



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 267 208**

51 Int. Cl.:
G08B 13/194 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **99102604 .8**

86 Fecha de presentación : **11.09.1992**

87 Número de publicación de la solicitud: **0924671**

87 Fecha de publicación de la solicitud: **23.06.1999**

54 Título: **Análisis de imágenes.**

30 Prioridad: **12.09.1991 GB 9119588**

45 Fecha de publicación de la mención BOPI:
01.03.2007

45 Fecha de la publicación del folleto de la patente:
01.03.2007

73 Titular/es:
ELECTRONIC DATA SYSTEMS CORPORATION
5400 Legacy Drive H3-3A-05
Plano, Texas 75024, US

72 Inventor/es: **Abbott, James P.;**
Gent, Christopher Robert;
Thomas, William G.;
Sermon, Gerald y
Stubbington, Brian D.

74 Agente: **Torner Lasalle, Elisabet**

ES 2 267 208 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

Análisis de imágenes.

5 La presente invención se refiere a un método para el análisis de imágenes y un aparato para aplicar este método. La invención es particularmente adecuada para aplicaciones de vigilancia, aunque no exclusivamente.

10 La patente US 4 249 207 se refiere a un sistema de vigilancia en el cual una imagen recibida por una cámara de televisión es dividida electrónicamente en una matriz de celdas o "baldosas". Las baldosas están dimensionadas teniendo en cuenta la perspectiva, de forma que cada una de ellas cubre sustancialmente la misma superficie sea cual sea la parte particular de la imagen a la que correspondan respectivamente. Esto se ilustra en la figura 1 de las ilustraciones adjuntas. El sistema monitoriza todas las celdas para determinar si en cada una de ellas se produce un acontecimiento potencialmente significativo. Si en una celda se detecta un acontecimiento tal, se comprueban las celdas adyacentes y si se produce un acontecimiento en, por ejemplo, dos de ellas, se considera indicativo de una traza de una persona u objeto de interés en movimiento.

15 La presencia de un acontecimiento potencialmente significativo en una celda se determina en función de cambios en la intensidad media de la luz recibida de esa celda. Cada celda se compone de una matriz de pixels. Estos se analizan línea por línea y se computa la intensidad de la luz incidente integrada para esa celda. A lo largo del tiempo, se realiza continuamente una comparación ponderada entre la intensidad media presente y la intensidad media en un cuadro anterior. Si el resultado de esta comparación supera un umbral predeterminado, se considera que está en esa celda ocurriendo un acontecimiento significativo.

20 La mencionada patente de Estados Unidos declara que este sistema aporta buena discriminación contra falsos acontecimientos como sombras de nubes en movimiento. Sin embargo, el sistema es más adecuado para escenas generalmente vacías como la superficie entre dos verjas de perímetro situadas en un paisaje abierto. El sistema es mucho menos capaz de discriminar falsos acontecimientos en una escena activa, como puede encontrarse en un entorno urbano.

25 Desventajas específicas de este sistema conocido son:

- 30 a. Es excesivamente sensible a ciertas pequeñas variaciones en la imagen, por ejemplo, debidas a cambios en el tiempo, nubes, movimientos de los árboles, etc. Esto conduce a un gran número de falsas alarmas resultantes de cambios normales en una escena;
- 35 b. tiene baja sensibilidad a ciertas grandes variaciones en la imagen, por ejemplo, un pequeño cambio de luz repartido en toda una baldosa tiene el mismo efecto que un gran cambio de luz en parte de una baldosa. Esto puede hacer que no se detecten algunos acontecimientos de interés; y
- 40 c. grandes cambios en la imagen desensibilizan el sistema durante un tiempo. Como resultado, un retorno a una escena normal provoca una alarma de actividad, y esto reduce la capacidad de seguir movimientos y de identificar la forma de objetos, ya que las baldosas abandonadas por un objeto lento tienen la misma actividad que una baldosa en la que entra el objeto.

45 Los problemas mencionados explican por qué este sistema existente sólo es eficaz en escenas con muy poca variación en los niveles de luz. De hecho, sólo es eficaz para monitorizar escenas de interiores o escenas al aire libre donde no hay expectativa de movimiento, por ejemplo, zonas muertas entre verjas alrededor de una cárcel.

50 La presente invención en sus diversos aspectos presenta una ventaja significativa sobre la descrita en la mencionada patente de Estados Unidos, en tanto que aporta la capacidad de "aprender" la naturaleza de un fondo complejo. Esto le permite discriminar mejor los auténticos acontecimientos de interés en una escena por lo demás activa o muy llena.

55 Otro sistema ya conocido denominado "WISARD" es descrito por I. Aleksander *et al.* en *Sensor Review*, julio de 1984, pp. 120-124. Los pixels del interior de una celda son interrogados pseudoaleatoriamente. En esta última presentación se expone que la intensidad de la luz incidente sobre un pixel puede convertirse en una señal de escala de grises o alternativamente en una señal binaria.

60 En este último caso, el nivel de luz detectado se compara con un umbral y se adjudica un 1 o un 0 según si la intensidad está por debajo del umbral o no. Estas señales binarias a continuación se combinan en grupos, que forman punteros a una matriz almacenado en una memoria de acceso aleatorio (RAM). Cada vez que un elemento de la matriz es direccionado por un puntero, su valor se define como "on", y de otro modo permanece sin cambio. Este proceso se lleva a cabo a lo largo de un cierto número de cuadros.

65 De esta manera, con WISARD, a lo largo de sucesivos cuadros, los valores de ciertos elementos de la matriz se vuelven definidos. En efecto se trata de un "periodo de entrenamiento". Una vez que han transcurrido suficientes cuadros, la matriz se considera adecuadamente preparada y el sistema se pone en modo de funcionamiento y se generan punteros de la misma manera. Después, si un puntero direcciona un elemento que está definido, esto quiere decir que las intensidades en los pixels que forman el puntero son semejantes a valores que se presentaron en la serie

de entrenamiento. Los punteros que direccionan elementos que no están definidos indican que la situación es distinta a la del periodo de entrenamiento. De esta manera WISARD detecta cuándo hay cambios en la escena.

