

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 534 250**

51 Int. Cl.:

**G08B 13/196** (2006.01)  
**H04N 7/18** (2006.01)  
**G06K 9/62** (2006.01)  
**G06F 17/30** (2006.01)  
**H04N 5/272** (2006.01)  
**G06T 7/20** (2006.01)  
**H04N 21/44** (2011.01)  
**H04N 21/234** (2011.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.04.2006 E 06749276 (9)**

97 Fecha y número de publicación de la concesión europea: **19.11.2014 EP 1872583**

54 Título: **Método de procesamiento de vídeo, medio legible por ordenador que contiene instrucciones que implementan dicho método y sistema de procesamiento de vídeo**

30 Prioridad:

**05.04.2005 US 98385**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**21.04.2015**

73 Titular/es:

**AVIGILON FORTRESS CORPORATION (100.0%)**  
**2900 - 550 Burrard Street**  
**Vancouver, British Columbia V6C 0A3, CA**

72 Inventor/es:

**VENETIANER, PETER L.;**  
**LIPTON, ALAN J.;**  
**CHOSAK, ANDREW J.;**  
**FRAZIER, MATTHEW F.;**  
**HAERING, NIELS;**  
**MYERS, GARY;**  
**YIN, WEIHONG y**  
**ZHANG, ZHONG**

74 Agente/Representante:

**FÚSTER OLAGUIBEL, Gustavo Nicolás**

ES 2 534 250 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**

Método de procesamiento de vídeo, medio legible por ordenador que contiene instrucciones que implementan dicho método y sistema de procesamiento de vídeo

5

**Antecedentes de la invención****Campo de la invención**

La invención se refiere a un sistema para videovigilancia automática que emplea primitivas de vídeo.

**Referencias**

Para mayor comodidad del lector, a continuación se enumeran las referencias citadas en el presente documento. En la memoria descriptiva, los números entre paréntesis se refieren a las respectivas referencias. Las referencias enumeradas se incorporan como referencia al presente documento.

Las siguientes referencias describen detección de objetivos en movimiento:

{1} A. Lipton, H. Fujiyoshi y R. S. Patil, "Moving Target Detection and Classification from Real-Time Video", Proceedings of IEEE WACV '98, Princeton, NJ, 1998, págs. 8-14.

{2} W.E.L. Grimson, *et al.*, "Using Adaptive Tracking to Classify and Monitor Activities in a Site", CVPR, págs. 22-29, junio de 1998.

{3} A. J. Lipton, H. Fujiyoshi, R.S. Patil, "Moving Target Classification and Tracking from Real-time Video", IUW, págs. 129-136, 1998.

{4} T.J. Olson y F.Z. Brill, "Moving Object Detection and Event Recognition Algorithm for Smart Cameras", IUW, págs. 159-175, mayo de 1997.

Las siguientes referencias describen detección y seguimiento de humanos:

{5} A. J. Lipton, "Local Application of Optical Flow to Analyse Rigid Versus Non-Rigid Motion", International Conference on Computer Vision, Corfú, Grecia, septiembre de 1999.

{6} F. Bartolini, V. Cappellini y A. Mecocci, "Counting people getting in y out of a bus by real-time image-sequence processing", IVC, 12(1):36-41, enero de 1994.

{7} M. Rossi y A. Bozzoli, "Tracking and counting moving people", ICIP94, págs. 212-216, 1994.

{8} C.R. Wren, A. Azarbayejani, T. Darrell y A. Pentland, "Pfinder: Real-time tracking of the human body", Vismod, 1995.

{9} L. Khoudour, L. Duviéubourg, J.P. Deparis, "Real-Time Pedestrian Counting by Active Linear Cameras", JEI, 5(4):452-459, octubre de 1996.

{10} S. Ioffe, D.A. Forsyth, "Probabilistic Methods for Finding People", IJCV. 43(1):45-68, junio de 2001.

{11} M. Isard y J. MacCormick, "BraMBLe: A Bayesian Multiple-Blob Tracker", ICCV, 2001.

Las siguientes referencias describen análisis de blobs:

{12} D.M. Gavrila, "The Visual Analysis of Human Movement: A Survey", CVIU, 73(1):82-98, enero de 1999.

{13} Niels Haering y Niels da Vitoria Lobo, "Visual Event Detection", Video Computing Series, Editor Mubarak Shah, 2001.

Las siguientes referencias describen análisis de blobs para camiones, coches y personas:

{14} Collins, Lipton, Kanade, Fujiyoshi, Duggins, Tsin, Tolliver, Enomoto y Hasegawa, "A System for Video Surveillance and Monitoring: VSAM Final Report", Technical Report CMU-RI-TR-00-12, Robotics Institute, Carnegie Mellon University, mayo de 2000.

{15} Lipton, Fujiyoshi y Patil, "Moving Target Classification and Tracking from Real-time Video", 98 Darea IUW. 20-23 nov. 1998.

La siguiente referencia describe el análisis de un blob de una única persona y sus contornos:

5 {16} C.R. Wren, A. Azarbayejani, T. Darrell, y A.P. Pentland. "Pfinder: Real-Time Tracking of the Human Body", PAMI, vol. 19, págs. 780-784, 1997.

La siguiente referencia describe movimiento interno de blobs, incluyendo cualquier segmentación basada en movimiento:

10 {17} M. Allmen y C. Dyer, "Long-Range Spatiotemporal Motion Understanding Using Spatiotemporal Flow Curves", Proc. IEEE CVPR. Lahaina, Maui, Hawaii, págs. 303-309, 1991.

15 {18} L. Wixson, "Detecting Salient Motion by Accumulating Directionally Consistent Flow", IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, págs. 774-781, agosto de 2000.

El documento US2003/0043160 A1 da a conocer un método y un dispositivo para gestionar información de imagen en un sistema de monitorización. El sistema de monitorización comprende al menos un módulo de monitorización para monitorizar una ubicación monitorizada y una unidad receptora remota. El método comprende las etapas de grabar una imagen de la ubicación monitorizada con el módulo de monitorización, segmentar una región diferenciada comparando la imagen grabada con una imagen de referencia en el módulo de monitorización, crear una forma de contorno exterior en el módulo de monitorización que representa el borde de la región diferenciada, clasificar la forma de contorno exterior y determinar si es interesante para su transmisión y, en su caso, transmitir datos que representan la forma de contorno exterior a la unidad receptora, recrear la forma de contorno exterior en la unidad receptora por medio de dichos datos transmitidos y presentar la forma de contorno exterior visualmente en la unidad receptora. La unidad receptora puede estar ubicada en una estación de monitorización asistida por una persona humana en la que un operario puede tomar una decisión en cuanto a si hay o no situación de alarma.

El documento WO 2004 006184 da a conocer un método para la detección de eventos en el que no se transmite información si no se detecta movimiento por debajo de un umbral. El método da a conocer transmitir en dos modos: transmitiendo un primer modo un flujo de características y transmitiendo el segunda modo las características y una parte de flujo de imágenes, usando ambos modos una resolución y una frecuencia de fotogramas diferentes.

El documento US 2003/0085992 A1 da a conocer un método y un aparato para proporcionar vigilancia por inmersión, en el que un guardia de seguridad remoto puede monitorizar una escena usando una variedad de fuentes de imagen que se representan sobre un modelo para proporcionar una vista conceptual tridimensional de la escena. Usando un selector de vista, el guardia de seguridad puede seleccionar dinámicamente una vista de cámara para su presentación visual en su modelo conceptual, realizar un recorrido de la escena, identificar objetos en movimiento y seleccionar la mejor vista de esos objetos en movimiento.

40 **Antecedentes de la invención**

La videovigilancia de espacios públicos se ha generalizado enormemente y aceptado por el público general. Por desgracia, los sistemas de videovigilancia convencionales producen volúmenes de datos tan ingentes que se obtiene como resultado un problema inextricable en el análisis de datos de videovigilancia.

45 Existe la necesidad de reducir la cantidad de datos de videovigilancia de modo que pueda llevarse a cabo el análisis de los datos de videovigilancia.

50 Existe la necesidad de filtrar los datos de videovigilancia para identificar partes deseadas de los datos de videovigilancia.

**Sumario de la invención**

55 Un objetivo de la invención es reducir la cantidad de datos de videovigilancia de modo que pueda llevarse a cabo el análisis de los datos de videovigilancia.

Un objetivo de la invención es filtrar los datos de videovigilancia para identificar partes deseadas de los datos de videovigilancia.

60 Un objetivo de la invención es producir una alarma en tiempo real basándose en una detección automática de un evento a partir de los datos de videovigilancia.

Un objetivo de la invención es integrar los datos procedentes de sensores de vigilancia que no sean de vídeo para obtener capacidades de búsqueda mejoradas.

65 Un objetivo de la invención es integrar datos procedentes de sensores de vigilancia que no sean de vídeo para

obtener capacidades de detección de eventos mejoradas.

La invención incluye dos métodos según las reivindicaciones 1 y 2.

- 5 La invención incluye un medio legible por ordenador según la reivindicación 3 y un sistema de procesamiento de vídeo según la reivindicación 4.

Además, los anteriores objetivos y ventajas de la invención son ilustrativos, y no exhaustivos, de los que pueden conseguirse mediante la invención, que se define mediante las reivindicaciones adjuntas. Por tanto, estos y otros objetivos y ventajas de la invención resultarán evidentes a partir de la descripción en el presente documento, tanto de la manera incorporada en el presente documento como de manera modificada en vista de cualquier variación que resulte evidente para los expertos en la técnica.

### Definiciones

15 Un "vídeo" se refiere a imágenes en movimiento representadas en forma analógica y/o digital. Ejemplos de vídeo incluyen: televisión, películas, secuencias de imágenes procedentes de una cámara de vídeo u otro observador, y secuencias de imágenes generadas por ordenador.

20 Un "fotograma" se refiere a una imagen particular u otra unidad diferenciada en un vídeo.

Un "objeto" se refiere a un elemento de interés en un vídeo. Ejemplos de un objeto incluyen: una persona, un vehículo, un animal y un sujeto físico.

25 Una "actividad" se refiere a una o más acciones y/o una o más composiciones de acciones de uno o más objetos. Ejemplos de una actividad incluyen: entrar; salir; detenerse; moverse; subir; bajar; aumentar; y disminuir.

30 Una "ubicación" se refiere a un espacio en el que puede tener lugar una actividad. Una ubicación puede estar, por ejemplo, basada en escena o basada en imagen. Ejemplos de una ubicación basada en escena incluyen: un espacio público; una tienda; un espacio de venta al por menor; una oficina; un almacén; una habitación de hotel; un vestíbulo de hotel; un vestíbulo de un edificio; un casino; una estación de autobuses; una estación de tren; un aeropuerto; un puerto; un autobús; un tren; un avión; y un barco. Ejemplos de una ubicación basada en imagen incluyen: una imagen de vídeo; una línea en una imagen de vídeo; una zona en una imagen de vídeo; una sección rectangular de una imagen de vídeo; y una sección poligonal de una imagen de vídeo.

35 Un "evento" se refiere a uno o más objetos ocupados en una actividad. Puede hacerse referencia al evento con respecto a una ubicación y/o a un tiempo.

40 Un "ordenador" se refiere a cualquier aparato que pueda aceptar una entrada estructurada, procesar la entrada estructurada según unas reglas preestablecidas y producir resultados del procesamiento como salida. Ejemplos de un ordenador incluyen: un ordenador; un ordenador de propósito general; un superordenador; un ordenador central (*mainframe*); un superminiordenador; un miniordenador; una estación de trabajo; un microordenador; un servidor; una televisión interactiva; una combinación híbrida de un ordenador y una televisión interactiva; y hardware de aplicación específica para emular un ordenador y/o software. Un ordenador puede tener un único procesador o múltiples procesadores, que pueden funcionar en paralelo y/o no en paralelo. Un ordenador también se refiere a dos o más ordenadores conectados entre sí a través de una red para transmitir o recibir información entre los ordenadores. Un ejemplo de un ordenador de este tipo incluye un sistema informático distribuido para procesar información a través de ordenadores conectados mediante una red.

50 Un "medio legible por ordenador" se refiere a cualquier dispositivo de almacenamiento usado para almacenar datos que sea accesible por un ordenador. Ejemplos de un medio legible por ordenador incluyen: un disco duro magnético; un disco flexible; un disco óptico, tal como un CD-ROM y un DVD; una cinta magnética; un chip de memoria; y una onda portadora usada para transportar datos electrónicos legibles por ordenador, tales como las que se usan para transmitir y recibir correo electrónico o para acceder a una red.

55 "Software" se refiere a reglas preestablecidas para hacer funcionar un ordenador. Ejemplos de software incluyen: software; segmentos de código; instrucciones; programas informáticos; y lógica programada.

60 Un "sistema informático" se refiere a un sistema que tiene un ordenador, en el que el ordenador comprende un medio legible por ordenador que incorpora software para hacer funcionar el ordenador.

65 Una "red" se refiere a varios ordenadores y dispositivos asociados que están conectados mediante instalaciones de comunicación. Una red implica conexiones permanentes tales como cables o conexiones temporales tales como las que se establecen por teléfono u otros enlaces de comunicación. Ejemplos de una red incluyen: una internet, tal como Internet; una intranet; una red de área local (LAN); una red de área amplia (WAN); y una combinación de redes, tales como una internet y una intranet.

**Breve descripción de los dibujos**

- 5 Los siguientes ejemplos de antecedentes se explican en mayor detalle por medio de dibujos, en los que los mismos números de referencia se refieren a las mismas características.
- La figura 1 ilustra una vista en planta de un sistema de videovigilancia.
- 10 La figura 2 ilustra un diagrama de flujo para un sistema de videovigilancia.
- La figura 3 ilustra un diagrama de flujo para asignar tareas al sistema de videovigilancia.
- La figura 4 ilustra un diagrama de flujo para hacer funcionar el sistema de videovigilancia.
- 15 La figura 5 ilustra un diagrama de flujo para extraer primitivas de vídeo para el sistema de videovigilancia.
- La figura 6 ilustra un diagrama de flujo para realizar una acción con el sistema de videovigilancia.
- 20 La figura 7 ilustra un diagrama de flujo para la calibración semiautomática del sistema de videovigilancia.
- La figura 8 ilustra un diagrama de flujo para la calibración automática del sistema de videovigilancia.
- La figura 9 ilustra un diagrama de flujo adicional para el sistema de videovigilancia de la invención.
- 25 Las figuras 10-15 ilustran ejemplos de un sistema de videovigilancia aplicado para monitorizar una tienda de comestibles.
- La figura 16a muestra un diagrama de flujo de un subsistema de análisis de vídeo.
- 30 La figura 16b muestra el diagrama de flujo del subsistema de detección de aparición de evento y respuesta ante el mismo.
- La figura 17 muestra consultas en bases de datos a modo de ejemplo.
- 35 La figura 18 muestra tres detectores de actividad a modo de ejemplo que detectan cruces por cables trampa (figura 18a), merodeo (figura 18b) y robo (figura 18c).
- La figura 19 muestra una consulta de detector de actividad.
- 40 La figura 20 muestra una consulta a modo de ejemplo usando detectores de actividad y operadores booleanos con modificadores.
- Las figuras 21a y 21b muestran una consulta a modo de ejemplo usando múltiples niveles de combinadores, detectores de actividad y consultas de propiedad.
- 45 La figura 22 muestra una configuración a modo de ejemplo de un sistema de videovigilancia.
- La figura 23 muestra otra configuración a modo de ejemplo de un sistema de videovigilancia.
- 50 La figura 24 muestra otra configuración a modo de ejemplo de un sistema de videovigilancia.
- La figura 25 muestra una red que puede usarse en configuraciones a modo de ejemplo de realizaciones de la invención.

**55 Descripción detallada de la invención**

Un sistema de videovigilancia automática puede servir para monitorizar una ubicación con fines, por ejemplo, de investigación de mercado o de seguridad. El sistema puede ser una instalación de videovigilancia dedicada con componentes de vigilancia construidos a propósito, o el sistema puede ser una retroinstalación en un equipo de videovigilancia existente que aprovecha las alimentaciones (*feed*) de vídeo de vigilancia. El sistema puede analizar datos de vídeo procedentes de fuentes en directo o de medios grabados. El sistema puede procesar los datos de vídeo en tiempo real y almacenar las primitivas de vídeo extraídas para permitir una detección de eventos forense a muy alta velocidad con posterioridad. El sistema puede tener una respuesta preestablecida al análisis, tal como registrar datos, activar un mecanismo de alarma o activar otro sistema sensor. El sistema también puede integrarse con otros componentes de sistema de vigilancia. El sistema puede usarse para producir, por ejemplo, informes de seguridad o investigación de mercado que pueden confeccionarse a medida según las necesidades de un operador

y, como opción, pueden presentarse a través de una interfaz basada en web interactiva, u otro mecanismo de elaboración de informes.

Un operador dispone de la máxima flexibilidad a la hora de configurar el sistema usando discriminadores de evento. Los discriminadores de evento se identifican con uno o más objetos (cuyas descripciones se basan en primitivas de vídeo), junto con uno o más atributos espaciales opcionales y/o uno o más atributos temporales opcionales. Por ejemplo, un operador puede definir un discriminador de evento (denominado evento de “merodeo” en este ejemplo) como un objeto “persona” en el espacio “cajero automático” durante “más de 15 minutos” y “entre las 10:00 p.m. y las 6:00 a.m.”. Los discriminadores de evento pueden combinarse con operadores booleanos modificados para formar consultas más complejas.

Este sistema de videovigilancia a modo de ejemplo hace uso de técnicas de visión por ordenador muy conocidas y de dominio público. Por ejemplo, los sistemas de videovigilancia actuales usan grandes volúmenes de imágenes de vídeo como materia prima principal de intercambio de información. El sistema en el ejemplo usa primitivas de vídeo como materia prima principal con el uso de imágenes de vídeo representativas como prueba colateral. El sistema del ejemplo también puede calibrarse (de manera manual, semiautomática o automática) y después puede inferir automáticamente primitivas de vídeo a partir de las imágenes de vídeo. El sistema puede además analizar vídeo previamente procesado sin tener que volver a procesar por completo el vídeo. Al analizar vídeo previamente procesado, el sistema puede realizar análisis de inferencia basándose en primitivas de vídeo previamente registradas, lo que mejora en gran medida la velocidad de análisis del sistema informático.

El uso de primitivas de vídeo también puede reducir significativamente los requisitos de almacenamiento para el vídeo. Esto se debe a que el subsistema de detección de eventos y respuesta ante los mismos usa el vídeo sólo para ilustrar las detecciones. Por consiguiente, el vídeo puede almacenarse o transmitirse con menor calidad. En una posible realización, el vídeo puede almacenarse o transmitirse sólo cuando se detecta actividad, no todo el tiempo. En otra posible realización, la calidad del vídeo almacenado o transmitido puede depender de si se detecta actividad: puede almacenarse o transmitirse vídeo con mayor calidad (mayor frecuencia de fotogramas y/o tasa de transmisión de bits) cuando se detecta actividad y con menor calidad el resto del tiempo. En otra realización a modo de ejemplo, el almacenamiento de vídeo y la base de datos pueden gestionarse por separado, por ejemplo, mediante un grabador de vídeo digital (DVR), y el subsistema de procesamiento de vídeo puede simplemente controlar si hay datos almacenados y con qué calidad. En otra realización, el sistema de videovigilancia (o los componentes del mismo) pueden estar en un dispositivo de procesamiento (tal como un procesador de propósito general, un DSP, un microcontrolador, un ASIC, una FPGA, u otro dispositivo) incorporado en un dispositivo de gestión de vídeo tal como una cámara de vídeo digital, un servidor vídeo en red, un DVR o un grabador de vídeo en red (NVR), y el ancho de banda de vídeo enviado en flujo continuo desde el dispositivo puede modularse por el sistema. Sólo es necesario transmitir vídeo de alta calidad (tasa de transmisión de bits o frecuencia de fotogramas alta) a través de una red de vídeo por IP cuando se detectan actividades de interés. En esta realización, pueden difundirse primitivas procedentes de dispositivos con inteligencia habilitada a través de una red a múltiples aplicaciones de inferencia de actividad en ubicaciones físicamente diferentes para permitir que una única red de cámaras proporcione aplicaciones versátiles a través de procesamiento descentralizado.

