

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 439 726**

51 Int. Cl.:

G08B 13/194 (2006.01)

G08B 15/00 (2006.01)

G06T 7/20 (2006.01)

H04N 7/18 (2006.01)

G08B 13/196 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **09.07.2003 E 03787855 (0)**

97 Fecha y número de publicación de la concesión europea: **21.08.2013 EP 1529268**

54 Título: **Detector de anomalías de movimiento de vídeo**

30 Prioridad:

15.08.2002 GB 0218982

18.11.2002 GB 0226832

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

24.01.2014

73 Titular/es:

**ROKE MANOR RESEARCH LIMITED (100.0%)
ROKE MANOR, OLD SALISBURY LANE
ROMSEY, HAMPSHIRE SO51 0ZN, GB**

72 Inventor/es:

EVANS, RICHARD JOHN

74 Agente/Representante:

SUGRAÑES MOLINÉ, Pedro

ES 2 439 726 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

Detector de anomalías de movimiento de vídeo

5 La presente invención se refiere a dispositivos y a métodos para procesar imágenes de vídeo para disparar una señal de alarma cuando se detecta un evento de interés.

10 La televisión de circuito cerrado (CCTV) se usa ampliamente para fines de seguridad, transporte y otros fines. Aplicaciones de ejemplo incluyen la observación de delitos o vandalismo en espacios o edificios abiertos al público (tales como hospitales y escuelas), intrusión en áreas prohibidas, monitorización del flujo libre de tráfico rodado, detección de incidentes de tráfico y retenciones, detección de vehículos que se desplazan en sentido contrario en carreteras de un solo sentido.

15 Sin embargo, la monitorización de pantallas de CCTV (por operadores humanos) es una tarea muy laboriosa y hay riesgo considerable de que eventos de interés puedan pasar desapercibidos. Esto es especialmente cierto cuando se exige a los operadores que monitoricen varias salidas de cámara de CCTV simultáneamente. Como resultado, en muchas instalaciones de CCTV se graban datos de vídeo y sólo se inspeccionan en detalle si se sabe que se ha producido un evento. Incluso en estos casos, el volumen de datos grabados puede ser grande y la inspección manual de los datos puede ser laboriosa. Por consiguiente, existe la necesidad de que dispositivos automáticos procesen imágenes de vídeo y disparen una señal de alarma cuando hay un evento de interés. La señal de alarma puede usarse para o bien atraer la atención inmediata de un operador hacia el evento, colocar una marca de índice en el vídeo grabado o bien activar una grabación selectiva de datos de CCTV.

25 Se han desarrollado algunos detectores de eventos automáticos para sistemas de CCTV, aunque pocos de ellos son muy satisfactorios. Los dispositivos más comunes se denominan detectores de movimiento de vídeo (VMD) o detectores de actividad, aunque se basan generalmente en algoritmos sencillos en relación a la detección de cambios en el brillo de la imagen de vídeo, no el movimiento real de objetos visualizados. Con vistas a detectar cambios en el brillo, la imagen de vídeo se divide generalmente en una rejilla de normalmente 16 bloques horizontal y verticalmente (es decir 256 bloques en total). Hay varias desventajas en estos algoritmos. Por ejemplo, son propensos a falsas alarmas, por ejemplo cuando hay cambios en los niveles globales de iluminación. Además, no pueden detectar el movimiento de objetos pequeños, debido al procesamiento basado en bloques. Además, no pueden aplicarse si la escena contiene normalmente objetos móviles que no son de interés. Estas desventajas pueden reducirse, en pequeña medida, mediante una lógica de procesamiento adicional, pero la eficacia de VMD convencionales está limitada de manera inherente por el uso de la detección de cambio como fase de procesamiento de imágenes inicial.

35 Hay otro tipo de dispositivo de detección, que se caracteriza por el uso de algoritmos complejos que implican segmentación de imágenes, reconocimiento y seguimiento de objetos y reglas de decisión de alarma. Aunque estos dispositivos pueden ser muy eficaces, son generalmente sistemas caros diseñados para su uso en aplicaciones específicas y no tienen un buen rendimiento sin un ajuste y configuración cuidadosos, y pueden no funcionar en absoluto fuera de una variedad limitada de aplicaciones para las que se desarrollaron originalmente.

40 La patente estadounidense n.º 6.081.606 de los inventores Wade & Jeffrey describe un aparato y un método para detectar movimiento dentro de una secuencia de imágenes. Ese documento da a conocer que el movimiento dentro de una imagen puede calcularse correlacionando áreas de una imagen con áreas de la siguiente imagen en el vídeo para generar un campo de flujo. El campo de flujo se analiza entonces y se dispara una alarma dependiendo de la magnitud y dirección observadas del flujo.

50 La patente europea n.º 0 986 912 describe un método para monitorizar una zona de vigilancia predeterminada. Una imagen de vídeo se divide en varios segmentos. Se determina una distribución estadística para el nivel de gris medio para cada segmento. Puede usarse un cambio en el nivel de gris medio para un segmento, fuera de la variación estadística habitual, para activar una alarma.

