

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 383 174**

51 Int. Cl.:
G06K 9/00 (2006.01)
G06T 7/20 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **03815787 .1**
96 Fecha de presentación: **23.12.2003**
97 Número de publicación de la solicitud: **1588317**
97 Fecha de publicación de la solicitud: **26.10.2005**

54 Título: **Mantenimiento de fondo de escena de video utilizando detección y clasificación de cambios**

30 Prioridad:
30.01.2003 US 354096

45 Fecha de publicación de la mención BOPI:
18.06.2012

45 Fecha de la publicación del folleto de la patente:
18.06.2012

73 Titular/es:
OBJECTVIDEO, INC.
SUITE 290, 11600 SUNRISE VALLEY DRIVE
RESTON, VA 20191, US

72 Inventor/es:
VENETIANER, Peter, L.;
LIPTON, Alan, J.;
CHOSAK, Andrew, J.;
HAERING, Niels y
ZHANG, Zhong

74 Agente/Representante:
Espiell Volart, Eduardo María

ES 2 383 174 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

Mantenimiento de fondo de escena de vídeo utilizando detección y clasificación de cambios.

Campo de la invención

5 La presente invención se dirige al campo general del procesamiento de vídeo y al campo más específico del procesamiento de vídeo segmentado. En particular, la invención se refiere al mantenimiento de modelos de fondo en vídeo segmentado y a la clasificación de cambios del modelo de fondo.

Técnica relacionada

10 Muchas aplicaciones de procesamiento de vídeo requieren segmentación de objetos de vídeo (es decir, la diferenciación de objetos que realmente se mueven de la escena de fondo estática representada en una secuencia de vídeo). Tales aplicaciones incluyen, por ejemplo, construcción de mosaico de vídeo, compresión de vídeo basada en objetos, edición de vídeo basada en objetos y vigilancia automatizada de vídeo. Muchos algoritmos de segmentación de objetos de vídeo utilizan modelos de fondo de escena de vídeo (los cuales pueden denominarse simplemente como “modelos de fondo”) como una ayuda. La idea general es que cada fotograma de una secuencia de vídeo puede registrarse con el modelo de fondo y compararse, píxel por píxel, con el modelo de fondo. Los píxeles que muestran una diferencia suficiente se consideran píxeles de primer plano o móviles. Sin embargo, existen una amplia gama de fenómenos que pueden causar cambios a nivel de píxel, tales como: fondos inestables (por ejemplo, agua con ondas, hojas agitadas por el viento, etc.); fenómenos de iluminación (por ejemplo, nubes moviéndose por delante del sol, sombras, etc.); y fenómenos de cámara (por ejemplo, control automático de ganancia (AGC), iris automático, auto-enfoque, etc.).

20 Utilizando segmentación de objetos de vídeo (o una variación de la misma), habitualmente pueden detectarse objetos, o partes de objetos, los cuales muestran un movimiento independiente. La patente US 6.424.370 describe un método para proporcionar indexado automático de vídeo en base al contenido a partir del movimiento del objeto. Los objetos en movimiento en un vídeo de una cámara de vigilancia son detectados en la secuencia de vídeo utilizando métodos de segmentación del movimiento mediante un segmentador de movimiento. Los objetos son rastreados mediante datos segmentados en un rastreador de objetos y se genera una representación simbólica del vídeo en forma de un gráfico anotado que describe los objetos y su movimiento. Un analizador del movimiento analiza los resultados del rastreo de objetos y anota el movimiento del gráfico con índices que describen varios eventos. El gráfico se indexa a continuación utilizando un esquema de clasificación basado en normas para identificar eventos de interés, tales como aparición/desaparición. Hay dos problemas básicos que surgen cuando los objetos en una escena permanecen estacionarios durante un largo período de tiempo, y cualquiera de estos dos fenómenos puede degradar el rendimiento de la segmentación de objetos de vídeo para cualquier aplicación.

30 En primer lugar, si un objeto permanece estacionario durante un largo período de tiempo, el objeto podría ser detectado “permanentemente” como un objeto de primer plano. Sin embargo, a todos los efectos prácticos, el objeto se ha convertido en parte del fondo. En la figura 1A, este problema se ilustra para un coche 11 que entra en la secuencia de vídeo y aparca en ella. El coche es monitorizado de forma continua como un objeto de primer plano 12 pero realmente se ha convertido en parte del fondo (es decir, segmentación “permanente”).

40 En segundo lugar, si un objeto, inicialmente estacionario, es parte del modelo de fondo (por ejemplo, se “incrusta”) y a continuación se mueve, el objeto deja expuesta una región del modelo de fondo (por ejemplo, fondo estático) que no ha sido modelada. La región expuesta del modelo de fondo es detectada erróneamente como un objeto de primer plano. En la figura 1B, este problema se ilustra para un coche aparcado 13 que sale de la secuencia de vídeo. El coche 13 deja un “agujero” en forma de coche 14 segmentado en el modelo de fondo.

45 Tal como se ha descrito, por ejemplo, en la Solicitud de Patente de Estados Unidos N° de Serie 09/472.162, titulada “Method, Apparatus, and System for Compressing/Decompressing Digital Video Data”, presentada el 27 de diciembre 1999, y la Patente de Estados Unidos N° 6738424, titulada “Scene Model Generation from Video for Use in Video Processing”, presentada el 3 de julio de 2000 (ambas de asignación común), cuando se construyen mosaicos de fotografías, mosaicos de vídeo o modelos de escena de vídeo, a menudo es deseable extraer aquellas partes de las imágenes fuente que representan fondo “verdadero”. Por ejemplo, un coche aparcado en una secuencia de vídeo (o cualquier otra colección de imágenes) que permanece aparcado durante la duración de la secuencia de vídeo puede considerarse fondo verdadero. Sin embargo, un coche en una secuencia de vídeo que inicialmente está aparcado y después se aleja en algún punto en la secuencia de vídeo debe considerarse apropiadamente “no fondo”.

50 Si no se tiene cuidado para identificar regiones de fondo verdadero, se producirán aberraciones. Si el objetivo es producir un mosaico o imagen de fondo, los objetos de primer plano pueden “incrustarse” y el modelo de fondo daría como resultado imágenes de aspecto innatural. Si el objetivo es construir un modelo de escena como base para segmentación de vídeo, los resultados pueden ser segmentaciones malas, donde partes de objetos de primer plano no

se detectan, y donde algunas regiones de fondo expuestas se detectan como objetos de primer plano.

La figura 2 ilustra un ejemplo de la técnica anterior de permitir que objetos de primer plano corrompan un modelo de fondo. La secuencia de vídeo representa a un golfista preparado para dar el golpe inicial. Un subconjunto 21 de las imágenes fuente de la secuencia de vídeo representa una parte de esta secuencia de vídeo. Las imágenes fuente se utilizan para generar un modelo de fondo 22 y objetos de primer plano 23. Sin embargo, el modelo de fondo 22 contiene objetos de primer plano 23 (por ejemplo, el golfista a la izquierda y parte de la camiseta del golfista a la derecha) incrustados en el modelo de fondo 22, y los objetos de primer plano 23 están segmentados de forma incompleta (por ejemplo, parte del torso del golfista y parte del palo de golf).

Resumen de la invención

La invención emplea la detección y clasificación de cambios para mantener un modelo de fondo de una secuencia de vídeo. Además, la invención mantiene un modelo de fondo de una secuencia de vídeo y clasifica cambios en el modelo de fondo.

La invención incluye un método según la reivindicación 1.

La invención incluye un sistema informático para llevar a cabo el método anterior. Un sistema para la invención incluye un sistema informático que incluye un medio legible por ordenador que presenta software para hacer funcionar un ordenador de acuerdo con la invención.

Según un aspecto adicional de la invención, se proporciona un sistema informático para procesar vídeo según la reivindicación 12.

Características y ventajas adicionales de la invención, así como la estructura y el funcionamiento de diversas realizaciones de la invención, se describen con detalle a continuación con referencia a los dibujos adjuntos.

Definiciones

Un "ordenador" se refiere a cualquier aparato que es capaz de aceptar una entrada estructurada, procesar la entrada estructurada según normas prescritas, y producir resultados del procesamiento como una salida. Los ejemplos de un ordenador incluyen: un ordenador; un ordenador de utilización general; un superordenador; una unidad principal; un super mini-ordenador; un mini-ordenador; una estación de trabajo; un micro-ordenador; un servidor; una televisión interactiva; un dispositivo de Web; un dispositivo de telecomunicaciones con acceso a Internet; una combinación híbrida de un ordenador y una televisión interactiva; y hardware específico de aplicación para emular a un ordenador y/o software. Un ordenador puede ser fijo o portátil. Un ordenador puede presentar un único procesador o múltiples procesadores, los cuales pueden funcionar en paralelo y/o no en paralelo. Un ordenador también se refiere a dos o más ordenadores conectados juntos mediante una red para transmitir o recibir información entre los ordenadores. Un ejemplo de tal ordenador incluye un sistema informático distribuido para procesar información mediante ordenadores enlazados mediante una red.

Un "medio legible por ordenador" se refiere a cualquier dispositivo de almacenamiento utilizado para almacenar datos, al que puede acceder un ordenador. Los 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 utilizada para portar datos electrónicos legibles por ordenador, tales como los utilizados para transmitir y recibir un e-mail o para acceder a una red.

"Software" se refiere a las normas prescritas para hacer funcionar un ordenador. Los ejemplos de software incluyen: software; segmentos de código; instrucciones; programas informáticos; y lógica programada.

Un "sistema informático" se refiere a un sistema que presenta un ordenador, en el cual el ordenador comprende un medio legible por ordenador que incorpora software para hacer funcionar al ordenador

Una "red" se refiere a una serie de 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 establecidas a través del teléfono, inalámbricas u otros enlaces de comunicación. Los ejemplos de una red incluyen: una internet, tal como Internet; una red interna; una red de área local (LAN); una red de área amplia (WAN); y una combinación de redes, tales como una internet y una red interna.

"Vídeo" se refiere a imágenes en movimiento representadas en forma analógica y/o digital. Los ejemplos de vídeo incluyen televisión, películas, secuencias de imágenes de una cámara u otro observador, y secuencias de imágenes generadas por ordenador. Éstas pueden obtenerse, por ejemplo, de una señal en directo, un dispositivo de almacenamiento, una interfaz basada en IEEE 1394, un digitalizador de vídeo, un motor gráfico de un ordenador o una

conexión de red.

“Procesamiento de vídeo” se refiere a cualquier manipulación de vídeo, incluyendo, por ejemplo, compresión y edición.

Un “fotograma” se refiere a una imagen particular u otra unidad discreta en un vídeo.

5 **Breve descripción de los dibujos**

Las anteriores y otras características y ventajas de la invención serán evidentes a partir de la siguiente y más particular descripción de una realización preferida de la invención, tal como se ilustra en los dibujos adjuntos. Los dígitos más a la izquierda en el número de referencia correspondiente indican el dibujo en el el cual un elemento aparece en primer lugar.

- 10 Las figuras 1A y 1B ilustran problemas de la técnica anterior con la utilización de segmentación de objetos de vídeo para detectar objetos, o partes de objetos, los cuales muestran movimiento independiente;
- La figura 2 ilustra un ejemplo de la técnica anterior de permitir que objetos de primer plano corrompan un modelo de fondo;
- La figura 3 ilustra un diagrama de flujo para una primera
- 15 realizaciónde la invención;
- La figura 4 ilustra modelización de fondo estadística de píxeles para detectar píxeles de primer plano;
- La figura 5 ilustra modelización de fondo estadística de píxeles para manejar cambios de iluminación;
- La figura 6 ilustra la utilización de diferenciación de tres fotogramas para la
- 20 detección del movimiento;
- La figura 7 ilustra la detección de píxeles en movimiento y píxeles cambiados;
- La figura 8 ilustra un diagrama de flujo para la detección y clasificación de una diana estacionaria;
- 25 La figura 9 ilustra detección de cambios de fondo;
- La figura 10 ilustra la inserción de un objeto de primer plano;
- La figura 11 ilustra la eliminación de una parte del fondo;
- La figura 12 ilustra un diagrama de flujo para detectar bordes intensos;
- La figura 13 ilustra otro diagrama de flujo para detectar bordes intensos;
- 30 La figura 14 ilustra un diagrama de flujo para determinar la intensidad de un borde;
- La figura 15 ilustra la determinación de la intensidad de un borde;
- La figura 16 ilustra un diagrama de flujo para una segunda realización de la invención.

35 **Descripción detallada de las realizaciones ejemplares de la invención**

Un ejemplo de realización de la invención se describe en detalle a continuación. Aunque se describen realizaciones ejemplares específicas, debe entenderse que esto se realiza sólo con fines de ilustración. Un experto en la

materia reconocerá que pueden utilizarse otros componentes y configuraciones sin alejarse del espíritu y alcance de la invención. Las realizaciones y ejemplos descritos en esta patente son ejemplos no limitativos.

La invención emplea detección y clasificación de cambios para mantener un modelo de fondo de una secuencia de vídeo. La invención puede utilizarse para aplicaciones de procesamiento de vídeo en tiempo real (por ejemplo, compresión basada en objetos en tiempo real, o vigilancia de vídeo), en las cuales la secuencia de vídeo no puede estar disponible en su totalidad en cualquier momento, y podrían requerirse cambios progresivos del modelo de fondo para mantener su utilidad. La invención también puede utilizarse para aplicaciones de procesamiento de vídeo no en tiempo real. Una secuencia de vídeo se refiere a parte o todo de un vídeo.