Se observará que WISARD sólo puede comparar la situación presente con lo ocurrido durante su periodo de “entrenamiento” finito y que no tiene memoria de acontecimientos ocurridos antes o después. Asimismo, acontecimientos esporádicos durante el periodo de entrenamiento tienen tanto efecto como los acontecimientos comunes. En consecuencia, WISARD no puede perfeccionar continuamente su registro del fondo y por tanto tiene limitaciones en su capacidad para discriminar entre acontecimientos significativos y fondo.

La presente invención es capaz de perfeccionar continuamente su registro del fondo y por ello puede aportar mejor discriminación que WISARD.

Reconocemos, a modo de antecedente, que US-A-4774570 revela un sistema para el procesamiento de una señal de vídeo que detecta un cambio en los datos de vídeo de la señal de vídeo proporcionada por una fuente comparando la señal de vídeo de la fuente con una señal de referencia variable derivada de un número predeterminado de campos muestreados de la señal de vídeo de la fuente pero que excluye los datos de vídeo de aquellos pixels cuyos datos de vídeo fluctúan durante los cuadros muestreados.

La invención aporta un método para analizar una señal de imagen electrónica, tal como se define en la reivindicación 1, y un dispositivo correspondiente tal como se reivindica en la reivindicación 10.

Los ejemplos de realización preferentes de la invención pueden incluir:

- a) procesar la señal de imagen para considerarla representación de una pluralidad de celdas cada una de las cuales comprende una pluralidad de pixels;
- b) para cuadros consecutivos de la imagen monitorizada, generar valores característicos según el contenido detallado de la imagen; y
- c) para cuadros de imagen consecutivos, acumular los valores característicos de forma que en un solo modo de operación se actualice una historia temporal de la imagen, almacenada en una memoria, y se produzca una indicación que señale la presencia o no presencia de un acontecimiento no asociado con el fondo de imagen.

El dispositivo puede constar de medios electrónicos para efectuar los pasos (a) – (c) del método mencionado.

La presente invención en general mitiga los citados problemas de la técnica anterior y puede extender las posibilidades de monitorización de escenas a una mayor variedad de escenas. En particular, la invención es adecuada para monitorizar escenas con una gran cantidad de actividad normal, por ejemplo, para monitorizar escenas de calle activas a fin de identificar acontecimientos específicos. Ejemplos de tales acontecimientos pueden ser:

- llegada de una obstrucción a una línea ferroviaria,
- llegada o partida de un vehículo a través de un portón,
- llegada o partida de una persona a través de una puerta,
- formación de aglomeración en un andén de vía férrea.

Aunque en su definición más amplia la invención requiere solamente una celda, es preferible que la imagen se represente como una pluralidad de celdas, y más preferiblemente ajustada según la perspectiva tal como aquí se describe.

De preferencia al menos un par de pixels es seleccionado de esta manera por cada una de una pluralidad de celdas. También es muy preferible seleccionar de la misma manera una pluralidad de pixels por cada celda del campo de imagen.

Una vez que se han seleccionado uno o más pares de pixels en una celda determinada, los mismos pixels pueden utilizarse después en cuadros subsiguientes hasta el fin de la sesión operativa. En una sesión subsiguiente, se podrían utilizar los mismos pares o se podrían seleccionar pares distintos para el resto de esa sesión.

Cuando en una celda determinada se utiliza una pluralidad de pares de pixels, es posible que un mismo pixel sea utilizado en más de un par, aunque esto no es absolutamente necesario. También es posible, aunque no obligatorio, que alguno o algunos de los pixels no se utilicen en ningún par.

Para celdas distintas, el par o los pares de pixels podrían elegirse en las mismas posiciones respectivas en cada celda, aunque tampoco esto es necesario.

ES 2 267 208 T3

También preferimos utilizar los valores característicos para generar punteros que direccionen la memoria que contiene la historia temporal de la imagen. Para este fin, un medio particularmente ventajoso consiste en generar una serie de punteros $a_1, \dots, a_m$, cada uno de ellos de n bits binarios, y disponer secuencialmente estos bits según los respectivos valores binarios de las señales que resultan de las comparaciones de los pares de pixels. En otras palabras, los pares $1, \dots, n$ de los pares de pixels se utilizan para definir los valores de $a_1 [1, \dots, n]$, después los pares $(n+1), \dots, 2n$ se utilizan para definir $a_2 [1, \dots, n]$ y así sucesivamente.

En consecuencia, un método preferido comprende:

- a) procesar la señal de imagen para considerarla representación de una o más celdas cada una de las cuales comprende una pluralidad de pixels;
- b) para cuadros consecutivos de la imagen monitorizada, generar valores característicos según el contenido detallado de la imagen; y
- c) generar una serie de punteros $a_1, \dots, a_m$, cada uno de n bits binarios, y disponer secuencialmente estos bits según los valores característicos.

El dispositivo para analizar una señal de imagen de preferencia comprende medios electrónicos para efectuar los pasos (a)-(c) del método.

Los valores de n y m pueden ser cualesquiera números convenientes y pueden ser los mismos o diferentes, pero $n \times m$ será igual al número de valores característicos. Por consiguiente, en este contexto $n \times m$ será igual a la mitad del número de pixels seleccionados en cada celda.