55 El documento EPO 933726 da a conocer un método para procesar imágenes para detectar un evento.

60 El detector de anomalías de movimiento de vídeo aborda el problema de detectar automáticamente eventos de interés para operadores de sistemas de CCTV usados en aplicaciones de seguridad, transporte y otras aplicaciones, procesando imágenes de CCTV. El detector puede usarse de varias maneras, por ejemplo para disparar una alarma, emplazando a un operador humano a que visualice datos de vídeo, o para activar una grabación selectiva de datos de vídeo o para insertar una marca de índice en grabaciones de datos de vídeo.

Por consiguiente, la presente invención proporciona un método y un aparato para procesar imágenes de vídeo para detectar un evento de interés, según las reivindicaciones.

65 La señal de alarma puede provocar al menos uno de los siguientes efectos: atraer la atención de un operador;

colocar una marca de índice en el lugar apropiado en datos de vídeo grabados; y activar una grabación selectiva de datos de vídeo.

5 El proceso de aprendizaje puede acumular datos que representan el comportamiento de la(s) pista(s) a lo largo de un periodo de tiempo en un histograma de cuatro dimensiones, representando dichas cuatro dimensiones posición x, posición y, velocidad x y velocidad y, de la(s) pista(s) dentro de la imagen de vídeo. Además, la fase de aprender el comportamiento puede segregar las pistas según un umbral de velocidad; en el que las pistas que se mueven a una velocidad por debajo del umbral de velocidad se consideran estacionarias mientras que las pistas que se mueven a una velocidad superior al umbral de velocidad se consideran móviles; en el que los datos relativos a las pistas móviles se almacenan en dicho histograma de cuatro dimensiones, almacenándose los datos relativos a las pistas estacionarias en un histograma bidimensional, representando dichas dos dimensiones posición x y posición y dentro de la imagen de vídeo. Además, el tamaño de una celda del histograma de cuatro dimensiones puede variar según una velocidad medida en la imagen de cada respectiva pista. El histograma puede desponderarse periódicamente con el fin de desviar el resultado del proceso de aprendizaje hacia eventos más recientes.

15 El proceso de comparación puede clasificar una pista según una comparación de la frecuencia de ocupación de la correspondiente celda de histograma con un umbral de ocupación. El proceso de comparación puede actuar para clasificar como comportamiento normal una pista adyacente o próxima a una celda que está por encima del umbral de ocupación, a pesar de que la pista aparezca en una celda por debajo del umbral de ocupación, en el que una celda se considera próxima a otra si la distancia entre las mismas está por debajo de un umbral de distancia predeterminado.

20 Pueden filtrarse pistas anómalas, mediante lo cual se genera una señal de alarma activa en respuesta a una pista anómala que se parece a otras varias pistas anómalas, en cuanto a al menos uno de posición, velocidad y tiempo.

25 Pueden filtrarse pistas anómalas, mediante lo cual se genera una señal de alarma activa en respuesta sólo a una pista anómala que se ha clasificado como anómala un número predeterminado de ocasiones.

30 Pueden filtrarse pistas anómalas, mediante lo cual se genera una señal de alarma activa en respuesta sólo a que se clasifique una pista como anómala por primera vez.

Pueden filtrarse pistas anómalas, mediante lo cual se genera una señal de alarma activa sólo en respuesta a que una versión filtrada de la clasificación supere un valor umbral predeterminado.

35 Pueden inhibirse señales de alarma activa posteriores durante un intervalo de tiempo predeterminado después de que se haya producido una primera señal de alarma activa.

Pueden inhibirse señales de alarma activa posteriores si están provocadas por una pista anómala a una distancia predeterminada de otra pista que ha generado previamente una alarma.

40 La presente invención también proporciona un aparato para procesar imágenes de vídeo para detectar un evento de interés, que comprende: una fuente de imágenes de vídeo, que produce una señal de vídeo que representa las imágenes de vídeo que van a procesarse; un dispositivo de extracción de características que recibe la señal de vídeo y que produce datos que representan al menos una característica de punto detectada dentro de la imagen; un dispositivo de seguimiento de características que recibe los datos que representan características de punto y que produce datos que representan pistas, que son representativas de la posición y velocidad de cada respectiva característica de punto, dentro de la imagen; un dispositivo de aprendizaje que recibe los datos que representan las pistas y que produce una señal que representa un margen de comportamiento considerado normal por el dispositivo de aprendizaje, en respuesta a la operación de un proceso de aprendizaje sobre los datos que representan las pistas; un dispositivo de clasificación que recibe tanto la señal que representa el margen normal de comportamiento de las pistas como los datos que representan las pistas, que está adaptado para comparar la señal y los datos y para emitir una señal de normal/anómalo según el resultado de tal comparación; y un dispositivo de generación de alarma que recibe la señal de normal/anómalo y que genera al menos una señal de alarma activa en respuesta a que la señal de normal/anómalo indique comportamiento anómalo de al menos una pista.