Con la invención, en primer lugar, cambios locales en el modelo de fondo son detectados y pueden utilizarse para mantener el modelo de fondo y, en segundo lugar; tales cambios detectados se clasifican y pueden procesarse adicionalmente. Los cambios detectados se clasifican en dos categorías principales: primera, un objeto que está situado en la escena y permanece estático durante un periodo de tiempo (es decir, una inserción); y segunda, un objeto que se mueve saliendo de la escena y deja expuesta una sección del modelo de fondo (por ejemplo, el fondo estático) (es decir, una eliminación). El aspecto común de estas dos categorías es que hay un cambio local permanente en el modelo de fondo.

La clasificación de los cambios en estas dos categorías puede ser muy importante en una amplia gama de aplicaciones, tales como, por ejemplo, aplicaciones de vigilancia de vídeo. Los ejemplos de la primera categoría (es decir, una inserción) para aplicaciones de vigilancia de vídeo incluyen: monitorizar zonas de estacionamiento prohibido (y, por ejemplo, activar una alarma si un coche pasa más de cierta cantidad de tiempo en las zonas de estacionamiento prohibido); detectar bolsas abandonadas en aeropuertos; y detectar objetos abandonados cerca de zonas sensibles, tales como instalaciones militares y centrales energéticas. Los ejemplos de la segunda categoría (es decir, una eliminación) para aplicaciones de vigilancia de vídeo incluyen: detectar la eliminación de un artículo de gran valor, tal como una aberración de un museo, una cara pieza de hardware o un coche de un parking.

La figura 3 ilustra un diagrama de flujo para una primera realización de la invención en un posible contexto de un sistema de procesamiento de vídeo general. Una secuencia de vídeo es introducida en el sistema, y se genera y se mantiene un modelo de fondo 31, 32 y 33. El vídeo de entrada es procesado mediante dos técnicas diferentes de clasificación de píxeles de nivel bajo: clasificación de píxeles basada en un modelo de fondo 31 y clasificación de píxeles basada en el movimiento 34. Estas dos técnicas producen máscaras de píxeles (por fotograma) que representan píxeles de interés. La clasificación de píxeles basada en un modelo de fondo 31 produce una máscara de cambio e imágenes, y la clasificación de píxeles basada en el movimiento 34 produce una máscara de movimiento. La máscara de cambio y la máscara de movimiento se proporcionan para la generación de blobs (objetos binarios grandes) 35, que convierte las máscaras en un conjunto de uno o más blobs individuales que representan el aspecto de cada objeto de primer plano visible en cada fotograma. En general, si no hay objetos de primer plano visibles, no se generan blobs. Los blobs se rastrean utilizando rastreo de blobs 36, el cual conecta los blobs de un fotograma con aquellos de otros fotogramas para generar una "diana" que representa cada objeto en la escena. Una diana es una descripción espacio-temporal de un objeto de vídeo en el tiempo. Las dianas se analizan mediante detección y clasificación de una diana estacionaria 37, la cual determina si cualquiera de las dianas representa un cambio "permanente" del modelo de fondo 33 y si ese cambio representa una "inserción" (por ejemplo, un objeto que entra en la escena) o una "eliminación" (por ejemplo, un objeto que sale y deja expuesta una sección del modelo de fondo). Finalmente, cualesquiera dianas estacionarias detectadas se insertan en el modelo de fondo 33 mediante la actualización local del modelo de fondo 38.

La generación y el mantenimiento de un modelo de fondo incluyen la clasificación de píxeles basada en un modelo de fondo 31, la actualización del modelo de fondo 32, y el modelo de fondo 33. Una opción para la estrategia basada en el modelo de fondo 31, 32 y 33 emplea modelización de píxeles estadística dinámica. La modelización de píxeles estadística dinámica mantiene una precisa representación del fondo de la imagen y diferencia píxeles de fondo de píxeles de primer plano. En un ejemplo de realización, la modelización de píxeles estadística dinámica se implementa con las técnicas descritas en la Patente de Estados Unidos de asignación común N° 6625310, titulada "Video Segmentation Using Statistical Pixel Modeling", presentada el 23 de marzo de 2001. La idea general de la técnica ejemplar es que se mantiene un historial de todos los píxeles durante varios fotogramas, incluyendo valores cromáticos (o de intensidad) de píxeles y sus estadísticas. Un píxel estable, sin cambios es tratado como fondo. Si las estadísticas de un píxel cambian de forma significativa, puede considerarse que el píxel es de primer plano. Si el píxel vuelve a su estado original, el píxel puede volver a ser considerado un píxel de fondo. Esta técnica sirve para aliviar el ruido del sensor y para dirigir automáticamente cambios lentos en el fondo debidos a condiciones de iluminación y al control automático de ganancia de la cámara (AGC). En lugar de modelización de píxeles estadística dinámica, la clasificación de píxeles basada en un modelo de fondo 31 puede implementarse utilizando modelos de fondo estadísticos, una mezcla de modelos de fondo gaussianos o mezcla adaptativa de forma dinámica de modelos gaussianos.

El modelo de fondo 33 es la representación interna de la escena estática representada en el vídeo en cualquier momento dado. Cada vez que se analiza un nuevo fotograma, el modelo de fondo 33 puede actualizarse de forma

progresiva mediante la actualización del modelo de fondo 32. Además de las actualizaciones progresivas, es necesario actualizar el modelo de fondo 33 cuando se detecta un cambio de fondo. Por ejemplo, la información cromática que representa la nueva región de fondo estático local debe “incrustarse” en el modelo de fondo 33, lo que puede conseguirse con la actualización local del modelo de fondo 38.

5 Las figuras 4 y 5 ilustran la utilización de la modelización de píxeles para generar y mantener un modelo de fondo. En la figura 4, se ilustra la modelización de fondo estadística de píxeles para detectar píxeles de primer plano. El fotograma 41 es un fotograma actual de un vídeo de un hombre que camina delante de sillas apiladas y que deja caer una maleta. En el fotograma 41, el hombre ha dejado caer la maleta y sigue adelante. Tal como se ilustra con el gráfico 42 que representa la intensidad y el tiempo para un píxel en el vídeo, La media y desviación estándar de la intensidad para cada píxel 43 se utilizan para modelizar el fondo 44. El modelo de fondo 33 contiene una media y desviación estándar para cada píxel. El algoritmo de clasificación de píxeles 31 compara cada píxel del fotograma actual 41 con el píxel correspondiente del modelo de fondo 33. Cuando un objeto se mueve “a través de” un píxel en el fotograma actual 41, su valor no se adaptará a las estadísticas capturadas en el modelo de fondo 33 y se considera primer plano 45. Una máscara de cambio de píxeles de primer plano se crea mediante la clasificación basada en el modelo de fondo 31 y se envía a la generación de blobs 35. Esta máscara de cambio y el fotograma actual 41 son enviados ambos a la actualización del modelo de fondo 32, ya que las estadísticas de píxeles que comprenden el modelo de fondo 33 pueden actualizarse.

20 En la figura 5, se ilustra la modelización de fondo estadística de píxeles para manejar cambios de iluminación. El fotograma 51 ilustra un cambio de iluminación lento en un vídeo. Tal como se ilustra con el gráfico 52 que representa la intensidad y el tiempo para un píxel en el vídeo, la media y la desviación estándar de intensidad para cada píxel 53 se utilizan para modelizar el fondo. Dado que la media y la desviación estándar para cada píxel se calcula a partir de sólo los últimos fotogramas, el modelo de fondo 33 está adaptado para seguir la intensidad de píxeles que cambia lentamente 54.

25 La clasificación de píxeles basada en el movimiento 34 determina si un píxel está experimentando realmente un movimiento independiente de fotograma a fotograma. Una realización potencial para la clasificación de píxeles basada en el movimiento 34 es la diferenciación de tres fotogramas, tal como se describe en la Patente de Estados Unidos de asignación común N° 6954498, presentada el 24 de octubre de 2000. Otras realizaciones potenciales para la clasificación de píxeles en movimiento 34 incluyen diferenciación de dos fotogramas y flujo óptico.

30 La figura 6 ilustra la utilización de diferenciación de tres fotogramas para la detección de movimiento en la clasificación de píxeles basada en el movimiento 34. Los fotogramas 61, 62 y 63 son fotogramas del pasado, actual y del futuro, respectivamente, de un vídeo de un hombre que camina delante de sillas apiladas y que deja caer una maleta. La máscara de diferencia 64 se obtiene comparando los fotogramas 61 y 62, y la máscara de diferencia 65 se obtiene comparando los fotogramas 62 y 63. La máscara de movimiento 66 se obtiene comparando las máscaras de diferencia 64 y 65 utilizando una conjunción lógica AND. La máscara de movimiento 66 es enviada a la generación de blobs 35.

40 Las salidas de la clasificación de píxeles basada en un modelo de fondo 31 y la clasificación de píxeles basada en el movimiento 34 pueden no detectar concurrentemente un nuevo objeto de primer plano. Por ejemplo, un coche recién aparcado puede aparecer como un objeto de primer plano según la clasificación de píxeles basada en un modelo de fondo 31. Sin embargo, dado que el coche aparcado no muestra ningún movimiento independiente real, la clasificación de píxeles basada en el movimiento 34 no puede detectar ningún objeto de primer plano.

45 Otro ejemplo de esta diferencia entre píxeles cambiados y píxeles en movimiento se ilustra en la figura 7. El fotograma 71 es un fotograma de un vídeo de un hombre que camina por delante de sillas apiladas y que deja caer una maleta. La máscara de movimiento 72 resulta de la clasificación de píxeles basada en el movimiento 34, la cual detecta al hombre pero no a la maleta. La máscara de cambio 73 resulta de la clasificación basada en el modelo de fondo 31, la cual detecta tanto al hombre como a la maleta. En este ejemplo, un objeto de primer plano recientemente insertado (es decir, la maleta) es detectado por la clasificación de píxeles basada en un modelo de fondo 31 pero no por la clasificación de píxeles basada en el movimiento 34.

50 La generación de blobs 35 y el rastreo de blobs 36 integran la máscara de movimiento y la máscara de cambio de píxeles por fotograma en dianas (descripciones espacio-temporales de objetos de vídeo). Para la generación de blobs 35, existen muchas técnicas convencionales para aglomerar píxeles en blobs, por ejemplo: componentes conectados, tal como se describe en D. Ballard y C. Brown, “Computer Vision”, Prentice-Hall, Mayo de 1982; y componente cuasi-conectados, como se describe en T.E. Boult, R.J. Micheals, X. Gao, P. Lewis, C. Power, W. Yin y A. Erkan, “Frame-Rate Omnidirectional Surveillance and Tracking of Camouflaged and Occluded Targets”, Proc. of the IEEE Workshop on Visual Surveillance, junio de 1999. Para el rastreo de blobs 36, existen muchas técnicas convencionales para rastrear blobs a lo largo del tiempo para formar dianas. Ejemplos de técnicas de rastreo se describen en la siguiente Patente de Estados Unidos de asignación común N° 6954498, titulada “Interactive Video

Manipulation”, presentada el 24 de octubre de 2000; los documentos Wren, C.R. et al., “Pfinder: Real-Time Tracking of the Human Body”, IEEE Trans. on Pattern Matching and Machine Intelligence, Vol. 19, págs. 780-784, 1997; Grimson, W.E.L. et al., “Using Adaptive Tracking to Classify and Monitor Activities in a Site”, CVPR, págs. 22-29, junio de 1998; y Olson, T.J. y Brill, F.Z., “Moving Object Detection and Event Recognition Algorithm for Smart Cameras, IUW, págs. 159-175, mayo de 1997.

La detección y clasificación de una diana estacionaria 37 analiza dianas generadas mediante el rastreo de blobs 36 para determinar si cada diana es estacionaria. Puede determinarse que una diana es estacionaria si la diana representa un cambio local en el modelo de fondo 33. Una diana puede representar un cambio en el modelo de fondo 33 si, por ejemplo, un objeto de vídeo ha dejado de moverse (es decir, una inserción) o un objeto de vídeo previamente estacionario ha dejado expuesta una sección de fondo estático que aparece como una diana (es decir, una eliminación).

Una vez que se ha detectado una diana estacionaria, esta información puede devolverse a la actualización local del modelo de fondo 38 para actualizar el modelo de fondo 33. Con esta retroalimentación, el modelo de fondo 33 puede mantenerse actualizado respecto a lo que constituye fondo estático y actividad de primer plano legítima.

La detección y clasificación de una diana estacionaria 37 determina si una diana es estacionaria y, si es así, si debe etiquetarse como una inserción, una eliminación o desconocida, si no es posible determinar la diferencia. A la hora de distinguir entre una inserción y una eliminación, la relación entre las escalas temporales para una inserción y una eliminación es importante. Una inserción puede implicar una escala temporal diferente de la de una eliminación, y estas escalas temporales pueden ser dependientes de la aplicación. Por ejemplo, una aplicación puede requerir que un objeto permanezca en su sitio durante una gran cantidad de tiempo antes de que se considere una inserción pero solamente una pequeña cantidad de tiempo antes de que se considere una eliminación. Como ejemplo específico, un coche aparcado en un bordillo en un aeropuerto durante cinco minutos puede no ser un asunto de preocupación y puede no considerarse una inserción, pero un coche aparcado en el bordillo durante quince minutos puede ser un asunto de preocupación y considerarse una inserción. Además, el mismo coche, tan pronto como se aleje del bordillo puede considerarse una eliminación. En este ejemplo, la escala temporal para una inserción es mayor que la escala temporal para una eliminación. Para otra aplicación, las escalas temporales relativas para una inserción y una eliminación pueden invertirse a partir del ejemplo anterior, de modo que la escala temporal para una eliminación es mayor que la escala temporal para una inserción. Además, las escalas temporales para una inserción y una eliminación pueden ser configuradas por un usuario.