En un ejemplo de realización preferente, estos punteros se utilizan para direccionar una memoria que comprende una matriz de direcciones que comprende las filas $1, \dots, m$ y las columnas $1, \dots, 2^n$. Cuando se utilizan estos valores característicos para definir los valores de los punteros, estos últimos direccionan la memoria como se indica a continuación. El puntero a_1 direcciona la fila 1. El puntero contiene un número binario de un valor situado entre 0 y 2^n . El puntero apunta a la columna de la fila 1 que tiene un número de columna correspondiente al valor en a_1 . El valor en esa dirección se incrementa a continuación en 1 (al principio todas las direcciones se habrán fijado en cero). Esto se hace para todas las filas hasta m . En el cuadro siguiente, se repite todo el proceso, incrementando el valor en 1 en una dirección particular apuntada en cada fila.

Por consiguiente, el método de preferencia comprende:

- a) procesar la señal de imagen para considerarla representación de una o más celdas cada una de las cuales comprende una pluralidad de pixels;
- b) para cuadros consecutivos de la imagen monitorizada, generar una serie de punteros cada uno de los cuales contenga un valor relacionado con el contenido detallado de la imagen; y
- c) utilizar los punteros para direccionar una matriz de memoria y para cuadros sucesivos, actualizando el valor de una dirección en cada fila de la matriz indicada por el puntero correspondiente a esa fila.

En un ejemplo de realización preferente, la serie de punteros comprende los punteros $a_1, \dots, a_m$, cada uno de ellos de n bits binarios, y la matriz de memoria tiene m filas y 2^n columnas.

Se apreciará que en términos de una matriz de memoria, los términos “filas” y “columnas” son puramente arbitrarios y por lo tanto pueden ser transpuestos.

Otra característica ventajosa adicional es la manera en que se utilizan los valores de la memoria para discriminar entre el fondo y acontecimientos desacostumbrados o inesperados. En cuadros sucesivos, para cada fila de la memoria, el valor máximo en cualquier dirección de la fila está determinado, y el valor actualmente apuntado también es conocido. Para todas las filas, se calcula la relación de la suma de valores en las direcciones actualmente apuntadas, dividida por la suma de los máximos. Esta relación se compara con una media ponderada en tiempo para la misma relación de cuadros anteriores. Si esta relación es menor que una proporción elegida de la media ponderada en tiempo, se considera indicativa de la ocurrencia de un acontecimiento desacostumbrado o inesperado en la celda pertinente.

Así pues, el método de preferencia comprende:

- a) procesar la señal de imagen para considerarla representación de una o más celdas cada una de las cuales comprende una pluralidad de pixels;
- b) para cuadros consecutivos de la imagen monitorizada, actualizar una dirección en cada fila de una matriz de memoria en correspondencia con el contenido detallado de la imagen; y

- c) en todas las filas, sumar los valores en las direcciones que son actualizadas y sumar el valor máximo de cada fila, sometiendo la relación de las dos sumas a una media ponderada en tiempo con respecto a las relaciones obtenidas de cuadros anteriores, comparar la relación con la media ponderada en tiempo y utilizar el resultado para determinar la presencia o ausencia de un acontecimiento significativo en la imagen.

5 Cuando la presente invención en cualquiera de sus aspectos se manifiesta como un dispositivo antes que como un método, el dispositivo puede implementarse como hardware/firmware [hardware con soporte lógico incorporado] dispuesto para operar bajo el control de un software apropiado, o puede realizarse sólo como hardware/firmware configurado mediante puertas lógicas, flip-flops, etc., dispuestos para realizar las funciones requeridas.

10 A continuación se ilustra la invención por medio de la siguiente descripción de un ejemplo de realización preferente y con referencia a las ilustraciones adjuntas, en las que:

la figura 1 muestra la disposición de los límites de las celdas para tener en cuenta la perspectiva;

15 la figura 2 es una representación esquemática de un dispositivo según la presente invención, para computar niveles de actividad;

la figura 3 es un diagrama de bloques del hardware del dispositivo ilustrado en la figura 2;

20 la figura 4 es un diagrama de flujo que muestra la lógica aplicada a cada celda para determinar la existencia o ausencia de una huella; y

25 la figura 5 muestra las relaciones espaciales de celdas locales para movimiento desde la parte superior derecha a la inferior izquierda.

Antes de describir en detalle el ejemplo de realización preferente es conveniente definir primero los siguientes términos, cuyas definiciones se aplican sólo al ejemplo de realización descrito y no limitan el alcance de la invención reivindicada:

30 *Pixel*

Un pixel (elemento de imagen) define un nivel de gris en un punto de una imagen. Típicamente se representan hasta 256 niveles de gris en 8 bits.

35 *Imagen*

Una imagen es una matriz de pixels, típicamente 512 x 512.

40 *Baldosa*

Una baldosa (celda) es un subconjunto de una imagen compuesto por N pixels. N debe ser un número par. Los pixels de una baldosa se denotan:

$$P_1, P_2, \dots, P_N.$$

Se define $M = N/2$

Ordenación de los pixels

50 Los pixels se ordenan aleatoriamente y se dividen en dos mitades:

$$P_1, P_2, \dots, P_M$$

$$Q_1, Q_2, \dots, Q_M$$

55 *Generación de punteros*

Generar bits $b_1, b_2, b_3, \dots, b_M$ según:

60 si $p_1 > q_1$ entonces $b_1 = 0$, si no $b_1 = 1$,

si $p_2 > q_2$ entonces $b_2 = 0$, si no $b_2 = 1$,

...

65 si $p_M > q_M$ entonces $b_M = 0$, si no $b_M = 1$.

Se define $m = M/n$.

ES 2 267 208 T3

Generar m direcciones binarias de n bits:

$$a_1 = [b_1 \ b_2 \ b_3 \ \dots \ b_n],$$

$$a_2 = [b_{n+1} \ b_{n+2} \ \dots \ b_{2n}],$$

...

$$a_m = [\dots b_{M-1} \ b_M],$$

Las m direcciones $a_1, a_2, \dots, a_m$ caracterizan la imagen instantánea en la baldosa y actúan como punteros a la matriz de memoria definida a continuación.

