



UNIVERSIDAD TÉCNICA PARTICULAR DE LOJA
La Universidad Católica de Loja

MODALIDAD CLÁSICA

ESCUELA DE ELECTRÓNICA Y TELECOMUNICACIONES

DESARROLLO DE UN SISTEMA SEMIAUTOMÁTICO DE ANÁLISIS DE SITUACIONES EN TAREAS DE VIGILANCIA BASADO EN TÉCNICAS DE ATENCIÓN VISUAL SELECTIVA Y DINÁMICA CON CAPACIDAD DE APRENDIZAJE.

Trabajo de fin de carrera previa a la obtención del título de Ingeniería en Electrónica y Telecomunicaciones

AUTORES: Priscila Maldonado Mendieta

DIRECTOR: Ing. Manuel Quiñones.

Loja-Ecuador 2011



ESCUELA DE ELECTRÓNICA Y TELECOMUNICACIONES

DESARROLLO DE UN SISTEMA SEMIAUTOMÁTICO DE ANÁLISIS DE SITUACIONES EN TAREAS DE VIGILANCIA BASADO EN TÉCNICAS DE ATENCIÓN VISUAL SELECTIVA Y DINÁMICA CON CAPACIDAD DE APRENDIZAJE.

AUTORES: Priscila Maldonado Mendieta

**Proyecto de Tesis previa obtención del Título de
Ingeniero en Electrónica y Telecomunicaciones**

DIRECTOR: Ing. Manuel Quiñones.

Loja - Ecuador

Enero 2011

ACTA DE CESIÓN DE DERECHOS EN TESIS DE GRADO

Yo, Priscila Galina Maldonado Mendieta, declaro ser autora del presente trabajo y eximo expresamente a la Universidad Técnica Particular de Loja y a sus representantes legales de posibles reclamos o acciones legales.

Adicionalmente declaro conocer y aceptar la disposición del Art. 67 del Estatuto Orgánico de la Universidad Técnica Particular de Loja que en su parte pertinente textualmente dice: "Forman parte del patrimonio de la Universidad la propiedad intelectual de investigaciones, trabajos científicos o técnicos y tesis de grado que se realicen a través, o con el apoyo financiero, académico o institucional (operativo) de la Universidad"

Priscila Galina Maldonado Mendieta.



AUTORÍA

Las ideas, opiniones, conclusiones y contenidos expuestos en el presente informe de investigación son de exclusiva responsabilidad de su autor.

Priscila Maldonado Mendieta.



DEDICATORIA

Este proyecto de fin de carrera va dedicado enteramente a mi padre Galo Maldonado, porque gracias a su esfuerzo, apoyo, y confianza he podido cumplir mis objetivos tanto en lo personal como en lo académico. Sencillamente eres la base de mi vida y te estaré agradecida.

Priscila.



AGRADECIMIENTO

Agradecer especialmente la colaboración de mis amigos, que me han ayudado en numerosas ocasiones.

Gracias a todas las personas que de alguna u otra forma se preocupaban del progreso de mi proyecto y sobre todo gracias a mi familia por apoyarme y animarme durante este trabajo.

Al Ing. Manuel Quiñones, por su paciencia y por dedicar su tiempo para el éxito del proyecto, a mis profesores que me han ido formando intelectual y moralmente, para poder lograr el objetivo de ser profesional.

Y finalmente, a todos aquellos que aportaron para la realización de las pruebas y con ello llegar a la culminación exitosa de este proyecto.

Priscila Maldonado Mendieta



INTRODUCCIÓN

Actualmente, el índice de inseguridad en el Ecuador ha crecido exponencialmente, lo que conlleva a que el número de compañías de seguridad privadas aumenten, aunque no son una garantía porque los trabajadores no presentan conocimiento, debido a la falta de capacitación que existe tanto a nivel de seguridad como al manejo de equipos.

Hoy en día, en un sistema de video vigilancia, el operador cumple con una función importante para cumplir con el objetivo de seguridad, cabe recalcar que las personas presentan un índice de atención visual no superior a un tiempo de 45 minutos, con lo que el monitoreo a lo largo del tiempo pierde exactitud, por lo tanto, se da como solución el desarrollo de un sistema de video vigilancia IP complementado con un software, que nos permita el tratamiento de imágenes, logrando así obtener escenas de relevancia.

El objetivo principal del sistema de vigilancia es la prevención de actuaciones que puedan violar la seguridad de un lugar supervisado, por ejemplo la intrusión, la monitorización y registro de determinados espacios.

Un sistema de vigilancia consiste en una serie de sensores situados en posiciones estratégicas del espacio a vigilar y conectados a través de medios de transmisión adecuados. En este trabajo únicamente se considerarán aquellos que proporcionan imágenes para su análisis.

En este contexto, según (1), los sistemas de vigilancia de segunda generación o avanzados, están basados en técnicas de procesado de la información como una alternativa ideal frente a los sistemas de primera generación. La característica principal de esta tecnología es la explotación de las técnicas de procesado para seleccionar automáticamente un pequeño porcentaje de la información disponible, el cual será probablemente el más interesante para la tarea de vigilancia requerida y para ser finalmente valoradas por el personal, caso contrario a lo que ocurre actualmente, en el



cual se almacena el video de forma masiva sin existir ningún tipo de tratamiento ni procesamiento del mismo.

Este trabajo está conformado por tres capítulos. En el primero se recopila información sobre las características de los sistemas de video vigilancia existentes. En el segundo, se describe los métodos, técnicas y procesamiento utilizados para el sistema de video vigilancia e implementación en Matlab. El tercero muestra los resultados experimentales obtenidos en la evaluación del sistema y en base a ellos se presenta las conclusiones y trabajos futuros.



OBJETIVOS

Establecer la línea base existente sobre las generalidades de los Sistemas de vídeo vigilancia.

Investigar métodos, procesos y herramientas para procesamiento de imágenes aplicadas a video vigilancia.

Determinar el método a utilizar en base a la investigación de las diferentes técnicas existentes en procesamiento de imágenes.

Desarrollar un Sistema de Video Vigilancia semiautomático mediante la utilización del software MATLAB.

Realizar pruebas de verificación del sistema de video vigilancia semiautomático.