55 Los anteriores y otros objetos, características y ventajas de la presente invención resultarán más evidentes a partir de la siguiente descripción de determinadas realizaciones de la misma, con referencia a la figura 1, que muestra un diagrama de bloques esquemático de un detector de anomalías de movimiento de vídeo según la presente invención.

60 La presente invención proporciona un detector de anomalías de movimiento de vídeo que se basa en características, y genera alarmas basándose en el comportamiento anómalo de una pista que representa el comportamiento de una característica de punto. Los sistemas conocidos no tienen estas características.

65 El detector de anomalías de movimiento de vídeo extrae y sigue características de tipo punto en imágenes de vídeo

y dispara una alarma cuando una o más pistas, que representan, cada una, una característica, se comportan de manera anómala en comparación con el comportamiento de las pistas observadas durante un periodo de tiempo. "Comportamiento" quiere decir el movimiento de pistas en diferentes partes de la imagen de vídeo. Por ejemplo, el movimiento rápido de características en una dirección particular en una parte del campo de visión puede ser normal, pero puede ser anómalo si se produjera en otra parte del campo de visión donde el comportamiento normal es un movimiento lento. De manera similar, el movimiento rápido en la misma parte del campo de visión puede ser anómalo si el movimiento es en una dirección diferente.

La figura 1 muestra las principales fases de procesamiento en el detector de anomalías de movimiento de vídeo.

Se proporciona una señal de vídeo entrante 10 por una fuente de imágenes de vídeo, por ejemplo una cámara de CCTV. La señal de vídeo 10 representa imágenes de vídeo generadas por la fuente de imágenes de vídeo. Normalmente, las imágenes de vídeo mostrarán una vista de un área vigilada por una cámara de seguridad de CCTV.

La señal de vídeo 10 se suministra a una fase de extracción de características 1. La fase de extracción de características 1 localiza características de tipo punto en cada imagen procesada en la secuencia de imágenes de vídeo. Un proceso de extracción de características adecuado se describe en la patente n.º GB2218507, "Digital Data Processing". Las características localizadas se identifican en una señal 12 proporcionada a la siguiente fase, fase de seguimiento de características 2.

La fase de seguimiento de características 2 sigue características de modo que cada característica de tipo punto puede describirse por su posición actual y su velocidad estimada en la imagen. La fase de seguimiento de características 2 proporciona una pista, es decir una señal 14 que indica la posición actual y la velocidad estimada en la imagen, para cada característica seguida, a una fase de aprender el comportamiento 3 y a una fase de clasificación de pista 4.

La fase de aprender el comportamiento 3 acumula información acerca del comportamiento de características durante un periodo de tiempo. Una manera de realizar esto es acumular un histograma de cuatro dimensiones, siendo las cuatro dimensiones del histograma posición x, posición y, velocidad x, velocidad y.

La fase de clasificación de pista 4 clasifica cada pista como "normal" o "anómala", en comparación con la información de comportamiento acumulada por la fase de aprender el comportamiento 3. Una manera de clasificar una pista es comparar la frecuencia de ocupación de la correspondiente celda de histograma con un umbral. Si la frecuencia de ocupación está por debajo del umbral, la pista se clasifica como anómala, de lo contrario se considera normal. La fase de clasificación de pista 4 envía una señal de normal/anómalo 18 a una fase de generación de alarma 5. Si cualquiera de las características seguidas muestra un comportamiento anómalo, entonces esta señal informará a la fase de generación de alarma 5 acerca de la anomalía.

La fase de generación de alarma 5 genera una señal de alarma 20 en respuesta a una señal de normal/anómalo activa 18, que indica que se ha descubierto que está presente al menos una pista anómala. Puede proporcionarse una lógica de procesamiento adicional para resolver situaciones tales como comportamiento anómalo intermitente o múltiples casos de comportamiento anómalo asociados con un evento del mundo real, y otras situaciones de este tipo.

El detector de anomalías de movimiento de vídeo de la presente invención proporciona preferiblemente al menos una de las siguientes características: el uso de extracción y seguimiento de características de punto en un detector de eventos; y la detección de eventos mediante clasificación del comportamiento de pistas como anómalo, en comparación con el comportamiento de pistas observadas a lo largo del tiempo.

En comparación con una detección de eventos basada en una detección de movimiento de vídeo conocida, el detector de anomalías de movimiento de vídeo de la presente invención proporciona al menos algunas de las siguientes ventajas.

- Se basa en extracción de características de punto en lugar de detectar cambios en el brillo de la imagen. Esto hace que el sistema de la invención sea relativamente insensible a cambios en niveles de iluminación de escena. Los niveles de iluminación de escena son la principal fuente de falsas alarmas en los detectores de movimiento de vídeo conocidos.

- Debido a que se basa en la detección de características de punto en lugar de en el procesamiento de bloques, el sistema de la invención puede detectar el movimiento de objetivos incluso relativamente pequeños, e indicar una situación de alarma si el movimiento es inusual.