La figura 8 ilustra un diagrama de flujo para la detección y clasificación de una diana estacionaria 37. Un pseudo-código ejemplar para implementar la detección y clasificación de una diana estacionaria 37 es el siguiente:

```

si (la diana es POTENCIALMENTE ESTACIONARIA)
  si (umbral_de tiempo_de inserción <
    umbral_de tiempo_de eliminación)
    1°_umbral de tiempo ← umbral_de tiempo_de inserción
    1ª_prueba ← prueba de inserción
    1ª_etiqueta ← INSERCIÓN
    2°_umbral de tiempo ← umbral_de tiempo_de eliminación
    2ª_prueba ← prueba de eliminación
    2ª_etiqueta ← ELIMINACIÓN
  o
    2°_umbral de tiempo ← umbral_de tiempo_de inserción
    2ª_prueba ← prueba de inserción
    2ª_etiqueta ← INSERCIÓN
    1°_umbral de tiempo ← umbral_de tiempo_de eliminación
    1ª_prueba ← prueba de eliminación
    1ª_etiqueta ← ELIMINACIÓN
  fin
  si (edad_de la diana > 1°_umbral de tiempo)
    si (1ª_prueba es cierta para la diana)
      etiqueta_de la diana ← 1ª_etiqueta
    o si (edad_de la diana > 2°_umbral de tiempo)
      si (2ª_prueba es cierta para la diana)
        etiqueta_de la diana ← 2ª_etiqueta
      o
        etiqueta_de la diana DESCONOCIDA
    fin
  fin
fin

```

5 En el bloque 81, cada diana proporcionada por la generación de blobs 35 es examinada para determinar si la diana es potencialmente estacionaria. Este bloque corresponde a la primera condición "si" en el pseudo-código anterior (es decir, if (target is POTENTIALLY_STATIONARY) [si la diana es POTENCIALMENTE_ESTACIONARIA]). Si la diana no es potencialmente estacionaria, el flujo avanza al bloque 82 y finaliza.

10 Una técnica ejemplar para determinar si una diana es potencialmente estacionaria utiliza diversas propiedades y características espacio-temporales de la diana. Si una diana no ha cambiado radicalmente su forma y tamaño durante un periodo de tiempo, la diana puede ser una diana estacionaria. Además, si una diana muestra una gran cantidad de cambio respecto al fondo (según lo determinado mediante detección de cambios 31, 32, 33), pero muy poco movimiento independiente (según lo determinado mediante detección de movimiento 34), la diana es casi con seguridad una diana estacionaria.

5 Dos ejemplos de una diana potencialmente estacionaria se ilustran en la figura 9. La imagen 91 es un
 10 fotograma actual de un vídeo de un hombre que camina por delante de sillas apiladas y que deja caer una maleta, y la
 15 imagen 94 es un fotograma actual de un vídeo de un hombre que retira una obra de arte de una habitación. Las
 máscaras de movimiento 92 y 95 resultan de la clasificación de píxeles basada en el movimiento 34 e ilustran máscaras
 de píxeles de píxeles “en movimiento” (es decir, píxeles que muestran movimiento). La máscara de movimiento 92
 detecta al hombre pero no a la maleta en el fotograma 91, y la máscara de movimiento 95 detecta al hombre que
 camina con la obra de arte, pero no su ausencia de la pared. La máscaras de cambio 93 y 96 resultan de la clasificación
 de píxeles basada en un modelo de fondo 31 e ilustran máscaras de píxeles, de píxeles “cambiados” (es decir, píxeles
 que difieren del modelo de fondo 33). La máscara de cambio 93 detecta tanto al hombre como a la maleta, y la máscara
 de cambio 96 detecta tanto al hombre que camina con la obra de arte como la ausencia en la pared. Tal como se indica
 con los cuadrados superpuestos en las máscaras de cambio 93 y 96, hay zonas que han cambiado claramente con
 respecto al modelo de fondo 33, pero no muestran ningún movimiento independiente. En la máscara de cambio 93, la
 inserción de la maleta no muestra ningún movimiento independiente, y en la máscara de cambio 96, la eliminación de la
 obra de arte de la pared no muestra ningún movimiento independiente. Estas zonas se determinan mediante la
 detección y clasificación de una diana estacionaria 37 como dianas potencialmente estacionarias.

20 En una realización de la invención para determinar una diana estacionaria, se determinan propiedades de
 diana cuantificables. Por ejemplo, $\mu_{\Delta C}$ y $\sigma_{\Delta C}$ pueden representar propiedades estadísticas de una trayectoria centroidal
 de la diana. Específicamente, $\mu_{\Delta C}$ puede representar la media (a lo largo del tiempo) de la diferencia en la posición del
 centroide (en píxeles) entre fotogramas consecutivos, y $\sigma_{\Delta C}$ puede representar la desviación estándar (a lo largo del
 tiempo) de la diferencia de la posición del centroide (en píxeles) entre fotogramas consecutivos. En general, $\mu_{\Delta C}$ y $\sigma_{\Delta C}$
 representan propiedades estadísticas de una trayectoria centroidal de la diana estacionaria.

25 Además, μ_R y σ_R representan propiedades estadísticas del área de píxeles de la diana. Específicamente, μ_R
 puede representar la media (a lo largo de algún periodo de tiempo reciente) de la proporción del área de la diana (en
 píxeles) entre fotogramas consecutivos, y σ_R puede representar la desviación estándar (a lo largo de algún periodo de
 tiempo reciente) de la proporción del área de la diana (en píxeles) entre fotogramas consecutivos. Estas cuatro
 propiedades de la diana ejemplares (es decir, $\mu_{\Delta C}$, $\sigma_{\Delta C}$, μ_R , y σ_R) capturan el movimiento general y el cambio de tamaño
 de una diana a lo largo del tiempo.

30 Además, μ_M y σ_M representan propiedades estadísticas de píxeles en movimiento de la diana estacionaria.
 Específicamente, μ_M puede representar la media (a lo largo de algún periodo de tiempo reciente) de la proporción del
 número de píxeles “en movimiento” con respecto al área de la diana (en píxeles), y σ_M puede representar la desviación
 estándar (a lo largo de algún periodo de tiempo reciente) de la proporción del número de píxeles “en movimiento” con
 respecto al área de la diana (en píxeles). Estas dos propiedades de la diana ejemplares (es decir, σ_M y μ_M) capturan el
 grado en el cual una diana está mostrando movimiento independiente, tal como según la descripción anterior.

35 Utilizando estas seis propiedades de la diana ejemplares, una posible técnica para determinar si una diana es
 potencialmente estacionaria se basa en el siguiente pseudo-código:

```

Si ( $\mu_{\Delta C}$  < UMBRAL1 &&  $\sigma_{\Delta C}$  < UMBRAL 2 &&  $\mu_R$  < UMBRAL3 &&
       $\sigma_R$  < UMBRAL4 &&  $\mu_M$  < UMBRAL5 &&  $\sigma_M$  <
      UMBRAL6
      diana potencialmente_estacionaria
fin
    
```

40 En el pseudo-código, seis umbrales (es decir, THRESHOLD1, THRESHOLD2, THRESHOLD3, THRESHOLD4,
 THRESHOLD5 y THRESHOLD6) se utilizan para realizar comparaciones de umbral con las propiedades de la diana
 ejemplares. Los seis umbrales pueden estar preestablecidos y/o establecidos arbitrariamente como parámetros del
 usuario.

Aunque se describen cuatro propiedades de la diana ejemplares (es decir, $\mu_{\Delta C}$, $\sigma_{\Delta C}$, μ_R y σ_R) como
 representando el movimiento general y el cambio de tamaño de una diana a lo largo del tiempo, pueden utilizarse otras
 propiedades, tal como será evidente para los expertos en la materia.

45 Aunque dos propiedades de la diana ejemplares (es decir, μ_M y σ_M) se describen como representando que se
 muestra movimiento independiente, pueden utilizarse otras propiedades, tal como será evidente para los expertos en la
 materia.

Aunque se describen las seis propiedades estadísticas anteriores (es decir, $\mu_{\Delta C}$, $\sigma_{\Delta C}$, μ_R , σ_R , μ_M y σ_M), pueden utilizarse otras combinaciones de estas propiedades estadísticas, otras propiedades estadísticas, y/o otras propiedades tal como será evidente para los expertos en la materia.

5 En el bloque 83, se determinan las relaciones entre un umbral de inserción y un umbral de eliminación. Este bloque corresponde al bloque de la segunda condición "si" en el pseudo-código anterior (es decir, "if (insertion_time_threshold < removal_time_threshold)" [si (umbral_de tiempo_de inserción < umbral_de tiempo_de retirada)]). El pseudo-código para clasificar las dianas detectadas depende de la relación entre el umbral de tiempo de inserción y el umbral de tiempo de eliminación. Esta relación determina cuál de las dos pruebas, concretamente una prueba de inserción o una prueba de eliminación, se realiza en primer lugar. El umbral de tiempo de inserción y el
10 umbral de tiempo de eliminación son puntos en el tiempo basados en las escalas temporales establecidas para una inserción y una eliminación, tal como se ha descrito anteriormente. En el pseudo-código, el umbral de tiempo de inserción y el umbral de tiempo de eliminación se comparan con la edad de la diana.

15 En el bloque 84, se aplica la prueba de inserción y/o la prueba de eliminación. Si la aplicación de estas pruebas determina que la diana es una inserción, el flujo avanza hasta el bloque 85, y la diana se clasifica como una inserción. Si la aplicación de estas pruebas determina que la diana es una eliminación, el flujo avanza hasta el bloque 86, y la diana se clasifica como una eliminación. Si la aplicación de estas pruebas es inconcluyente en cuanto a si la diana es una inserción o una eliminación, el flujo avanza hasta el bloque 87, y la diana se clasifica como desconocida. Los bloques 84-86 corresponden al bloque de la tercera condición "si" en el pseudo-código anterior (es decir, "if (target_age > 1st_time_threshold)" [si (edad_de la diana > 1º_umbral_de tiempo)]). Una vez que se considera que una diana
20 potencialmente estacionaria es estacionaria al pasar la prueba de inserción y/o la prueba de eliminación, su descripción es enviada a la actualización local del modelo de fondo 38, el cual modifica el modelo de fondo 33 en la ubicación de la diana potencialmente estacionaria. Este proceso implica sustituir las estadísticas del modelo de fondo (media y variable) de los píxeles que representan la diana potencialmente estacionaria. Los valores de la media y la variable de los píxeles que representan la diana estacionaria se modificarán para representar la media y la variable de los píxeles de
25 fotografías más recientes que representan la diana potencialmente estacionaria.

La prueba de inserción y la prueba de eliminación se ilustran con las figuras 10 y 11. La teoría detrás de la técnica de clasificación ejemplar de la invención es que una inserción puede caracterizarse como una región que muestra bordes intensos alrededor de su periferia en una imagen actual pero no muestra bordes intensos alrededor de la periferia de la misma región en el modelo de fondo. A la inversa, una eliminación puede caracterizarse como una
30 región que muestra bordes intensos alrededor de su periferia en el modelo de fondo pero no muestra bordes intensos alrededor de su periferia en una imagen actual.

La figura 10 ilustra la clasificación de una inserción. El vídeo en este ejemplo es de un hombre que camina por delante de sillas apiladas y deja caer una maleta. La imagen 101 ilustra una imagen del modelo de fondo, y la imagen del borde de fondo 102 ilustra los bordes correspondientes de la imagen 101 determinados utilizando un detector de bordes Sobel. La imagen 103 ilustra una imagen del fotograma actual, y la imagen del borde del fotograma actual 104
35 ilustra los bordes correspondientes de la imagen 103 determinados utilizando un detector de bordes Sobel. Tal como puede verse, la maleta muestra bordes muy intensos en el fotograma actual (es decir, la imagen del borde del fotograma actual 104), pero no en el modelo de fondo (es decir, imagen del borde de fondo 102). La máscara de cambio 105 muestra los píxeles cambiados detectados, incluyendo el objeto estacionario (es decir, la maleta). La imagen 106 es un primer plano de la región de la maleta en la máscara de cambio 105, y la imagen 107 es un primer plano de una sección en la periferia de la región de la maleta en la imagen 106. Las imágenes 108 y 109 muestran los bordes correspondientes a la sección de la imagen 107 tanto para la imagen del borde de fondo 102 como la imagen del borde del fotograma actual 104, respectivamente. Tal como puede verse, la intensidad del borde en la imagen 109 para el
40 fotograma actual es mayor que la intensidad del borde en la imagen 108 para el modelo de fondo. Por lo tanto, la diana (es decir, la maleta) se clasifica como una inserción.