TABLA DE CONTENIDO

ACTA DE CESIÓN DE DERECHOS EN TESIS DE GRADO.....	I
AUTORÍA.....	II
DEDICATORIA.....	III
AGRADECIMIENTO.....	IV
INTRODUCCIÓN.....	V
OBJETIVOS.....	VII
TABLA DE CONTENIDO.....	VIII
LISTA DE FIGURAS.....	X
LISTA DE TABLAS.....	XII
ANEXOS.....	XIII
CAPÍTULO I.....	1
1. CARACTERÍSTICAS DEL SISTEMA DE VIDEO VIGILANCIA.....	1
1.1 SISTEMA DE VIDEO VIGILANCIA.....	1
1.2 CLASIFICACION DE LOS SISTEMAS DE VIDEO VIGILANCIA.....	1
1.2.1 Sistemas de video vigilancia de primera generación.....	2
1.2.2 Sistemas de video vigilancia de segunda generación.....	5
1.2.3 Sistemas de video vigilancia de tercera generación.....	7
CAPÍTULO II.....	11
2. DESCRIPCIÓN DEL SISTEMA DE VIDEO VIGILANCIA.....	11
2.1 Descripción general del sistema de video vigilancia.....	11
2.1.1 Adquisición de la imagen.....	11
2.1.2 Preprocesamiento.....	16
2.1.3 Segmentación.....	24
2.1.4 Morfología.....	29
CAPÍTULO III.....	32
3. DESCRIPCIÓN DEL SOFTWARE Y RESULTADOS.....	32



3.1	Modelo de desarrollo del software.....	32
3.1.1	Requisitos funcionales.	32
3.1.2	Requisitos no funcionales.	33
3.2	Resultados obtenidos y análisis.	33
3.2.1	Algoritmo de detección de personas.	33
3.2.2	Algoritmo para la detección de objetos abandonados.	42
3.3	Rendimiento del sistema de video vigilancia.....	51
CONCLUSIONES.....		54
TRABAJOS FUTUROS.....		57
REFERENCIA.....		58
ANEXOS.....		60



LISTA DE FIGURAS

Fig. 1.1 Esquema de la plataforma de video vigilancia analógica.....	3
Fig. 1.2 Esquema de un sistema analógico con DVR.....	4
Fig. 1.3 Sistema de video vigilancia IP con servidor de video.....	5
Fig. 1.4 Sistema de video vigilancia digital IP.....	7
Fig. 1.5 Sistema de video vigilancia digital IP distribuido.....	8
Fig. 2.1 Diagrama del sistema de video vigilancia.....	11
Fig. 2.2 Cámara de red IP7137 de VIVOTEK (13).....	12
Fig. 2.3 Configuración de video.....	14
Fig. 2.4 Configuración de red LAN.....	15
Fig. 2.5 Configuración de acceso o denegación de redes.....	15
Fig. 2.6 Configuración WIRELEES LAN.....	16
Fig. 2.7 Representación de Pixel (Picture element).....	17
Fig. 2.8 Imagen digital.....	18
Fig. 2.9 Imagen binaria.....	18
Fig. 2.10 Imagen de Intensidad.....	19
Fig. 2.11 Imagen a color.....	19
Fig. 2.12 a) Imagen original. b) Resultado de la imagen al aplicar filtro suavizado gaussiano. ...	21
Fig. 2.13 Detección movimiento.....	22
Fig. 2.14 Sustracción de background.....	23
Fig. 2.15 Píxeles Adyacentes.....	25
Fig. 2.16 Vecindades.....	26
Fig. 2.17 Detección de contornos mediante algoritmo de Roberts.....	28
Fig. 2.18 Detección de contornos mediante algoritmo de Prewitt.....	29
Fig. 2.19 Detección de contornos mediante algoritmo de Sobel.....	29
Fig. 2.20 Esquema del procesado morfológico.....	30
Fig. 3.1 Esquema del modelo de software.....	32
Fig. 3.2 Diagrama de flujo del algoritmo de detección de personas.....	34
Fig. 3.3 Individuo a 3 metros de la cámara.....	35
Fig. 3.4 Individuo a 1 metro de la cámara.....	36
Fig. 3.5 (a) Individuos captados en el borde derecho de la imagen (b) Individuos captados en el borde izquierdo de la imagen.....	37
Fig. 3.6 Individuos a 3 metros de la cámara.....	38
Fig. 3.7 Individuos a mínima distancia entre pixeles.....	39



Fig. 3.8 Actualización de background, detección de movimiento, segmentación y técnicas de morfología.....	40
Fig. 3.9 Cambio de iluminación en la escena.....	41
Fig. 3.10 Cambio de iluminación en el entorno.....	42
Fig. 3.11 Diagrama de flujo del algoritmo de detección de objetos abandonados.....	43
Fig. 3.12 Individuo abandonando objeto.....	44
Fig. 3.13 Individuo retira el objeto.....	45
Fig. 3.14 Objeto a 1 metro de la cámara.....	46
Fig. 3.15 Objeto retirado de la escena.....	47
Fig. 3.16 Objeto de mayor tamaño a 2 metros de la cámara.....	48
Fig. 3.17 Objeto a 2 metros de la cámara es retirado de la escena.....	49
Fig. 3.18 Objeto de mayor tamaño a 1 metro de la cámara.....	50
Fig. 3.19 Objeto a 1 metro de la cámara es retirado de la escena.....	51
Fig. 3.20 Características técnicas del computador.....	52
Fig. 3.21 Tiempo de respuesta obtenido para el algoritmo de detección de personas.....	52
Fig. 3.22 Tiempo de respuesta obtenido para el algoritmo de detección de objetos abandonados.....	53



LISTA DE TABLAS

Tabla 1.1 Tipos de sistemas de video vigilancia.....	1
--	---



ANEXOS.

ANEXO 1 Especificaciones de la Cámara IP137.

ANEXO 2 Manual de usuario.

ANEXO 3 Artículo.



CAPÍTULO I

CARACTERÍSTICAS DEL SISTEMA DE VIDEO VIGILANCIA.

1.1 SISTEMA DE VIDEO VIGILANCIA.

Las técnicas de visión por computador tienen como objetivo extraer un conjunto de propiedades del entorno a través de imágenes. Mediante estas propiedades se pueden construir sistemas inteligentes que realicen determinadas acciones respondiendo a eventos del entorno, por lo que su aplicación en video vigilancia resulta más apropiada.