Memoria

La memoria es una matriz de $m \times 2n$ de valores de datos que contiene información sobre la historia pasada de la escena. Esta matriz se actualiza continuamente según se desarrolla el proceso.

Además, el valor máximo de cada fila de Memoria se mantiene en un máximo de la matriz. La notación es que, para $j = 1, \dots, m$:

$$\max_j = \text{máximo de Memoria } [j, i], \ i = 1, \dots, 2^n.$$

La dirección a_j apunta a una celda en la “jota-ésima” fila de la memoria. Esta celda se denota:

$$\text{Memoria } [j, a_j].$$

Velocidad de cuadros

Los cuadros de imagen se procesan a una velocidad determinada, que será específica según la aplicación. El tiempo en que se procesa el cuadro actual se denota con t , y el tiempo en que se procesó el cuadro anterior se denota $t-1$.

Parámetros

Los parámetros f , Límite Superior [o “UpperLimit”] y k se seleccionan para determinar las constantes de tiempo sobre las que se requiere un historial.

Umbral [o “Threshold”] es un parámetro que determina la sensibilidad de detección de la técnica.

El propósito de los mecanismos es computar, para cada baldosa, un nivel de actividad que indique si la imagen se encuentra en un estado normal.

Los mecanismos se implementan por los siguientes medios:

- a. una cámara de TV que captura imágenes y las suministra como salida en un formato estándar,
- b. un capturador de cuadros de imagen que captura las imágenes a la velocidad de cuadros de la cámara, digitaliza cada cuadro y lo almacena en un buffer de cuadros digitales,
- c. un procesador de ordenador que realiza la generación de punteros por medio de
 - acceder al buffer de cuadros digitales y
 - computar punteros en una matriz de memoria,
- d. un almacén digital, la matriz de memoria, para acumular estadísticas sobre las imágenes.
- e. un procesador de ordenador que utiliza los punteros para incrementar la matriz de memoria (actualización de la memoria),
- f. un procesador de ordenador que accede a la matriz de memoria para computar mediciones de actividad en la imagen.

Los procesadores de ordenador en c), e) y f) pueden ser un mismo procesador, o bien estas actividades pueden distribuirse entre varios procesadores.

La ejecución del proceso va precedida por la definición de baldosas y órdenes de pixels. La elección y las dimensiones de las baldosas son específicas según la aplicación. Durante la ejecución, se capturan imágenes en el buffer de cuadros a la velocidad de cuadros de la cámara. A continuación, los procesadores de ordenador realizan las siguientes acciones con respecto a cada cuadro analizado. La computación se describe a continuación y se realiza para cada baldosa y la relación entre los procesos separados se representa en la figura 2 y la naturaleza precisa del hardware de ordenador se muestra en la figura 3.

Generación de punteros

Generar punteros a la memoria, de la siguiente manera.

Para cada $j = 1, \dots, m$

Generar dirección a_j , como ya se ha descrito.

Actualización de la memoria

Realizar actualización de la memoria para mantener un historial de imágenes pasadas de la siguiente manera:

Para cada $j = 1, \dots, m$

Incrementar memoria $[j, a_j]$ en 1, y

actualizar $\max_j$ para que sea el máximo elemento en la “jota-ésima” fila de memoria.

Si memoria $[j, a_j] > \text{LímiteSuperior}$ entonces escalar la jota-ésima fila de memoria, a saber:

Computar memoria $[j, i] = f \cdot \text{memoria}[j, i]$, $i = 1, \dots, 2^n$, y computar $\max_j = f \cdot \max_j$.

Evaluación de la imagen

Realizar una evaluación de la imagen para determinar el estado actual de la imagen computando una puntuación, que se comparará con una puntuación media.

Sumar los valores memoria $[j, a_j]$, $j = 1, \dots, m$, puntuación denotada.

Sumar los valores $\max_j$, $j = 1, \dots, m$, máximo denotado.

Calcular un valor A_t que define la actividad instantánea en la imagen en el tiempo t , computando: $A_t = \text{puntuación}/\text{máximo}$,

(nota: A_t puede asumir valores entre 0 y 1).

Calcular la actividad media en la imagen en el tiempo t , computando:

$$\text{Media}_t = k \cdot \text{Media}_{t-1} + (1-k) \cdot A_t$$

(nota: Media_t puede asumir valores entre 0 y 1).

Calcular la Actividad_t medida en el tiempo t según:

$$\text{Actividad}_t = (\text{Media}_t - A_t)$$

(nota: Actividad_t puede asumir valores entre -1 y 1).

Si Actividad_t > Umbral, entonces indica que la baldosa está activa.

La técnica general aquí descrita deriva un único valor (Actividad_t) que indica si la escena actual ha cambiado desde las escenas recientes. El marco de tiempo en el cual se determina el cambio se especifica mediante la elección de los parámetros f y k . Por ejemplo: para una escena con un fondo que cambia lentamente, se podrían utilizar valores de f y k próximos a 1. Si la escena de fondo cambia rápidamente, serían necesarios valores más pequeños.

A diferencia de otros sistemas ya conocidos, la técnica es sensible a auténticos cambios en la escena y discrimina entre estos y los cambios típicos aceptables en la escena. Esto es así porque la matriz de memoria recuerda las características de una escena normal.

El cálculo de los punteros tal como se ha descrito es muy sensible a cambios locales en una baldosa, pero es insensible a cambios de luz globales en una baldosa. Cada puntero hace que se incremente una dirección de la memoria. Así, escenas que son habituales harán que ciertas direcciones de la memoria se incrementen con frecuencia, mientras que otras direcciones se incrementarán rara vez.

Mientras una escena se encuentra en un estado común (advuértase que los niveles de luz pueden estar cambiando significativamente y que el contenido de la escena también puede estar cambiando), los punteros apuntarán a direcciones de la memoria que contienen valores altos. Entonces los valores de actividad derivados serán pequeños o negativos.