La figura 11 ilustra la clasificación de una eliminación. El vídeo en este ejemplo es de un hombre que retira una obra de arte de una habitación. La imagen 111 ilustra una imagen del modelo de fondo, y la imagen del borde de fondo 112 ilustra los bordes correspondientes de la imagen 111 determinados utilizando un detector de bordes Sobel. La imagen 113 ilustra una imagen del fotograma actual, y la imagen del borde del fotograma actual 114 ilustra los bordes correspondientes de la imagen 113 determinados utilizando un detector de bordes Sobel. Tal como puede verse, la obra
50 de arte muestra bordes muy intensos en el modelo de fondo (es decir, imagen del borde del modelo de fondo 112), pero no en el fotograma actual (es decir, imagen del fotograma actual 114). La máscara de cambio 115 muestra los píxeles cambiados detectados, incluyendo el objeto estacionario (es decir, la obra de arte). La imagen 116 es un primer plano de la región de la obra de arte en la máscara de cambio 115, y la imagen 117 es un primer plano de una sección en la periferia de la región de la obra de arte en la imagen 116. Las imágenes 118 y 119 muestran los bordes correspondientes a la sección de la imagen 117 tanto para la imagen del borde de fondo 112 como la imagen del borde del fotograma actual 114, respectivamente. Tal como puede verse, la intensidad del borde en la imagen 118 para el
55 modelo de fondo es mayor que la intensidad del borde en la imagen 119 para el fotograma actual. Por lo tanto, la diana

(es decir, la obra de arte) se clasifica como una eliminación.

5 Las figuras 12 y 13 ilustran dos realizaciones para los bloques 84-87 en la figura 8. La figura 12 ilustra la realización para el caso en el que el umbral de tiempo de inserción es menor que el umbral de tiempo de eliminación, y la figura 13 ilustra el otro caso correspondiente en el que el umbral de tiempo de inserción no es menor que el umbral de tiempo de eliminación.

En la figura 12, para el bloque 1201, la intensidad del borde E_B del fondo se determina a lo largo de el límite de la diana potencialmente estacionaria (es decir, el cambio detectado).

En el bloque 1202, la intensidad del borde E_F del fotograma actual se determina a lo largo del límite de la diana estacionaria.

10 En el bloque 1203, se determina la diferencia entre la intensidad del borde E_B del fondo y la intensidad del borde E_F del fotograma actual (es decir, $\Delta E = E_B - E_F$).

En el bloque 1204, la edad de la diana se compara con el umbral de tiempo de inserción. Si la edad de la diana es mayor que el umbral de tiempo de inserción, el flujo avanza hasta el bloque 1205. En caso contrario, el flujo avanza hasta el bloque 1211 y finaliza.

15 En el bloque 1205, la diferencia ΔE se compara con un umbral de inserción TH_I . Para la formulación en este caso, si $\Delta E < TH_I$ (donde $TH_I < 0$), la diana estacionaria es una inserción, y el flujo avanza hasta el bloque 1206. En caso contrario, el flujo avanza hasta el bloque 1207.

En el bloque 1206, la diana estacionaria se clasifica como una inserción.

20 En el bloque 1207, la edad de la diana se compara con el umbral de tiempo de eliminación. Si la edad de la diana es mayor que el umbral de tiempo de eliminación, el flujo avanza hasta el bloque 1208. En caso contrario, el flujo avanza hasta el bloque 1211 y finaliza.

En el bloque 1208, la diferencia ΔE se compara con un umbral de eliminación TH_R . Para la formulación en este caso, si $\Delta E > TH_R$, la diana estacionaria es una eliminación, y el flujo avanza hasta el bloque 1209. En caso contrario, el flujo avanza hasta el bloque 1210.

25 En el bloque 1209, la diana estacionaria se clasifica como una eliminación.

En el bloque 1210, la diana estacionaria no puede clasificarse como una inserción o una eliminación y, en su lugar, se clasifica como desconocida.

30 Después de los bloques 1206, 1208 y 1210, la descripción de la diana estacionaria es enviada a la actualización local del modelo de fondo 38, el cual modifica el modelo de fondo 33 para reflejar el cambio causado por la diana estacionaria detectada. Incluso aunque la diana estacionaria no pueda clasificarse como inserción o eliminación (bloque 1210), el modelo de fondo sigue estando actualizado.

Para aumentar la robustez, las intensidades de los bordes E_B y E_F pueden determinarse en los bloques 1201 y 1202 a lo largo de una serie de fotogramas y promediarse a lo largo del tiempo.

35 La figura 13 es igual que la figura 12, excepto por el cambio de lugares en el diagrama de flujo para los bloques 1204-1206 y los bloques 1207-1209.

40 La figura 14 ilustra un diagrama de flujo para un ejemplo técnico para determinar las intensidades de los bordes E_B y E_F para los bloques 1201 y 1202. Otras técnicas están disponibles, tal como será evidente para los expertos en la materia. La figura 14 se describe en relación con la figura 15, la cual ilustra un ejemplo de diana estacionaria en la cual se determinan las intensidades de los bordes. Con el ejemplo técnico de la figura 14, se da cabida a cierta incertidumbre en el límite del cambio detectado, y se ignoran agujeros y lagunas en el objeto.

En el bloque 141, se selecciona una banda de la imagen. Por ejemplo, la banda Y se selecciona en una imagen YCrCb. Otras bandas, aparte de la banda Y, pueden seleccionarse. También, como una opción, pueden seleccionarse múltiples bandas. Además, también se puede dar cabida a otros tipos de imágenes con la invención, tales como una imagen RGB o CMYK.

45 En el bloque 142, se selecciona una línea a lo largo de un píxel del perímetro P_p y el centroide P_c de la diana. En la figura 15, el centroide P_c de la diana 151 se designa con una estrella, y los píxeles del perímetro ejemplares 152, 153 y 154 se designan con círculos claros a lo largo del perímetro de la diana 151. Tres píxeles del perímetro ejemplares se identifican en la figura 15, y para cada píxel del perímetro, se elige una línea a través del píxel del

perímetro P_p y el centroide P_c ,

En el bloque 143, dos píxeles P_1 y P_2 en la línea se seleccionan a una distancia +/- equivalente del píxel del perímetro P_p . En la figura 15, los dos píxeles para cada línea se designan con círculos oscuros.

5 En el bloque 144, si ambos píxeles a distancia están dentro o fuera de la diana, el flujo avanza hasta el bloque 145. En caso contrario, si un píxel a distancia está dentro de la diana y el otro píxel a distancia está fuera de la diana, el flujo avanza hasta el bloque 146. En la figura 15, los píxeles del perímetro 152 y 153 presentan ambos píxeles a distancia dentro de la diana 151, y el píxel del perímetro 154 presenta un píxel a distancia dentro de la diana y el otro píxel a distancia fuera de la diana.

10 En el bloque 145, si ambos píxeles a distancia están dentro o fuera de la diana, se ignora el píxel del perímetro, y el flujo avanza hasta el bloque 147. En la figura 15, se ignoran los píxeles del perímetro 152 y 153.

En el bloque 146, un contraste C_p del píxel del perímetro que presenta un píxel a distancia dentro de la diana y el otro píxel a distancia fuera de la diana se determina en base a la intensidad de los dos píxeles a distancia I_{p1} e I_{p2} de la siguiente manera: $C_p = |I_{p1} - I_{p2}|$.

15 En el bloque 147, si se comprobaron todos los píxeles del perímetro, el flujo avanza hasta el bloque 148. En caso contrario, el flujo avanza hasta el bloque 142 para seguir comprobando los píxeles del perímetro.

En el bloque 148, el contraste promedio se determina en todos píxeles del perímetro para los cuales se determinó un contraste C_p en el bloque 146. Este contraste promedio puede utilizarse como las intensidades de los bordes E_B y E_F en los bloques 1201 y 1202, respectivamente.

20 La descripción anterior para la figura 15 abordaba los tres píxeles del perímetro ejemplares 151, 152 y 153 de modo concurrente. Sin embargo, en el examen de los píxeles del perímetro según la figura 14, cada píxel del perímetro es examinado individualmente hasta que todos los píxeles del perímetro se hayan examinado, como mediante el bucle de vuelta del bloque 147 al bloque 142.

Además, con el ejemplo técnico de la figura 14, se ignoran los agujeros y pequeñas lagunas en la diana 151 en la figura 15, aumentando, de este modo, la robustez de la técnica.

25 En otra realización de la invención, las dianas detectadas se monitorizan adicionalmente para determinar si una diana nuevamente detectada se había detectado previamente mediante la detección y clasificación de una diana estacionaria 37 como un cambio en el modelo de fondo. Por ejemplo, en una aplicación de vigilancia, puede ser de interés detectar cuándo una diana entró en una escena y cuándo dejó de moverse (por ejemplo, un coche aparcando) y seguidamente monitorizar la diana (o el área de la escena donde la diana dejó de moverse) para determinar si, y
30 cuándo, vuelve a moverse la diana (por ejemplo, un coche aparcado que sale).

La figura 16 ilustra un diagrama de flujo para la segunda realización de la invención. La figura 16 es igual que la figura 3, excepto por la adición de un monitor de la diana estacionaria 161. El monitor de la diana estacionaria 161 recibe descripciones de la diana estacionaria de la detección y clasificación de una diana estacionaria 37 y proporciona una reactivación de la diana al rastreo de blobs 36. Si la diana estacionaria se clasifica como una inserción, el monitor de la diana estacionaria 161 registra la diana (por ejemplo, tiempo, tamaño, color y ubicación) y monitoriza la diana para cualquier actividad adicional. En este punto, la diana es "olvidada" por el resto del sistema como si estuviera integrada en el modelo de fondo 33 y, en efecto, pasa a hibernación. Si, en cualquier momento posterior, una diana estacionaria se detecta como una eliminación y es descrita por la detección y clasificación de una diana estacionaria 37 en las proximidades de la inserción anterior, el monitor de la diana estacionaria 161 registra la eliminación con la diana estacionaria en hibernación y ordena al rastreo de blobs 36 para reactivar esa diana.
35
40

Las realizaciones de la invención pueden implementarse con un sistema informático. Un ejemplo de sistema informático incluye un ordenador y un medio legible por ordenador. En referencia a las figuras 3 y 16, los bloques 31-38 y 161 pueden implementarse con software que reside en uno o más medios legibles por ordenador del sistema informático. El vídeo y/o las imágenes que se procesarán con la invención pueden residir en uno o más medios legibles por ordenador o proporcionarse, por ejemplo, mediante una entrada de vídeo o imágenes o una red.
45

Aunque anteriormente se han descrito diversas realizaciones de la presente invención, debe entenderse que éstas se han presentado sólo a modo de ejemplo, y no de limitación. Por lo tanto, la amplitud y alcance de la presente invención no deben estar limitados por ninguno de los ejemplos de las realizaciones anteriormente descritas, sino que, en su lugar, deben definirse sólo de acuerdo con las siguientes reivindicaciones y sus equivalentes.
50

REIVINDICACIONES

1. Un método para el procesamiento de vídeo, que comprende las etapas de:
- 5 mantener un modelo de fondo para dicho vídeo;
- detectar una diana en tal vídeo; detectar si dicha diana es una diana estacionaria, que comprende las etapas de:
- 10 determinar propiedades de movimiento general y de cambio de tamaño de dicha diana; **caracterizado por**
- determinar propiedades de movimiento independiente de dicha diana, en el que tal movimiento independiente es un movimiento de fotograma a fotograma en el que la determinación de las propiedades de movimiento independiente de dicha diana comprende además:
- 15 determinar una proporción media del número de píxeles en movimiento de dicha diana con respecto al área de píxeles de la diana; y
- determinar una desviación estándar de la proporción del número de tales píxeles en movimiento de dicha diana con respecto al área de píxeles de la diana,
- en el que tales píxeles en movimiento experimentan un movimiento independiente, y
- 20 determinar si dicha diana es estacionaria en base a tales propiedades de movimiento general y cambio de tamaño y tales propiedades de movimiento independiente; y
- clasificar dicha diana estacionaria como una inserción en tal modelo de fondo o eliminación de tal modelo de fondo.
2. Un método según la reivindicación 1, en el que tales propiedades de movimiento general y cambio de tamaño comprenden propiedades estadísticas de una trayectoria centroidal de dicha diana y propiedades estadísticas de un área de dicha diana.
3. Un método según la reivindicación 1, en el que tal movimiento independiente comprende propiedades estadísticas de píxeles en movimiento de dicha diana.
- 25 4. Un método según la reivindicación 1, en el que la determinación de si dicha diana es estacionaria se basa, además, en el envejecimiento de dicha diana.
5. Un método según la reivindicación 1, en el que la clasificación de dicha diana estacionaria comprende las etapas de:
- 30 determinar una intensidad del borde en tal modelo de fondo a lo largo de un límite de dicha diana estacionaria;
- determinar una intensidad del borde en un fotograma actual de la secuencia de vídeo a lo largo de tal límite de dicha diana estacionaria; y
- determinar si dicha diana estacionaria es tal inserción o tal eliminación en base a tal intensidad del borde en tal modelo de fondo y tal intensidad del borde en tal fotograma actual.
- 35 6. Un método según la reivindicación 5, en el que la determinación de tal intensidad del borde en tal modelo de fondo se basa en un píxel centroide de dicha diana estacionaria y píxeles del perímetro a lo largo de tal límite de dicha diana estacionaria.
7. Un método según la reivindicación 5, en el que la determinación de tal intensidad del borde en tal fotograma actual se basa en un píxel centroide de dicha diana estacionaria y píxeles del perímetro a lo largo de tal límite de dicha diana estacionaria.
- 40 8. Un método según la reivindicación 1, en el que la clasificación de dicha diana estacionaria comprende, además, clasificar dicha diana estacionaria como una inserción en tal modelo de fondo, una eliminación de tal modelo de fondo, o como inidentificable como tal inserción o tal eliminación.
9. Un método según la reivindicación 1, que comprende además la etapa de:

determinar si dicha diana se detectó previamente como una diana estacionaria.