- El sistema de la invención acumula información acerca del comportamiento de cada característica seguida, lo que le permite construir una definición de comportamiento "normal" para cada característica. El sistema de la invención

puede detectar entonces un movimiento de interés, es decir, un movimiento que se aleja de un comportamiento "normal" para esa característica particular, incluso en presencia de otras características que se mueven "normalmente".

5 - El sistema de la invención detecta un comportamiento anómalo, es decir, un comportamiento que se aleja del comportamiento "normal" calculado para una característica particular, en lugar de un comportamiento específico predefinido. La invención puede aplicarse por consiguiente a una variedad muy amplia de diferentes aplicaciones con poca configuración especial, y puede adaptarse a cambios a largo plazo o permanentes en la imagen visualizada.

10 En comparación con los sistemas de detección de eventos existentes basados en soluciones de software complejas, el detector de anomalías de movimiento de vídeo de la presente invención es un sistema relativamente sencillo adecuado para su implementación en hardware relativamente barato.

15 A continuación se describen en más detalle las principales fases de procesamiento del detector de anomalías de movimiento de vídeo, con referencia a la figura 1.

20 La figura 1 muestra esquemáticamente las principales fases de procesamiento en un detector de anomalías de movimiento de vídeo de la presente invención, que recibe información de vídeo 10, y que produce una señal de alarma 20.

25 La fase de extracción de características 1 localiza características de tipo punto en cada imagen procesada en la secuencia de imágenes de vídeo. Un extractor de características preferido se describe en la patente n.º GB2218507, "Digital Data Processing" y esto se describe adicionalmente por Harris y Stephens en los *Proceedings of the 4th Alvey Vision Conference*, septiembre de 1988, Universidad de Manchester, "A combined corner and edge detector". Un aspecto importante de este extractor de características es que proporciona atributos de características, es decir descripciones cuantitativas de las características extraídas además de su posición en la imagen. En comparación con otros algoritmos de extracción de características de punto el método de Harris es particularmente robusto.

30 Se han desarrollado otras técnicas de extracción de características de punto, y podrían usarse en la presente invención. Por ejemplo, una técnica de este tipo se describe por Moravec en Tech Report CMU-RI-TR-3, Carnegie-Mellon University, Robotics Institute, septiembre de 1980 "Obstacle avoidance and navigation in the real world by a seeing robot rover". Pueden concebirse otros esquemas *ad hoc*.

35 En efecto, cualquier algoritmo que pueda usarse para extraer características de imagen que puedan estar asociadas con la localidad puede usarse como extractor de características de tipo punto en la presente invención. Como ejemplos, a puntos de nudo, es decir puntos de alta curvatura en características de borde, se les puede asignar una posición. A una característica tal como una región de imagen, por ejemplo un vehículo completo o una persona completa, segmentada mediante técnicas de detección de bordes o de crecimiento de regiones, se le puede asignar la posición de su centroide.

40 El extractor de características 1 puede emplear un umbral de nivel de variación de características fijo, en cuyo caso el número de características extraídas puede variar de fotograma a fotograma y puede variar de una aplicación a otra. Alternativamente, el umbral puede ser variable de modo que se extraiga un número fijo de características.

45 La fase de seguimiento de características 2 sigue características entre fotogramas de imagen de modo que cada característica de tipo punto puede describirse por su posición actual y su velocidad estimada en la imagen.

50 En la presente aplicación se requiere un algoritmo de seguimiento de múltiples objetivos ya que una escena puede contener un gran número de objetos móviles y cada uno puede dar lugar a varias características extraídas. Por ejemplo, un coche que pasa por el campo de visión puede generar varias características extraídas en cada fotograma de vídeo, dependiendo de la resolución espacial usada, y una escena de tráfico puede contener varios coches que se mueven en diferentes partes del campo de visión. Los algoritmos de seguimiento en sí mismos se conocen y se entienden relativamente bien. Blackman & Popoli realizaron un tratado sobre el tema en "Design & Analysis of Modern Tracking Systems", Artech House 1999.

55 El seguimiento es un proceso cíclico. En cualquier momento, varios objetos pueden estar siguiéndose. Se conocen su posición y velocidad actuales y a medida que se presenta cada nuevo fotograma de datos de vídeo, hay que actualizar la información de pista. Un algoritmo de seguimiento de este tipo consiste normalmente en las siguientes fases.

60 - Asociación de trazado (característica) a pista: se trata del proceso de decidir cuál de las características, extraídas del fotograma de vídeo más reciente, es el mismo objeto que el representado por cualquier pista particular. Dado que "trazado" (*plot*) en lugar de "característica" (*feature*) es el término normal usado en los documentos sobre algoritmos de seguimiento, se usará "trazado" en esta sección de la descripción. Un enfoque convencional es sólo