Sin embargo, cuando se produce un cambio desacostumbrado en la imagen, ello afecta a varios punteros que pasarán a direccionar direcciones de memoria que contienen valores pequeños. Entonces el valor de actividad resultante será alto (es decir, cercano a 1).

De esta manera, la técnica identifica de inmediato los cambios en la imagen. Un cambio local, causado por ejemplo por una persona moviéndose en la escena, puede afectar únicamente a un pequeño número de pixels. Supongamos, por ejemplo, que diez pixels resultan afectados; en tal caso, por la forma en que se generan los punteros, este cambio afectará típicamente a diez punteros. Así, un pequeño cambio local ejerce un gran efecto en la puntuación de actividad.

Asimismo, debido a la manera en que la memoria recuerda escenas pasadas, cuando la actividad termina la actividad computada revierte de inmediato a un valor pequeño. De este modo, la presencia o ausencia de un acontecimiento en una baldosa puede monitorizarse con precisión.

El resultado de los cálculos es derivar, para cada imagen procesada, un atributo Activo, que puede asumir el valor verdadero o falso, para cada baldosa. Estos valores de Activo pueden ser la entrada de un proceso que determina la naturaleza de un acontecimiento en la escena. Dicho proceso puede seguir la pista de acontecimientos y computar atributos de forma y tamaño de objetos de la escena. La fiabilidad del procedimiento reivindicado para identificar baldosas activas, en contraposición a cambios en la luz del ambiente, permite realizar con éxito un Reconocimiento de Acontecimientos.

En general, el número de pixels de una baldosa es específico para cada aplicación. El tamaño de la baldosa dependerá del tamaño del objeto por identificar en la escena. Asimismo, debido a la perspectiva, variará el número de pixels en cada baldosa.

De la misma manera, el número de pares de pixels en una baldosa, M , también es específico de la aplicación y específico de la baldosa. En general, no es necesario considerar todos los pixels de una baldosa y por lo general se puede submuestrear los pixels. Sin embargo, dada la definición del presente ejemplo de realización preferente, M no debería ser menor que 8.

A fin de aprovechar el direccionamiento de memoria de 8 bits de los ordenadores modernos, n por lo general se establece en 8. Asimismo, los elementos de la memoria son típicamente direcciones de 8 bits, de modo que LímiteSuperior se establecería normalmente en 255.

Cada baldosa se procesa separadamente y genera n punteros hacia la matriz de memoria. En la práctica, puede obtenerse una mayor sensibilidad repitiendo la generación de pixels cierto número de veces para definir más punteros. Típicamente se utilizan tres reordenamientos de pixels, para generar $3n$ punteros, lo que implica una matriz de memoria de $3m \times 2^n$ elementos.

El ritmo a que se procesan las imágenes es típicamente de 25 cuadros por segundo, aunque podría ser menor según el ritmo en que se producen acontecimientos en la escena. El mecanismo incorpora flexibilidad de forma que el sistema pueda implementarse en un ordenador o en varios ordenadores según la velocidad de cuadros requerida.

Los mecanismos de procesamiento previamente definidos muestran cómo se evalúa cada baldosa para ver si es activa. A continuación se puede realizar la determinación de acontecimientos en la escena mediante la consideración de cómo varían temporalmente y espacialmente los valores de actividad.

A diferencia de otros sistemas anteriormente conocidos, la determinación de acontecimientos puede realizarse con éxito aplicando una lógica sencilla basada en los valores de actividad. Esto es así porque los mecanismos reivindicados reducen la tasa de falsas alarmas, es decir, la tasa de baldosas que se indican como activas cuando no lo son.

Los valores activos pueden utilizarse para seguir la pista de objetos que pasan a través de la escena. En este caso, una baldosa tiene el atributo adicional Seguir ["Track"], que puede asumir los valores verdadero o falso, para indicar si un objeto está pasando actualmente por esa baldosa.

La lógica aplicada a cada baldosa se indica en el diagrama de flujo presentado en la figura 4. Las pruebas se aplican en el paso temporal t en base al valor Activo en el tiempo t y los valores Seguir en el tiempo $t-1$.

ES 2 267 208 T3

Los seguimientos (huellas) en una dirección arbitraria se determinan definiendo como baldosas locales:

- a) todas las baldosas contiguas a la baldosa de interés, incluyendo
- b) la baldosa misma.

Para detectar movimiento en una dirección particular, “baldosas locales” se define en consecuencia. Por ejemplo, como se indica en la figura 5, para seguir un movimiento desde la parte superior derecha hacia la parte inferior izquierda de una escena, se elegirían como baldosas locales:

- a) baldosas contiguas encima y a la derecha de la baldosa de interés, y
- b) la baldosa misma.

Las baldosas que tienen un valor de Seguir verdadero pueden indicarse en una imagen para mostrar objetos en movimiento. Asimismo, se puede utilizar una combinación de dirección de desplazamiento para detectar objetos que presentan un comportamiento específico, por ejemplo:

- a) movimiento hacia abajo seguido por un movimiento a través de la imagen,
- b) un objeto que se mueve en una dirección específica y que después se detiene en la imagen (ver más adelante),
- c) un objeto que se mueve a través de la escena y que después pasa por una puerta, o viceversa.

Los valores Activo y Seguir pueden utilizarse para detectar objetos que se mueven hasta el interior de la escena y después se detienen. Esto se consigue contando para cada baldosa el número de cuadros consecutivos en que la baldosa tiene un valor de Seguir verdadero. El acontecimiento se identifica cuando este número supera un umbral prescrito.

Este proceso puede identificar, por ejemplo, un vehículo que aparca en una calle con mucho movimiento, o una persona que se demora en una acera concurrida.

Varios aspectos de la presente invención también pueden incorporarse en el sistema WISARD ya conocido y anteriormente descrito. A continuación se explicará cómo se puede adaptar WISARD de esta manera.