En este punto resulta más eficiente la detección de elementos móviles (activos), pero a pesar de lo antes mencionado es importante centrarse en la zona inmóvil (pasiva), ya que no siempre varía con el tiempo, pero a pesar de ello se ve afectada por las condiciones de iluminación (1) lo mismo que da lugar a cambios en la percepción de la imagen en el computador y por lo tanto en el procesamiento de las mismas.

1.2 CLASIFICACION DE LOS SISTEMAS DE VIDEO VIGILANCIA.

Existen diversos tipos de sistemas de video-vigilancia en la Tabla 1.1, los mostrados son considerados de mayor importancia.

Tabla 1.1 Tipos de sistemas de video vigilancia.

SISTEMA DE VIDEOVIGILANCIA DE PRIMERA GENERACIÓN	Sistemas de circuito cerrado de televisión (CCT)
	Sistemas de video IP que utilizan servidores de video
SISTEMA DE VIDEOVIGILANCIA DE SEGUNDA GENERACIÓN	Sistemas de video vigilancia IP
SISTEMA DE VIDEOVIGILANCIA DE TERCERA GENERACIÓN	Sistemas de video vigilancia IP con sensores heterogéneos.



1.2.1 Sistemas de video vigilancia de primera generación.

Los sistemas de video vigilancia de primera generación necesitan de operadores humanos que son encargados de evaluar en todo momento la actividad del espacio vigilado y tomar acciones frente a los eventos ocurridos.

En este tipo de sistemas se requiere del aporte humano para el procesado de la información y por ello presenta una serie de limitaciones y problemas pese a que su implementación es sencilla.

Las limitaciones y problemas son causadas debido a que el rendimiento del operador humano decrece con el paso del tiempo, ya sea por el cansancio y la cantidad de información que este debe procesar.

1.2.1.1 Sistemas de circuito cerrado de televisión (CCTV).

Circuito cerrado de televisión o CCTV (siglas en inglés de *closed circuit television*) es una tecnología de vídeo vigilancia visual la cual es utilizada para la monitorización de diferentes tipos de entornos. Se la denomina circuito cerrado debido a que todos sus componentes están enlazados (2).

El sistema puede estar conformado por una o varias cámaras de vigilancia conectadas a monitores de video o televisores que permiten la reproducción de las imágenes adquiridas por las cámaras.

Sus características analógicas encarecen su precio y no proporcionan: flexibilidad, escalabilidad, redundancia y tolerancia a fallas (3).

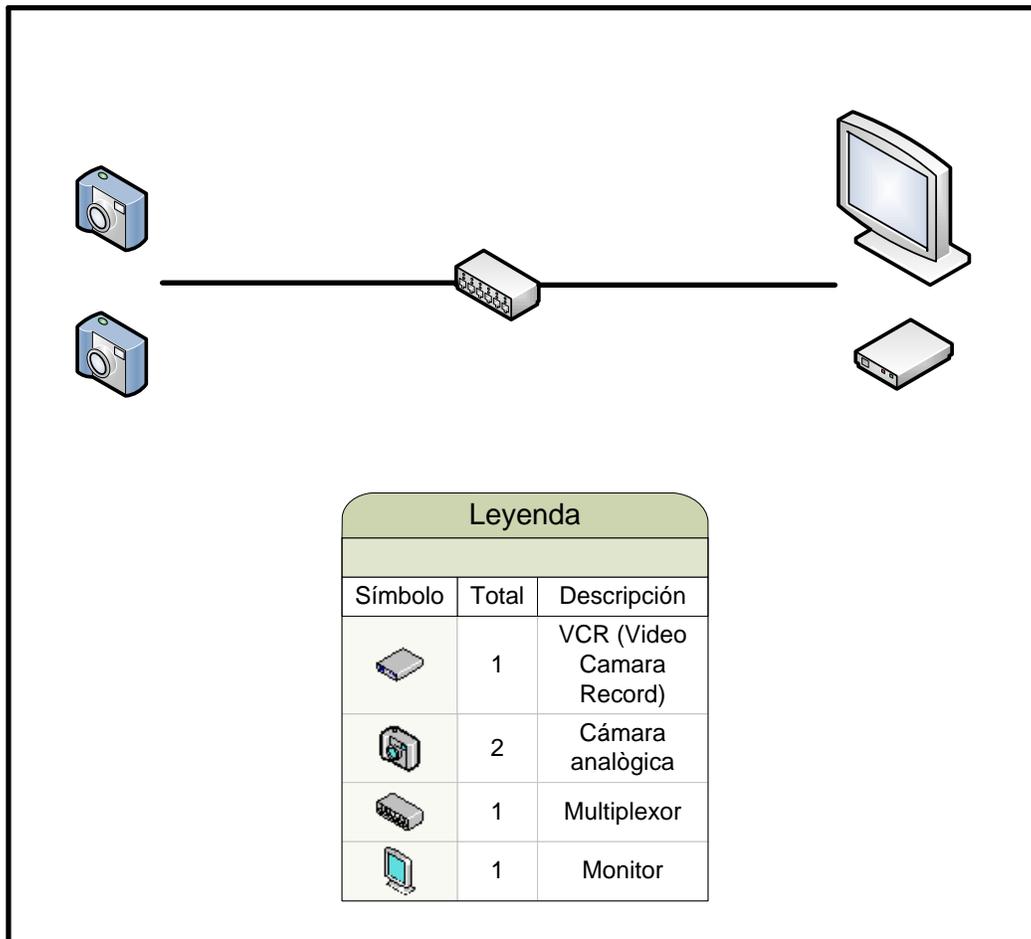


Fig. 1.1 Esquema de la plataforma de video vigilancia analógica.

En la Fig. 1.1 se presenta un sistema de circuito cerrado de televisión en el cual se emplea cámaras analógicas y el multiplexor o concentrador de señales de video, además se utiliza el VCR (Video Cámara Record) conectado al multiplexor para grabar el video proveniente de las cámaras.