considerar para la asociación trazados que entran dentro de una ventana o compuerta, que está centrada en la posición predicha del trazado (véase a continuación). Muchos algoritmos de seguimiento emplean entonces reglas sencillas para gestionar situaciones en las que más de un trazado entra dentro de la compuerta de aceptación. Reglas de ejemplo incluyen “asociar el trazado más próximo a la posición predicha”, o “no asociar ninguno”. Estas opciones se comentarán a continuación. En la presente aplicación, la densidad de trazados es normalmente alta y la posibilidad de error de asociación de trazado a pista es alta, de modo que el enfoque preferido hace uso de la similitud de atributos de trazados (característica) así como de posición de trazados a la hora de tomar una decisión de asociación de trazado a pista. Son posibles otros esquemas para resolver ambigüedades, por ejemplo coincidencia probabilística, seguimiento de múltiples hipótesis y toma de decisiones diferida. Pueden emplearse otros refinamientos para mejorar el rendimiento, por ejemplo correlación cruzada de parches de imagen para confirmar la similitud de características visualizadas, o variación del tamaño de la ventana de aceptación según la precisión esperada de la predicción. Esta última opción permitirá que pistas bien establecidas y de movimiento lento tengan una ventana de aceptación más pequeña que pistas de movimiento rápido o formadas recientemente. También son posibles esquemas de coincidencia bidireccional.

- Mantenimiento de pista: ese trata del proceso de iniciar pistas nuevas, normalmente porque se ve un objeto nuevo, y de borrar pistas, normalmente porque deja de verse una característica seguida. La mayoría de los algoritmos de seguimiento también tendrán un proceso de confirmación de pista para decidir si una pista está lo suficientemente bien establecida como para que sea la base para una toma de decisiones posterior. En una implementación preferida, se inician pistas nuevas a partir de trazados que no pueden explicarse después de una asociación de trazado a pista con las pistas existentes. Debido a incertidumbres en el rendimiento del extractor de características, no se borran inmediatamente pistas en la implementación preferida si no se asocian con ningún trazado. En cambio, una pista puede “extenderse” por varios fotogramas antes del borrado. Se confirman pistas una vez que se han asociado satisfactoriamente con trazados en varios fotogramas.

El seguimiento de características con un bajo umbral de nivel de variación de características es problemático, porque no se detectan de manera fiable en cada fotograma de vídeo. Los errores de seguimiento se vuelven más comunes cuando las pistas están estrechamente espaciadas. Una manera de reducir este problema es ignorar trazados que no tienen coincidencia con un bajo nivel de variación de características, y sólo iniciar pistas nuevas para trazados que no tienen coincidencia con un mayor nivel de variación de características.

- Filtrado y predicción de pistas: ese trata del proceso de estimar la actual posición y velocidad de trazados, y predecir la posición de trazados esperada en fotogramas de imagen posteriores. Este proceso es necesario porque las mediciones de posiciones de característica pueden no ser perfectas debido a cuantificación de píxeles y ruido de imagen, y los objetos pueden moverse de manera apreciable entre fotogramas de imagen. Varios métodos son aplicables, por ejemplo filtrado de Kalman o Alpha-Beta recursivo, ajuste de curvas polinómicas a trozos (*splines*) para datos de trazado recientes, etc. El rendimiento en este caso puede mejorarse mediante varios esquemas, por ejemplo extracción de valor atípico y variación del orden de una curva polinómica a trozos según la longitud del historial de una pista.

En general, los algoritmos de seguimiento siguen un patrón común aunque las implementaciones individuales pueden variar en cuanto al detalle según factores específicos de la aplicación y otras cuestiones.

La fase de aprender el comportamiento 3 acumula información acerca del comportamiento de las pistas, que representan característica seguidas, a lo largo de un periodo de tiempo. La manera preferida de hacer esto es acumular un histograma de cuatro dimensiones, siendo las cuatro dimensiones del histograma posición x, posición y, velocidad x, velocidad y. Alternativamente, podrían usarse otros sistemas de coordenadas tales como coordenadas polares o coordenadas espaciales de transformada de Hough, etc.

El histograma de comportamiento, que es de cuatro dimensiones, puede requerir una gran cantidad de almacenamiento de datos y esto puede no ser práctico para la implementación en hardware de procesamiento. La manera preferida de superar este problema es dividir el histograma en dos secciones. La primera es para pistas estacionarias, es decir que se mueven muy despacio. Esta sección es un histograma bidimensional y por tanto puede usar un tamaño de celda fino sin requerir grandes cantidades de almacenamiento. La segunda sección de histograma es para objetos móviles de diferentes velocidades. Aunque se trata de un histograma de cuatro dimensiones, puede usarse un tamaño de celda más basto para tales objetos de modo que tampoco sea necesario que el tamaño de esta sección del histograma sea grande. El tamaño de la segunda división también puede reducirse usando un tamaño de celda que varía con la velocidad.

Hay varias maneras alternativas de construir el histograma para minimizar requisitos de memoria. Éstas incluyen un árbol cuádruple y otras técnicas de representación de matriz dispersa.