La figura 22 muestra una configuración de una implementación del sistema de videovigilancia. El bloque 221 representa una entrada de vídeo digital sin procesar (sin comprimir). Esto puede obtenerse, por ejemplo, mediante captura analógico a digital de una señal de vídeo analógica o decodificación de una señal de vídeo digital. El bloque 222 representa una plataforma de hardware que aloja los principales componentes del sistema de videovigilancia (análisis de contenido de vídeo - bloque 225 - e inferencia de actividad - bloque 226). La plataforma de hardware puede contener otros componentes tales como un sistema operativo (bloque 223); un codificador de vídeo (bloque 224) que comprime vídeo digital sin procesar para envío en flujo continuo o almacenamiento de vídeo usando cualquier esquema de compresión disponible (JPEG, MJPEG, MPEG1, MPEG2, MPEG4, H.263, H.264, Wavelet o cualquier otro); un mecanismo de almacenamiento (bloque 227) para mantener datos tal como vídeo, vídeo comprimido, alertas y primitivas de vídeo – este dispositivo de almacenamiento puede ser, por ejemplo, un disco duro, RAM incorporada, memoria FLASH incorporada u otro medio de almacenamiento; y una capa de comunicaciones (bloque 228) que puede, por ejemplo, empaquetar y/o digitalizar datos para su transmisión por un canal de comunicación (bloque 229).

Puede haber otros componentes de software que residen en plataformas computacionales en otros nodos de una red a la que se conecta el canal 229 de comunicaciones. El bloque 2210 muestra una herramienta de gestión de reglas que es una interfaz de usuario para crear reglas de videovigilancia. El bloque 2211 muestra una consola de alertas para presentar visualmente alertas e informes a un usuario. El bloque 2212 muestra un dispositivo de almacenamiento (tal como DVR, NVR o PC) para almacenar alertas, primitivas y vídeo para su procesamiento adicional *a posteriori*.

Los componentes en la plataforma de hardware (bloque 222) pueden implementarse en cualquier hardware de procesamiento (procesador de propósito general, microcontrolador, DSP, ASIC, FPGA u otro dispositivo de procesamiento) en cualquier dispositivo de captura, procesamiento o gestión de vídeo tal como una cámara de vídeo, una cámara de vídeo digital, una cámara de vídeo por IP, un servidor de vídeo por IP, un grabador de vídeo

digital (DVR), un grabador de vídeo en red (NVR), un PC, un ordenador portátil, u otro dispositivo. Hay varios modos diferentes posibles de funcionamiento para esta configuración.

5 En un modo, el sistema está programado para buscar eventos específicos. Cuando tienen lugar estos eventos, se transmiten alertas a través del canal de comunicación (bloque 229) a otros sistemas.

En otro modo, se envía vídeo en flujo continuo desde el dispositivo de vídeo mientras éste está analizando los datos de vídeo. Cuando se producen eventos, se transmiten alertas a través del canal de comunicación (bloque 229).

10 Según la invención la codificación y envío en flujo continuo de vídeo se modula mediante el análisis de contenido y la inferencia de actividad. Cuando no hay actividad presente (no están generándose primitivas), no se envía en flujo continuo vídeo (o de baja calidad, tasa de transmisión de bits, frecuencia de fotogramas, resolución). Cuando hay presente actividad (están generándose primitivas), se envía en flujo continuo vídeo de mayor calidad, tasa de transmisión de bits, frecuencia de fotogramas, resolución. Cuando se detectan eventos de interés mediante la inferencia de eventos, se envía en flujo continuo vídeo de muy alta calidad, tasa de transmisión de bits, frecuencia de fotogramas, resolución.

20 En otro modo de funcionamiento, se almacena información en el dispositivo de almacenamiento incorporado (bloque 227). Los datos almacenados pueden consistir en vídeo digital (sin procesar o comprimido), primitivas de vídeo, alertas, u otra información. La calidad del vídeo almacenado también puede controlarse por la presencia de primitivas o alertas. Cuando hay primitivas y alertas, puede almacenarse vídeo de mayor calidad, tasa de transmisión de bits, frecuencia de fotogramas, resolución.

25 La figura 23 muestra otra configuración de una implementación del sistema de videovigilancia. El bloque 231 representa una entrada de vídeo digital sin procesar (sin comprimir). Esto puede obtenerse, por ejemplo, a través de captura analógico a digital de una señal de vídeo analógica o decodificación de una señal de vídeo digital. El bloque 232 representa una plataforma de hardware que aloja el componente de análisis del sistema de videovigilancia (bloque 235). La plataforma de hardware puede contener otros componentes tales como un sistema operativo (bloque 233); un codificador de vídeo (bloque 234) que comprime vídeo digital sin procesar para envío en flujo continuo o almacenamiento de vídeo usando cualquier esquema de compresión disponible (JPEG, MJPEG, MPEG1, MPEG2, MPEG4, H.263, H.264, Wavelet, o cualquier otro); un mecanismo de almacenamiento (bloque 236) para mantener datos tales como vídeo, vídeo comprimido, alertas y primitivas de vídeo - este dispositivo de almacenamiento puede ser, por ejemplo, un disco duro, RAM incorporada, memoria FLASH incorporada, u otro medio de almacenamiento; y una capa de comunicaciones (bloque 237) que puede, por ejemplo, empaquetar y/o digitalizar datos para su transmisión por un canal de comunicación (bloque 238). En el ejemplo mostrado en la figura 23, el componente de inferencia de actividad (bloque 2311) se muestra en un componente de hardware independiente (bloque 239) conectado a una red a la que se conecta el canal 238 de comunicación.

40 También puede haber otros componentes de software que residen en plataformas computacionales en otros nodos de esta red (bloque 239). El bloque 2310 muestra una herramienta de gestión de reglas, que es una interfaz de usuario para crear reglas de videovigilancia. El bloque 2312 muestra una consola de alertas para presentar visualmente alertas e informes a un usuario. El bloque 2313 muestra un dispositivo de almacenamiento que puede estar físicamente ubicado en la misma plataforma de hardware (tal como un disco duro, disk flexible, otro disco magnético, un CD, un DVD, otro disco óptico, un MD u otro disco magnetoóptico, un dispositivo de almacenamiento de estado sólido tal como RAM o FLASH RAM, u otro dispositivo de almacenamiento) o puede ser un dispositivo de almacenamiento independiente (tal como una unidad de disco externa, un PC, un ordenador portátil, un DVR, un NVR, u otro dispositivo de almacenamiento).

50 Los componentes en la plataforma de hardware (bloque 222) pueden implementarse en cualquier plataforma de procesamiento (procesador de propósito general, microcontrolador, DSP, FPGA, ASIC o cualquier otra plataforma de procesamiento) en cualquier dispositivo de captura, procesamiento o gestión de vídeo tal como una cámara de vídeo, cámara de vídeo digital, cámara de vídeo por IP, servidor de vídeo por IP, grabador de vídeo digital (DVR), grabador de vídeo en red (NVR), PC, ordenador portátil, u otro dispositivo. Los componentes en la plataforma de hardware de *back-end* (bloque 239) pueden implementarse en cualquier hardware de procesamiento (procesador de propósito general, microcontrolador, DSP, FPGA, ASIC, o cualquier otro dispositivo) en cualquier dispositivo de procesamiento tal como PC, ordenador portátil, ordenador monoplaca, DVR, NVR, servidor de vídeo, encaminador de red, dispositivo de mano (tal como videoteléfono, busca o PDA). Hay varios modos diferentes posibles de funcionamiento para esta configuración.

60 En un modo, el sistema está programado en el dispositivo de *back-end* (o cualquier otro dispositivo conectado al dispositivo de *back-end*) para buscar eventos específicos. El módulo de análisis de contenido (bloque 235) en la plataforma de procesamiento de vídeo (bloque 232) genera primitivas que se transmiten a la plataforma de procesamiento de *back-end* (bloque 239). El módulo de inferencia de eventos (bloque 2311) determina si se han infringido las reglas y genera alertas que pueden presentarse visualmente en una consola de alertas (bloque 2312) o almacenarse en un dispositivo de almacenamiento (bloque 2313) para su análisis posterior.

65

En otro modo, pueden almacenarse primitivas de vídeo y vídeo en un dispositivo de almacenamiento en la plataforma de *back-end* (2313) para su análisis posterior.

5 En otro modo, la calidad, tasa de transmisión de bits, frecuencia de fotogramas, resolución del vídeo almacenado pueden modularse por alertas. Cuando hay una alerta, puede almacenarse vídeo con mayor calidad, tasa de transmisión de bits, frecuencia de fotogramas, resolución.

10 En otro modo, pueden almacenarse primitivas de vídeo en el dispositivo de procesamiento de vídeo (bloque 236 en el bloque 232) para su análisis posterior a través del canal de comunicación.

15 En otro modo, la calidad del vídeo almacenado en el dispositivo de procesamiento de vídeo (en bloque 236 en bloque 232) puede modularse por la presencia de primitivas. Cuando hay primitivas (cuando está sucediendo algo) pueden aumentarse la calidad, tasa de transmisión de bits, frecuencia de fotogramas, resolución del vídeo almacenado.

20 En otro modo, puede enviarse en flujo continuo vídeo desde el procesador de vídeo a través del codificador (234) a otros dispositivos en la red, a través del canal 238 de comunicación.

25 En otro modo, la calidad de vídeo puede modularse por el módulo de análisis de contenido (235). Cuando no hay primitivas (no está sucediendo nada), no se envía en flujo continuo vídeo (o de baja calidad, tasa de transmisión de bits, frecuencia de fotogramas, resolución). Cuando hay actividad, se envía en flujo continuo vídeo de mayor calidad, tasa de transmisión de bits, frecuencia de fotogramas, resolución.

30 En otro modo, la calidad, tasa de transmisión de bits, frecuencia de fotogramas, resolución del vídeo enviado en flujo continuo puede modularse por la presencia de alertas. Cuando el módulo de inferencia de eventos de *back-end* (bloque 2311) detecta un evento de interés, puede enviar una señal u orden al componente de procesamiento de vídeo (bloque 232) solicitando el vídeo (o vídeo de mayor calidad, tasa de transmisión de bits, frecuencia de fotogramas, resolución). Cuando se recibe esta petición, el componente de compresión de vídeo (bloque 234) y la capa de comunicación (bloque 237) pueden cambiar los parámetros de compresión y envío en flujo continuo.

35 En otro modo, la calidad del vídeo almacenado incorporado en el dispositivo de procesamiento de vídeo (bloque 236 en el bloque 232) puede modularse por la presencia de alertas. Cuando se genera una alerta por el módulo de inferencia de eventos (bloque 2311) en el procesador de *back-end* (bloque 239), puede enviar un mensaje a través del canal de comunicación (bloque 238) al hardware procesador de vídeo (bloque 232) para aumentar la calidad, tasa de transmisión de bits, frecuencia de fotogramas, resolución del vídeo almacenado en el dispositivo de almacenamiento incorporado (238).

40 La figura 24 muestra una extensión de la configuración descrita en la figura 23. Al separar la funcionalidad de análisis de contenido de vídeo y la inferencia de actividad de *back-end*, es posible habilitar un sistema de videovigilancia inteligente y versátil a través del proceso de enlace de aplicaciones posterior. Una única red de cámaras con inteligencia habilitada puede difundir un único flujo de primitivas de vídeo a aplicaciones de *back-end* independientes en diferentes partes de una organización (en diferentes ubicaciones físicas) y conseguir múltiples funciones. Esto es posible porque el flujo de primitivas contiene información sobre todo lo que está sucediendo en la escena y no está vinculado a áreas de aplicación específicas. El ejemplo representado en la figura 24 corresponde a un entorno de venta al por menor pero es ilustrativo del principio en general y es aplicable a cualquier otra área de aplicación y cualquier otra funcionalidad de vigilancia. El bloque 241 muestra una red con inteligencia habilitada de una o más cámaras de vídeo en una instalación o en múltiples instalaciones. El componente o componentes de análisis de contenido pueden residir en un dispositivo de procesamiento dentro de las cámaras, en servidores de vídeo, en encaminadores de red, en DVR, en NVR, en PC, en ordenadores portátiles o cualquier otro dispositivo de procesamiento de vídeo conectado a la red. A partir de estos componentes de análisis de contenido, se difunden flujos de primitivas a través de redes convencionales a módulos de inferencia de actividad en procesadores de *back-end* (bloques 242-245) que residen en zonas físicamente diferentes usadas para diferentes fines. Los procesadores de *back-end* pueden estar en ordenadores, ordenadores portátiles, DVR, NVR, encaminadores de red, dispositivos de mano (teléfonos, buscas, PDA) u otros dispositivos informáticos. Una ventaja de esa descentralización es que no es necesario que haya una aplicación de procesamiento central que deba programarse para hacer todo el procesamiento para todas las aplicaciones posibles. Otra ventaja es la seguridad de modo que una parte de una organización puede realizar inferencia de actividad basándose en reglas que están almacenadas localmente de modo que nadie más en la red tiene acceso a esa información.

60 En el bloque 242, el flujo de primitivas procedente de la red de cámaras inteligentes se analiza en cuanto a aplicaciones de seguridad física: para determinar si ha habido un franqueo de perímetro, vandalismo y para proteger bienes críticos. Evidentemente, estas aplicaciones son meramente a modo de ejemplo y es posible cualquier otra aplicación.

65 En el bloque 243, el flujo de primitivas procedente de la red de cámaras inteligente se analiza en cuanto a aplicaciones de prevención de pérdidas: para monitorizar un muelle de carga; para vigilar el robo de clientes o

empleados, para monitorizar un almacén y para hacer un seguimiento de mercancías. Evidentemente, estas aplicaciones son meramente a modo de ejemplo y es posible cualquier otra aplicación.

5 En el bloque 244, el flujo de primitivas procedente de la red de cámaras inteligente se analiza en cuanto a aplicaciones de seguridad pública y responsabilidad: para monitorizar el movimiento de personas o vehículos demasiado rápido en aparcamientos, para vigilar resbalones y caídas de personas y para monitorizar aglomeraciones de personas en y alrededor de la instalación. Evidentemente, estas aplicaciones son meramente a modo de ejemplo y es posible cualquier otra aplicación.

10 En el bloque 245, el flujo de primitivas procedente de la red de cámaras inteligente se analiza en cuanto a aplicaciones de inteligencia empresarial: para vigilar la longitud de las colas, para realizar un seguimiento del comportamiento de los consumidores, para aprender patrones de comportamiento, para realizar tareas de gestión de edificios, tales como control de iluminación y calefacción cuando no hay personas presentes. Evidentemente, estas aplicaciones son meramente a modo de ejemplo y es posible cualquier otra aplicación.

15 La figura 25 muestra una red (bloque 251) con varios posibles dispositivos con inteligencia habilitada conectados a la misma. El bloque 252 es una cámara IP con componentes de análisis de contenido incorporados que pueden enviar en flujo continuo primitivas por una red. El bloque 253 es una cámara IP con componentes tanto de análisis de contenido como de inferencia de actividad incorporados que pueden programarse directamente con reglas y que  
20 generarán alertas en red directamente. El bloque 254 es una cámara analógica convencional sin componentes inteligentes incorporados; pero está conectada a una plataforma de gestión de vídeo por IP (bloque 256) que realiza digitalización y compresión de vídeo así como análisis de contenido e inferencia de actividad. Puede programarse con reglas específicas de la vista y puede transmitir flujos de primitivas y alertas a través de una red. El bloque 255 es un DVR con componentes de inferencia de actividad que puede admitir flujos de primitivas procedentes de otros dispositivos y generar alertas. El bloque 257 es una PDA de mano habilitada con comunicaciones en red inalámbricas que tiene algoritmos de inferencia de actividad incorporados y puede aceptar primitivas de vídeo procedentes de la red y presentar visualmente alertas. El bloque 258 es un sistema de análisis de vídeo inteligente completo que puede aceptar flujos de vídeo analógico o digital, realizar análisis de contenido e inferencia de actividad y presentar visualmente alertas en una serie de consolas de alertas.

30 Como otro ejemplo, la invención proporciona asignación de tareas única al sistema. Usando directivas de control de equipo, los sistemas de vídeo actuales permiten a un usuario colocar sensores de vídeo y, en algunos sistemas convencionales sofisticados, enmascarar regiones de interés o de no interés. Las directivas de control de equipo son instrucciones para controlar la posición, orientación y enfoque de cámaras de vídeo. En lugar de directivas de control de equipo, la invención usa discriminadores de evento basados en primitivas de vídeo como mecanismo de  
35 asignación de tareas primario. Con discriminadores de evento y primitivas de vídeo, se proporciona a un operador un enfoque mucho más intuitivo sobre los sistemas convencionales para extraer información útil del sistema. En lugar de asignar tareas a un sistema con directivas de control de equipo, tales como “cámara A desplazamiento de 45 grados hacia la izquierda”, con la invención pueden asignarse tareas de manera intuitiva para los humanos con uno o más discriminadores de evento basados en primitivas de vídeo, tales como “una persona entra en la zona restringida A”.

40 Usando la invención para investigación de mercado, se dan los siguientes ejemplos del tipo de videovigilancia que puede realizarse con la invención: contar personas en una tienda; contar personas en una parte de una tienda; contar personas que se detienen en un lugar particular en una tienda; medir cuánto tiempo pasan personas en una tienda; medir cuánto tiempo pasan personas en una parte de una tienda; y medir la longitud de una cola en una tienda.

45 Usando la invención para seguridad, se dan los siguientes ejemplos del tipo de videovigilancia que puede realizarse con la invención: determinar cuándo entra alguien a una zona restringida y almacenar las imágenes asociadas; determinar cuándo entra una persona a una zona un número inusual de veces; determinar cuándo se producen cambios en el espacio de estanterías y el espacio de almacenamiento que podrían no estar autorizados; determinar cuándo pasajeros a bordo de una aeronave se acercan a la cabina de mando; determinar cuándo personas se cuelan detrás de otras por un portal de acceso protegido; determinar si hay una bolsa abandonada en un aeropuerto; y determinar si hay un robo de un bien.

50 Un área de aplicación a modo de ejemplo puede ser control de acceso, que puede incluir, por ejemplo: detectar si una persona trepa por una valla, o entra en una zona prohibida; detectar si alguien se mueve en la dirección equivocada (por ejemplo, en un aeropuerto, entrando en una zona protegida por la salida); determinar si un número de objetos detectados en una zona de interés no coincide con un número esperado basándose en etiquetas RFID o pases de tarjeta de entrada, indicar la presencia de personal no autorizado. Esto puede ser útil también en una aplicación residencial, donde puede que el sistema de videovigilancia pueda diferenciar entre el movimiento una persona y un animal de compañía, eliminando así la mayor parte de falsas alarmas. Obsérvese que en muchas aplicaciones residenciales, la privacidad puede ser un problema; por ejemplo, puede que el propietario de una casa  
60 no desee que otra persona monitorice de manera remota la casa y que pueda ver lo que hay en la casa y lo que sucede en la casa. Por tanto, en algunas realizaciones usadas en tales aplicaciones, el procesamiento de vídeo

puede realizarse localmente, y pueden enviarse vídeo o instantáneas opcionales a una o más estaciones de monitorización remotas sólo cuando sea necesario (por ejemplo, pero sin limitarse a, detección de actividad criminal u otras situaciones de peligro).

5 Otra área de aplicación a modo de ejemplo puede ser la monitorización de bienes. Esto puede significar detectar si se saca un objeto de la escena, por ejemplo, si se retira un artefacto de un museo. En un entorno de venta al por menor la monitorización de bienes puede tener diversos aspectos asociados y puede incluir, por ejemplo: detectar si una sola persona coge un número sospechosamente grande de un artículo dado; determinar si una persona sale por la entrada, particularmente si lo hace empujando un carrito de la compra; determinar si una persona aplica a un artículo una etiqueta cuyo precio no coincide, por ejemplo, llenar una bolsa con el tipo de café más caro pero usar una etiqueta de precio de un tipo más barato; o detectar si una persona sale de un muelle de carga con cajas grandes.