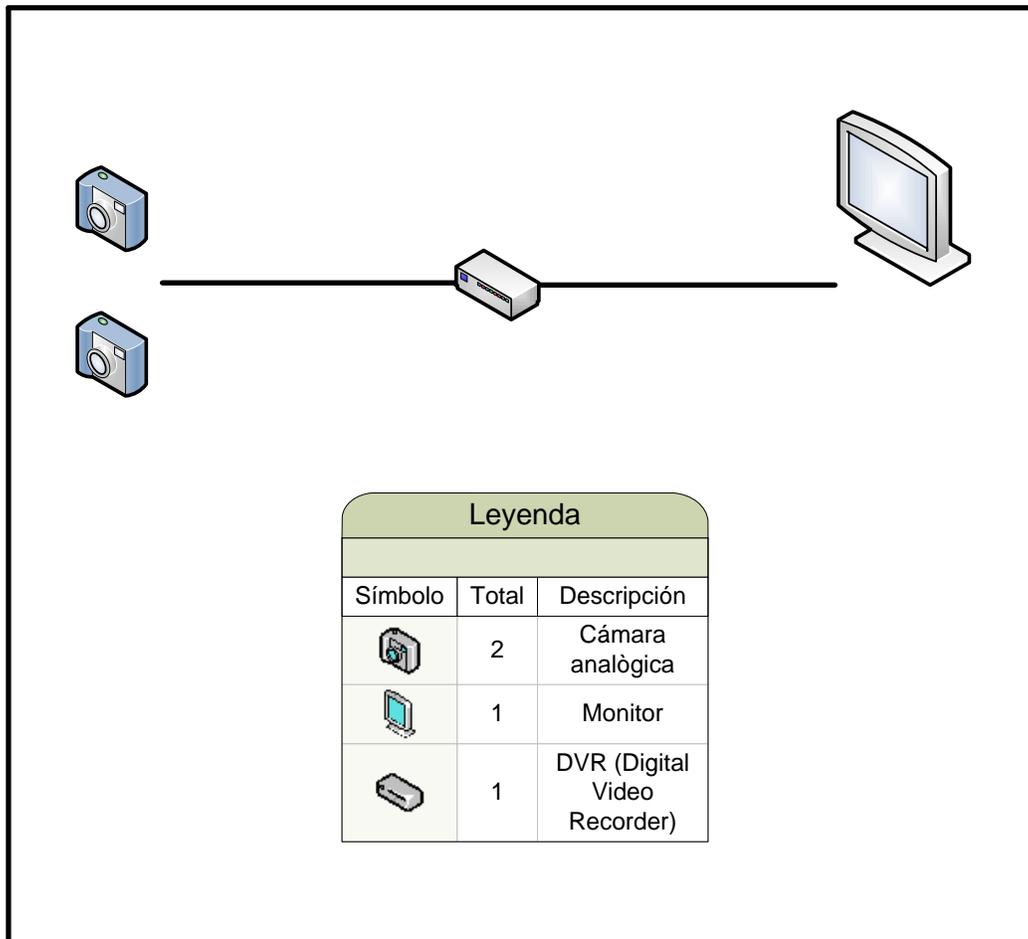


Fig. 1.2 Esquema de un sistema analógico con DVR.

En la Fig. 1.2 se muestra un sistema CCTV con DVR (Digital Video Recorder) el cual es el encargado de almacenar el video digitalizado proveniente de las cámaras analógicas. El disco duro sustituye a las cintas magnéticas y es necesario comprimir el video de manera que se logra almacenar un mayor porcentaje de secuencia de video.

1.2.1.2 Sistemas de video IP que utilizan servidores de video.

Estos sistemas a diferencia del anterior se requieren de un servidor de video, el mismo que es el encargado de digitalizar, comprimir y distribuir las secuencias de video a la



red, como se observa en la Fig. 1.3, permitiendo así un sistema más eficiente y flexible convirtiéndose en un vínculo entre los sistemas analógicos y de video vigilancia IP.

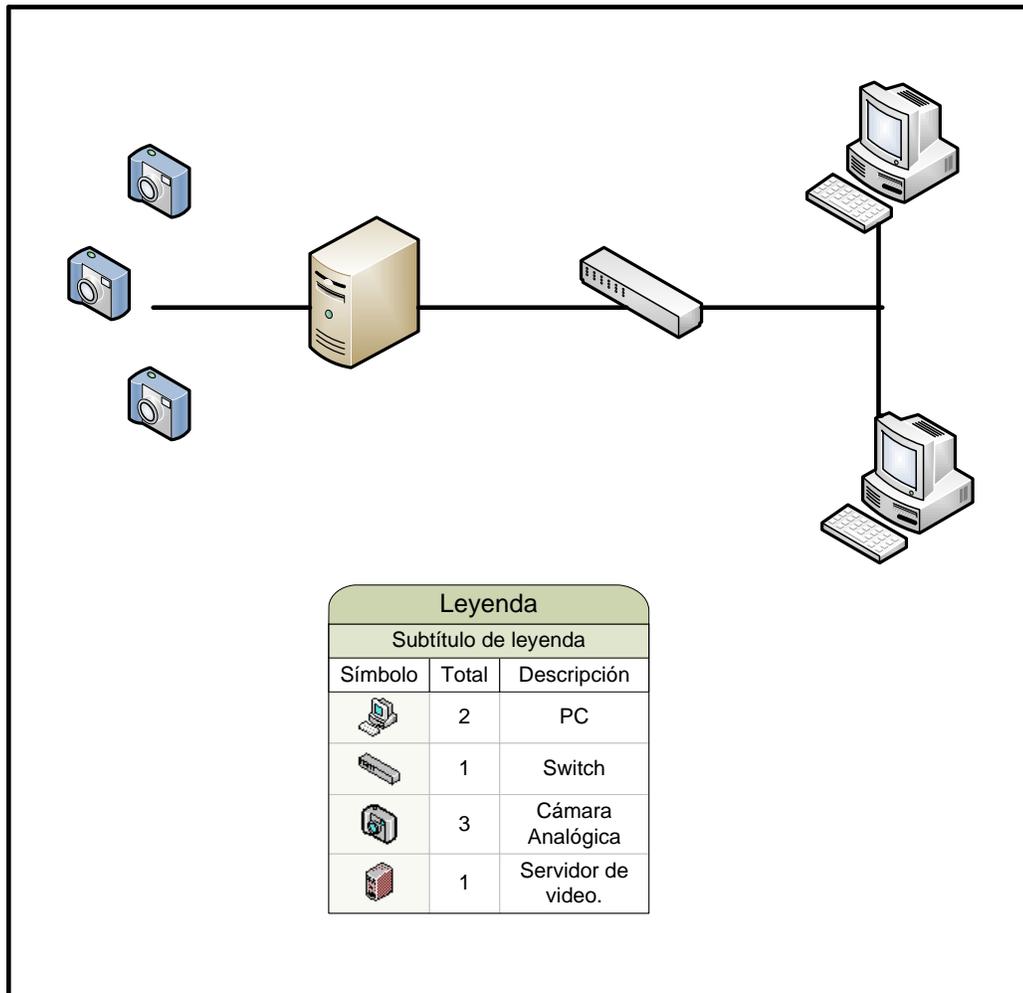


Fig. 1.3 Sistema de video vigilancia IP con servidor de video.