El histograma que describe el comportamiento puede construirse a lo largo un periodo fijo de tiempo y permanecer sin cambios después. Esto tiene la desventaja de que cambios lentos en el comportamiento real, deriva en la electrónica de la cámara o desplazamiento del soporte de la cámara, etc., pueden dar como resultado en última

instancia que un comportamiento normal se clasifique como anómalo. Esto puede superarse de varias maneras. En el método preferido, el histograma se despondera periódicamente mediante un factor próximo a la unidad, con el resultado de que el sistema tiene una memoria de desvanecimiento, es decir su memoria se desvía hacia los eventos más recientes. Dependiendo de las limitaciones de procesador, puede ser necesario implementar la memoria de desvanecimiento con vistas a reducir la carga de procesador. Por ejemplo, la desponderación podría aplicarse en un intervalo más grande, o sólo una parte del histograma podría desponderarse a intervalos más cortos. Otra manera de desviar la memoria del comportamiento hacia eventos recientes, pero sin una memoria de desvanecimiento, es usar dos almacenes de histogramas, construyéndose uno mientras que el otro se usa para clasificación de pistas.

Aunque la creación de histogramas es el enfoque preferido, hay otros posibles métodos de aprendizaje y clasificación, y algunos de ellos se comentan a continuación.

La fase de clasificación de pista 4 clasifica cada pista como "normal" o "anómala" y el método usado depende de cómo se aprende el comportamiento en la fase de aprender el comportamiento 3. En el método basado en histogramas preferido, una pista se clasifica comparando la frecuencia de ocupación de la correspondiente celda de histograma con un umbral de ocupación. Si la frecuencia de ocupación está por debajo del umbral, la pista se clasifica como anómala, de lo contrario se considera normal.

Si se requiere una tasa de falsas alarmas muy baja, puede ser necesario adoptar etapas adicionales para impedir una respuesta de sistema adversa, particularmente si se ha limitado el tiempo de entrenamiento. Como ejemplo de tales etapas, una pista puede clasificarse como normal, incluso aunque la ocupación de la correspondiente celda de histograma sea baja, si es adyacente o próxima a una celda por encima del umbral. La distancia desde la correspondiente celda de histograma a la celda por encima del umbral de ocupación más próxima (medida dentro del histograma mediante una métrica de ciudad, cartesiana o alguna otra métrica que pueda ponderarse por ocupación) puede compararse con un umbral de distancia. Si la distancia es menor que el umbral, la pista se clasifica como normal, de lo contrario la pista se considera anómala. La tasa de falsas alarmas puede ajustarse variando el umbral de ocupación y el umbral de distancia. Como alternativa al uso de un umbral de distancia, los datos de histograma podrían difuminarse para suprimir la clasificación de pistas como anómalas en celdas próximas a celdas de alta ocupación. De manera similar, podrían aplicarse otros operadores morfológicos.

La fase de generación de alarma 5 genera una señal de alarma 20 cuando se descubre que están presentes pistas anómalas, con sujeción a una lógica de procesamiento adicional para resolver situaciones tales como comportamiento anómalo intermitente o múltiples casos de comportamiento anómalo asociados con un evento del mundo real, y otras situaciones tales como cuando se generan datos de pista espurios por un error de seguimiento.

El riesgo de que se generen alarmas por datos espurios puede reducirse limitando el procesamiento a pistas confirmadas o pistas cuyo historial de pista muestra un historial sistemático de trazados asociadas. Las alarmas generadas por datos espurios también pueden reducirse imitando el procesamiento a o bien: pistas anómalas que se producen en estrecha proximidad (en cuanto de cualquiera de: posición, velocidad o tiempo) a otras varias pistas anómalas, y/o: pistas que se clasifican como anómalas en varias ocasiones.

Para evitar alarmas intermitentes puede dispararse una señal de alarma 20 (con sujeción a otra lógica) sólo cuando una pista se clasifica como anómala por primera vez, o cuando una versión filtrada de la clasificación supera un umbral.

Podrían generarse múltiples alarmas, por ejemplo, si un vehículo visualizado por el sistema de CCTV adopta una trayectoria anómala y genera varias pistas anómalas independientes en el proceso, generándose las diferentes pistas por diferentes partes del vehículo. Estas múltiples alarmas serían confusas y no deseadas en un sistema práctico. Éstas pueden suprimirse inhibiendo alarmas durante un periodo de segundos después de una primera alarma. Alternativamente, podrían suprimirse alarmas si la pista que provoca la alarma está a una cierta distancia de una pista que ha generado previamente una alarma.

Pueden concebirse varios métodos diferentes de lógica de generación de alarma, con diferentes formulaciones *ad hoc*.

Las secciones anteriores describen métodos preferidos para su uso dentro de la presente invención, aunque pueden usarse métodos de aprendizaje y clasificación alternativos. A continuación se describe una selección de tales alternativas.