10. Un sistema informático que comprende un medio legible por ordenador que presenta un software para hacer funcionar un ordenador según el método de la reivindicación 1.
- 5 11. Un medio legible por ordenador que presenta software para hacer funcionar un ordenador según el método de la reivindicación 1.
12. Un sistema informático para el procesamiento de vídeo, que comprende:
- un modelo de fondo de tal vídeo;
- una unidad de clasificación de píxeles basada en un modelo de fondo adaptada para producir una máscara de cambio e imágenes en base a tal vídeo y tal modelo de fondo;
- 10 una unidad de actualización del modelo de fondo adaptada para actualizar tal modelo de fondo en base a tal máscara de cambio y tales imágenes;
- una unidad de clasificación de píxeles basada en el movimiento adaptada para producir una máscara de movimiento;
- 15 una unidad de generación de blobs adaptada para producir al menos un blob en base a tal máscara de cambio y tal máscara de movimiento;
- una unidad de rastreo de blobs adaptada para producir al menos una diana en base a tales blobs;
- una unidad de detección y clasificación de una diana estacionaria para producir una descripción de la diana estacionaria en base a cada diana, tal descripción de la diana estacionaria para identificar a cada diana como una inserción en tal modelo de fondo o una eliminación de tal modelo de fondo; y
- 20 una unidad de actualización local del modelo de fondo adaptada para actualizar tal modelo de fondo en base a cada descripción de dicha diana estacionaria;
- caracterizado porque** tal unidad de clasificación de píxeles basada en el movimiento está adaptada para determinar propiedades de movimiento independiente de dicha diana, en el que tal movimiento independiente es un movimiento de fotograma a fotograma y en el que tal unidad de clasificación de píxeles basada en el movimiento está adaptada para determinar propiedades de movimiento independiente de dicha diana, determinando una proporción media del número de píxeles en movimiento de dicha diana con respecto al área de píxeles de la diana y determinar una desviación estándar de la proporción del número de tales píxeles en movimiento de dicha diana con respecto al área del píxel, en el que tales píxeles en movimiento experimentan un movimiento independiente, estando tal unidad de clasificación de píxeles basada en el movimiento adaptada, además, para producir tal máscara de movimiento en base a tal vídeo y tales propiedades de movimiento independiente.
- 25 30
13. Un sistema informático según la reivindicación 12, en el que tales descripciones de la diana estacionaria identifican, además, a cada una de dichas dianas como una inserción en tal modelo de fondo, una eliminación de tal modelo de fondo o como inidentificable como tal inserción o tal eliminación.
- 35 14. Un sistema informático según la reivindicación 12, que comprende, además, una unidad monitora de la diana estacionaria adaptada para ordenar a tal unidad de rastreo de blobs que rastree una diana que fue detectada como una inserción y a continuación detectada como una eliminación.