1.2.2 Sistemas de video vigilancia de segunda generación.

Los sistemas de vigilancia de segunda generación o avanzados, son una alternativa frente a los de primera generación; ya que son basados en técnicas de procesado de la información (2).



Lo que destaca a esta tecnología es que permite seleccionar automáticamente un pequeño porcentaje de la información disponible, lo cual permitirá realizar la tarea de vigilancia de manera más eficiente.

Mediante la explotación de las técnicas de procesado permite alcanzar objetivos importantes:

- Disminuir los costes del sistema permitiendo evaluar solo eventos importantes y descartar información no necesaria dando lugar a la eficiencia.
- Optimizar el rendimiento del operador mejorando las condiciones de trabajo liberándolo de eventos monótonos y repetitivos.

1.2.2.1 Sistemas de video vigilancia IP.

El sistema consiste en una transmisión unidireccional de video en tiempo real, como se observa en la Fig. 1.4, por este motivo la red debe ofrecer un nivel de calidad de servicio.

La información de video debe transmitirse independientemente de la red y del medio empleado. El sistema debe tener la capacidad de adaptarse a la infraestructura de la red y causar el menor impacto en ella.

En los sistemas de video vigilancia IP (protocolo de internet) se presenta una gran ventaja debido a que se puede estar informado en caso de problemas en tiempo real y desde cualquier terminal conectado a Internet.

La información obtenida puede ser encontrada, manipulada y transferida fácilmente. Además, al conectarse de manera remota permite que una o más personas puedan acceder a las imágenes capturadas y de esta manera evaluar los eventos de mayor importancia.

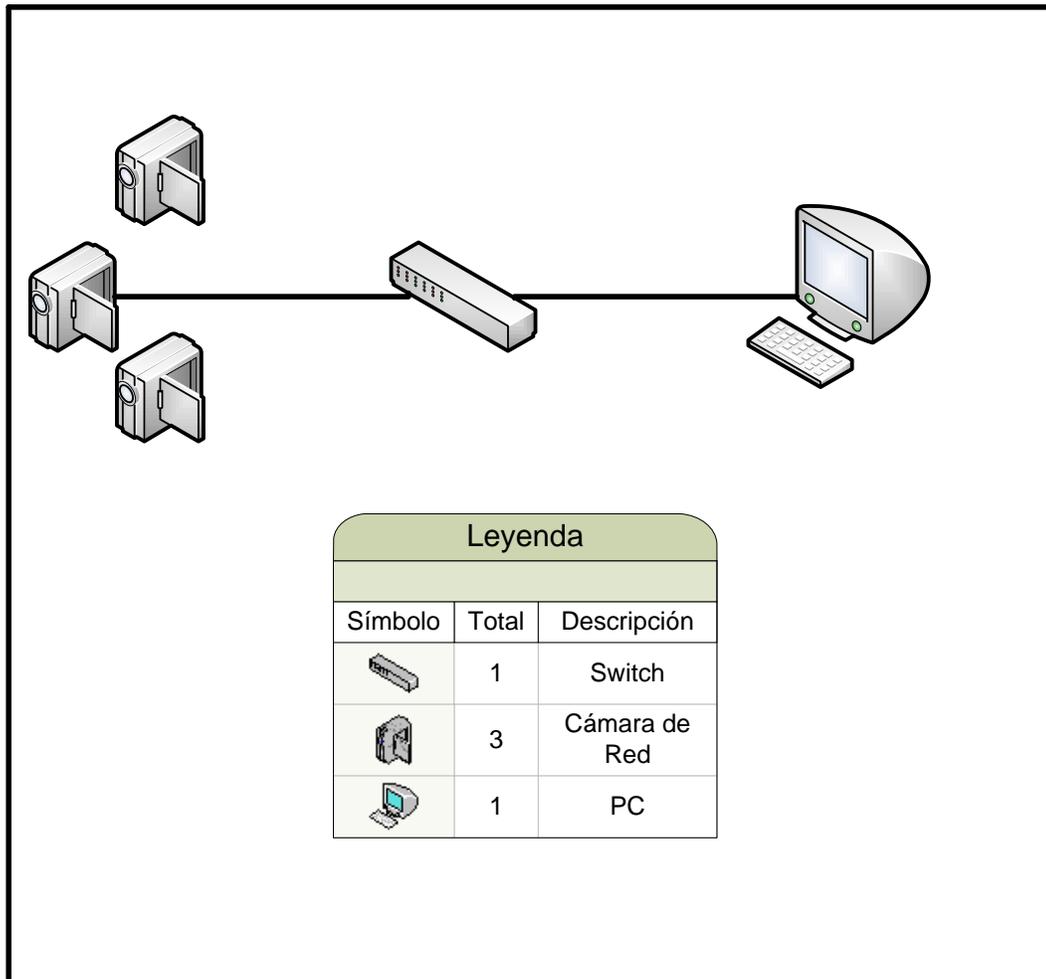


Fig. 1.4 Sistema de video vigilancia digital IP.