El sistema WISARD original implica la generación de punteros hacia una matriz de memoria binaria durante un periodo fijo, es decir, un número fijo de cuadros. El sistema no tiene éxito porque:

- a) no se puede adaptar a escenas cambiantes, y por lo tanto sólo es aplicable a escenas fijas.
- b) Es muy sensible a ciertos pequeños cambios de luz a causa de la manera en que sustituye una imagen de escala de grises por una imagen binaria (blanco/negro).

Adecuadamente modificada, una combinación de la Generación de Punteros WISARD junto con la Actualización de Memoria y la Evaluación de Imagen de la presente invención puede ser más eficaz. La siguiente descripción identifica estas modificaciones presentando un sistema de Generación de Punteros que puede ser muy eficaz siempre y cuando las imágenes procesadas tengan un alto contraste (es decir, una amplia gama de niveles de luz). La notación es esencialmente la misma que la utilizada aquí anteriormente y tiene los mismos significados.

Dado un umbral T_2 , se generan bits $b_1, b_2, \dots, b_n$ de la siguiente manera:

Si $P_1 > T_2$ entonces $b_1 = 0$, si no $b_1 = 1$,

si $P_2 > T_2$ entonces $b_2 = 0$, si no $b_2 = 1$,

...

si $P_N > T_2$ entonces $b_N = 0$, si no $b_N = 1$.

Generar l direcciones binarias de n bits, donde $l = N/n$

$a_1 = [b_1 \ b_2 \ \dots \ b_n]$,

$a_2 = [b_{n+1} \ b_{n+2} \ \dots \ b_{2n}]$,

$a_l = [\dots \ b_{N-1} \ b_N]$,

que apunten a una matriz de memoria de $l \times 2^n$.

ES 2 267 208 T3

Este método para la generación de punteros es sensible a la elección del umbral T_2 . En particular, un objeto que sea de un nivel de gris similar al fondo puede no tener ningún efecto sobre los punteros. Esta deficiencia se puede reducir manteniendo una segunda matriz de memoria de 1×2^n y derivando una segunda serie de punteros de la siguiente manera:

Si $T_1 < P_1 < T_3$ entonces $b_1 = 0$, si no $b_1 = 1$,

si $T_1 < P_2 < T_3$ entonces $b_2 = 0$, si no $b_2 = 1$,

si $T_1 < P_N < T_3$ entonces $b_N = 0$, si no $b_N = 1$,

De la misma manera descrita, se genera una segunda serie de 1 direcciones de n bits a partir de esta segunda serie de bits $b_1, b_2, \dots, b_n$, que actúan como punteros hacia la segunda matriz de memoria. A continuación se promedian los valores de Actividad derivados de las dos matrices de memoria para obtener una Actividad acumulada.

Después, si los umbrales T_1, T_2, T_3 se eligen cuidadosamente, un sistema basado en estos punteros también puede tener éxito.

Los umbrales deben elegirse de manera que se adapten a variaciones en el nivel de luz en toda la imagen. Una manera eficaz consiste ajustarlos de manera que, en toda la imagen (es decir, en todas las baldosas):

25% de los valores de pixels sean menores que T_1 ,

25% de los valores de pixels estén entre T_1 y T_2 ,

25% de los valores de pixels estén entre T_2 y T_3 ,

25% de los valores de pixels sean mayores que T_3 .

Después, T_1, T_2 y T_3 deben ajustarse para cada imagen que se procese.

Obviamente, este sistema puede ampliarse para que se adapte a más niveles de luz introduciendo más umbrales. Estas técnicas de Generación de Punteros son ejemplos específicos de un proceso general que genera bits b_j (que asumen los valores 0 o 1) como función de pares de pixels y parámetros. Así, los ejemplos citados adoptan la forma:

si $p > q$ entonces $b = 0$, si no $b = 1$

si $p > T$ entonces $b = 0$, si no $b = 1$

Otro ejemplo es:

si $p > q + \epsilon$ entonces $b = 0$, si no $b = 1$

donde ϵ es un valor de parámetro fijo determinado por el nivel de ruido presente en la imagen (debido a efectos de cámara y de transmisión).

REIVINDICACIONES

1. Un método para analizar una señal de imagen electrónica asociada con una imagen que comprende una pluralidad de pares de pixels para monitorizar, método que consta de:

procesar la señal de imagen como una o más celdas cada una de las cuales consta de una pluralidad de pixels;

caracterizado por generar, para cada cuadro consecutivo de la imagen que está siendo monitorizada, valores característicos seleccionando aleatoriamente o pseudoaleatoriamente pares de pixels en la celda o celdas dentro de un cuadro y, para cada par, establecer una señal binaria en dependencia de si la intensidad de la luz incidente en un pixel predeterminado del par de pixels es mayor o no que la intensidad de la luz incidente en el otro.

2. Un método según la reivindicación 1, en el que se selecciona al menos un par de pixels por cada celda de una pluralidad de celdas.

3. Un método según la reivindicación 2, en el que se selecciona una pluralidad de pixels por cada celda de un campo de imagen.

4. Un método según la reivindicación 1, en el que se utilizan los mismos pixels en cada cuadro hasta un fin de una subsiguiente sesión operativa.

5. Un método según la reivindicación 1, que además consta de:

para cuadros de imagen consecutivos, acumular los valores característicos de manera que en un solo modo operativo una historia temporal de la imagen almacenada en una memoria, en la que cada celda tiene su propio bloque de memoria, se actualiza durante la monitorización de la imagen y se produce una indicación que indica la presencia o la ausencia de un acontecimiento no asociado con el fondo de la imagen.

6. Un método según la reivindicación 1, en el que la imagen se representa como una pluralidad de celdas, y dicha pluralidad de celdas está ajustada para la perspectiva.