15 Otra área de aplicación a modo de ejemplo puede ser con fines de seguridad. Esto puede incluir, por ejemplo: detectar si una persona resbala y cae, por ejemplo, en una tienda o en un aparcamiento; detectar si un coche va demasiado rápido en un aparcamiento; detectar si una persona está demasiado cerca del borde de un andén en una estación de tren o de metro cuando no hay ningún tren en la estación; detectar si una persona se encuentra en las vías; detectar si una persona queda atrapada en la puerta de un tren cuando éste empieza a moverse; o contar el número de personas que entran y salen de una instalación, manteniendo por tanto un recuento preciso, lo que puede ser muy importante en caso de emergencia.

20 Otra área de aplicación a modo de ejemplo puede ser la monitorización de tráfico. Esto puede incluir detectar si un vehículo se ha parado, especialmente en lugares como un puente o un túnel, o detectar si un vehículo estaciona en una zona en la que no está permitido el estacionamiento.

25 Otra área de aplicación a modo de ejemplo puede ser la prevención del terrorismo. Esto puede incluir, además de algunas de las aplicaciones mencionadas anteriormente, detectar si un objeto ha sido abandonado en un vestíbulo de aeropuerto, si se arroja un objeto por una valla o si se abandona un objeto en una vía de ferrocarril; detectar el merodeo de una persona o un vehículo dando vueltas por una infraestructura crítica; o detectar una lancha que se desplaza a gran velocidad y se aproxima a un barco en un puerto o en mar abierto.

30 Otra área de aplicación a modo de ejemplo puede ser el cuidado de enfermos y mayores, incluso en su domicilio. Esto puede incluir, por ejemplo, detectar si la persona se cae; o detectar un comportamiento inusual, como que la persona no entra en la cocina durante un periodo prolongado de tiempo.

35 La figura 1 ilustra una vista en planta de una videovigilancia a modo de ejemplo. Un sistema 11 informático comprende un ordenador 12 que tiene un medio 13 legible por ordenador que incorpora software para hacer funcionar el ordenador 12 según la invención. El sistema 11 informático está acoplado a uno o más sensores 14 de vídeo, uno o más grabadores 15 de vídeo y uno o más dispositivos 16 de entrada/salida (E/S). Los sensores 14 de vídeo también pueden estar acoplados opcionalmente a los grabadores 15 de vídeo para dirigir la grabación de datos de videovigilancia. El sistema informático está acoplado opcionalmente a otros sensores 17.

40 Los sensores 14 de vídeo proporcionan vídeo de origen al sistema 11 informático. Cada sensor 14 de vídeo puede estar acoplado al sistema 11 informático usando, por ejemplo, una conexión directa (por ejemplo, una interfaz de cámara digital *firewire*) o una red. Los sensores 14 de vídeo pueden existir antes de la instalación de la invención. Ejemplos de sensor 14 de vídeo incluyen: una cámara de vídeo; una cámara de vídeo digital; una cámara a color; una cámara monocroma; una cámara; un videocámara de mano, una cámara de PC; una cámara web; una cámara de vídeo de infrarrojos; y una cámara CCTV.

45 Los grabadores 15 de vídeo reciben datos de videovigilancia desde el sistema 11 informático para grabar y/o proporcionar vídeo de origen al sistema 11 informático. Cada grabador 15 de vídeo puede estar acoplado al sistema 11 informático usando, por ejemplo, una conexión directa o una red. Los grabadores 15 de vídeo pueden existir antes de la instalación de la invención o pueden instalarse como parte de la invención. El sistema de videovigilancia en el sistema 11 informático puede controlar cuándo y con qué ajuste de calidad graba vídeo un grabador 15 de vídeo. Ejemplos de grabador 15 de vídeo incluyen: un grabador de cinta de vídeo; un grabador de vídeo digital; un grabador de vídeo en red; un disco de vídeo; un DVD; y un medio legible por ordenador. El sistema también puede modular el ancho de banda y la calidad del vídeo enviado en flujo continuo por una red controlando un protocolo de codificador de vídeo y de envío en flujo continuo. Cuando se detectan actividades de interés, pueden codificarse y enviarse en flujo continuo imágenes a mayor tasa de transmisión de bits, frecuencia de fotogramas o resolución.

50 Los dispositivos 16 de E/S proporcionan entrada a y reciben salida del sistema 11 informático. Los dispositivos 16 de E/S pueden usarse para asignar tareas al sistema 11 informático y producir informes desde el sistema 11 informático. Ejemplos de dispositivos 16 de E/S incluyen: un teclado; un ratón; un lápiz; un monitor; una impresora; otro sistema informático; una red; y una alarma.

55 Los otros sensores 17 proporcionan entrada adicional al sistema 11 informático. Cada uno de los otros sensores 17

5 puede estar acoplado al sistema 11 informático usando, por ejemplo, una conexión directa o una red. Los otros sensores 17 pueden existir antes de la instalación de la invención. Ejemplos de otro sensor 17 incluyen, aunque no se limitan a: un sensor de movimiento; un cable trampa óptico; un sensor biométrico; un sensor de RFID; y un sistema de autorización basado en tarjeta o basado en teclado. Las salidas de los otros sensores 17 pueden grabarse por el sistema 11 informático, dispositivos de grabación y/o sistemas de grabación.

La figura 2 ilustra un diagrama de flujo para un sistema de videovigilancia a modo de ejemplo. Las figuras 10-15 ilustran ejemplos de un sistema de videovigilancia aplicado a la monitorización de una tienda de comestibles.

10 En el bloque 21, el sistema de videovigilancia está configurado tal como se comentó en cuanto a la figura 1. Cada sensor 14 de vídeo está orientado hacia una ubicación para la videovigilancia. El sistema 11 informático está conectado a las alimentaciones de vídeo procedentes de los equipos 14 y 15 de vídeo. El sistema de videovigilancia puede implementarse usando equipos existentes o nuevos equipos instalados para la ubicación.

15 En el bloque 22, se calibra el sistema de videovigilancia. Una vez que el sistema de videovigilancia está listo a partir del bloque 21, se produce la calibración. El resultado del bloque 22 es la capacidad del sistema de videovigilancia de determinar un tamaño y rapidez absolutos aproximados de un objeto particular (por ejemplo, una persona) en diversos lugares en la imagen de vídeo proporcionada por el sensor de vídeo. El sistema puede calibrarse usando calibración manual, calibración semiautomática y calibración automática. La calibración se describe adicionalmente tras la explicación del bloque 24.

20 En el bloque 23 de la figura 2, se asignan tareas al sistema de videovigilancia. La asignación de tareas se produce tras la calibración en el bloque 22 y es opcional. La asignación de tareas al sistema de videovigilancia implica especificar uno o más discriminadores de evento. Sin asignación de tareas, el sistema de videovigilancia funciona detectando y archivando primitivas de vídeo e imágenes de vídeo asociadas sin realizar ninguna acción, como en el bloque 45 en la figura 4.

25 La figura 3 ilustra un diagrama de flujo para asignar tareas al sistema de videovigilancia para determinar discriminadores de evento. Un discriminador de evento se refiere a uno o más objetos que opcionalmente interaccionan con uno o más atributos espaciales y/o uno o más atributos temporales. Un discriminador de evento se describe en términos de primitivas de vídeo (también denominadas metadatos de descripción de actividad). Algunos de los criterios de diseño de primitivas de vídeo incluyen los siguientes: capacidad de extraerse del flujo de vídeo en tiempo real; inclusión de toda la información relevante del vídeo; y concisión de representación.

30 La extracción en tiempo real de las primitivas de vídeo del flujo de vídeo es deseable para permitir que el sistema pueda generar alertas en tiempo real, y para hacerlo, puesto que el vídeo proporciona un flujo de entrada continua, el sistema no puede quedar rezagado.

35 Las primitivas de vídeo también deberían contener toda la información relevante del vídeo, puesto que en el momento de la extracción de las primitivas de vídeo, el sistema no conoce las reglas definidas por el usuario. Por tanto, las primitivas de vídeo deberían contener información para poder detectar cualquier evento especificado por el usuario, sin necesidad de regresar al vídeo y analizarlo de nuevo.

40 Una representación concisa también es deseable por múltiples motivos. Una finalidad de la invención propuesta puede ser ampliar el tiempo de reciclado de almacenamiento de un sistema de vigilancia. Esto puede conseguirse sustituyendo el almacenamiento de vídeo de buena calidad todo el tiempo por el almacenamiento de metadatos de descripción de actividad y vídeo con calidad dependiente de la presencia de actividad, como se comentó anteriormente. Por tanto, cuanto más concisas sean las primitivas de vídeo, más datos podrán almacenarse. Además, cuanto más concisa sea la representación de primitivas de vídeo, más rápido se vuelve el acceso a los datos, y esto, a su vez, puede acelerar la investigación forense.

45 El contenido exacto de las primitivas de vídeo puede depender de la aplicación y de posibles eventos de interés. A continuación se describen algunas realizaciones a modo de ejemplo.

50 Una realización a modo de ejemplo de las primitivas de vídeo puede incluir descriptores de escena/vídeo, que describen la escena y vídeo global. En general, esto puede incluir una descripción detallada del aspecto de la escena, por ejemplo, la ubicación del cielo, follaje, objetos artificiales, agua, etc.; y/o condiciones meteorológicas, por ejemplo, la presencia/ausencia de precipitación, niebla, etc. Para una aplicación de videovigilancia, por ejemplo, un cambio en la vista global puede ser importante. Descriptores a modo de ejemplo pueden describir cambios repentinos de iluminación; pueden indicar movimiento de cámara, especialmente el hecho de que la cámara empiece o deje de moverse, y en este último caso, si ha vuelto a su vista previa o al menos a una previamente conocida; pueden indicar cambios en la calidad de la alimentación de vídeo, por ejemplo, si de repente se ha vuelto más ruidosa o se ha oscurecido, indicando en potencia manipulación de la alimentación; o pueden mostrar un nivel de agua que cambia a lo largo de una masa de agua (para más información sobre enfoques específicos a este último problema, puede consultarse, por ejemplo, la patente estadounidense US 7 424 167, presentada el 1 de octubre de 2004).

- Otra realización a modo de ejemplo de las primitivas de vídeo puede incluir descriptores de objeto que se refieren a un atributo observable de un objeto visto en una alimentación de vídeo. La información que se almacena sobre un objeto puede depender del área de aplicación y de las capacidades de procesamiento disponibles. Descriptores de objeto a modo de ejemplo pueden incluir propiedades genéricas incluyendo, pero sin limitarse a, tamaño, forma, perímetro, posición, trayectoria, rapidez y dirección de movimiento, destacabilidad del movimiento y sus características, color, rigidez, textura y/o clasificación. El descriptor de objeto también puede contener más información específica de la aplicación y el tipo: para humanos, puede incluir la presencia y proporción del tono de piel, información de sexo y raza, algún modelo de cuerpo humano que describa la forma y pose humana; o para vehículos, puede incluir el tipo (por ejemplo, camión, utilitario, sedán, bicicleta, etc.), marca, modelo, número de matrícula. El descriptor de objeto también puede contener actividades, incluyendo, pero sin limitarse a, transportar un objeto, correr, caminar, estar de pie o subir los brazos. Algunas actividades, tales como hablar, discutir o colisionar, también pueden referirse a otros objetos. El descriptor de objeto también puede contener información de identificación, incluyendo, pero sin limitarse a, cara o modo de andar.
- Otra realización a modo de ejemplo de las primitivas de vídeo puede incluir descriptores de flujo que describen la dirección de movimiento de cada zona del vídeo. Tales descriptores pueden usarse, por ejemplo, para detectar eventos de reutilización de acreditación de acceso, al detectar cualquier movimiento en una dirección prohibida (para más información sobre enfoques específicos para este último problema, puede consultarse, por ejemplo, la patente estadounidense en tramitación junto con la presente US 7 646 401, presentada el 30 de enero de 2004, e incorporada como referencia al presente documento).
- Las primitivas también pueden proceder de fuentes no de vídeo, tales como sensores de audio, sensores de calor, sensores de presión, lectores de tarjetas, etiquetas RFID, sensores biométricos, etc.
- Clasificación se refiere a una identificación de un objeto como perteneciente a una categoría o clase particular. Ejemplos de clasificación incluyen: una persona; un perro; un vehículo; un coche de policía; una persona individual; y un tipo específico de objeto.
- Tamaño se refiere a un atributo dimensional de un objeto. Ejemplos de tamaño incluyen: grande; mediano; pequeño; plano; altura superior a 6 pies; altura inferior a 1 pie; anchura superior a 3 pies; anchura inferior a 4 pies; aproximadamente de tamaño humano; mayor que un humano; menor que un humano; aproximadamente el tamaño de un coche; un rectángulo en una imagen con dimensiones aproximadas en píxeles; y número de píxeles de imagen.
- Posición se refiere a un atributo espacial de un objeto. La posición puede ser, por ejemplo, una posición de imagen en coordenadas de píxel, una posición absoluta en el mundo real en algún sistema de coordenadas del mundo, o una posición relativa a una marca de referencia o a otro objeto.
- Color se refiere a un atributo cromático de un objeto. Ejemplos de color incluyen: blanco; negro; gris; rojo; un rango de valores HSV; un rango de valores YUV; un rango de valores RGB; un valor RGB promedio; un valor YUV promedio; y un histograma de valores RGB.
- Rigidez se refiere a un atributo de consistencia de forma de un objeto. La forma de objetos no rígidos (por ejemplo, personas o animales) puede cambiar de un fotograma a otro, mientras que la de objetos rígidos (por ejemplo, vehículos o casas) puede permanecer sin cambiar en gran medida de un fotograma a otro (excepto, tal vez, ligeros cambios debido a giros).
- Textura se refiere a un atributo de patrón de un objeto. Ejemplos de características de textura incluyen: autosemejanza; potencia espectral; linealidad; y aspereza.
- Movimiento interno se refiere a una medida de la rigidez de un objeto. Un ejemplo de objeto bastante rígido es un coche, que no muestra gran cantidad de movimiento interno. Un ejemplo de objeto bastante poco rígido es una persona cuyos brazos y piernas oscilan, lo que muestra gran cantidad de movimiento interno.
- Movimiento se refiere a cualquier movimiento que pueda detectarse automáticamente. Ejemplos de movimiento incluyen: aparición de un objeto; desaparición de un objeto; movimiento vertical de un objeto; movimiento horizontal de un objeto; y movimiento periódico de un objeto.
- Movimiento destacado se refiere a cualquier movimiento que pueda detectarse automáticamente y del que pueda realizarse un seguimiento durante un cierto periodo de tiempo. Tal objeto en movimiento muestra un movimiento aparentemente intencionado. Ejemplos de movimiento destacado incluyen: moverse de un lado a otro; y moverse para interactuar con otro objeto.
- Característica de un movimiento destacado se refiere a una propiedad de un movimiento destacado. Ejemplos de característica de un movimiento destacado incluyen: trayectoria; longitud de una trayectoria en el espacio de imagen;

- longitud aproximada de una trayectoria en una representación tridimensional del entorno; posición de un objeto en el espacio de imagen en función del tiempo; posición aproximada de un objeto en una representación tridimensional del entorno en función del tiempo; duración de una trayectoria; velocidad (por ejemplo, rapidez y dirección) en el espacio de imagen; velocidad aproximada (por ejemplo, rapidez y dirección) en una representación tridimensional del entorno; duración de tiempo a una velocidad; cambio de velocidad en el espacio de imagen; cambio aproximado de velocidad en una representación tridimensional del entorno; duración de un cambio de velocidad; cese de movimiento; y duración de cese de movimiento. Velocidad se refiere a la rapidez y dirección de un objeto en un momento particular. Trayectoria se refiere a un conjunto de pares (posición, velocidad) para un objeto mientras pueda realizarse un seguimiento del objeto o durante un periodo de tiempo.
- Cambio de escena se refiere a cualquier región de una escena que pueda detectarse como que cambia a lo largo de un periodo de tiempo. Ejemplos de cambio de escena incluyen: un objeto estacionario que abandona una escena; un objeto que entra en una escena y pasa a ser estacionario; un objeto que cambia de posición en una escena; y un objeto que cambia de aspecto (por ejemplo color, forma o tamaño).
- Característica de un cambio de escena se refiere a una propiedad de un cambio de escena. Ejemplos de característica de un cambio de escena incluyen: tamaño de un cambio de escena en el espacio de imagen; tamaño aproximado de un cambio de escena en una representación tridimensional del entorno; momento en el que se produjo un cambio de escena; ubicación de un cambio de escena en el espacio de imagen; y ubicación aproximada de un cambio de escena en una representación tridimensional del entorno.
- Modelo predefinido se refiere a un modelo conocido *a priori* de un objeto. Ejemplos de modelo predefinido pueden incluir: un adulto; un niño; un vehículo; y un semirremolque.
- La figura 16a muestra una parte de análisis de vídeo a modo de ejemplo de un sistema de videovigilancia. En la figura 16a, un sensor 1601 de vídeo (por ejemplo, pero sin limitarse a, una cámara de vídeo) puede proporcionar un flujo 1602 de vídeo a un subsistema 1603 de análisis de vídeo. El subsistema 1603 de análisis de vídeo puede realizar entonces un análisis del flujo 1602 de vídeo para derivar primitivas de vídeo, que pueden almacenarse en el almacenamiento 1605 de primitivas. El almacenamiento 1605 de primitivas puede usarse para almacenar primitivas no de vídeo también. El subsistema 1603 de análisis de vídeo puede controlar además el almacenamiento de todo o de partes del flujo 1602 de vídeo en el almacenamiento 1604 de vídeo, por ejemplo, calidad y/o cantidad de vídeo, tal como se comentó anteriormente.
- Con referencia ahora a la figura 16b, una vez que están disponibles las primitivas 161 de vídeo y, si hay otros sensores, no de vídeo, el sistema puede detectar eventos. El usuario asigna tareas al sistema definiendo reglas 163 y correspondientes respuestas 164 usando la interfaz 162 de definición de reglas y respuestas. Las reglas se traducen en discriminadores de evento, y el sistema extrae correspondientes apariciones 165 de evento. Las apariciones 166 de evento detectadas disparan respuestas 167 definidas por el usuario. Una respuesta puede incluir una instantánea de un vídeo del evento detectado desde el almacenamiento 168 de vídeo (que puede ser o no el mismo que el almacenamiento 1604 de vídeo en la figura 16a). El almacenamiento 168 de vídeo puede formar parte del sistema de videovigilancia, o puede ser un dispositivo 15 de grabación independiente. Ejemplos de respuesta pueden incluir, pero no se limitan necesariamente a, los siguientes: activar una alerta visual y/o de audio en una pantalla de sistema; activar un sistema de alarma visual y/o de audio en la ubicación; activar una alarma silenciosa; activar un mecanismo de respuesta rápida; bloquear una puerta; ponerse en contacto con un servicio de seguridad; reenviar o enviar en flujo continuo datos (por ejemplo, datos de imagen, datos de vídeo, primitivas de vídeo; y/o datos analizados) a otro sistema informático a través de una red, tal como, pero sin limitarse a, Internet; guardar tales datos en un medio legible por ordenador designado; activar algún otro sensor o sistema de vigilancia; asignar tareas al sistema 11 informático y/u otro sistema informático; y/o dirigir el sistema 11 informático y/u otro sistema informático.
- Los datos de primitiva pueden considerarse como datos almacenados en una base de datos. Para detectar apariciones de evento en los mismos, es necesario un lenguaje de consulta eficaz. Las realizaciones del sistema de la invención pueden incluir un lenguaje de inferencia de actividad, que se describirá más adelante.
- Los esquemas de consulta de bases de datos relacionales tradicionales a menudo siguen una estructura de árbol binario booleano para permitir a los usuarios crear consultas flexibles respecto a datos almacenados de diversos tipos. Los nodos de hoja habitualmente tienen la forma de "propiedad relación valor", donde una propiedad es alguna característica clave de los datos (tal como tiempo o nombre); una relación es habitualmente un operador numérico (">", "<", "=", etc.); y un valor es un estado válido para esa propiedad. Los nodos de rama habitualmente representan operadores lógicos booleanos unarios o binarios como "y", "o" y "no".
- Esto puede constituir la base de un esquema de formulación de consulta de actividad. En el caso de una aplicación de videovigilancia, las propiedades pueden ser características del objeto detectado en el flujo de vídeo, tal como tamaño, rapidez, color, clasificación (ser humano, vehículo), o las propiedades pueden ser propiedades de cambio de escena. La figura 17 da ejemplos del uso de tales consultas. En la figura 17a, se realiza la consulta, "Muéstrame cualquier vehículo rojo", 171. Ésta puede descomponerse en dos consultas de "propiedad relación valor" (o

5 simplemente “propiedad”), que comprueban si la clasificación de un objeto es vehículo 173 y si su color es predominantemente rojo 174. Estas dos subconsultas pueden combinarse con el operador booleano “y” 172. De manera similar, en la figura 17b, la consulta, “Muéstreme cuándo una cámara empieza a o deja de moverse”, puede expresarse como la combinación booleana “o” 176 de las subconsultas de propiedad, “ha empezado la cámara a moverse” 177 y “ha dejado la cámara de moverse” 178.