1.2.3 Sistemas de video vigilancia de tercera generación.

Los sistemas avanzados de vigilancia se encuentran constituidos por una amplia gama de sensores heterogéneos distribuidos a lo largo de escenario a vigilar como se observa en la Fig. 1.5

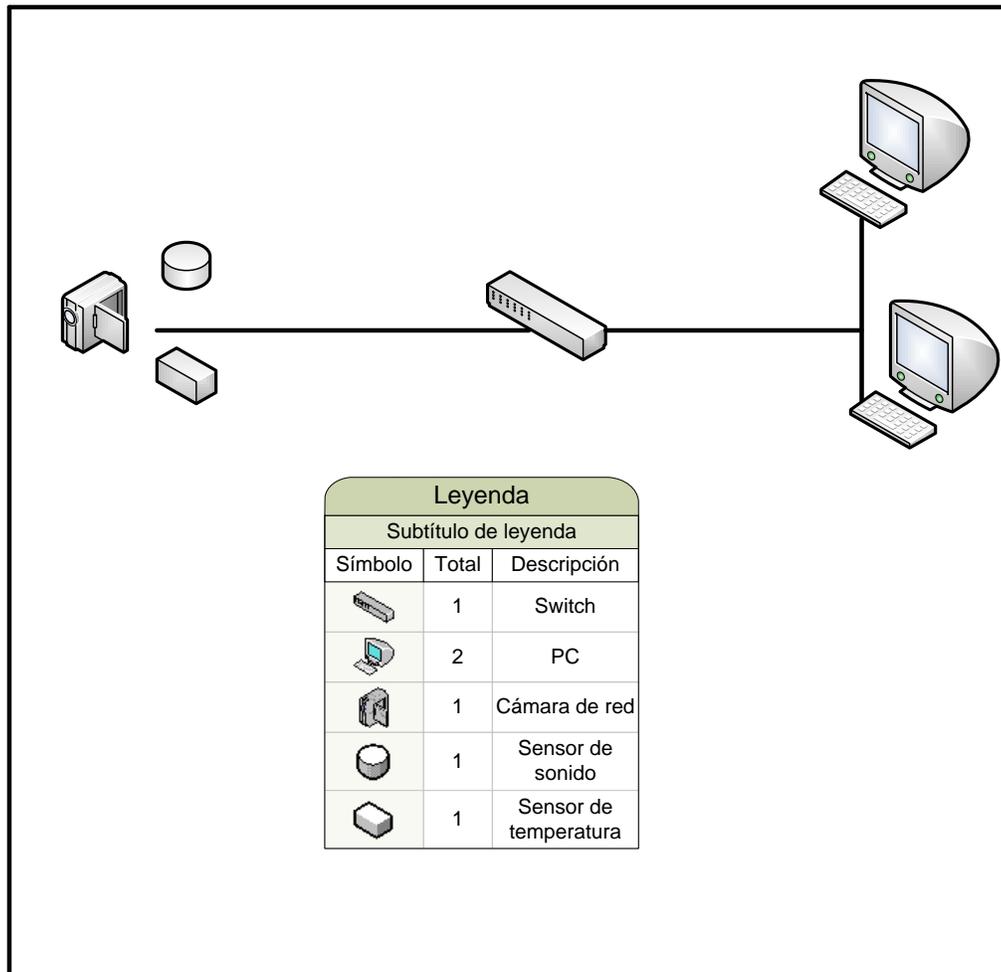


Fig. 1.5 Sistema de video vigilancia digital IP distribuido.

Además de las ventajas que proporcionan los sistemas de vigilancia de segunda generación, hay que añadir la mejora en el proceso de interpretación, ya que cuenta con información procedente de un mayor tipo de sensores.

La obtención de los datos se realiza manera descentralizada, considerando la comunicación entre los dispositivos heterogéneos y la interpretación de la información procedente de las respectivas fuentes para la activación de alarmas. (2)

1.2.3.1.1 Aplicaciones.



Los sistemas de seguridad y su implementación actualmente presentan una alta demanda. Es así que alrededor del mundo existen varios grupos de investigación que dan aporte a lo que a seguridad se refiere.

- El grupo de investigación ORETO, en su proyecto Hesperia, tiene como objeto el desarrollo de tecnologías que permitan la creación de sistemas punteros de seguridad, vídeo vigilancia y control de operaciones de infraestructuras y espacios públicos. Con lo antes mencionado el grupo pretende incrementar de forma sustancial los niveles de seguridad de grandes espacios públicos como: aeropuertos, estaciones de ferrocarril, puertos, centros de ciudades especialmente en zonas peatonales, centros comerciales, etc (4).
- El grupo de investigación ORETO, en su proyecto Argos, desarrollan modelos y técnicas de aprendizaje para expresar y verificar la normalidad de los eventos y acciones que se desarrollen en entornos abiertos (estacionamientos, pasos de peatones, etc), de manera que se procesa la información procedente del entorno y como consecuencia, se modela las desviaciones en relación con dichos comportamientos y se genera alarmas o medidas correctoras en su caso (5).
- El grupo de investigación GAVAB, en su proyecto ReActivS (Reconocimiento de actividades humanas basado en seguimiento visual), desarrollan sistemas computacionales capaces de interpretar una secuencia de imágenes y extraer de ella información veraz, útil y compacta. Se basan en métodos de reconocimiento de patrones adecuados para el desarrollo de algoritmos que permitan reconocer acciones humanas en un contexto de aplicaciones de vigilancia y de seguridad a través de la inyección de conocimiento explícito de un observador externo al sistema. Desarrollan mecanismos de aprendizaje que permite al sistema adaptarse a las condiciones del entorno para intentar detectar situaciones anómalas de potencial riesgo (6).



- El grupo de investigación GAVAB, en su proyecto V-ATRAP (Video vigilancia asistida en tiempo real mediante hardware de altas prestaciones), trata sobre la alta capacidad de cómputo que exigen ciertas aplicaciones relativas a la visión artificial, especialmente aquellas dedicadas a las tareas de bajo y medio nivel tales como la detección y seguimiento de objetos en tiempo real para la video vigilancia. A partir de esta premisa, en este proyecto se pretende encontrar soluciones eficientes y de bajo coste a dicho problema para un conjunto de escenarios diferentes (seguimiento de objetos simples, de múltiples objetos, detección de intrusos, etc) (7).
- El grupo de investigación GAVAB-CAPO, desarrolla sistemas de visión artificial que entre otras múltiples aplicaciones. Establecen medidas estadísticas en entornos de tráfico que permiten, por ejemplo, conocer con más exactitud la utilidad de ciertas infraestructuras (8).
- En la Universidad de Reading en el Reino Unido, desarrollan sistemas basados en el comportamiento humano, y detección de objetos. Este proyecto se lo ha realizado con el fin anti-terrorista y se espera que para el mes de julio del año 2012 se encuentre en óptimo funcionamiento (9).



CAPÍTULO II

DESCRIPCIÓN DEL SISTEMA DE VIDEO VIGILANCIA.

2.1 Descripción general del sistema de video vigilancia.

Se tienen cuatro fases principales del procesamiento de imagen:

- La primera etapa es la adquisición de imágenes digitales usando un dispositivo óptico (cámara IP).
- En la segunda etapa se preprocesa la imagen mediante una serie de operaciones y filtros.
- La tercera etapa, conocida como segmentación, realiza divisiones en la imagen.
- En la última etapa se clasifican los elementos segmentados a través del estudio de sus propiedades (10).