7. Un dispositivo para analizar una señal de imagen electrónica asociada con una imagen que comprende una pluralidad de pares de pixels para monitorizar, dispositivo que consta de:

medios para procesar la señal de imagen como una pluralidad de celdas cada una de las cuales consta de una pluralidad de pixels;

caracterizado por unos medios para generar valores característicos, para cada cuadro consecutivo de la imagen que está siendo monitorizada, seleccionando aleatoriamente o pseudoaleatoriamente pares de pixels en la celda o celdas dentro de un cuadro y, para cada par, establecer una señal binaria en dependencia de si la intensidad de la luz incidente en un pixel predeterminado del par de pixels es mayor o no que la intensidad de la luz incidente en el otro.

8. Un dispositivo según la reivindicación 7, en el que dichos medios para generar valores característicos están adaptados para seleccionar al menos un par de pixels por cada celda de una pluralidad de celdas.

9. Un dispositivo según la reivindicación 7, en el que dichos medios para generar valores característicos están adaptados para seleccionar una pluralidad de pixels por cada celda de un campo de imagen.

10. Un dispositivo según la reivindicación 7, en el que dichos medios para generar valores característicos están adaptados para utilizar los mismos pixels en cada cuadro hasta un fin de una subsiguiente sesión operativa.

11. Un dispositivo según la reivindicación 7, en el que dichos medios para generar valores característicos consta de un capturador de cuadros de imagen, una memoria y uno o más procesadores de ordenador.

12. Un dispositivo según la reivindicación 7, que además consta de:

medios para acumular los valores característicos, para cuadros de imagen consecutivos, de manera que en un solo modo operativo una historia temporal de la imagen almacenada en una memoria, en la que cada celda tiene su propia matriz de memoria, se actualiza durante la monitorización de la imagen y se produce una indicación que indica la presencia o la ausencia de un acontecimiento no asociado con el fondo de la imagen.

13. Un dispositivo según la reivindicación 7, en el que la imagen se representa como una pluralidad de celdas, y dicha pluralidad de celdas está ajustada para la perspectiva.

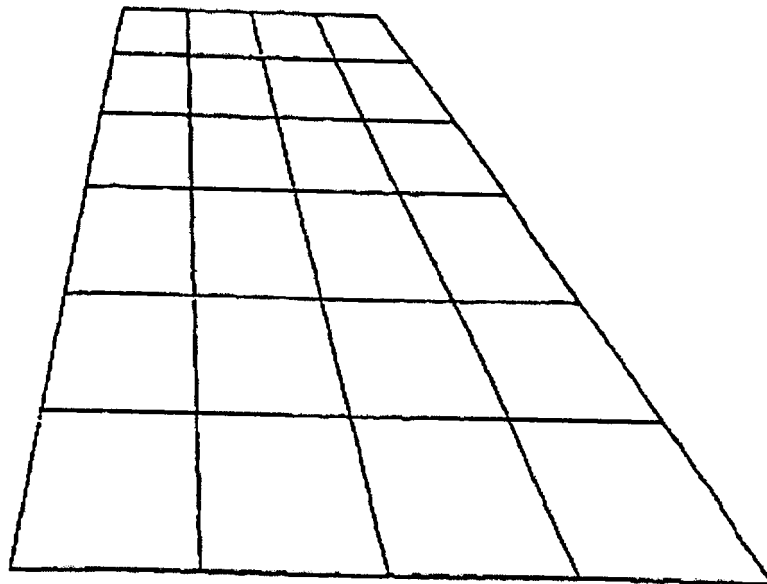
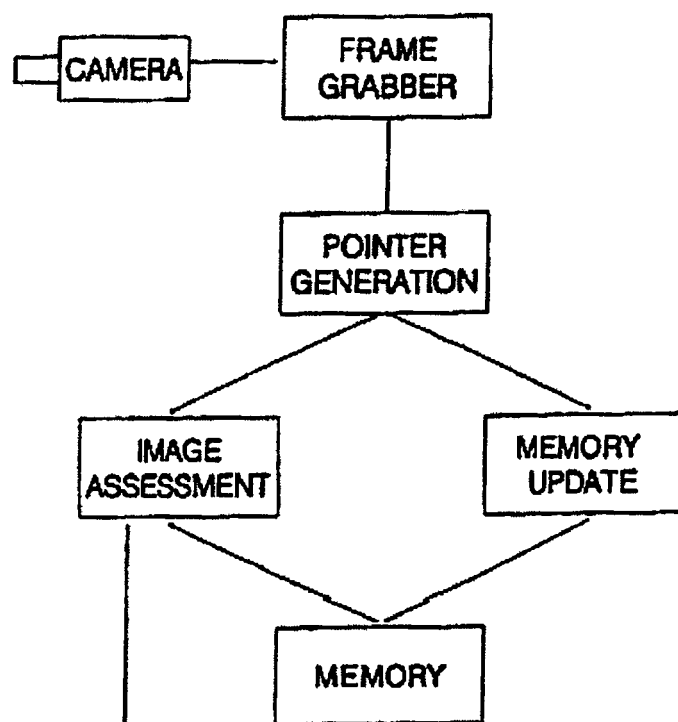