10 Este tipo de esquema de consulta de base de datos puede ampliarse de dos maneras a modo de ejemplo: (1) los nodos de hoja básicos pueden aumentarse con detectores de actividad que describen actividades espaciales dentro de una escena; y (2) los nodos de rama de operador booleano pueden aumentarse con modificadores que especifican interrelaciones espaciales, temporales y de objeto.

15 Los detectores de actividad corresponden a un comportamiento relacionado con una zona de la escena de vídeo. Describen cómo un objeto puede interactuar con una ubicación en la escena. La figura 18 ilustra tres detectores de actividad a modo de ejemplo. La figura 18a representa el comportamiento de cruzar un perímetro en una dirección particular usando un cable trampa de vídeo virtual (para información adicional sobre cómo pueden implementarse tales cables trampa de vídeo virtuales, puede consultarse, por ejemplo, la patente estadounidense n.º 6.696.945). La figura 18b representa el comportamiento de merodeo durante un periodo de tiempo por una vía de ferrocarril. La figura 18c representa el comportamiento de llevarse algo de una sección de una pared (para enfoques a modo de ejemplo de cómo puede hacerse esto puede consultarse la solicitud de patente estadounidense n.º 20 10/331.778, titulada, “Video Scene Background Maintenance - Change Detection & Classification”, presentada el 30 de enero de 2003). Otros detectores de actividad a modo de ejemplo pueden incluir detectar que una persona se cae, detectar que una persona cambia de dirección o rapidez, detectar que una persona entra en una zona o detectar que una persona va en la dirección equivocada.

25 La figura 19 ilustra un ejemplo de cómo un nodo de hoja de detector de actividad (en este caso, cruce por cable trampa) puede combinarse con consultas de propiedad simples para detectar si un vehículo rojo cruza un cable trampa de vídeo 191. Las consultas de propiedad 172, 173, 174 y el detector 193 de actividad se combinan con un operador “y” booleano 192.

30 La combinación de consultas con operadores booleanos modificados (combinadores) puede añadir flexibilidad adicional. Modificadores a modo de ejemplo incluyen modificadores espaciales, temporales, de objeto y de contador.

35 Un modificador espacial puede hacer que el operador booleano opere sólo en actividades hijo (es decir, los argumentos del operador booleano, como se muestran bajo un operador booleano, por ejemplo, en la figura 19) que están próximas/no están próximas dentro de la escena. Por ejemplo, “y - a 50 píxeles de” puede usarse para que signifique que “y” sólo se aplica si la distancia entre actividades es inferior a 50 píxeles.

40 Un modificador temporal puede hacer que el operador booleano opere sólo en actividades hijo que se producen dentro de un periodo de tiempo especificado una respecto a otra, fuera de tal periodo de tiempo o dentro de un rango de tiempos. También puede especificarse el orden en el tiempo de los eventos. Por ejemplo “y - primero a 10 segundos del segundo” puede usar para que signifique que la “y” sólo se aplica si la segunda actividad hijo tiene lugar no más de 10 segundos después de la primera actividad hijo.

45 Un modificador de objeto puede hacer que el operador booleano opere sólo en actividades hijo que se producen implicando los mismos objetos u objetos diferentes. Por ejemplo “y - implicando el mismo objeto” puede usarse para que signifique que la “y” sólo se aplica si las dos actividades hijo implican el mismo objeto específico.

50 Un modificador de contador puede hacer que el operador booleano se dispare sólo si la condición o las condiciones se cumplen un número preestablecido de veces. Un modificador de contador puede incluir generalmente una relación numérica, tal como “al menos  $n$  veces”, “exactamente  $n$  veces”, “como mucho  $n$  veces”, etc. Por ejemplo, “o - al menos dos veces” puede usarse para que signifique que al menos dos de las subconsultas del operador “o” tienen que ser verdaderas. Otro uso del modificador de contador puede ser implementar una regla como “alertar si la misma persona coge al menos cinco artículos de una estantería”.

55 La figura 20 ilustra un ejemplo de usar combinadores. En este caso, la consulta de actividad requerida es “hallar un vehículo rojo que realice un giro ilegal a la izquierda” 201. El giro ilegal a la izquierda puede captarse a través de una combinación de descriptores de actividad y operadores booleanos modificados. Puede usarse un cable trampa virtual para detectar objetos que salen de la calle secundaria 193 y puede usarse otro cable trampa virtual para detectar objetos que se desplazan hacia la izquierda a lo largo de la carretera 205. Pueden combinarse mediante un operador 202 “y” modificado. El operador “y” booleano convencional garantiza que ambas actividades 193 y 205 tienen que detectarse. El modificador de objeto 203 comprueba que el mismo objeto cruzó ambos cables trampa, mientras que el modificador temporal 204 comprueba que el cable 193 trampa de abajo arriba se cruza en primer lugar, seguido por el cruce del cable 205 trampa de derecha a izquierda no más de 10 segundos después.

65 Este ejemplo también indica el poder de los combinadores. En teoría es posible definir un detector de actividad independiente para un giro a la izquierda sin basarse en combinadores y detectores de actividad simples. Sin

embargo, ese detector no sería flexible, haciendo difícil adaptarse a direcciones y ángulos de giro arbitrarios y también sería complicado escribir un detector independiente para todos los posibles eventos. Por el contrario, el uso de los combinadores y detectores simples proporciona una gran flexibilidad.

5 Otros ejemplos de actividades complejas que pueden detectarse como una combinación de unas más simples pueden incluir estacionamiento de un coche y una persona que sale del coche o múltiples personas que forman un grupo y montan una fiesta. Estos combinadores también pueden combinar primitivas de diferentes tipos y orígenes. Ejemplos pueden incluir reglas tales como “mostrar a una persona dentro de una sala antes de que se apaguen las  
10 luces”; “mostrar a una persona entrando por una puerta sin haber pasado previamente su tarjeta”; o “mostrar si una zona de interés tiene más objetos de lo esperado mediante un lector de etiquetas RFID”, es decir, un objeto ilegal sin etiqueta RFID se encuentra en la zona.

Un combinador puede combinar cualquier número de subconsultas e incluso puede combinar otros combinadores, hasta profundidades arbitrarias. Un ejemplo, ilustrado en las figuras 21a y 21b, puede ser una regla para detectar si  
15 un coche gira a la izquierda 2101 y a continuación gira a la derecha 2104. El giro a la izquierda 2101 puede detectarse con los cables 2102 y 2103 trampa direccionales, mientras que el giro a la derecha 2104 puede detectarse con los cables 2105 y 2106 trampa direccionales. El giro a la izquierda puede expresarse como detectores 2112 y 2113 de actividad de cable trampa, correspondientes a los cables 2102 y 2103 trampa, respectivamente, unidos con el combinador “y” 2111 con el modificador de objeto “el mismo” 2117 y el modificador temporal “2112 antes de 2113” 2118. De manera similar, el giro a la derecha puede expresarse como los detectores 2115 y 2116 de actividad de cable trampa, correspondientes a los cables 2105 y 2106 trampa, respectivamente, unidos con el combinador “y” 2114 con el modificador de objeto “el mismo” 2119 y el modificador temporal “2115 antes de 2116” 2120. Para detectar que el mismo objeto giró primero a la izquierda y luego a la derecha, el detector 2111 de giro a la izquierda y el detector 2114 de giro a la derecha se unen con el combinador “y” 2121 con el modificador de objeto “el mismo” 2122 y el modificador temporal “2111 antes de 2114” 2123. Finalmente, para  
25 garantizar que el objeto detectado es un vehículo, se usa un operador booleano “y” 2125 para combinar el detector 2121 de giro a la izquierda y a la derecha y la consulta de propiedad 2124.

Todos estos detectores pueden combinarse opcionalmente con atributos temporales. Ejemplos de atributo temporal incluyen: cada 15 minutos; entre 9:00 p.m. y 6:30 a.m.; menos de 5 minutos; más de 30 segundos; y durante el fin de semana.

En el bloque 24 de la figura 2 se hace funcionar el sistema de videovigilancia. El sistema de videovigilancia funciona automáticamente, detecta y archiva primitivas de vídeo de objetos en la escena, y detecta apariciones de evento en tiempo real usando discriminadores de evento. Además, se realiza una acción en tiempo real, según proceda, tal como activar alarmas, generar informes y generar una salida. Los informes y la salida pueden presentarse visualmente y/o almacenarse localmente en el sistema o en cualquier otro lugar a través de una red, tal como Internet. La figura 4 ilustra un diagrama de flujo para hacer funcionar el sistema de videovigilancia.

40 En el bloque 41, el sistema 11 informático obtiene vídeo de origen desde los sensores 14 de vídeo y/o los grabadores 15 de vídeo.

En el bloque 42 se extraen primitivas de vídeo en tiempo real a partir del vídeo de origen. Como opción, pueden obtenerse y/o extraerse primitivas no de vídeo desde de uno o más otros sensores 17 y usarse con la invención. La extracción de primitivas de vídeo se ilustra con la figura 5.

La figura 5 ilustra un diagrama de flujo para extraer primitivas de vídeo para el sistema de videovigilancia. Los bloques 51 y 52 funcionan en paralelo y pueden realizarse en cualquier orden o de manera simultánea. En el bloque 51 se detectan objetos a través de movimiento. Para este bloque puede usarse cualquier algoritmo de detección de movimiento para detectar el movimiento entre fotogramas a nivel de píxel. Como ejemplo, puede usarse la técnica de diferenciación de tres fotogramas, que se comenta en {1}. Los objetos detectados se reenvían al bloque 53.

En el bloque 52, los objetos se detectan a través de cambio. Para este bloque puede usarse cualquier algoritmo de detección de cambios para detectar cambios de un modelo de fondo. Se detecta un objeto en este bloque si se considera que uno o más píxeles en un fotograma están en el primer plano del fotograma porque los píxeles no se ajustan a un modelo de fondo del fotograma. Como ejemplo, puede usarse una técnica de moldeado de fondo estocástica, tal como la sustracción de fondo dinámicamente adaptativa, que se describe en {1} y en la solicitud de patente estadounidense n.º 09/694.712 presentada el 24 de octubre de 2000. Los objetos detectados se reenvían al  
55 bloque 53.

La técnica de detección de movimiento del bloque 51 y la técnica de detección de cambio del bloque 52 son técnicas complementarias, abordando cada técnica ventajosamente deficiencias de la otra técnica. Como opción, pueden usarse esquemas de detección adicionales y/o alternativos para las técnicas comentadas para los bloques 51 y 52. Ejemplos de un esquema de detección adicional y/o alternativo incluyen los siguientes: el esquema de detección de Pfister para encontrar personas como se describe en {8}; un esquema de detección de tono de piel; un esquema de detección facial; y un esquema de detección basado en modelo. Los resultados de tales esquemas de detección  
60 65

adicionales y/o alternativos se proporcionan al bloque 53.

Como opción, si el sensor 14 de vídeo tiene movimiento (por ejemplo, una cámara de vídeo que realiza un barrido, un zum y/o una traslación), puede insertarse un bloque adicional antes de que los bloques entre los bloques 51 y 52 para proporcionar entradas a los bloques 51 y 52 para la estabilización de vídeo. La estabilización de vídeo puede lograrse mediante compensación de movimiento global de proyección o afín. Por ejemplo, para obtener estabilización de vídeo puede usarse la alineación de imagen descrita en la solicitud de patente estadounidense n.º 09/609.919, presentada el 3 de julio de 2000, ahora patente estadounidense n.º 6.738.424, que se incorpora como referencia al presente documento.

En el bloque 53 se generan blobs. En general, un blob es cualquier objeto en un fotograma. Ejemplos de blob incluyen: un objeto en movimiento, tal como una persona o un vehículo; y un producto de consumo, tal como un mueble, un artículo de ropa o un artículo de estantería de venta al por menor. Los blobs se generan usando los objetos detectados a partir de los bloques 32 y 33. Para este bloque puede usarse cualquier técnica para generar blobs. Una técnica a modo de ejemplo para generar blobs a partir de detección de movimiento y detección de cambio usa un esquema de componentes conectados. Por ejemplo, puede usarse el algoritmo de componentes conectados y morfología que se describe en {1}.

En el bloque 54 se realiza un seguimiento de blobs. Para este bloque puede usarse cualquier técnica para realizar un seguimiento de blobs. Por ejemplo, puede usarse filtrado de Kalman o el algoritmo de condensación. Como otro ejemplo, puede usarse una técnica de coincidencia de plantillas, tal como se describe en {1}. Como ejemplo adicional, puede usarse un dispositivo de seguimiento de Kalman de múltiples hipótesis, que se describe en {5}. Como otro ejemplo más, puede usarse la técnica de seguimiento por fotogramas descrita en la patente estadounidense 6 954 498 presentada el 24 de octubre de 2000. Para el ejemplo en el que una ubicación es una tienda de comestibles, ejemplos de objetos de los que puede realizarse un seguimiento incluyen personas en movimiento, artículos de inventario y dispositivos de movimiento de inventario, tales como carritos o carros de la compra.

Como opción, los bloques 51-54 pueden sustituirse por cualquier esquema de detección y seguimiento, como conocen los expertos. Un ejemplo de tal esquema de detección y seguimiento se describe en {11}.

En el bloque 55 se analiza cada trayectoria de los objetos de los que se realiza un seguimiento para determinar si la trayectoria es destacada. Si la trayectoria no es destacada, la trayectoria representa un objeto que muestra movimiento inestable o representa un objeto de tamaño o color inestable, y el correspondiente objeto se rechaza y el sistema ya no lo analiza más. Si la trayectoria es destacada, la trayectoria representa un objeto que es potencialmente de interés. Se determina que una trayectoria es destacada o no destacada aplicando una medida de destacabilidad a la trayectoria. En {13} y {18} se describen técnicas para determinar si una trayectoria es destacada o no destacada.

En el bloque 56 se clasifica cada objeto. El tipo general de cada objeto se determina como la clasificación del objeto. La clasificación puede realizarse mediante varias técnicas, y ejemplos de tales técnicas incluyen usar un clasificador de red neuronal {14} y usar un clasificador de discriminación lineal {14}. Ejemplos de clasificación son los mismos que los comentados para el bloque 23.

En el bloque 57 se identifican primitivas de vídeo usando la información de los bloques 51-56 y procesamiento adicional según sea necesario. Ejemplos de primitivas de vídeo identificadas son los mismos que los comentados para el bloque 23. Como ejemplo, para el tamaño, el sistema puede usar información obtenida a partir de la calibración en el bloque 22 como primitiva de vídeo. A partir de la calibración, el sistema tiene información suficiente para determinar el tamaño aproximado de un objeto. Como otro ejemplo, el sistema puede usar la velocidad medida a partir del bloque 54 como primitiva de vídeo.

En el bloque 43 se archivan las primitivas de vídeo del bloque 42. Las primitivas de vídeo pueden archivar en el medio 13 legible por ordenador o en otro medio legible por ordenador. Junto con las primitivas de vídeo, pueden archivar fotogramas asociados o imágenes de vídeo del vídeo de origen. Esta etapa de archivado es opcional; si el sistema va a usarse sólo para detección de eventos en tiempo real, puede omitirse la etapa de archivado.

En el bloque 44 se extraen apariciones de evento a partir de las primitivas de vídeo usando discriminadores de evento. Las primitivas de vídeo se determinan en el bloque 42 y los discriminadores de evento se determinan a partir de la asignación de tareas al sistema en el bloque 23. Los discriminadores de evento se usan para filtrar las primitivas de vídeo para determinar si se ha producido alguna aparición de evento. Por ejemplo, un discriminador de evento puede ser buscar un evento de "camino equivocado" definido por una persona que se desplaza por el "camino equivocado" hacia una zona entre las 9:00 a.m. y las 5:00 p.m. El discriminador de evento comprueba todas las primitivas de vídeo que se generan según la figura 5 y determina si existe alguna primitiva de vídeo que tenga las siguientes propiedades: un sello de fecha y hora entre las 9:00 a.m. y las 5:00 p.m., una clasificación de "persona" o "grupo de personas", una posición dentro de la zona y una dirección o movimiento "incorrecto". Los discriminadores de evento también pueden usar otros tipos de primitivas, tal como se comentó anteriormente, y/o combinar primitivas

de vídeo de múltiples fuentes de vídeo para detectar apariciones de evento.

En el bloque 45 se realiza una acción para cada aparición de evento extraída en el bloque 44, según proceda. La figura 6 ilustra un diagrama de flujo para realizar una acción con el sistema de videovigilancia.

5 En el bloque 61 se inician respuestas según dicten los discriminadores de evento que detectaron las apariciones de evento. Las respuestas, en caso de existir, se identifican para cada discriminador de evento en el bloque 34.

10 En el bloque 62 se genera un registro de actividad para cada aparición de evento que se haya producido. El registro de actividad incluye, por ejemplo: detalles de una trayectoria de un objeto; el momento de detección de un objeto; la posición de detección de un objeto y una descripción o definición del discriminador de evento empleado. El registro de actividad puede incluir información, tal como primitivas de vídeo, que el discriminador de evento necesita. El registro de actividad también puede incluir vídeo representativo o imágenes sin movimiento representativas del objeto o los objetos implicados y/o la zona o las zonas implicadas en la aparición de evento. El registro de actividad se almacena en un medio legible por ordenador.

15 En el bloque 63 se genera una salida. La salida se basa en las apariciones de evento extraídas en el bloque 44 y en una alimentación directa del vídeo de origen del bloque 41. La salida se almacena en un medio legible por ordenador, se presenta visualmente en el sistema 11 informático o en otro sistema informático, o se reenvía a otro sistema informático. Mientras el sistema está en funcionamiento, se recopila información respecto a las apariciones de evento, y el operador puede visualizar la información en cualquier momento, incluyendo en tiempo real. Ejemplos de formatos para recibir la información incluyen: una presentación visual en un monitor de un sistema informático; una copia física; un medio legible por ordenador; y una página web interactiva.

20 La salida puede incluir una presentación visual de la alimentación directa del vídeo de origen del bloque 41 transmitida o bien a través de medios de transmisión de vídeo analógicos o bien a través de envío en flujo continuo de vídeo en red. Por ejemplo, el vídeo de origen puede presentarse visualmente en una ventana del monitor de un sistema informático o en un monitor de circuito cerrado. Además, la salida puede incluir vídeo de origen marcado con gráficos para resaltar los objetos y/o zonas implicados en la aparición de evento. Si el sistema está funcionando en modo de análisis forense, el vídeo puede proceder del grabador de vídeo.

25 La salida puede incluir uno o más informes para un operador basándose en los requisitos del operador y/o las apariciones de evento. Ejemplos de informe incluyen: el número de apariciones de evento que se han producido; las posiciones en la escena en las que se produjo la aparición de evento; los momentos en los que se produjeron las apariciones de evento; imágenes representativas de cada aparición de evento; vídeo representativo de cada aparición de evento; datos estadísticos sin procesar; estadísticas de apariciones de evento (por ejemplo, cuántas, con qué frecuencia, dónde y cuándo); y/o presentaciones visuales gráficas legibles por el ser humano.