A continuación en la Fig. 2.1 se muestra el diagrama del sistema de video vigilancia.

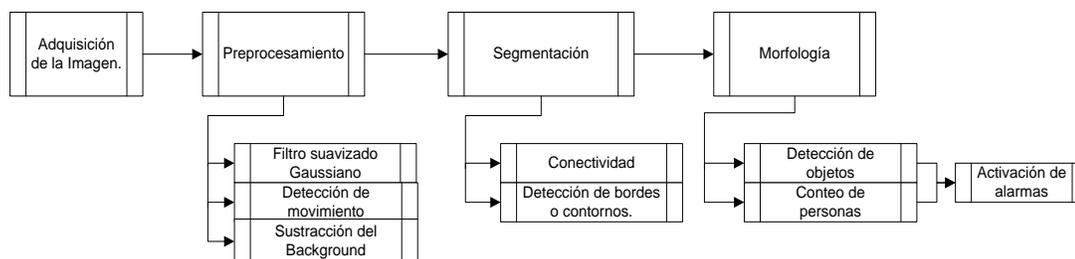


Fig. 2.1 Diagrama del sistema de video vigilancia.

2.1.1 Adquisición de la imagen.

En la fase de adquisición, los sensores captan información sobre el exterior. De esta manera se va obteniendo cada fotograma de la secuencia y es almacenada en una estructura de datos. Su aportación a la siguiente fase es una representación cuantitativa y digitalizada de la escena.



Con la instalación de un sistema de vigilancia de red, se puede monitorizar de forma local o remota la seguridad de las personas y las propiedades en cualquier lugar y en cualquier momento (11).

La IP7137 de VIVOTEK como se puede observar en la Fig. 2.2, da la posibilidad de crear y mantener un sistema de vigilancia IP efectivo y seguro, el mismo que puede monitorizarse ya sea por cable o inalámbricamente (12). Para mayor información ver Anexo1.



Fig. 2.2 Cámara de red IP7137 de VIVOTEK (13)

La integración eficiente de componentes tanto de hardware como de software para el funcionamiento de video vigilancia a través de la red, es sin duda el eje principal para obtener un sistema eficaz de video vigilancia IP.

El protocolo de comunicaciones más común en la actualidad es TCP/IP (Protocolo de control de transmisión/Protocolo de Internet) debido a su escalabilidad. En un sistema de video vigilancia la disposición de cámaras IP en diferentes escenarios a monitorear es fundamental, con lo cual independientemente del medio de transmisión a utilizarse, es decir que puede estar conectada a una red Ethernet, Red de Área Local (LAN) o LAN inalámbrica, se logra un sistema confiable y que abarca diferentes áreas a controlar.

En general para la transmisión de video se tiene consideraciones como el ancho de banda, el cual está dado por:

$$AB = \text{Tamaño del cuadro de la imagen (Kb)} * \text{Cuadros por segundo} * \text{Canal} \quad (2.1)$$



Las cámaras IP utilizan recursos de red para la transmisión de imágenes, lo que se traduce en consumo de ancho de banda. Para este sistema se ha estimado un ancho de banda de 7 Mbps, esto dependerá del tamaño de cuadro principalmente, ya que siempre se está tomando 30 fps (fotogramas por segundo) además utiliza 1 canal para video.

La secuencia de video soporta diversos protocolos sobre la red entre los principales UDP (Protocolo de Datagrama de Usuario), TCP (Protocolo de Control de Transmisión), HTTP (Protocolo de Transferencia de Hipertexto) o HTTPS (Protocolo Seguro de Transferencia de Hipertexto), además soporta vigilancia móvil a través de 3GPP (Generation Partnership Project), y compresión MPEG-4 bajo el estándar formal de ISO/IEC 14496. Las características antes mencionadas permiten que la secuencia de video tenga la capacidad de transmitirse de manera segura, garantizando que los datos sean recibidos en el mismo orden como fueron transmitidos y sin errores.

La cámara ofrece transmisión vía TCP/IP es por ello que no necesita del DVR para transmitir, ya que tiene un servidor interno el cual proporciona directamente, vía DHCP (Protocolo de configuración dinámica de host), una IP de manera automática ya sea estática o dinámica.

La red debe estar en condiciones de ofrecer conectividad a internet a través de su DMZ (Zona desmilitarizada), para dar salida con IP pública, la que ofrecerá el ISP (Proveedor de servicios de Internet) al sistema de video vigilancia y de esta manera se poder acceder remotamente desde cualquier sitio y hora.

La seguridad en este tipo de sistemas es fundamental por lo que la red debe brindar seguridad al acceso del sistema. Sin embargo, en caso de ser burlado el acceso se puede realizar la configuración de red, ya sea a nivel de usuario o por lista de acceso de redes o IP.

Los principales parámetros a configurar en la cámara IP7137 para su funcionamiento en el sistema de video vigilancia son:

- Configuración de Video



- Configuración de red
- Configuración WLAN¹

2.1.1.1 CONFIGURACIÓN VIDEO.

Como se observa en la Fig. 2.3 la configuración de video será establecida a color, el tamaño del fotograma es de 320x240, la tasa máxima de fotogramas es de 30 porque se requiere que la sucesión continua de los fotogramas se reproduzca de tal manera que cree la ilusión de movimiento.

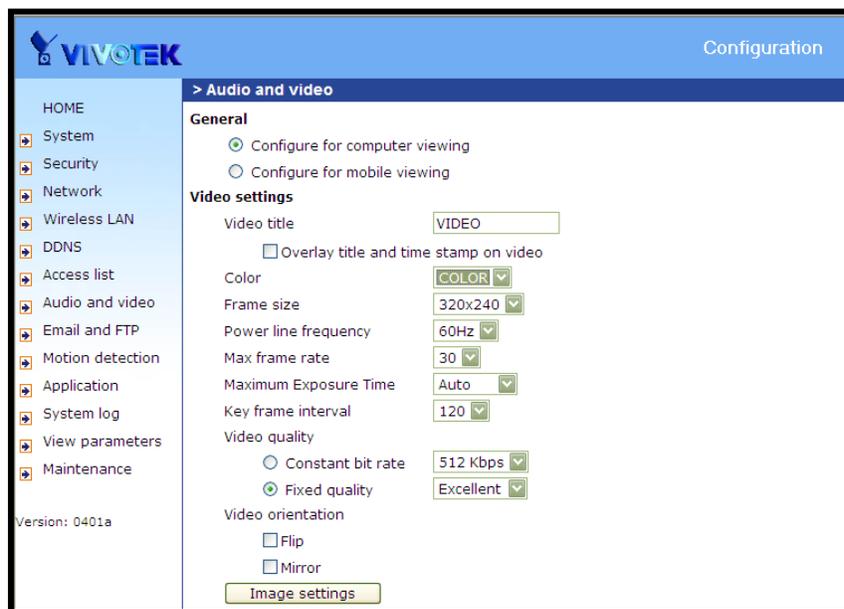


Fig. 2.3 Configuración de video.