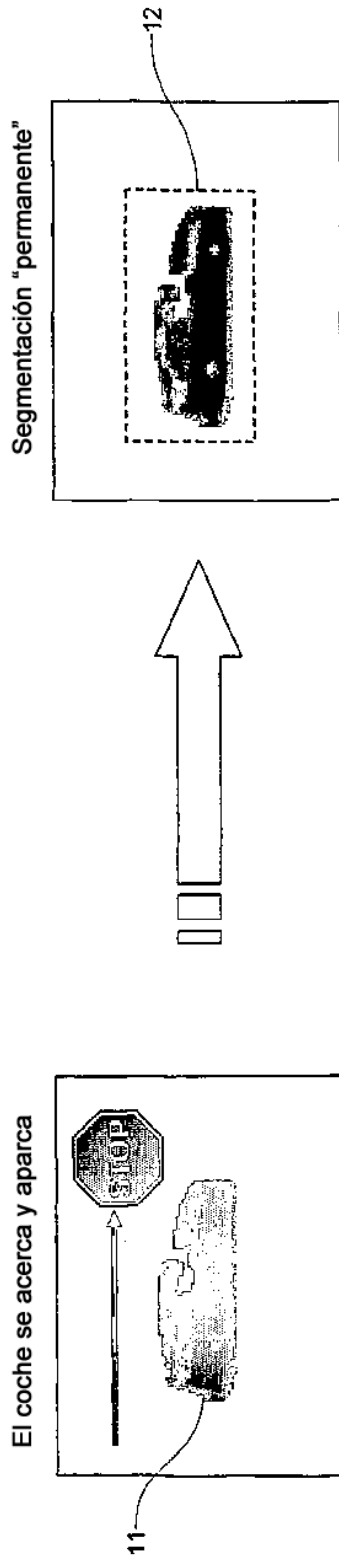


FIG. 1A

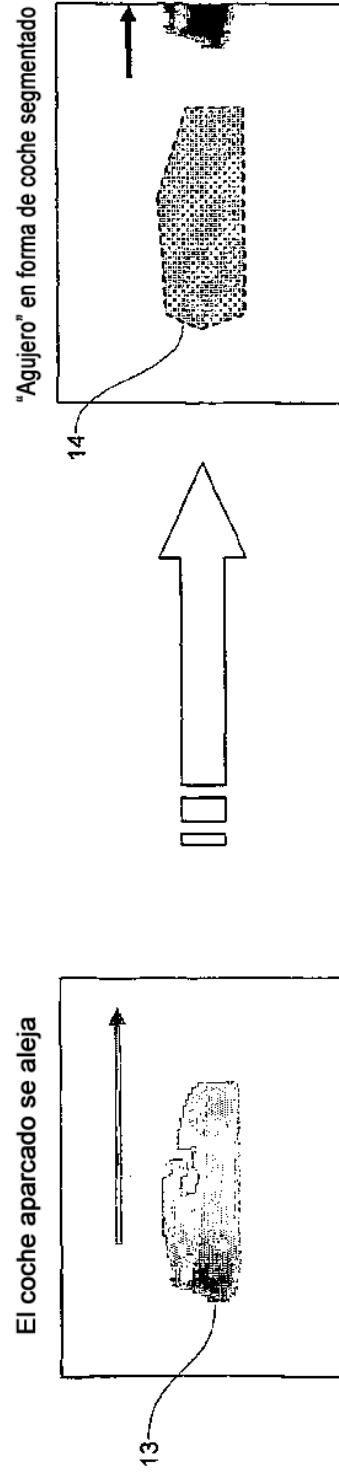


FIG. 1B

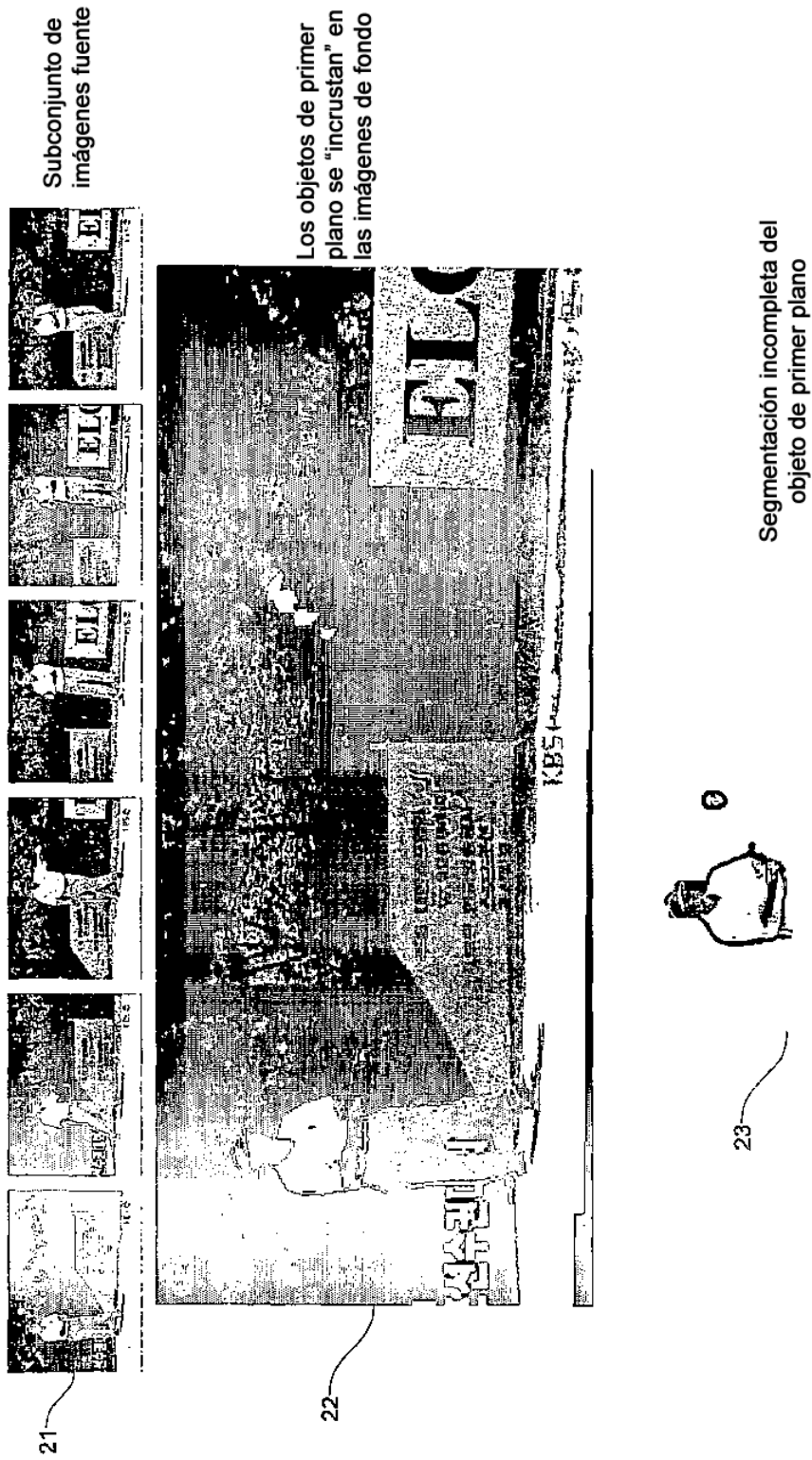


FIG. 2

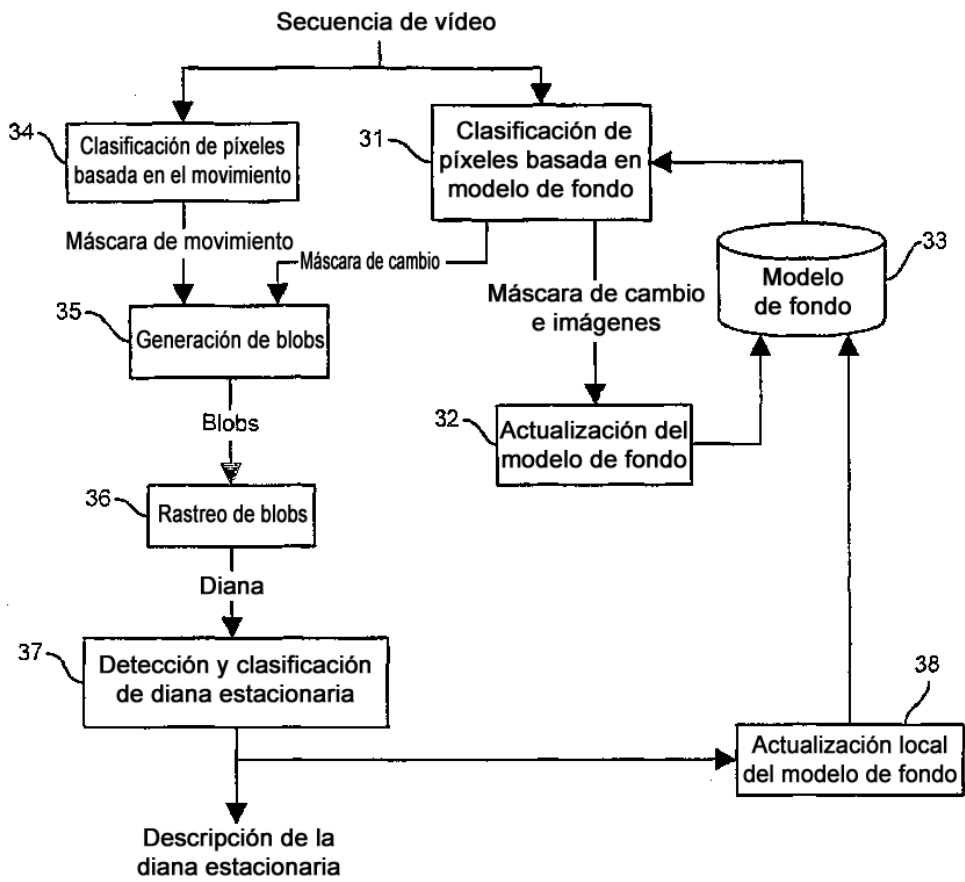


FIG. 3

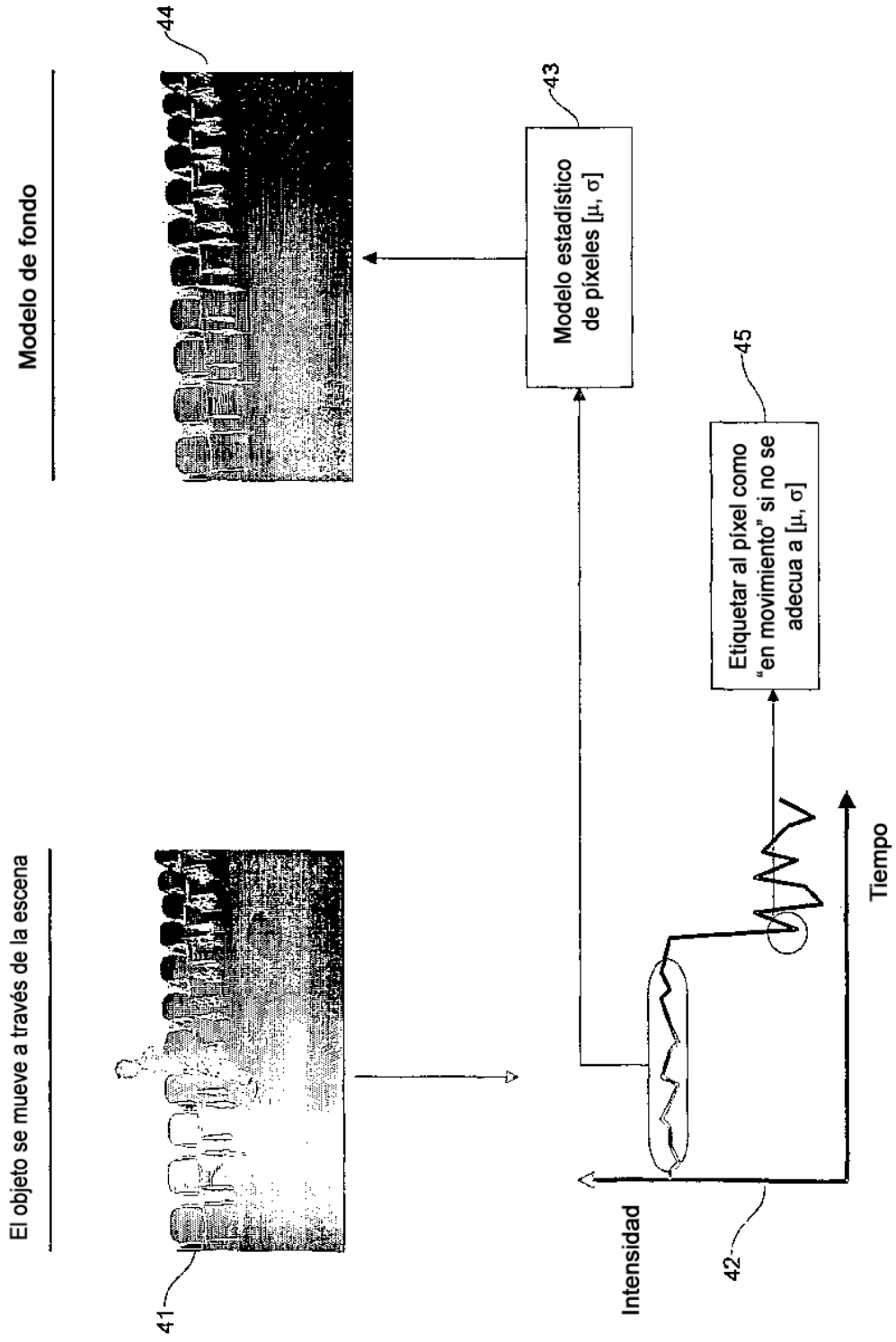


FIG. 4

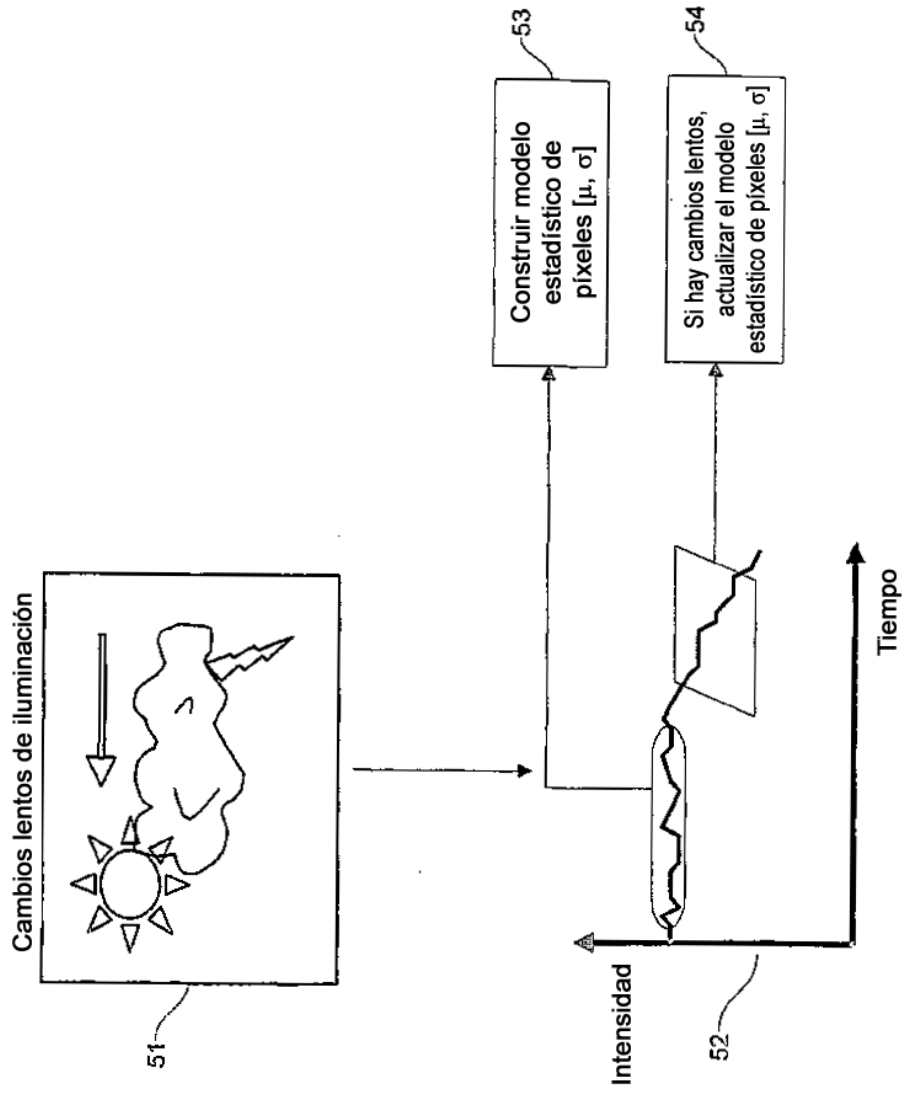


FIG. 5

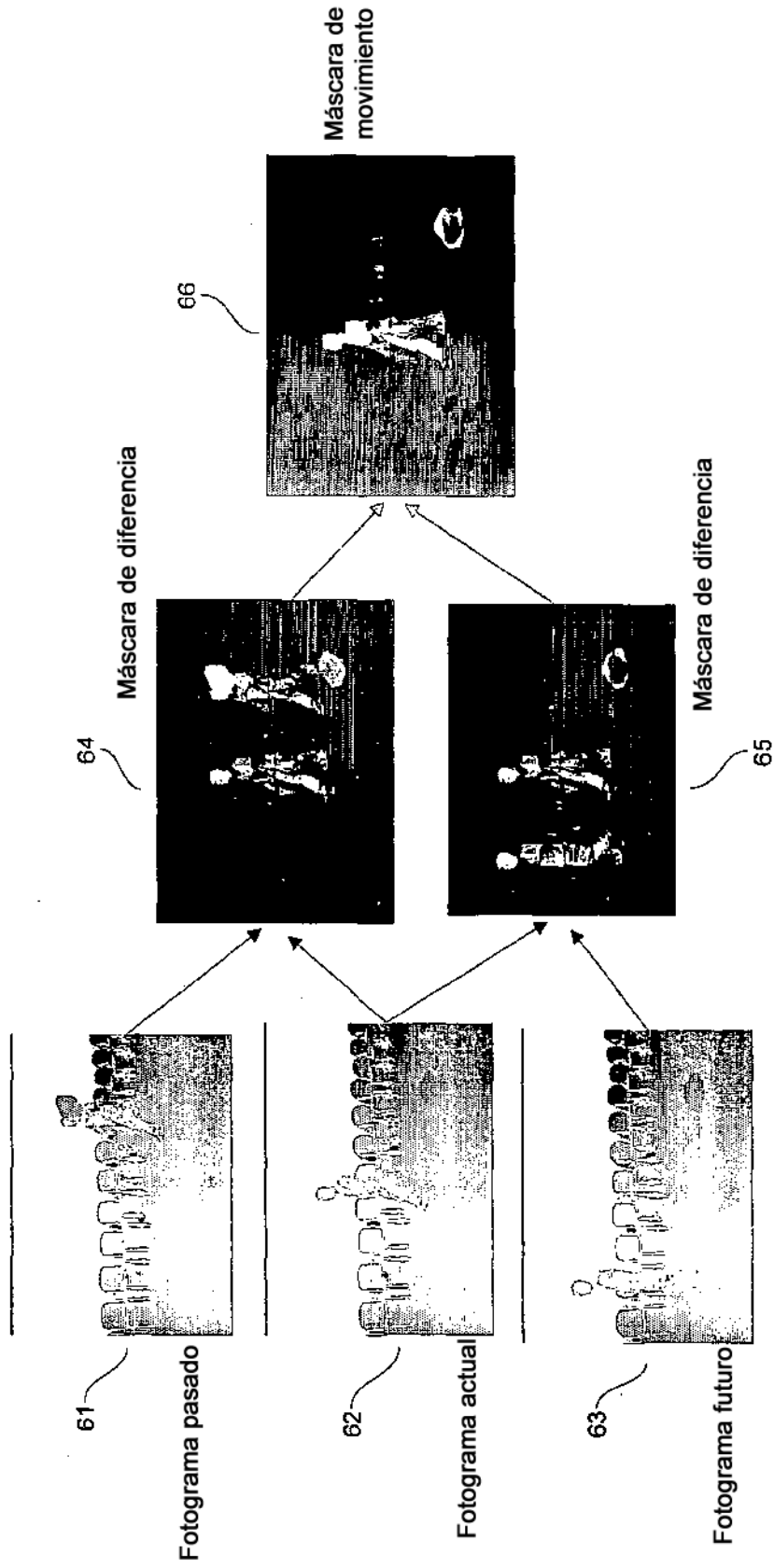


FIG. 6

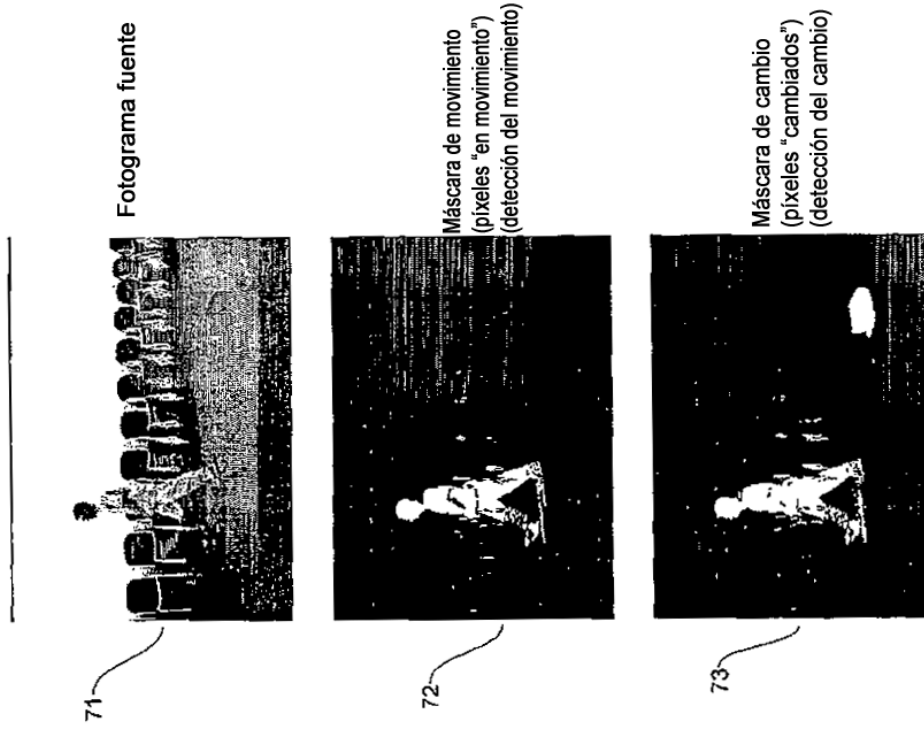


FIG. 7

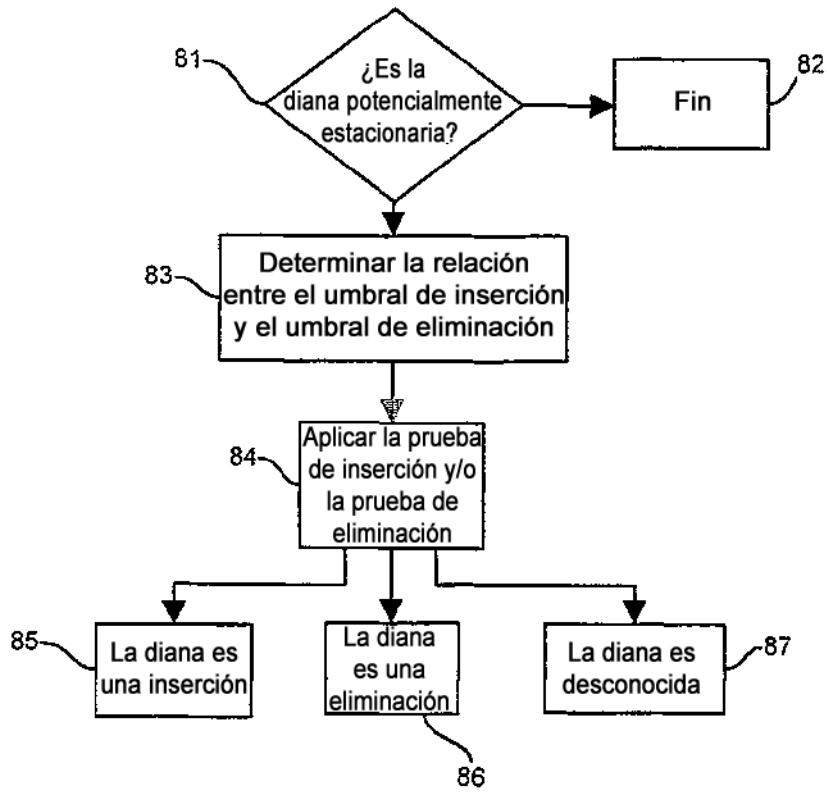


FIG. 8

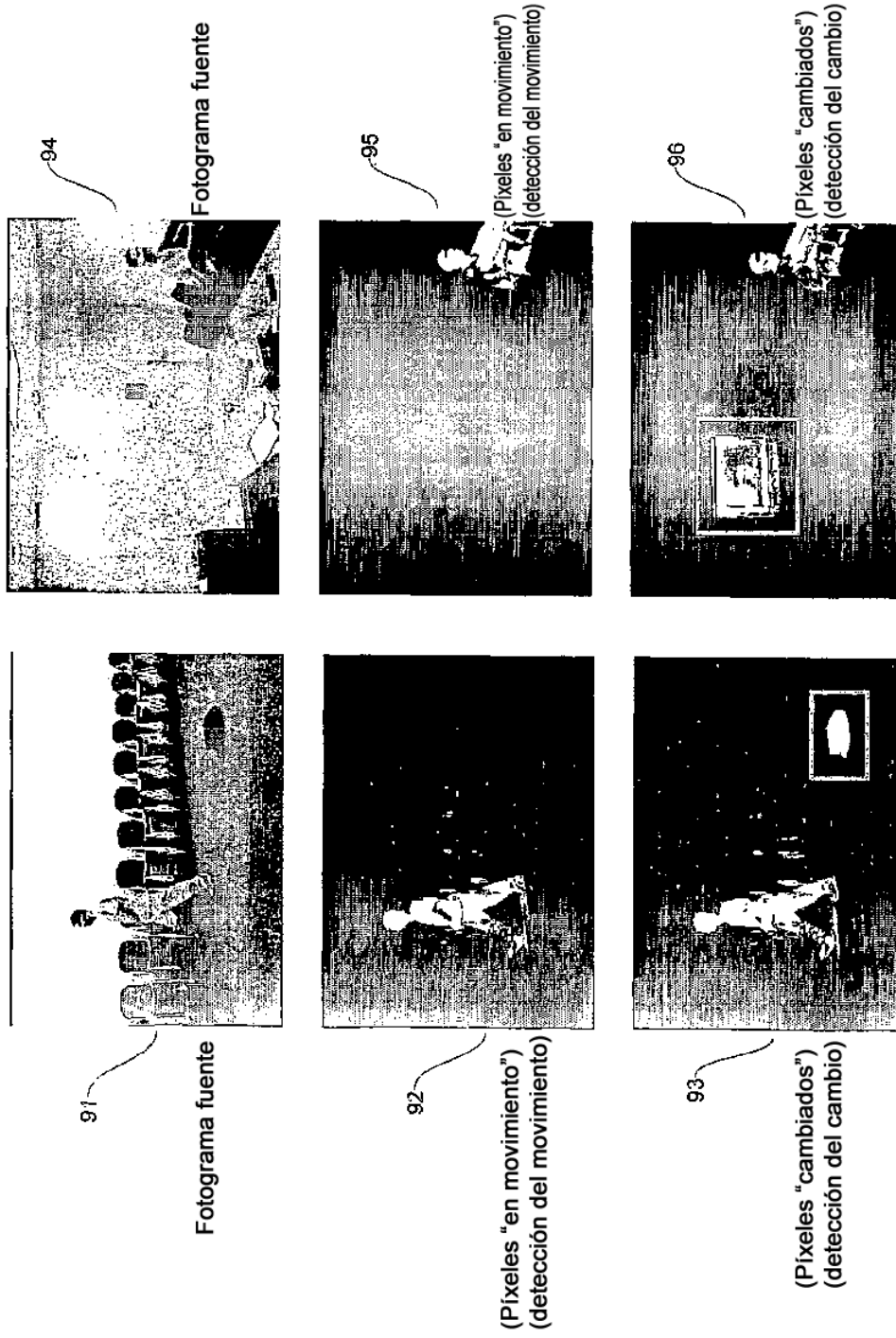


FIG. 9

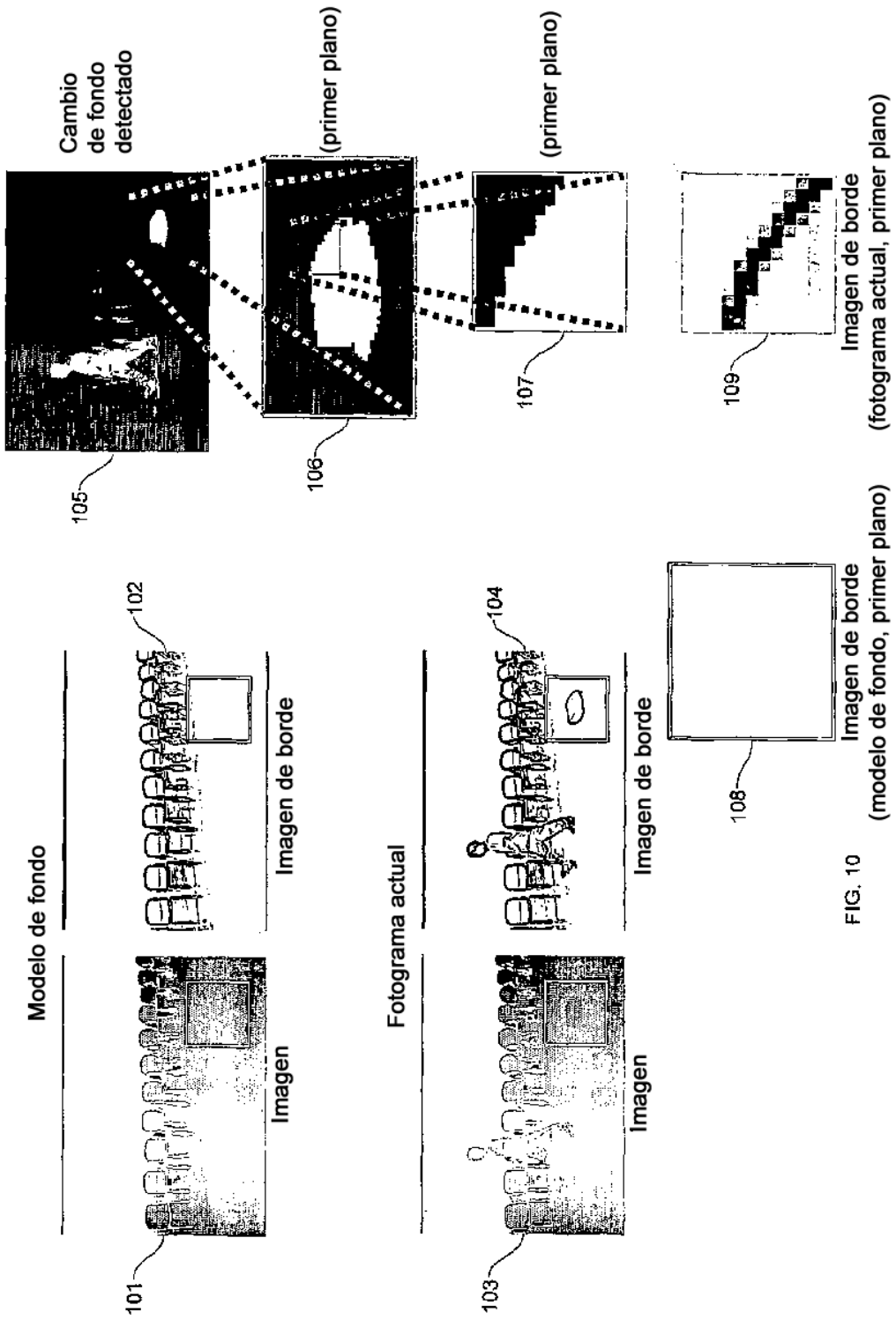


FIG. 10

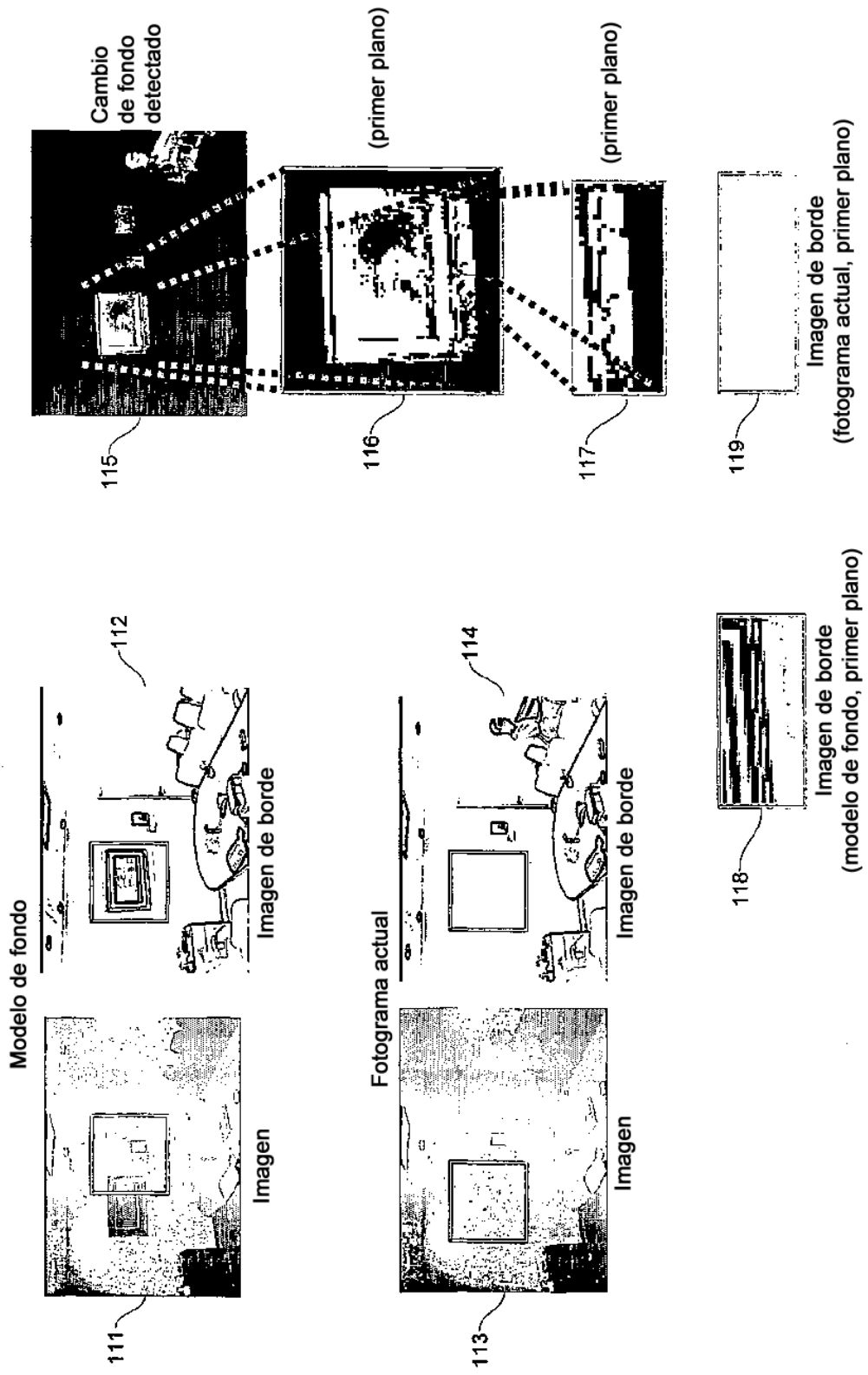


FIG. 11

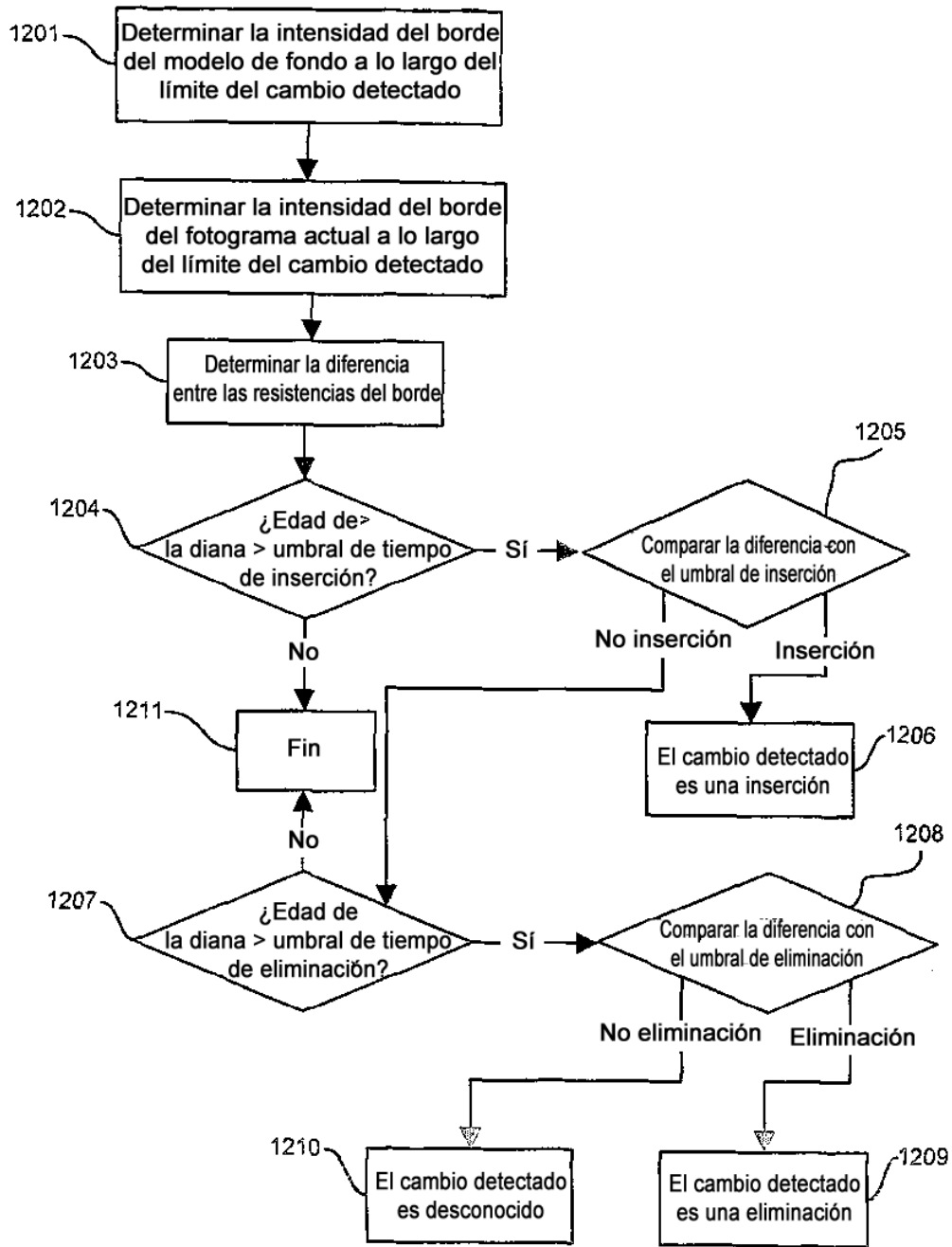


FIG. 12

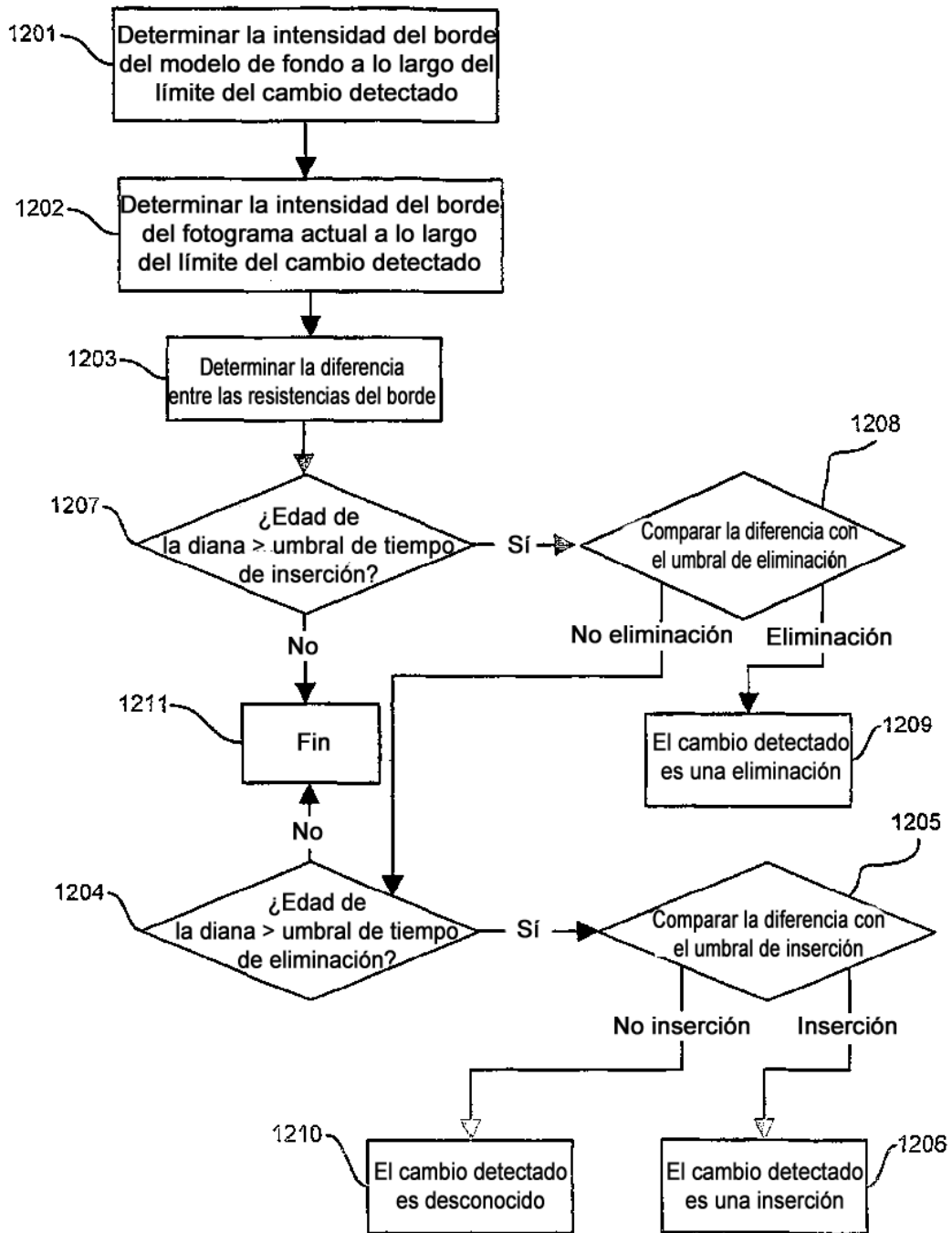


FIG. 13