Figura 1



Activity Levels

Figura 2

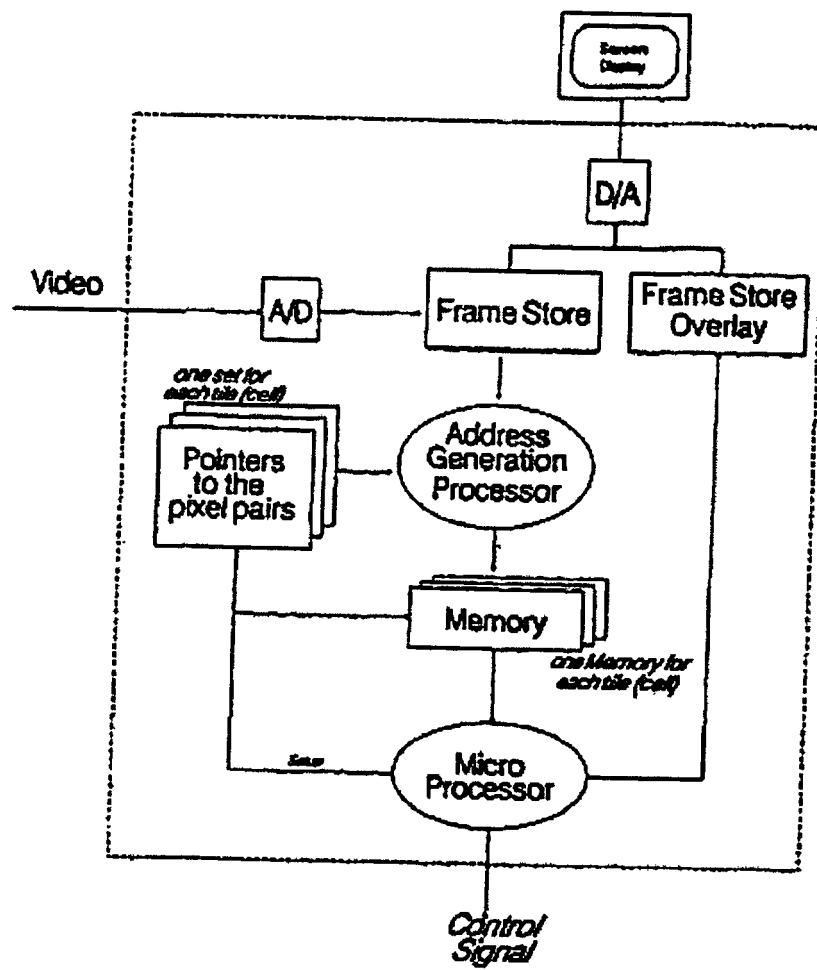


Figura 3

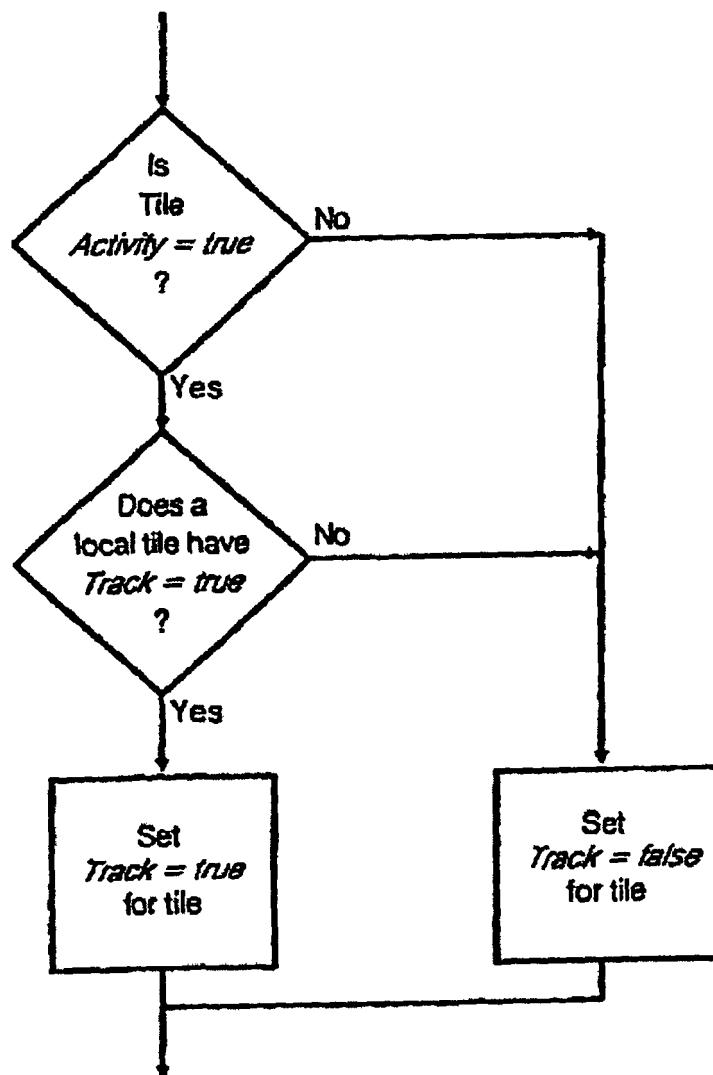


Figura 4

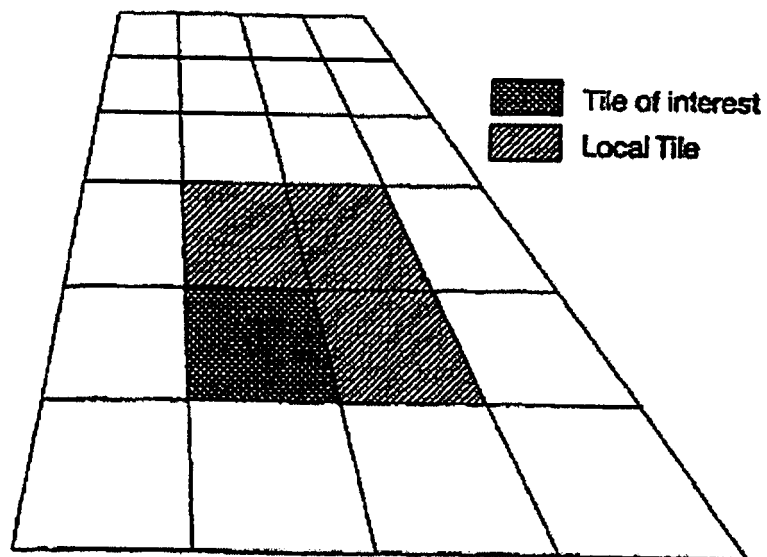


Figura 5