Red neuronal (MLP) con abanico de entrada/salida (*fan-in/fan-out*). De manera esquemática, la idea es usar un perceptrón de múltiples capas con 4 nodos de entrada (para datos de seguimiento X, Y, Vx, Vy) y 4 nodos de salida, pero una capa intermedia de 2 ó 3 nodos. El perceptrón es un programa que aprende conceptos, es decir puede aprender a responder con Verdadero (1) o Falso (0) para entradas que se le presentan, "estudiando" de manera repetida ejemplos que se le presentan. El perceptrón es una red neuronal cuyos pesos y desviaciones pueden

5 entrenarse para producir un vector objetivo correcto cuando se le presenta el correspondiente vector de entrada. La técnica de entrenamiento usada se denomina regla de aprendizaje de perceptrón. El perceptrón puede generalizar a partir de sus vectores de entrenamiento y funcionar con conexiones distribuidas aleatoriamente. Los perceptrones son especialmente adecuados para problemas sencillos en la clasificación de patrones. La red se entrenaría para reconstruir su propia entrada. Debido a la constricción interna, la reconstrucción debe ser mejor para pistas "normales" y peor para pistas "anómalas". Por tanto, la precisión de reconstrucción podría usarse para evaluar la cualidad de normal de la pista.

10 Vecino más próximo. En este caso, se retienen datos de característica seguida para fotogramas recientes para crear una base de datos de historial de pista. Cada pista nueva se somete a prueba en cuanto a la cualidad de normal buscando en esa base de datos para encontrar alguna pista previa similar.

15 Vecino más próximo reducido. Se trata de una variación de la técnica de vecino más próximo completa. La base de datos de historial se reduce en tamaño omitiendo duplicados o casi duplicados de datos anteriores.

20 Red de Kohonen. Aunque se trata de una técnica de red neuronal, puede considerarse similar al método de vecino más próximo reducido. El comportamiento de objetos seguidos se describe mediante un conjunto de nodos situados en el espacio de entrada de cuatro dimensiones. Las posiciones reales de los nodos se determinan mediante un proceso de entrenamiento iterativo. Este método también está relacionado con métodos de generación de libros de códigos adaptativos en sistemas de compresión de datos.

25 Comprobación probabilística. Se trata de un enfoque lateral para buscar en bases de datos de historial los algoritmos basados en el vecino más próximo. En este caso, se busca en la base de datos de historial eligiendo entradas de comparación en una secuencia aleatoria hasta que se encuentren varias coincidencias. Si la pista que se comprueba es muy normal, se encontrarán varias coincidencias muy rápidamente.

30 Por consiguiente, la presente invención proporciona un detector de anomalías de movimiento de vídeo que aborda el problema de detectar automáticamente eventos de interés para operadores de sistemas de CCTV usados en aplicaciones de seguridad, transporte y otras aplicaciones, procesando imágenes de CCTV. El detector puede usarse, por ejemplo, para activar una alarma para emplazar a un operador humano a que visualice datos de vídeo, para activar una grabación selectiva de datos de vídeo o para insertar una marca de índice en grabaciones de datos de vídeo. El detector de anomalías de movimiento de vídeo de la presente invención extrae y sigue características de tipo punto en imágenes de vídeo y dispara una alarma cuando una característica se comporta de manera anómala, en comparación con el compartimiento "normal" de esas características, derivado de observaciones de la característica a lo largo un periodo de tiempo.

40 Los detectores de movimiento de vídeo existentes son dispositivos que se basan esencialmente en detectar cambios de brillo de imagen promediados a través de subbloques de imagen. El detector de anomalías de movimiento de vídeo de la presente invención tiene la ventaja de ser menos propenso a falsas alarmas provocadas por cambios en niveles de iluminación de la escena. El detector de la presente invención también puede detectar el movimiento de objetos más pequeños y detectar movimientos de interés, incluso en presencia de otros objetos móviles. Además, puede aplicarse a una variedad muy amplia de diferentes aplicaciones con un pequeño ajuste especial. En comparación con otros sistemas de detección de eventos existentes basados en soluciones de software complejas, el detector de anomalías de movimiento de vídeo puede implementarse en hardware relativamente barato.

