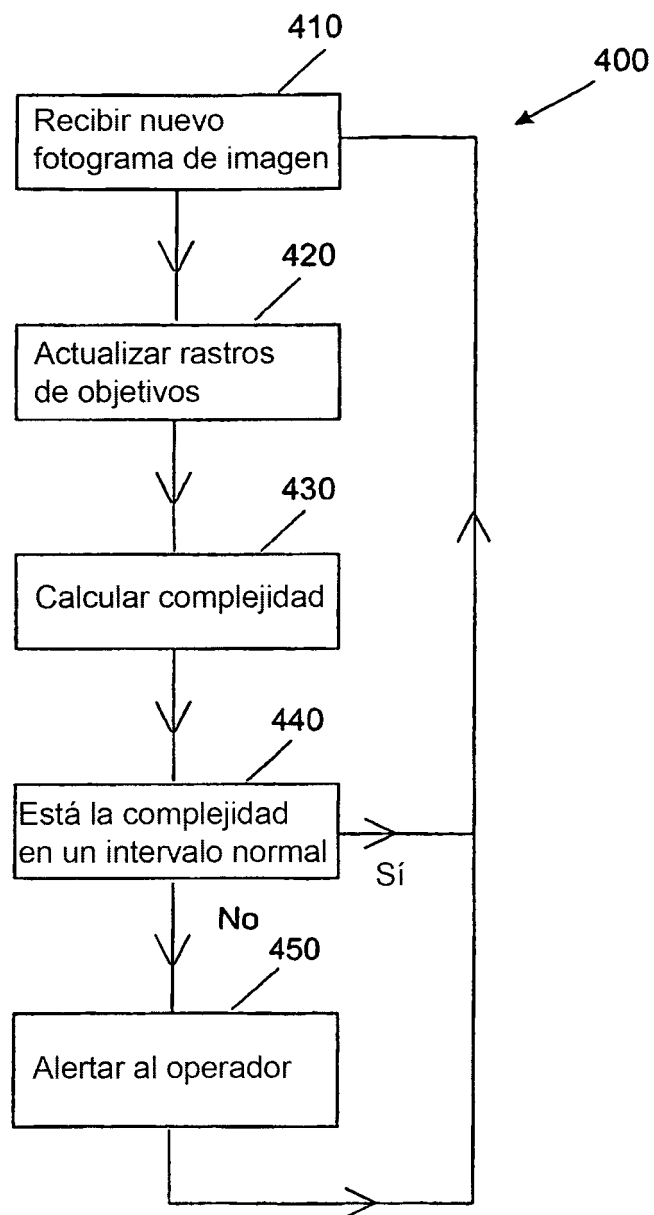


FIGURA 4

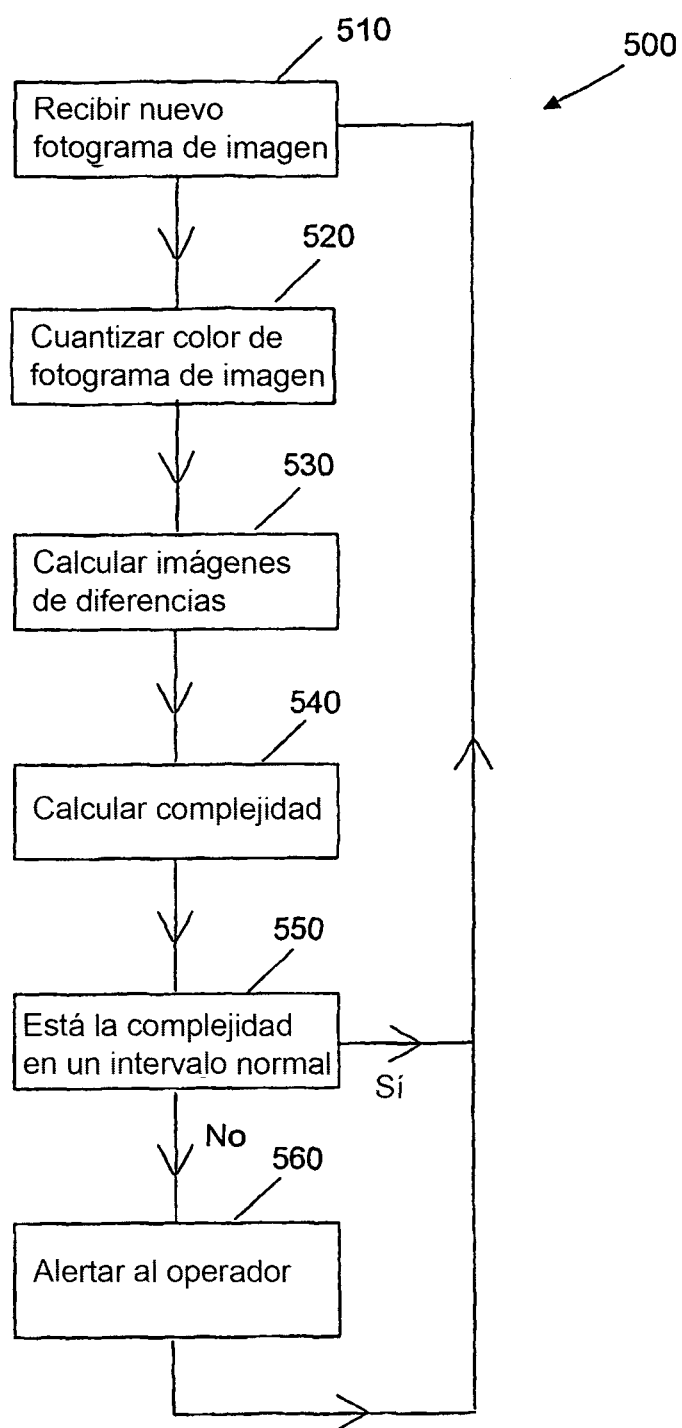


FIGURA 5