30 Las figuras 13 y 14 ilustran un informe a modo de ejemplo para el pasillo en la tienda de comestibles de la figura 15. En las figuras 13 y 14, se identifican varias zonas en el bloque 22 y se etiquetan de manera correspondiente en las imágenes. Las zonas en la figura 13 coinciden con las de la figura 12 y las zonas en la figura 14 son diferentes. Al sistema se le asigna la tarea de buscar personas que se detienen en la zona.

35 En la figura 13, el informe a modo de ejemplo es una imagen de un vídeo marcado para incluir etiquetas, gráficos, información estadística y un análisis de la información estadística. Por ejemplo, la zona identificada como café tiene información estadística de un número promedio de clientes en la zona de 2/hora y un tiempo de permanencia promedio en la zona de 5 segundos. El sistema determinó que esta zona era una región "fría", lo que significa que no hay mucha actividad comercial por esta región. Como otro ejemplo, la zona identificada como refrescos tiene información estadística de un número promedio de clientes en la zona de 15/hora y un tiempo de permanencia promedio en la zona de 22 segundos. El sistema determinó que esta zona era una región "caliente", lo que significa que hay una gran cantidad de actividad comercial en esta región.

40 En la figura 14, el informe a modo de ejemplo es una imagen de un vídeo marcado para incluir etiquetas, gráficos, información estadística y un análisis de la información estadística. Por ejemplo, la zona en la parte posterior del pasillo tiene un número promedio de clientes de 14/hora y se determina que tiene poco tráfico. Como otro ejemplo, la zona en la parte anterior del pasillo tiene un número promedio de clientes de 83/hora y se determina que tiene mucho tráfico

45 Para la figura 13 o la figura 14, si el operador desea más información sobre cualquier zona particular o cualquier zona particular, una interfaz de apuntar y pinchar permite al operador navegar a través de imágenes de vídeo y sin movimiento representativas de regiones y/o actividades que el sistema ha detectado y archivado.

50 La figura 15 ilustra otro informe a modo de ejemplo para un pasillo en una tienda de comestibles. El informe a modo de ejemplo incluye una imagen de un vídeo marcado para incluir etiquetas e indicaciones de trayectoria y texto que describe la imagen marcada. Al sistema del ejemplo se le asigna la tarea de buscar en varias zonas: longitud, posición y tiempo de una trayectoria de un objeto; tiempo y ubicación en los que un objeto estuvo inmóvil;

65

correlación de trayectorias con zonas, según especifique el operador; y clasificación de un objeto como no una persona, una persona, dos personas y tres o más personas.

La imagen de vídeo de la figura 15 es de un periodo de tiempo en el que se grabaron las trayectorias. De los tres objetos, dos objetos se clasifican cada uno como una persona, y un objeto se clasifica como no una persona. A cada objeto se le asigna una etiqueta, concretamente ID de persona 1032, ID de persona 1033 e ID de 32001 objeto. Para el ID de persona 1032, el sistema determinó que la persona pasó 52 segundos en la zona y 18 segundos en la posición designada por el círculo. Para el ID de persona 1033, el sistema determinó que la persona pasó 1 minuto y 8 segundos en la zona y 12 segundos en la posición designada por el círculo. Las trayectorias para el ID de persona 1032 y el ID de persona 1033 se incluyen en la imagen marcada. Para el ID de objeto 32001, el sistema no analizó adicionalmente el objeto e indicó la posición del objeto con una X.

Con referencia de nuevo al bloque 22 en la figura 2, la calibración puede ser (1) manual, (2) semiautomática usando imágenes procedentes de un sensor de vídeo o un grabador de vídeo o (3) automática usando imágenes procedentes de un sensor de vídeo o un grabador de vídeo. Si se requieren imágenes, se supone que el vídeo de origen que va a analizar el sistema 11 informático procede de un sensor de vídeo que obtuvo el vídeo de origen usado para la calibración.

Para la calibración manual, el operador proporciona al sistema 11 informático la orientación y parámetros internos para cada uno de los sensores 14 de vídeo y la colocación de cada sensor 14 de vídeo con respecto a la ubicación. El sistema 11 informático puede mantener opcionalmente un mapa de la ubicación y la colocación de los sensores 14 de vídeo puede indicarse en el mapa. El mapa puede ser una representación bidimensional o tridimensional del entorno. Además, la calibración manual proporciona al sistema información suficiente para determinar el tamaño aproximado y la posición relativa de un objeto.

Alternativamente, para la calibración manual, el operador puede marcar una imagen de vídeo procedente del sensor con un gráfico que representa la aparición de un objeto de tamaño conocido, tal como una persona. Si el operador puede marcar una imagen en al menos dos ubicaciones diferentes, el sistema puede inferir información de calibración de cámara aproximada.

Para la calibración semiautomática y automática, no se requiere conocimiento de los parámetros de cámara o la geometría de escena. A partir de la calibración semiautomática y automática se genera una tabla de consulta para realizar una aproximación del tamaño de un objeto en diversas zonas en la escena, o se infieren los parámetros de calibración de cámara internos o externos de la cámara.

Para la calibración semiautomática, el sistema de videovigilancia se calibra usando una fuente de vídeo combinada con una entrada del operador. Una única persona se coloca en el campo de visión del sensor de vídeo que va a calibrarse de manera semiautomática. El sistema 11 informático recibe vídeo de origen con respecto a la única persona y automáticamente infiere el tamaño de persona basándose en estos datos. A medida que se aumenta el número de ubicaciones en el campo de visión del sensor de vídeo en el que se visualiza a la persona y a medida que se aumenta el periodo de tiempo en el que se visualiza a la persona en el campo de visión del sensor de vídeo, se aumenta la precisión de la calibración semiautomática.

La figura 7 ilustra un diagrama de flujo para la calibración semiautomática del sistema de videovigilancia. El bloque 71 es igual al bloque 41, excepto porque un objeto típico se mueve por la escena en diversas trayectorias. El objeto típico puede tener diversas velocidades y puede estar estacionario en diversas posiciones. Por ejemplo, el objeto típico se mueve acercándose al sensor de vídeo lo más posible y a continuación alejándose del sensor de vídeo lo más posible. Este movimiento por parte del objeto típico puede repetirse según sea necesario.

Los bloques 72-75 son iguales a los bloques 51-54, respectivamente.

En el bloque 76, el objeto típico se monitoriza por toda la escena. Se supone que el único objeto estable (o al menos el más estable) del que se realiza un seguimiento es el objeto de calibración en la escena (es decir, el objeto típico que se mueve por la escena). El tamaño del objeto estable se recopila para cada punto en la escena en la que se observa y se usa esta información para generar información de calibración.

En el bloque 77, el tamaño del objeto típico se identifica para diferentes zonas por toda la escena. El tamaño del objeto típico se usa para determinar los tamaños aproximados de objetos similares en diversas zonas en la escena. Con esta información se genera una tabla de consulta correlacionando los tamaños aparentes típicos del objeto típico en diversas zonas en la imagen, o se infieren parámetros de calibración de cámara internos o externos. Como salida de muestra, una visualización de figuras básicas en diversas zonas de la imagen indica lo que el sistema determinó como altura apropiada. Una figura básica de este tipo se ilustra en la figura 11.

Para la calibración automática, se lleva a cabo una fase de aprendizaje en la que el sistema 11 informático determina información con respecto a la ubicación en el campo de visión de cada sensor de vídeo. Durante la calibración automática, el sistema 11 informático recibe vídeo de origen de la ubicación durante un periodo de

tiempo representativo (por ejemplo, minutos, horas o días) que es suficiente para obtener un muestreo estadísticamente significativo de objetos típicos de la escena y así inferir ubicaciones y tamaños aparentes típicos.

5 La figura 8 ilustra un diagrama de flujo para la calibración automática del sistema de videovigilancia. Los bloques 81-86 son iguales a los bloques 71-76 en la figura 7.

10 En el bloque 87 se identifican regiones de las que puede realizarse un seguimiento en el campo de visión del sensor de vídeo. Una región de la que puede realizarse un seguimiento se refiere a una región en el campo de visión de un sensor de vídeo en la que puede realizarse un seguimiento sencillo y/o preciso de un objeto. Una región de la que no puede realizarse un seguimiento se refiere a una región en el campo de visión de un sensor de vídeo en la que no puede realizarse un seguimiento sencillo y/o preciso de un objeto y/o es difícil realizar un seguimiento del mismo. Una región de la que no puede realizarse un seguimiento puede designarse como región inestable o no destacada. Puede ser difícil realizar un seguimiento de un objeto porque el objeto es demasiado pequeño (por ejemplo, más pequeño que un umbral predeterminado), aparecer durante un tiempo muy corto (por ejemplo, inferior a un umbral predeterminado) o mostrar un movimiento que no es destacado (por ejemplo, no intencionado). Una región de la que puede realizarse un seguimiento puede identificarse usando, por ejemplo, las técnicas descritas en {13}.

20 La figura 10 ilustra regiones de las que puede realizarse un seguimiento determinadas para un pasillo en una tienda de comestibles. La zona en el extremo más alejado del pasillo se determina como no destacada porque aparecen demasiados elementos de confusión en esta zona. Un elemento de confusión se refiere a algo en un vídeo que confunde un esquema de seguimiento. Ejemplos de un elemento de confusión incluyen: hojas volando; lluvia; un objeto parcialmente oculto; y un objeto que aparece durante un tiempo demasiado corto para realizarse un seguimiento preciso del mismo. Por el contrario, la zona en el extremo más cercano del pasillo se determina como destacada porque se determinan seguimientos adecuados para esta zona.

25 En el bloque 88 se identifican los tamaños de los objetos para diferentes zonas por toda la escena. Los tamaños de los objetos se usan para determinar los tamaños aproximados de objetos similares en diversas zonas en la escena. Se usa una técnica, tal como usar un histograma o una mediana estadística, para determinar la anchura y altura aparentes típicas de objetos en función de la ubicación en la escena. En una parte de la imagen de la escena, los objetos típicos pueden tener una anchura y altura aparentes típicas. Con esta información, se genera una tabla de consulta correlacionando los tamaños aparentes típicos de objetos en diversas zonas en la imagen, o pueden inferirse los parámetros de calibración de cámara internos o externos.

35 La figura 11 ilustra la identificación de tamaños típicos para objetos típicos en el pasillo de la tienda de comestibles de la figura 10. Se supone que los objetos típicos son personas y se identifican mediante una etiqueta de manera correspondiente. Los tamaños típicos de personas se determinan a través de gráficas de la altura promedio y anchura promedio para las personas detectadas en la región destacada. En el ejemplo, la gráfica A se determina para la altura promedio de una persona promedio y la gráfica B se determina para la anchura promedio para una persona, dos personas y tres personas.

40 Para la gráfica A, el eje x representa la altura del blob en píxeles y el eje y representa el número de instancias de una altura particular, identificadas en el eje x, que se han producido. El pico de la línea para la gráfica A corresponde a la altura más común de blobs en la región designada en la escena y, para este ejemplo, el pico corresponde a la altura promedio de una persona situada en la región designada.

45 Suponiendo que las personas se desplazan en grupos abiertos, se genera un gráfico similar a la gráfica A para la anchura como gráfica B. Para la gráfica B, el eje x representa la anchura de los blobs en píxeles y el eje y representa el número de instancias de una anchura particular, identificada en el eje x, que se han producido. Los picos de la línea para la gráfica B corresponden a la anchura promedio de varios blobs. Suponiendo que la mayoría de grupos contienen sólo una persona, el pico más grande corresponde a la anchura más común, que corresponde a la anchura promedio de una única persona en la región designada. De manera similar, el segundo pico más grande corresponde a la anchura promedio de dos personas en la región designada y el tercer pico más grande corresponde a la anchura promedio de tres personas en la región designada.

55 La figura 9 ilustra un diagrama de flujo adicional para la videovigilancia. En este ejemplo adicional, el sistema analiza primitivas de vídeo archivadas con discriminadores de evento para generar informes adicionales, por ejemplo, sin necesidad de revisar todo el vídeo de origen. En cualquier momento después de que se haya procesado una fuente de vídeo según la invención, se archivan las primitivas de vídeo para el vídeo de origen en el bloque 43 de la figura 4. El contenido de vídeo puede volver a analizarse con la realización adicional en un tiempo relativamente corto porque sólo se revisan las primitivas de vídeo y porque la fuente de vídeo no vuelve a procesarse. Esto proporciona una mejora de gran eficacia con respecto a los sistemas del estado de la técnica actuales porque el procesamiento de datos de imágenes de vídeo es extremadamente costoso desde el punto de vista computacional, mientras que el análisis de las primitivas de vídeo de pequeño tamaño condensadas a partir del vídeo es extremadamente económico desde el punto de vista computacional. Como ejemplo, puede generarse el siguiente discriminador de evento: "El número de personas que se detienen durante más de 10 minutos en la zona A en los últimos dos meses". Con la realización adicional, no es necesario revisar los últimos dos meses de vídeo de origen. En su lugar,

sólo es necesario revisar las primitivas de vídeo de los últimos dos meses, lo que es un proceso significativamente más eficaz.

El bloque 91 es igual al bloque 23 en la figura 2.

5 En el bloque 92 se accede a las primitivas de vídeo archivadas. Las primitivas de vídeo se archivan en el bloque 43 de la figura 4.

Los bloques 93 y 94 son iguales a los bloques 44 y 45 en la figura 4.

10 Como aplicación a modo de ejemplo, la invención puede usarse para analizar un espacio de mercado de venta al por menor midiendo la eficacia de expositor de venta al por menor. Se inyectan grandes sumas de dinero en expositores de venta al por menor en un intento por que resulten lo más atractivos posible para promover las ventas tanto de los artículos en el expositor como de artículos subsidiarios. La invención puede configurarse para medir la  
15 eficacia de estos expositores de venta al por menor.

Para esta aplicación a modo de ejemplo, el sistema de videovigilancia se configura orientando el campo de visión de un sensor de vídeo hacia el espacio alrededor del expositor de venta al por menor deseado. Durante la asignación de tareas, el operador selecciona una zona representativa del espacio alrededor del expositor de venta al por menor deseado. Como discriminador, el operador define que desea monitorizar objetos con tamaño de persona que entran en la zona y o bien muestran una reducción medible en la velocidad o bien se detienen durante una cantidad de  
20 tiempo apreciable.

Tras funcionar durante un determinado periodo de tiempo, el sistema de videovigilancia puede proporcionar informes para análisis de mercado. Los informes pueden incluir: el número de personas que aminoraron el paso alrededor del expositor de venta al por menor; el número de personas que se detuvieron en el expositor de venta al por menor; el desglose de personas que se interesaron en el expositor de venta al por menor en función del tiempo, tal como cuántas se interesaron en fines de semana y cuántas se interesaron por las tardes; e instantáneas de vídeo de las personas que mostraron interés en el expositor de venta al por menor. La información de investigación de mercado  
25 obtenida a partir del sistema de videovigilancia puede combinarse con información de ventas de la tienda y registros de clientes de la tienda para mejorar la comprensión de los analistas de la eficacia del expositor de venta al por menor.

Las realizaciones y los ejemplos comentados en el presente documento son ejemplos no limitativos de la enseñanza general expresada por las reivindicaciones.  
35

**REIVINDICACIONES**

1. Método de procesamiento de vídeo que comprende:
- 5 detectar si hay o no una o más actividades en una secuencia de vídeo;
- si se detectan una o más actividades, generar una o más primitivas;
- 10 si se generan una o más primitivas, inferir si hay o no uno o más eventos de interés en la secuencia de vídeo basándose en las primitivas generadas;
- si no se generan primitivas, codificar la secuencia de vídeo para obtener vídeo codificado a una primera tasa de transmisión de bits, a una primera frecuencia de fotogramas y a una primera resolución;
- 15 si se generan una o más primitivas y no se infieren eventos de interés, codificar la secuencia de vídeo para obtener vídeo codificado a una segunda tasa de transmisión de bits, a una segunda frecuencia de fotogramas y a una segunda resolución, siendo al menos una de la segunda tasa de transmisión de bits, la segunda frecuencia de fotogramas y la segunda resolución superior a la primera tasa de transmisión de bits, la primera frecuencia de fotogramas y la primera resolución, respectivamente;
- 20 si se infieren uno o más eventos de interés, codificar la secuencia de vídeo para obtener vídeo codificado a una tercera tasa de transmisión de bits, a una tercera frecuencia de fotogramas y a una tercera resolución, siendo al menos una de la tercera tasa de transmisión de bits, la tercera frecuencia de fotogramas y la tercera resolución superior a la segunda tasa de transmisión de bits, la segunda frecuencia de fotogramas y la segunda resolución, respectivamente; y
- 25 transmitir dicho vídeo codificado.
2. Método de procesamiento de vídeo que comprende:
- 30 detectar si hay o no una o más actividades en una secuencia de vídeo;
- si se detectan una o más actividades, generar una o más primitivas;
- 35 si se generan una o más primitivas, inferir si hay o no uno o más eventos de interés en la secuencia de vídeo basándose en las primitivas generadas;
- si se generan una o más primitivas y no se infieren eventos de interés, codificar la secuencia de vídeo para obtener vídeo codificado a una primera tasa de transmisión de bits, a una primera frecuencia de fotogramas y a una primera resolución;
- 40 si se infieren uno o más eventos de interés, codificar la secuencia de vídeo para obtener vídeo codificado a una segunda tasa de transmisión de bits, a una segunda frecuencia de fotogramas y a una segunda resolución, siendo al menos una de la segunda tasa de transmisión de bits, la segunda frecuencia de fotogramas y la segunda resolución superior a la primera tasa de transmisión de bits, la primera frecuencia de fotogramas y la primera resolución, respectivamente;
- 45 si no se generan primitivas, no se transmite la secuencia de vídeo;
- 50 si se generan una o más primitivas, transmitir dicho vídeo codificado.
3. Medio legible por ordenador que contiene instrucciones que, cuando se ejecutan por un sistema informático, hacen que dicho sistema informático implemente el método según la reivindicación 1 ó 2.
- 55 4. Sistema de procesamiento de vídeo que comprende:
- el medio legible por ordenador según la reivindicación 3; y
- 60 un ordenador acoplado a dicho medio legible por ordenador para ejecutar las instrucciones contenidas en dicho medio legible por ordenador.