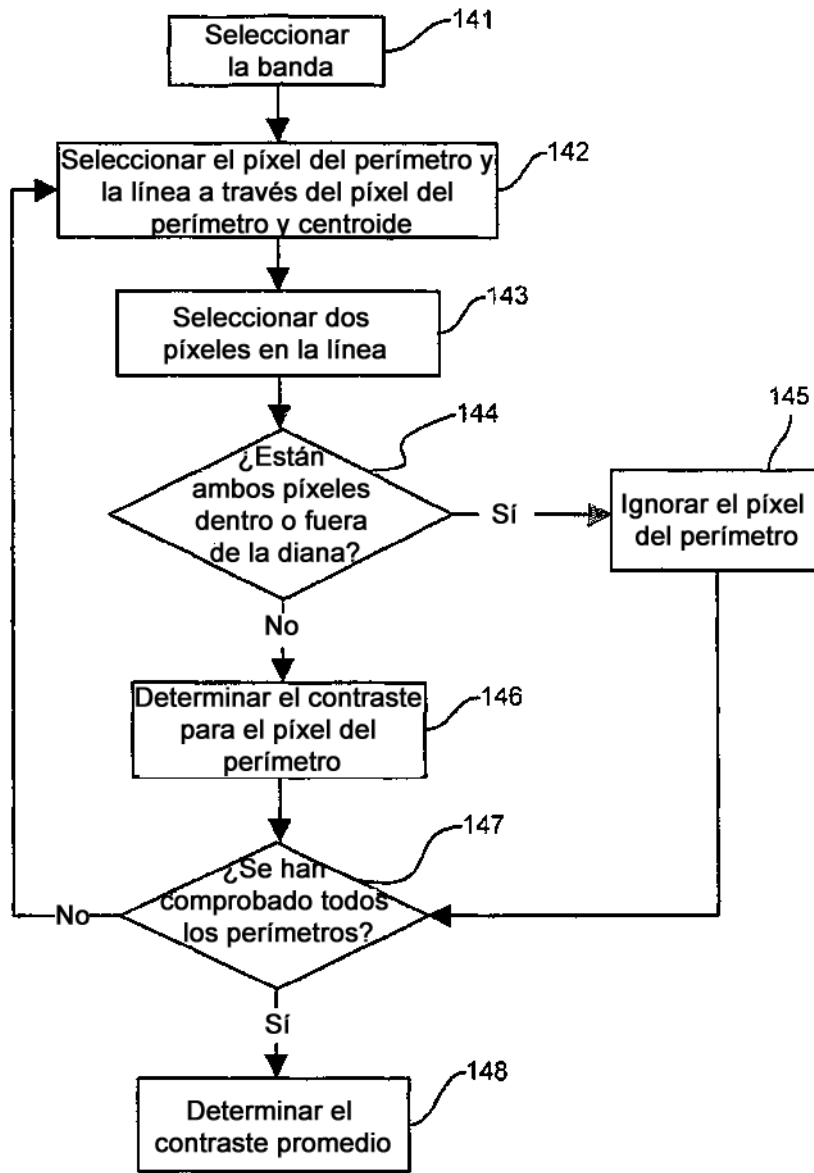


FIG. 14

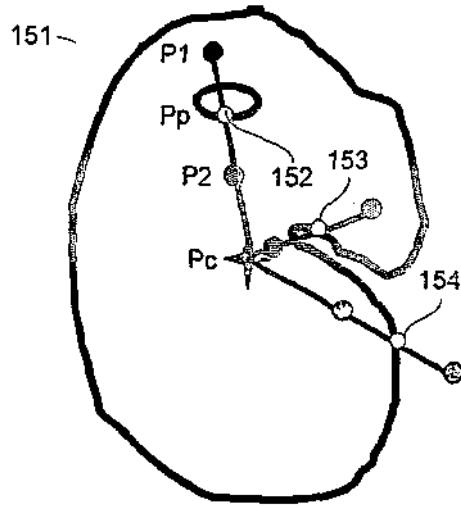


FIG. 15

DOCUMENTOS INDICADOS EN LA DESCRIPCIÓN

5 En la lista de documentos indicados por el solicitante se ha recogido exclusivamente para información del lector, y no es parte constituyente del documento de patente europeo. Ha sido recopilada con el mayor cuidado; sin embargo, la EPA no asume ninguna responsabilidad por posibles errores u omisiones.

Documentos de patente indicados en la descripción

- 10
- US 6424370 B [0003]
 - US 47216299 A [0006]
 - US 6738424 B [0006]
 - US 6625310 B [0028]
 - US 6954498 B [0032] [0036]

LITERATURA NO CITADA EN LA DESCRIPCION

- 15
- **D. Ballard ; C. Brown.** Computer Vision. Prentice-Hall, May 1982 [0036]
 - **T.E. Boulton ; R.J. Micheals ; X. Gao ; P. Lewis ; C.Power ; W. Yin ; A. Erkan.** Frame-Rate Omnidirectional Surveillance and Tracking of Camouflaged and Occluded Targets. *Proc. of the IEEE Workshop on Visual Surveillance*, 1999 [0036]
 - **Wren, C.R. et al.** Pfinder: Real-Time Tracking of the Human Body. *IEEE Trans. on Pattern Matching and Machine Intelligence*, 1997, vol. 19, 780-784 [0036]
 - **Grimson, W.E.L. et al.** Using Adaptive Tracking to Classify and Monitor Activities in a Site. *CVPR*, June 1998, 22-29 [0036]
 - **Olson, T.J. ; Brill, F.Z.** Moving Object Detection and Event Recognition Algorithm for Smart Cameras. *IJW*, May 1997, 159-175 [0036]