45

REIVINDICACIONES

1. Método para procesar imágenes de vídeo para detectar un evento de interés, que comprende las etapas de:
 - 5 - recibir una señal de vídeo (10) que representa las imágenes de vídeo que van a procesarse;
 - extraer (1) al menos una característica de punto de la señal de vídeo;
 - 10 - seguir (2) la posición y movimiento de la al menos una característica de punto dentro de las imágenes de vídeo para generar una correspondiente al menos una pista, representando cada una, una correspondiente característica de punto;
 - 15 - usar (3) un proceso de aprendizaje iterativo para derivar un patrón normal de comportamiento para cada pista, en el que el proceso de aprendizaje (3) acumula datos que representan el comportamiento de la(s) pista(s) a lo largo de un periodo de tiempo en un histograma de cuatro dimensiones, representando dichas cuatro dimensiones posición x, posición y, velocidad x y velocidad y, de la(s) pista(s) dentro de la imagen de vídeo;
 - 20 - comparar (4) el comportamiento presente de la al menos una pista con el respectivo patrón normal de comportamiento, en el que el proceso de comparación (4) clasifica una pista como normal según una comparación de la frecuencia de ocupación de la correspondiente celda de histograma con un umbral de ocupación; y
 - 25 - en respuesta a que el comportamiento presente caiga fuera del patrón normal de comportamiento, generar (5) una señal de alarma (20).
2. Método según la reivindicación 1, en el que la señal de alarma (20) provoca al menos uno de los siguientes efectos:
 - 30 - atraer la atención de un operador;
 - colocar una marca de índice en el lugar apropiado en datos de vídeo grabados; y
 - 35 - activar una grabación selectiva de datos de vídeo.
3. Método según cualquier reivindicación anterior, en el que la fase de aprender el comportamiento segrega las pistas según un umbral de velocidad; en el que las pistas que se mueven a una velocidad por debajo del umbral de velocidad se consideran estacionarias mientras que las pistas que se mueven a una velocidad superior al umbral de velocidad se consideran móviles; en el que los datos relativos a las pistas móviles se almacenan en dicho histograma de cuatro dimensiones, almacenándose los datos relativos a las pistas estacionarias en un histograma bidimensional, representando dichas dos dimensiones posición x y posición y dentro de la imagen de vídeo.
4. Método según cualquier reivindicación anterior, en el que un tamaño de celda del histograma de cuatro dimensiones varía según una velocidad medida en la imagen de cada respectiva pista.
5. Método según cualquier reivindicación anterior, en el que el histograma se despondera periódicamente con el fin de desviar el resultado del proceso de aprendizaje (3) hacia eventos más recientes.
6. Método según cualquier reivindicación anterior, en el que el proceso de comparación (4) actúa para clasificar como comportamiento normal una pista adyacente o próxima a una celda que está por encima del umbral de ocupación, a pesar de que la pista aparezca en una celda por debajo del umbral de ocupación, en el que una celda se considera próxima a otra si la distancia entre las mismas está por debajo de un umbral de distancia predeterminado.
7. Método según cualquier reivindicación anterior, en el que se filtran pistas anómalas, mediante lo cual se genera una señal de alarma activa (20) en respuesta a una pista anómala que se parece a otras varias pistas anómalas, en cuanto a al menos uno de posición, velocidad y tiempo.
8. Método según cualquier reivindicación anterior, en el que se filtran pistas anómalas, mediante lo cual se genera una señal de alarma activa (20) en respuesta sólo a una pista anómala que se ha clasificado como anómala un número predeterminado de ocasiones.
9. Método según cualquier reivindicación anterior, en el que se filtran pistas anómalas, mediante lo cual se

genera una señal de alarma activa (20) en respuesta sólo a que se clasifique una pista como anómala por primera vez.

- 5 10. Método según cualquier reivindicación anterior, en el que se filtran pistas anómalas, mediante lo cual se genera una señal de alarma activa (20) sólo en respuesta a que una versión filtrada de la clasificación supere un valor umbral predeterminado.
- 10 11. Método según cualquier reivindicación anterior, en el que se inhiben señales de alarma activa (20) posteriores durante un intervalo de tiempo predeterminado después de que se haya producido una primera señal de alarma activa (20).
- 15 12. Método según cualquier reivindicación anterior, en el que se inhiben señales de alarma activa (20) posteriores si están provocadas por una pista anómala a una distancia predeterminada de otra pista que ha generado previamente una alarma.
- 20 13. Aparato para procesar imágenes de vídeo para detectar un evento de interés, que comprende:
- una fuente de imágenes de vídeo, que produce una señal de vídeo (10) que representa las imágenes de vídeo que van a procesarse;
 - un dispositivo de extracción de características (1) que recibe la señal de vídeo y que produce datos (12) que representan al menos una característica de punto detectada dentro de la imagen;
 - un dispositivo de seguimiento de características (2) que recibe los datos (12) que representan características de punto y que produce datos (14) que representan pistas, que son representativas de la posición y velocidad de cada respectiva característica de punto, dentro de la imagen;
 - un dispositivo de aprendizaje (3) que recibe los datos (14) que representan las pistas y que produce una señal (16) que representa un margen de comportamiento considerado normal por el dispositivo de aprendizaje, en respuesta a la operación de un proceso de aprendizaje sobre los datos (14) que representan las pistas, en el que el dispositivo de aprendizaje (3) acumula datos que representan el comportamiento de la(s) pista(s) a lo largo de un periodo de tiempo en un histograma de cuatro dimensiones, representando dichas cuatro dimensiones posición x, posición y, velocidad x y velocidad y, de la(s) pista(s) dentro de la imagen de vídeo;
 - un dispositivo de clasificación (4) que recibe tanto la señal (16) que representa el margen normal de comportamiento de las pistas como los datos (14) que representan las pistas, que está adaptado para comparar la señal (16) y los datos (14) y para emitir una señal de normal/anómalo (18) según el resultado de tal comparación, en el que el proceso de comparación (4) clasifica una pista como normal según una comparación de la frecuencia de ocupación de la correspondiente celda de histograma con un umbral de ocupación; y
 - un dispositivo de generación de alarma (5) que recibe la señal de normal/anómalo (18) y que genera al menos una señal de alarma activa (20) en respuesta a que la señal de normal/anómalo indique un comportamiento anómalo de al menos una pista.
- 25
- 30
- 35
- 40
- 45