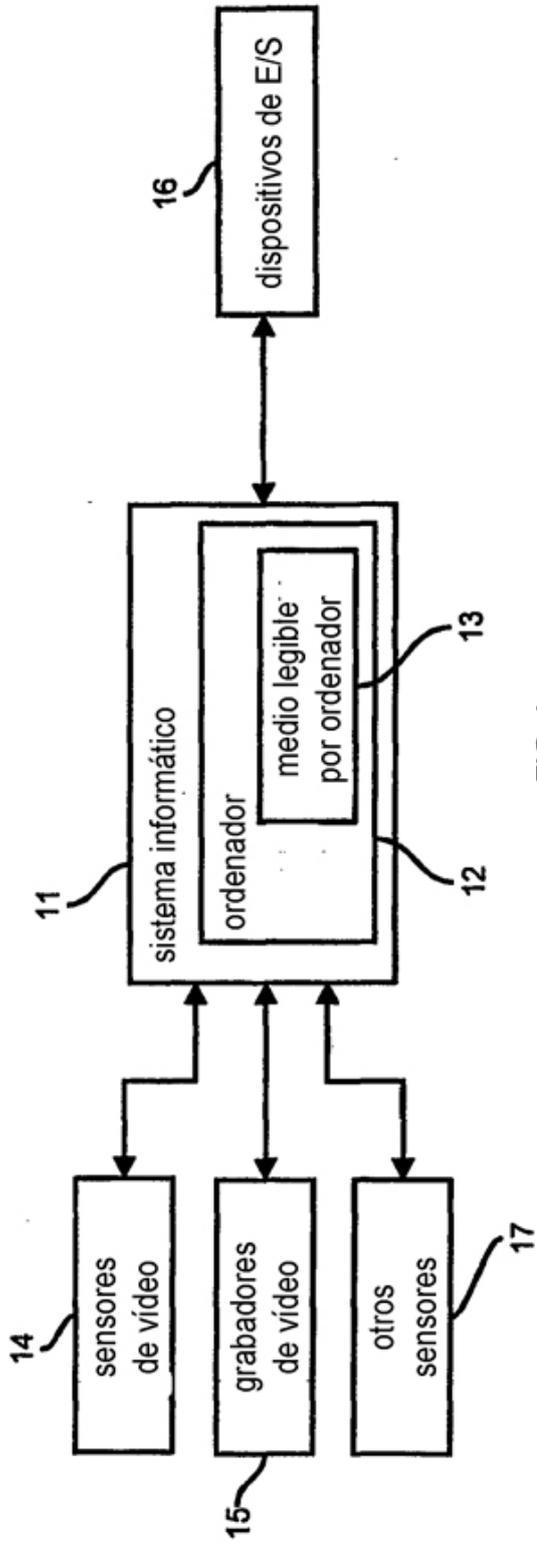


FIG. 1

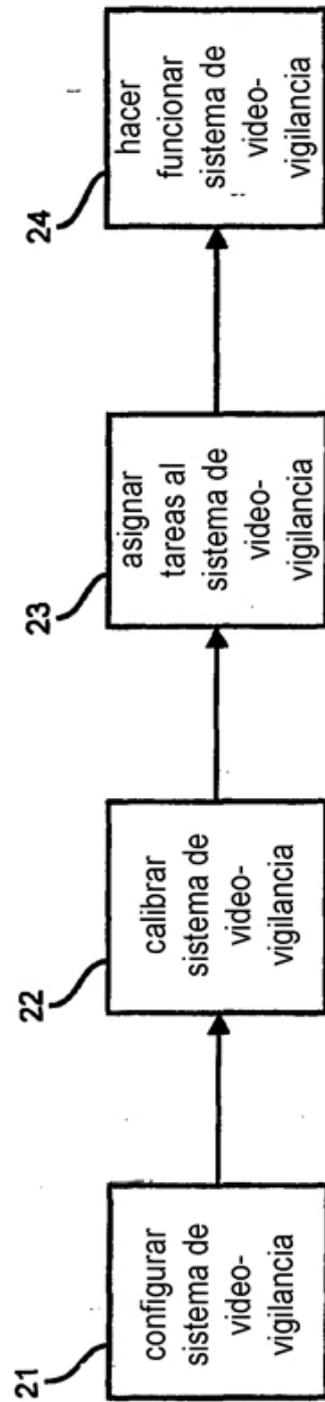


FIG. 2

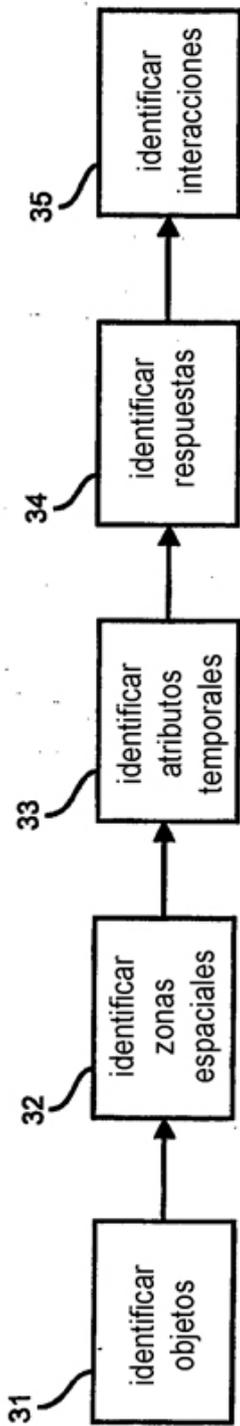


FIG. 3

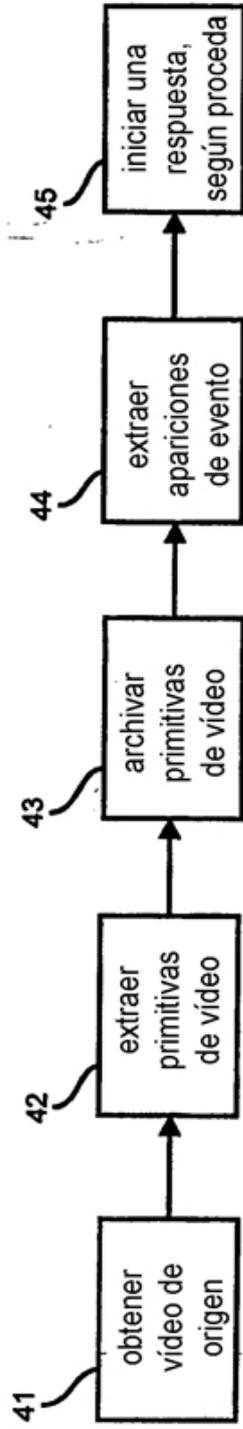


FIG. 4

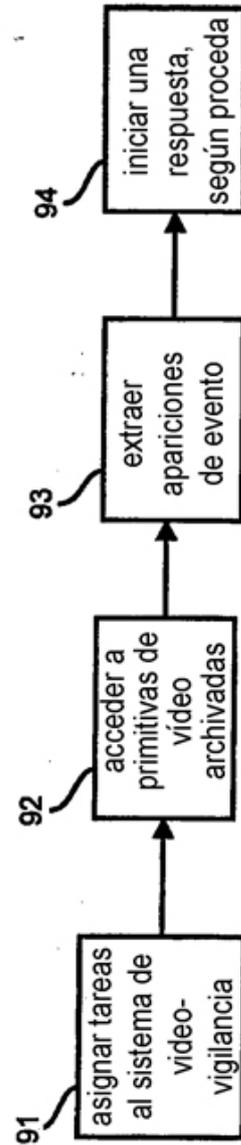
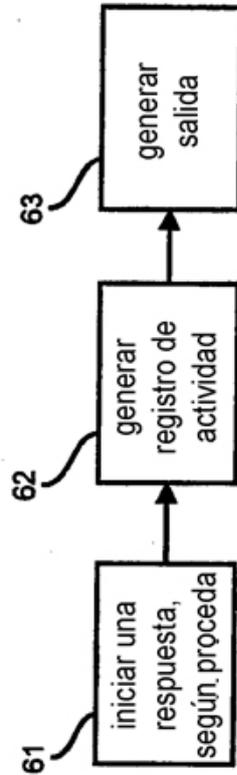
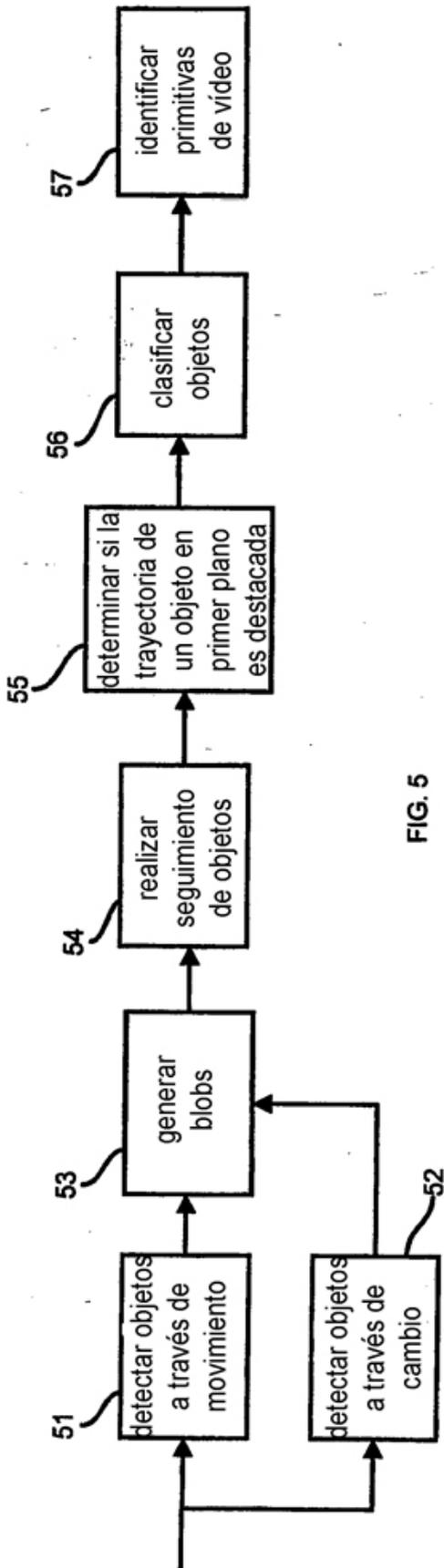


FIG. 9



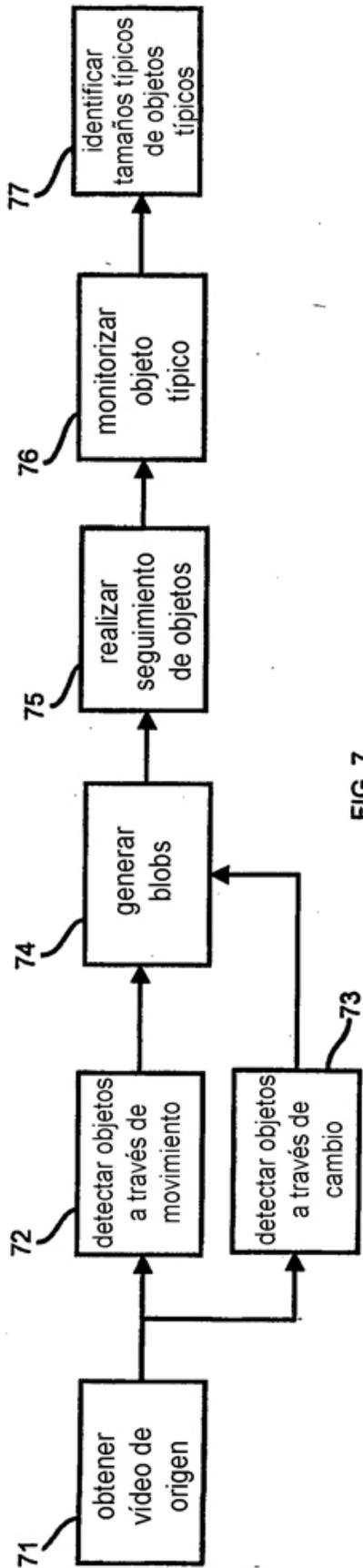


FIG. 7

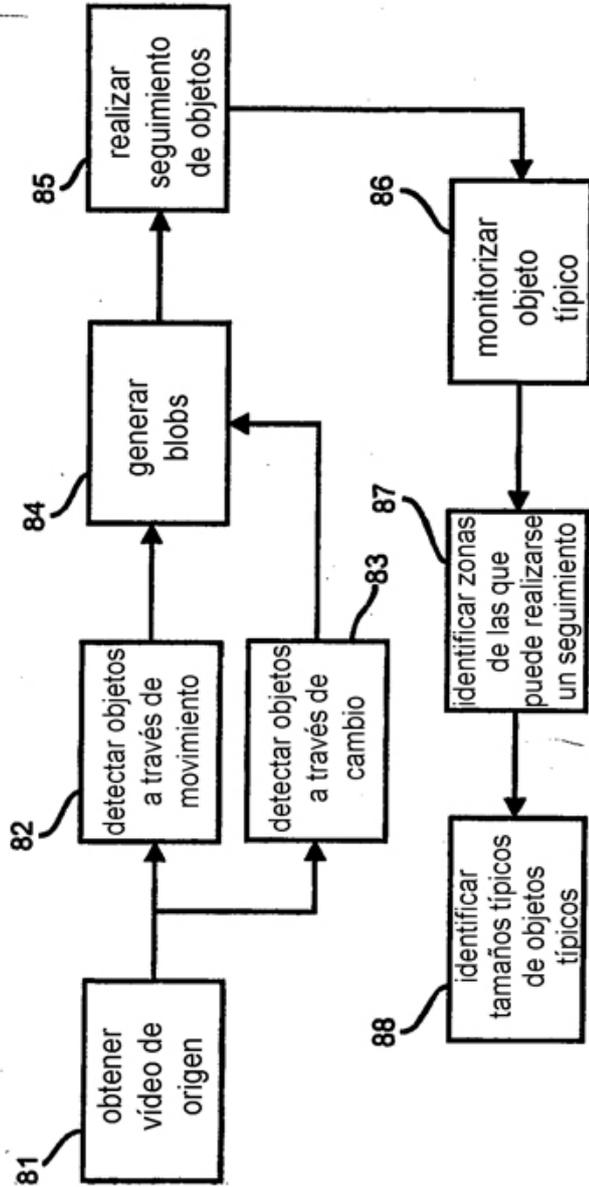


FIG. 8

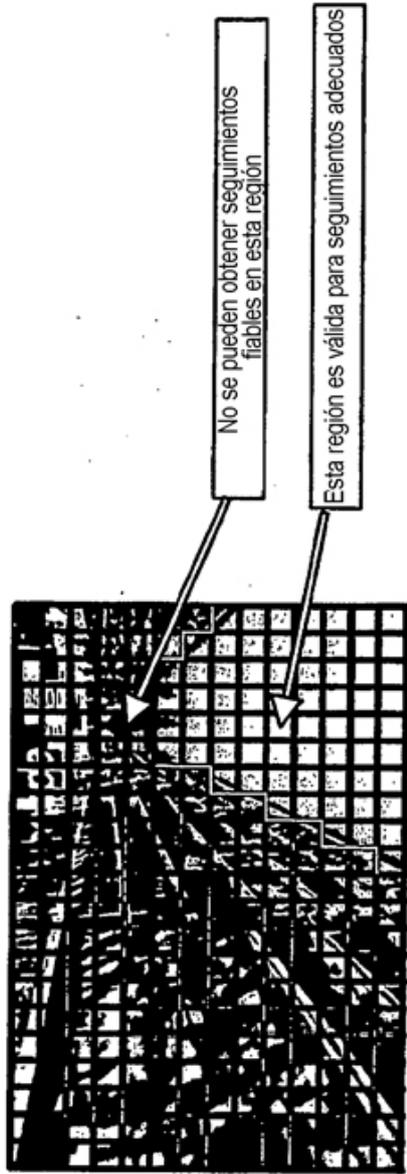


FIG. 10

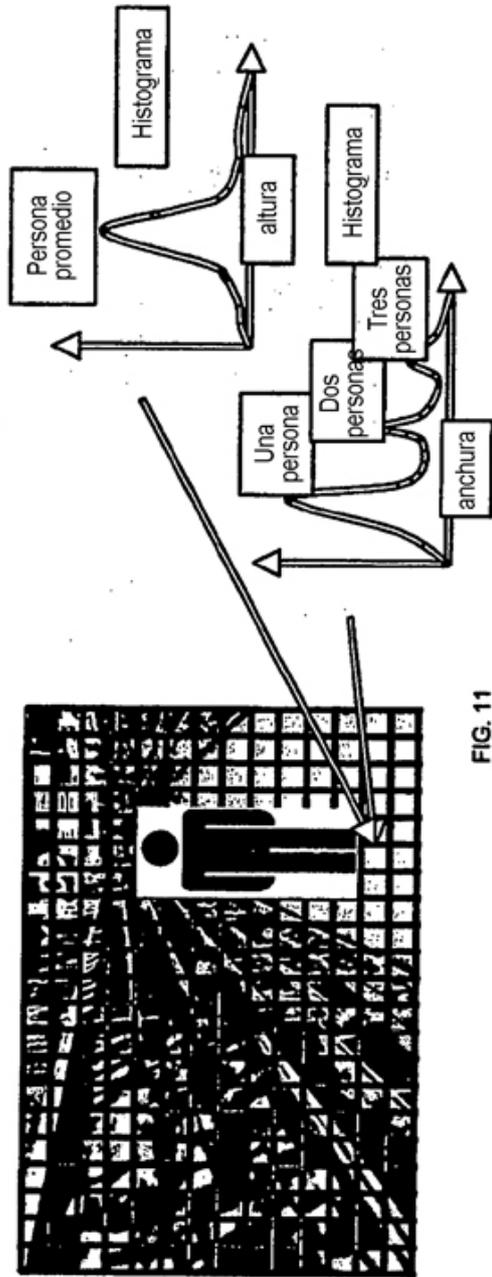


FIG. 11

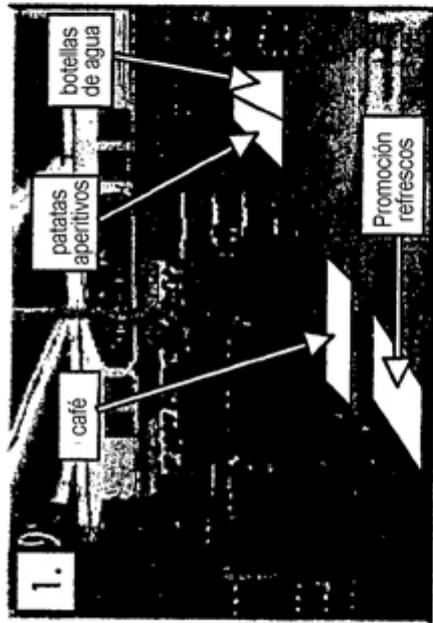


FIG. 12

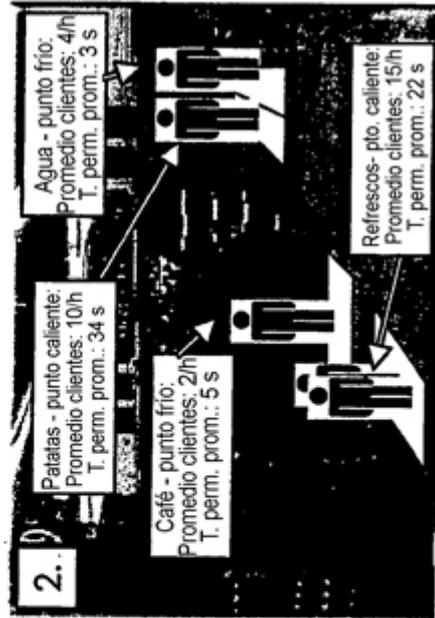


FIG. 13

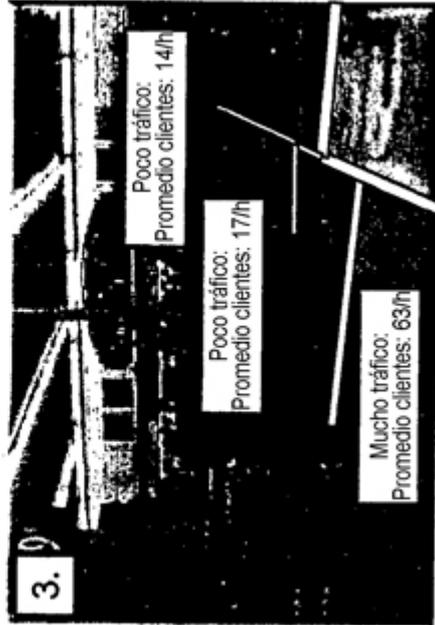


FIG. 14

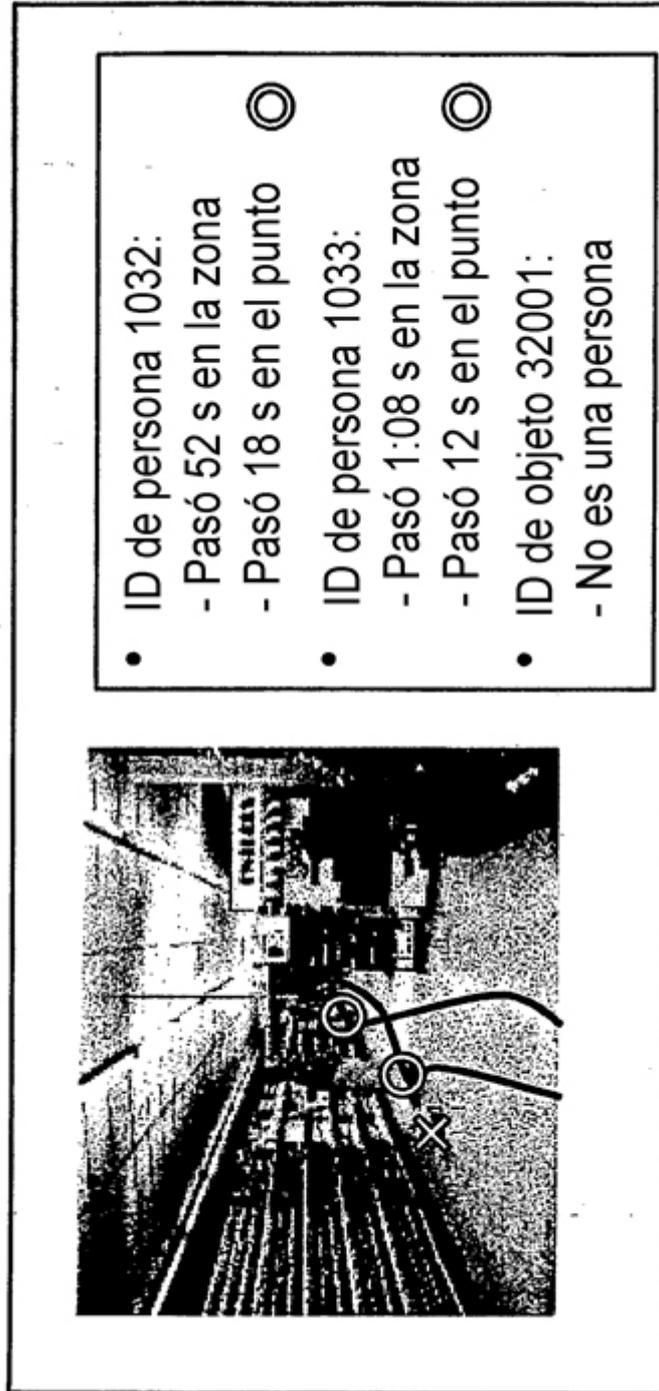


FIG. 15

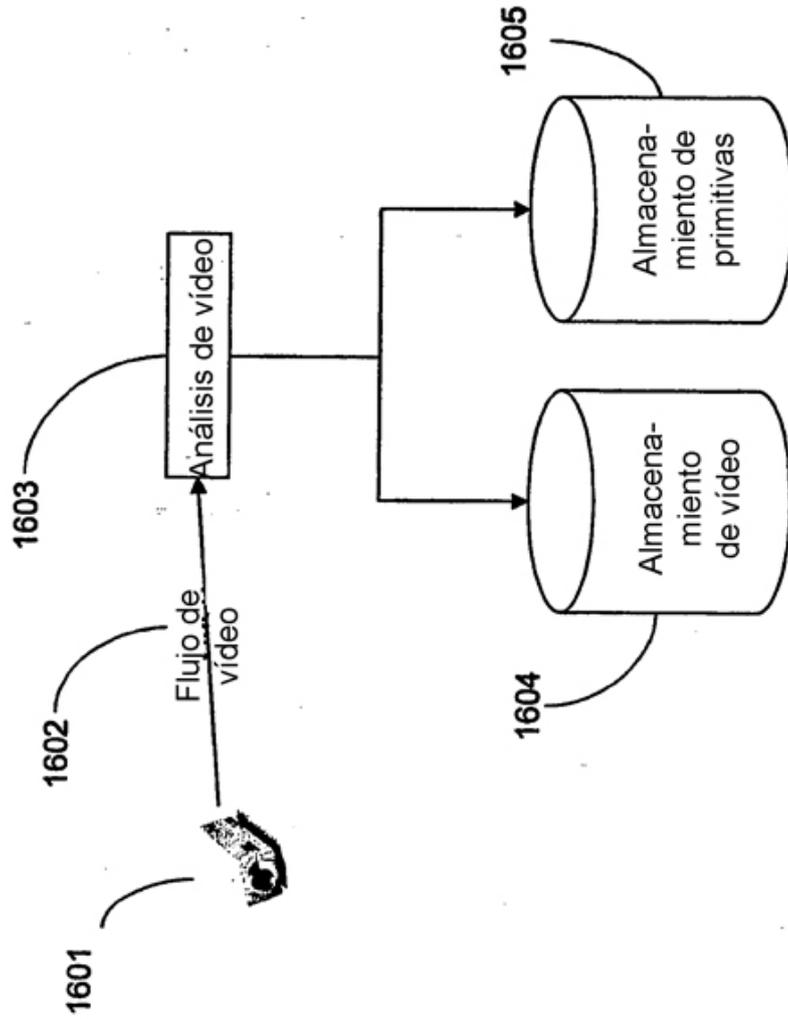


Figura 16a

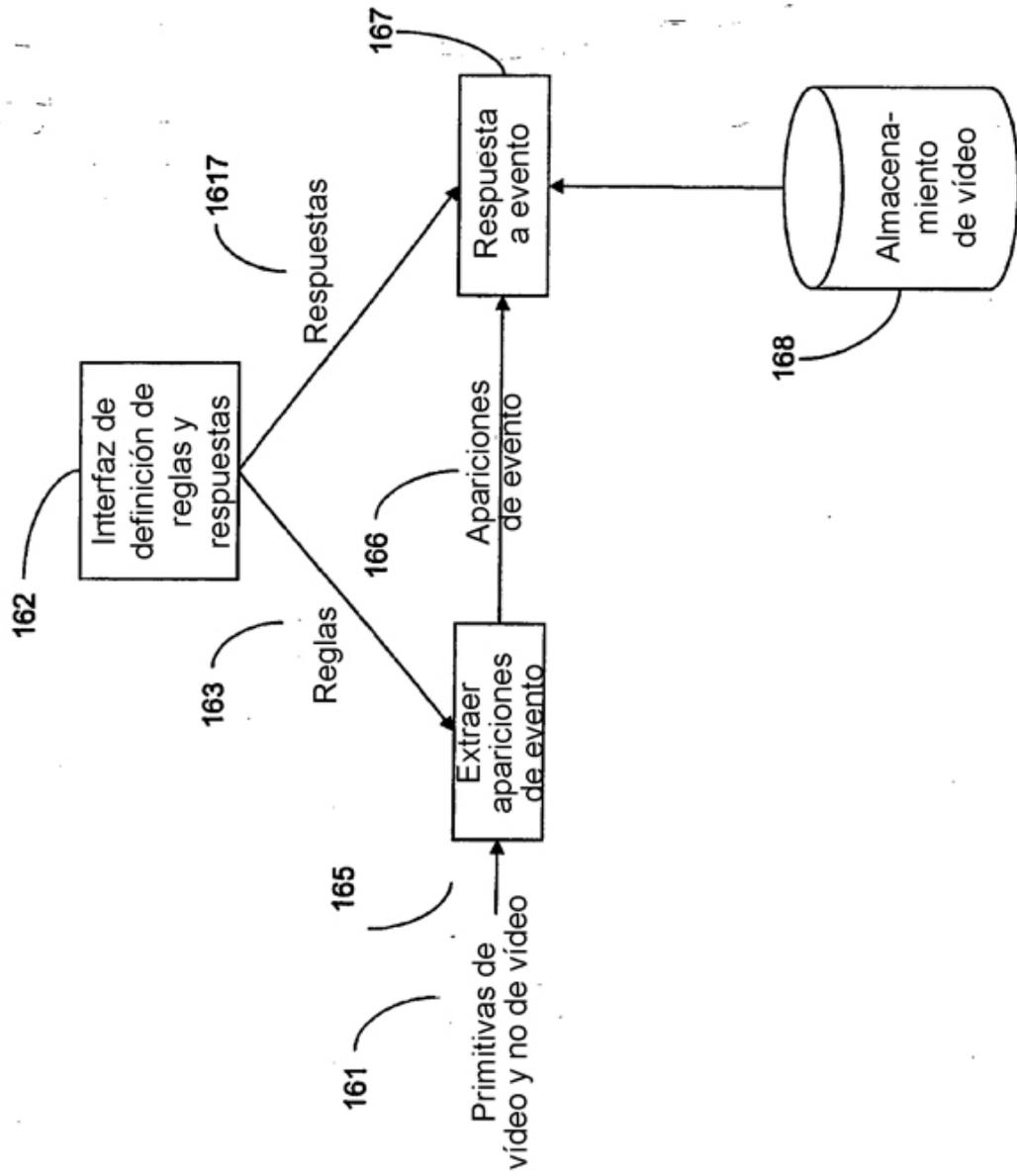


Figura 16b

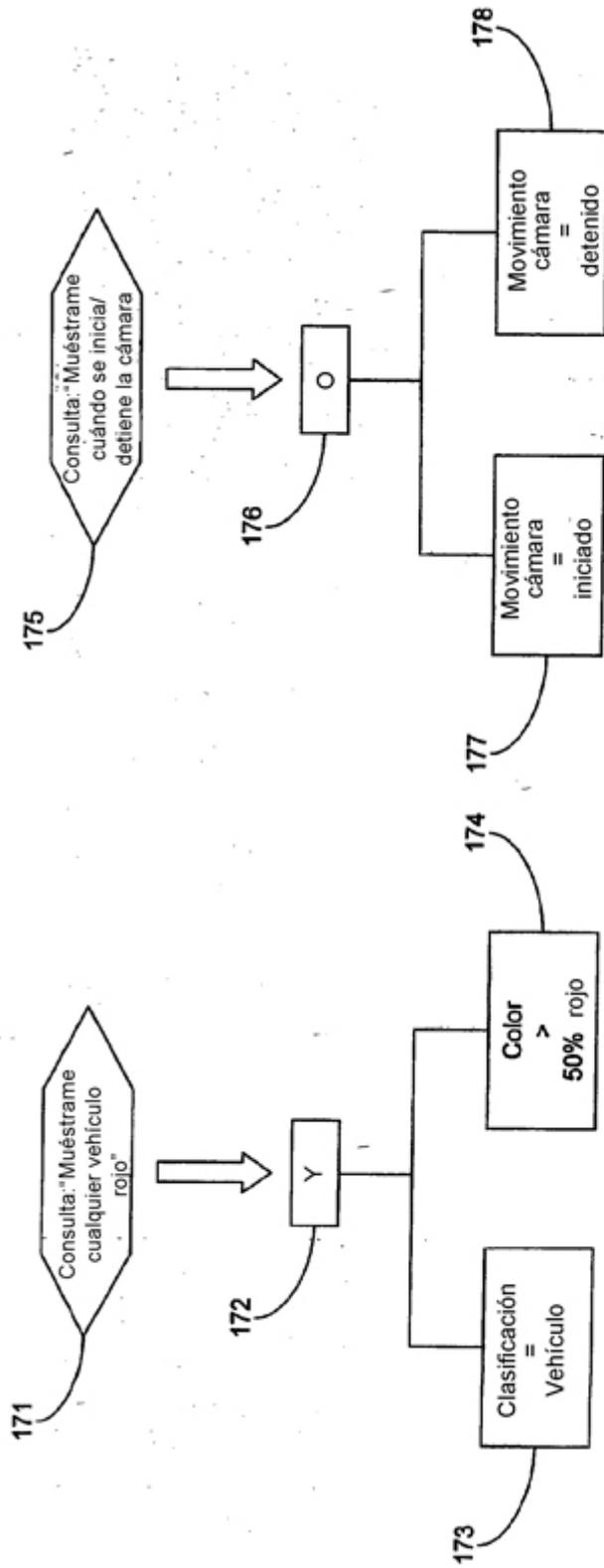


Figura 17a

Figura 17b



Figura 18a

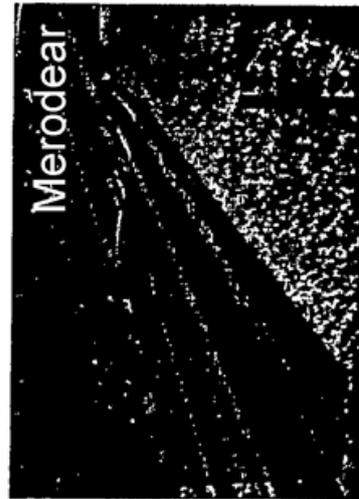


Figura 18b



Figura 18c

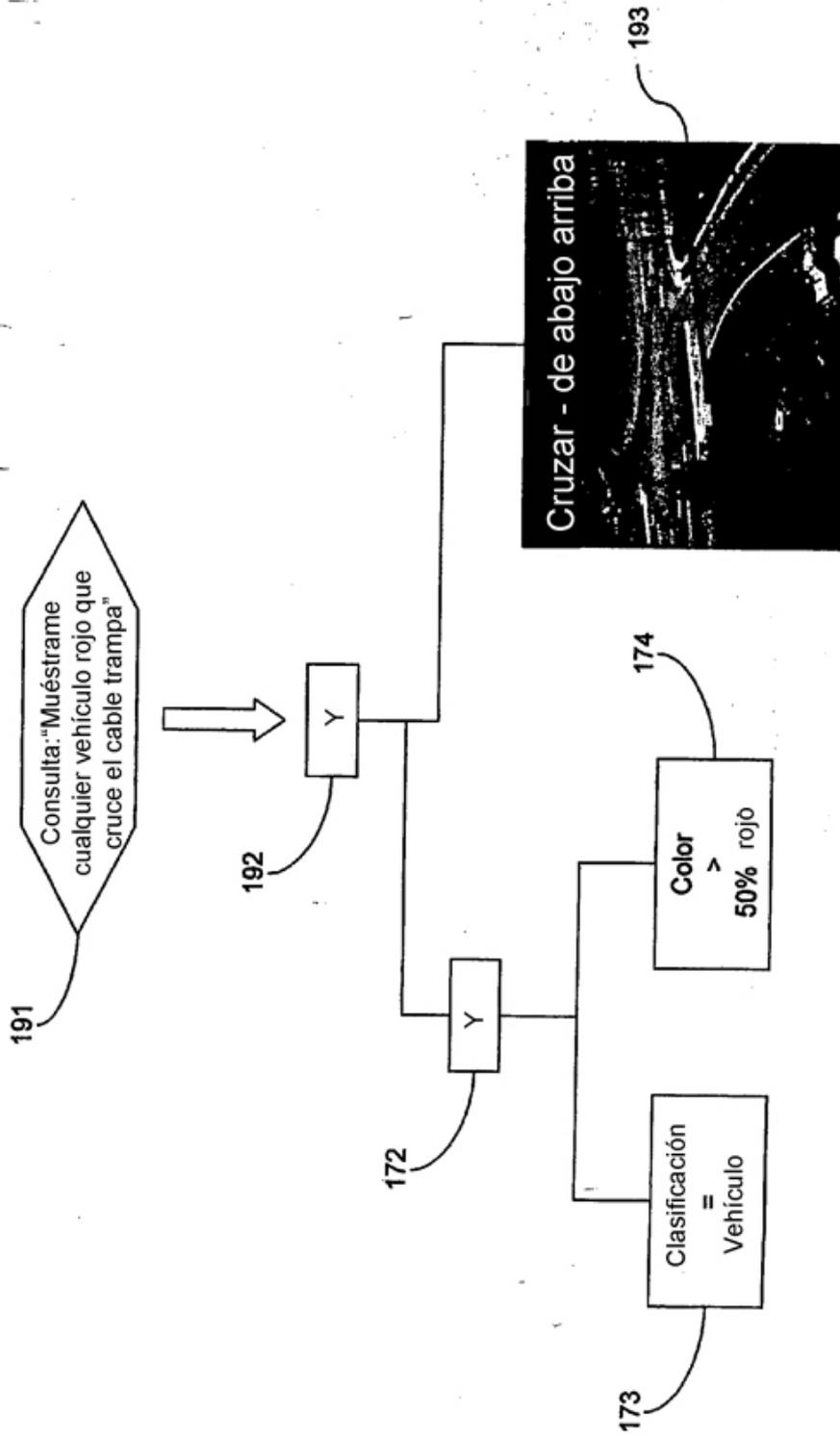


Figura 19

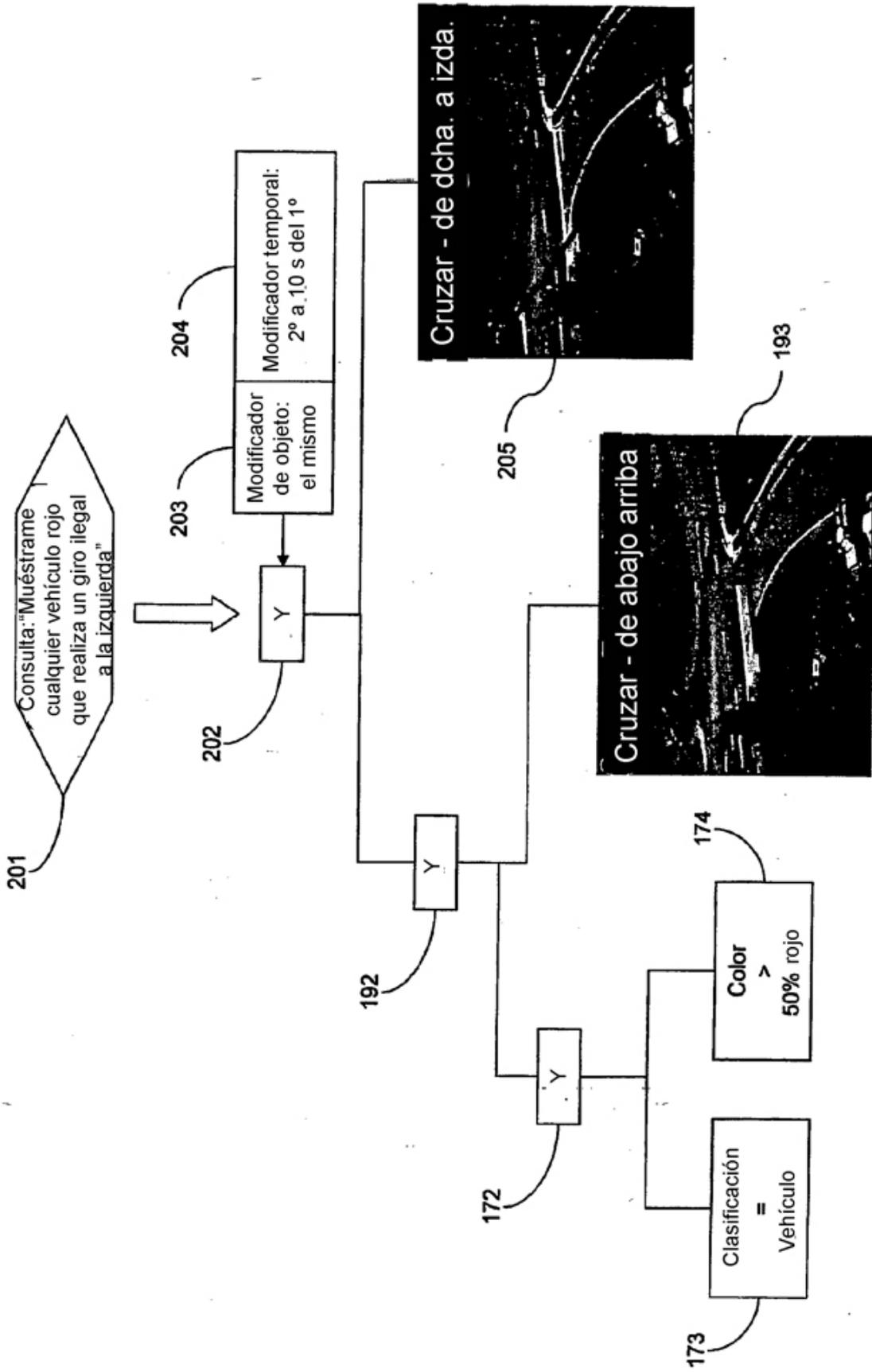


Figura 20

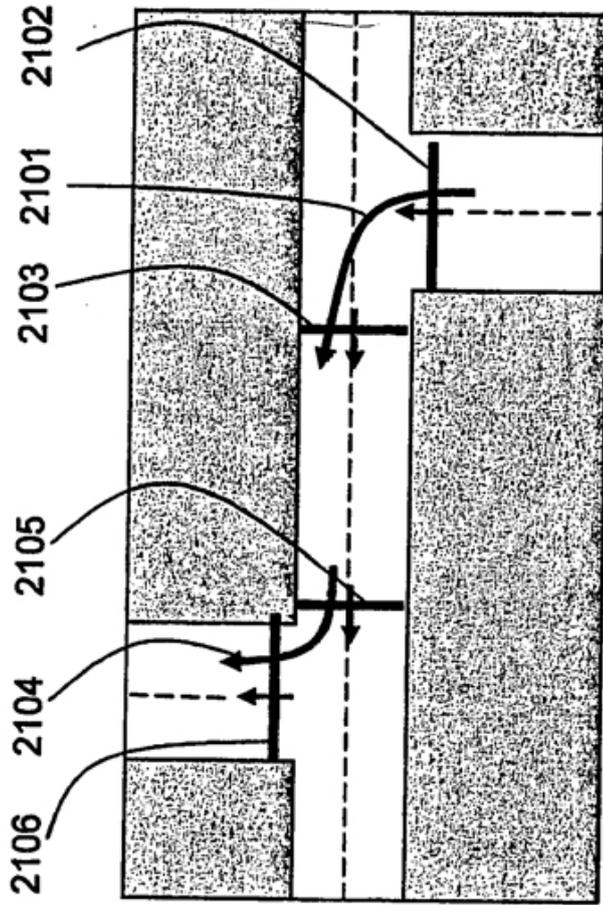


Figura 21a

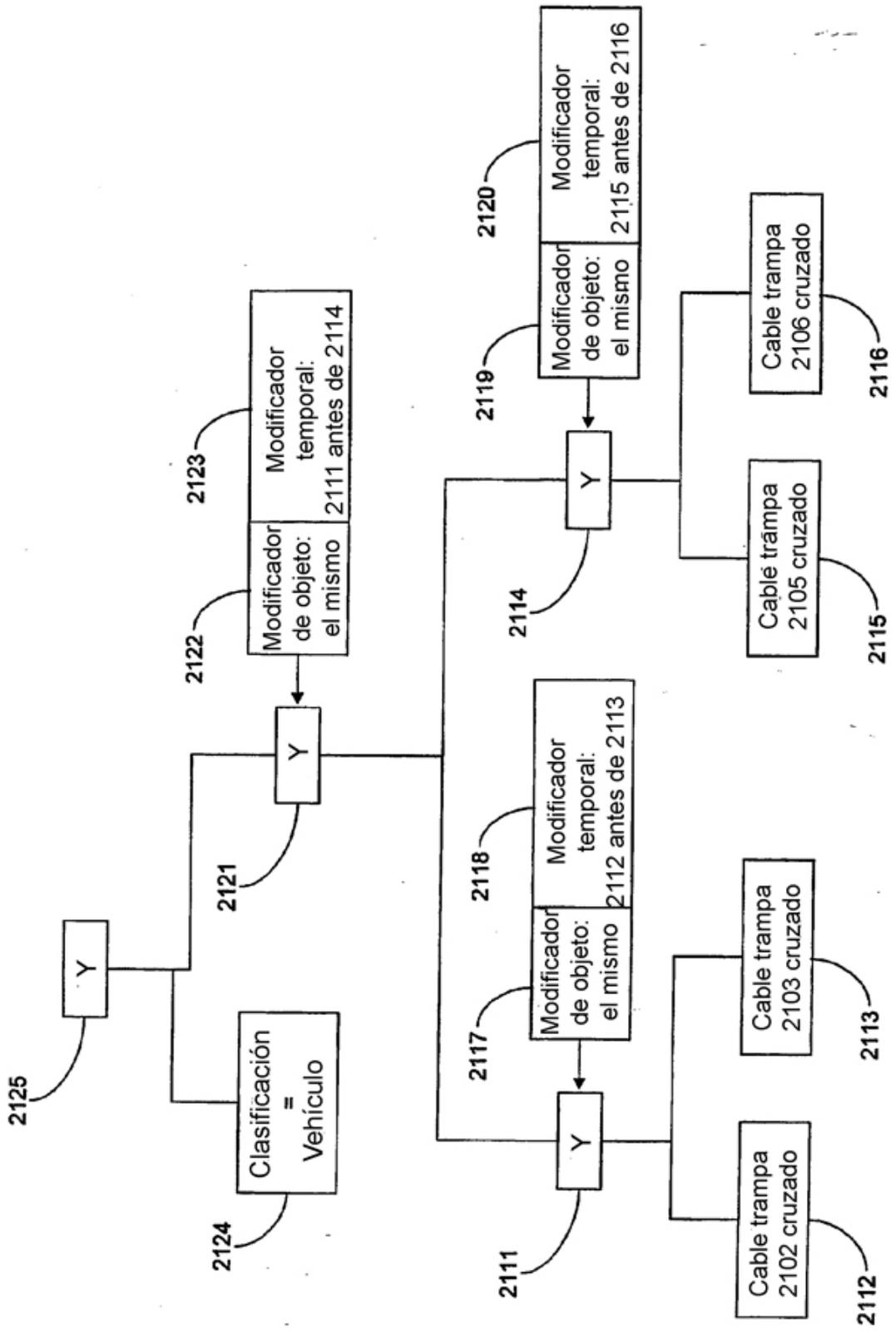


Figura 21b

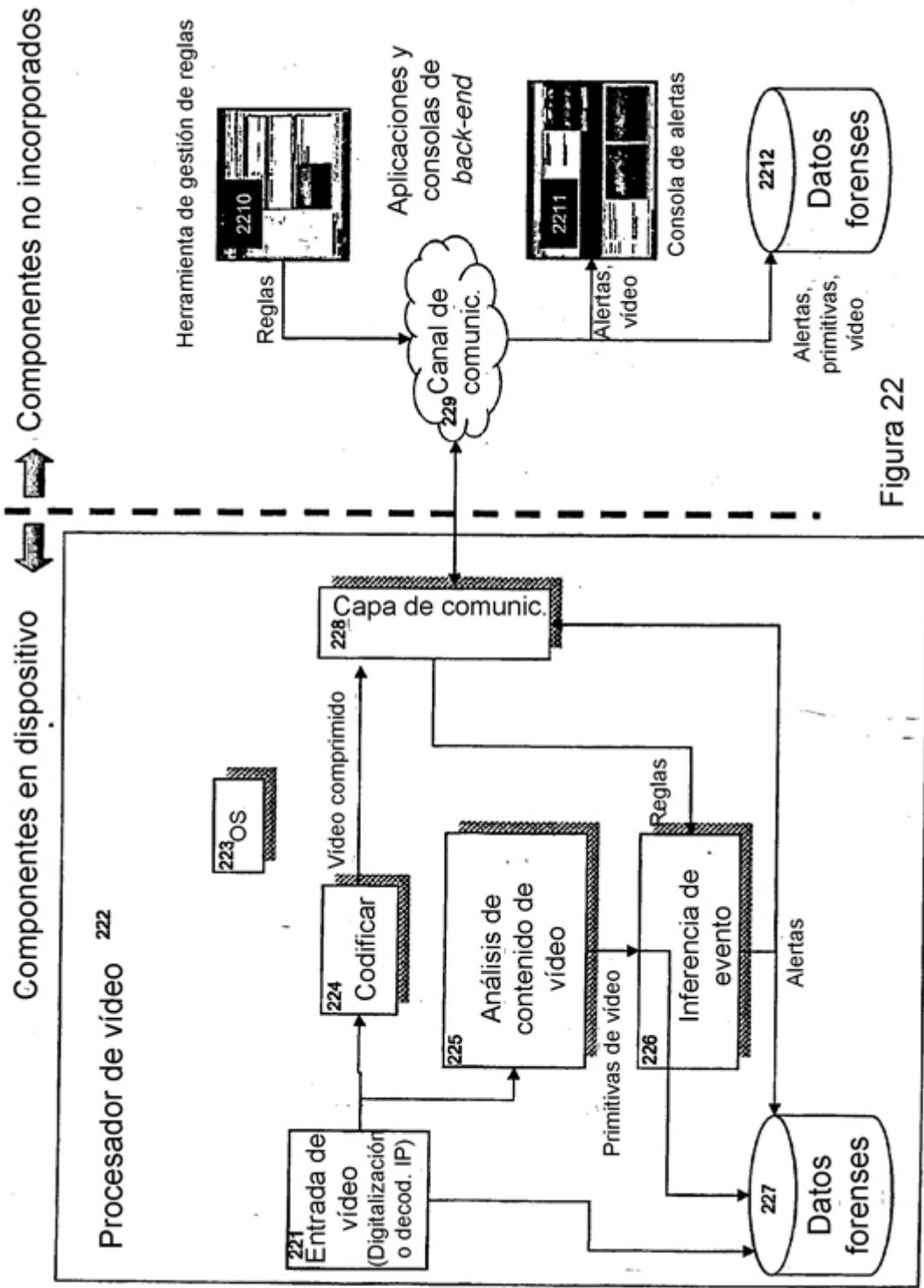


Figura 22

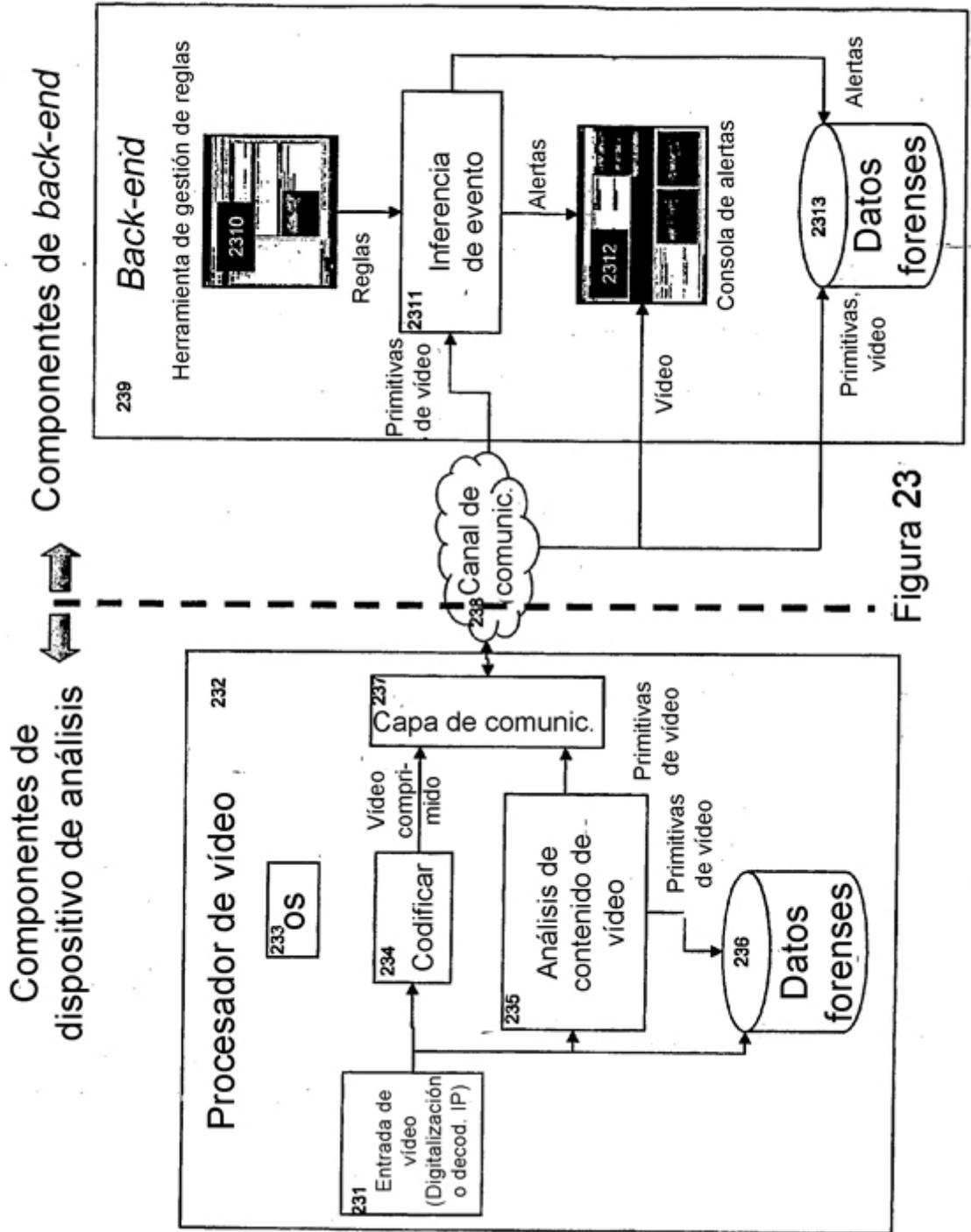


Figura 23

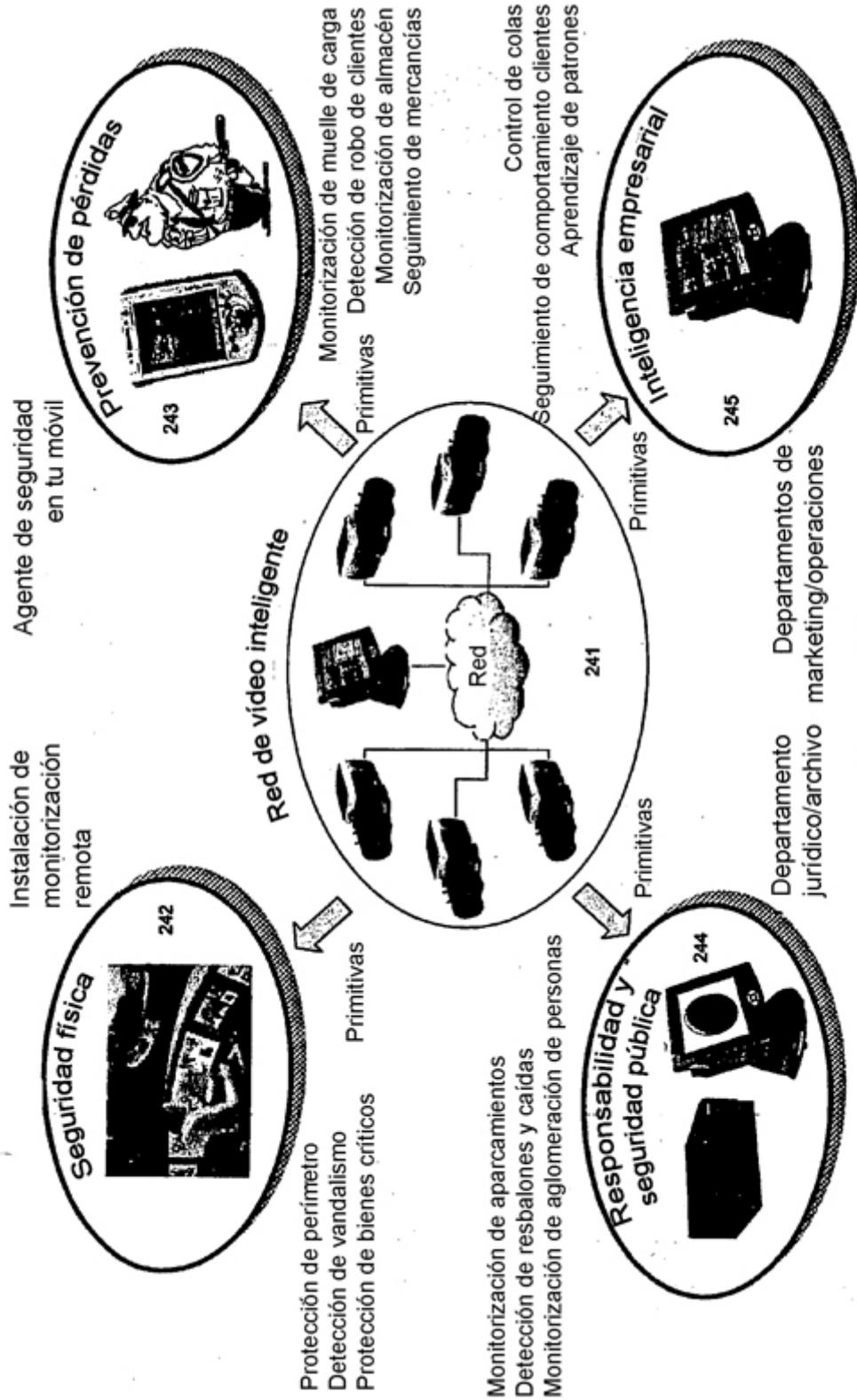


Figura 24

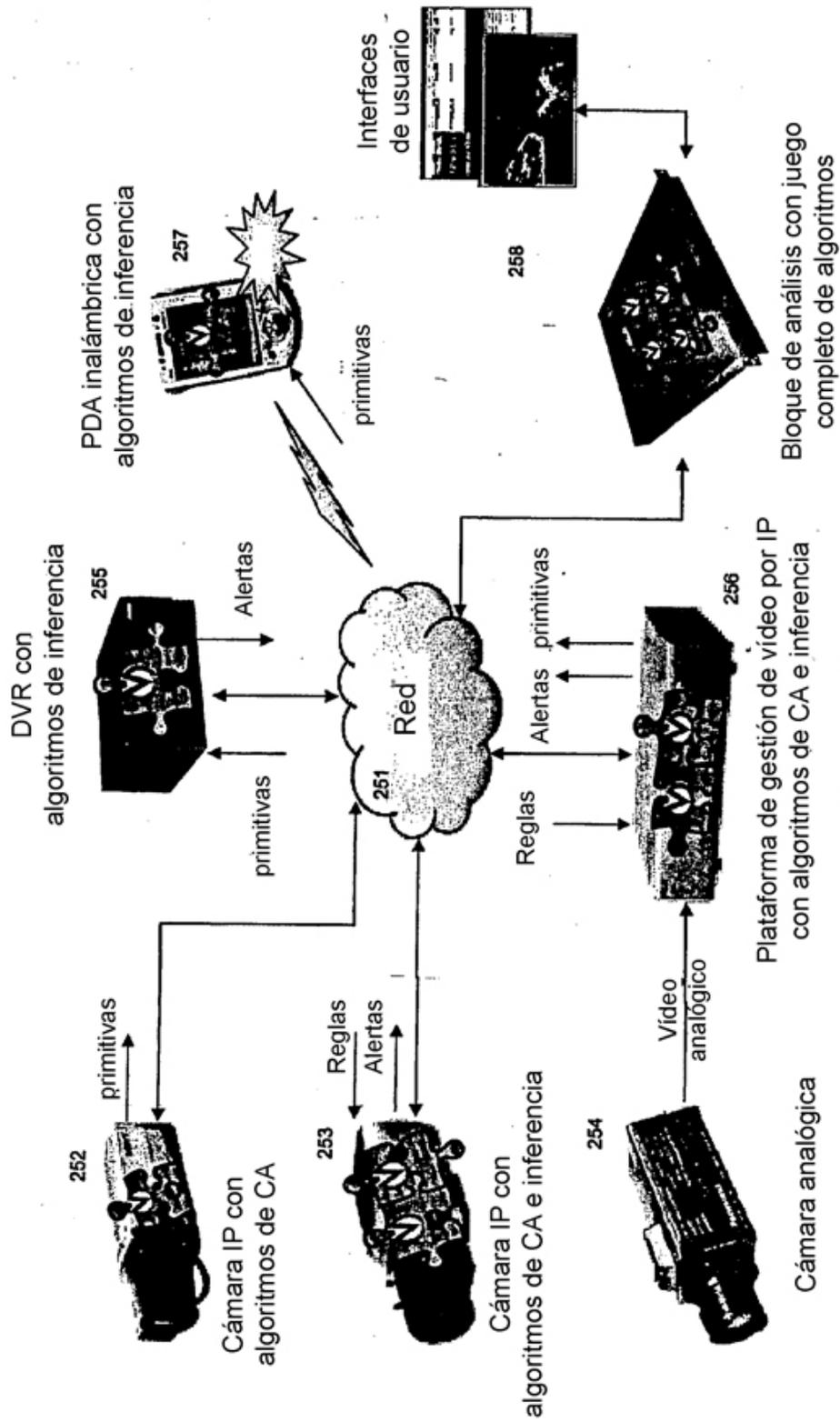


Figura 25