2.1.1.2 CONFIGURACION DE RED.

La Fig. 2.4 muestra la configuración de red mediante una red de área local para lo cual se establecen parámetros como IP, la máscara de subred y la puerta de enlace por defecto.

¹ WLAN (Wireless Local Area Network) en español, red de área local inalámbrica es un sistema de comunicación de datos inalámbrico flexible que utiliza tecnologías de radiofrecuencia para mayor movilidad a los usuarios.



The screenshot shows the VIVOTEK Configuration interface. The left sidebar contains a menu with options: HOME, System, Security, Network, Wireless LAN, DDNS, Access list, Audio and video, Email and FTP, and Motion detection. The main content area is titled '> Network' and 'Network type'. Under 'Network type', there are two radio buttons: 'LAN' (selected) and 'Get IP address automatically'. Below this, there are four input fields: 'IP address' (192.168.50.1), 'Subnet mask' (255.255.255.0), 'Default router' (192.168.50.2), 'Primary DNS' (empty), and 'Secondary DNS' (empty).

Fig. 2.4 Configuración de red LAN.

También se puede observar en la Fig. 2.5 que podemos configurar listas de acceso ya sea para denegar o dar acceso a redes.

The screenshot shows the VIVOTEK Configuration interface. The left sidebar contains a menu with options: HOME, System, Security, Network, Wireless LAN, DDNS, Access list, Audio and video, Email and FTP, Motion detection, Application, System log, View parameters, and Maintenance. The main content area is titled '> Access list' and 'Access list'. Under 'Access list', there are two input fields: 'Start IP address' and 'End IP address', with an 'Add' button to the right. Below this, there is a 'Delete allow list' section with a dropdown menu showing '1.0.0.0 ~ 255.255.255.255' and a 'Delete' button. Under 'Deny list', there are two input fields: 'Start IP address' and 'End IP address', with an 'Add' button to the right. Below this, there is a 'Delete deny list' section with a dropdown menu showing '-- none --' and a 'Delete' button. The bottom left corner of the page shows 'Version: 0401a'.

Fig. 2.5 Configuración de acceso o denegación de redes.

2.1.1.3 CONFIGURACION WIRELEES LAN.

Mediante la configuración WIRELEES LAN se puede realizar el monitoreo remoto. Es necesario, como se observa en la Fig. 2.6, dar un nombre a la red mediante el SSID²

² SSID (Service Set Identifier) es un nombre incluido en todos los paquetes de una red inalámbrica para identificarlos como parte de esa red.



(Identificador de servicios), configurar el canal en este caso es el 11 viene configurado por defecto, sin embargo, se puede establecer otro canal diferente. En este caso la seguridad no está configurada, pero se puede establecer una clave de acceso en caso de ser necesario.

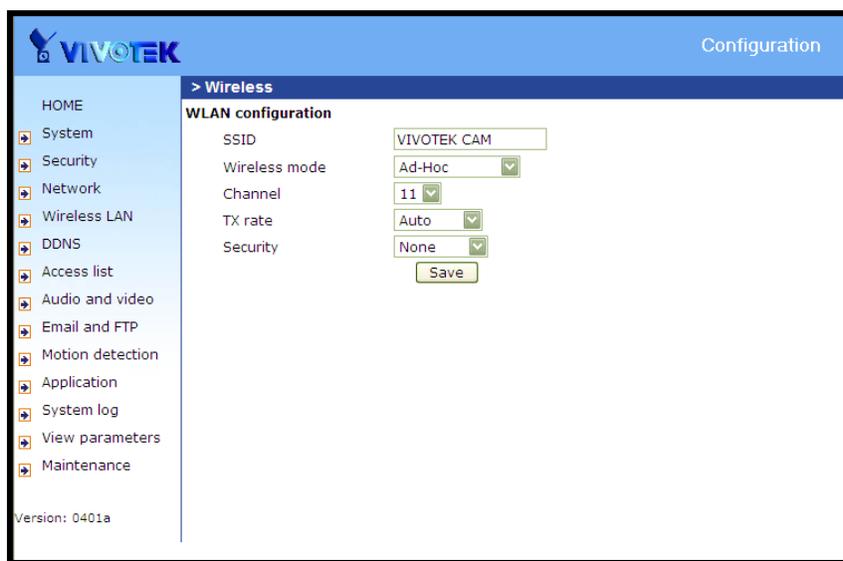


Fig. 2.6 Configuración WIRELEES LAN.

2.1.2 Preprocesamiento.

La fase de preprocesamiento o tratamiento abarca aspectos como corrección de la imagen original en una nueva, donde se hayan eliminado los problemas de ruido e iluminación espacialmente variable.

La utilización de estas técnicas permiten el mejoramiento de las imágenes digitales adquiridas de acuerdo a los objetivos planteados en el sistema.

2.1.2.1 Pixel.

Un pixel (acrónimo del inglés *picture element*, elemento de imagen), Fig. 2.7, es la unidad mínima homogénea en color que forma parte de una imagen digital (14), la cual puede ser una fotografía o un fotograma de video, el cual ofrece información que servirá posteriormente al procesamiento.



Fig. 2.7 Representación de Pixel (Picture element).

2.1.2.2 Imagen digital.

Es una representación bidimensional de una imagen utilizando bits (unos y ceros), de forma que cada elemento de la matriz le corresponde con cada píxel en la imagen, como se observa en la Fig. 2.8.

Puede tratarse de un gráfico rasterizado o de un gráfico vectorial. Los gráficos rasterizados se distinguen de los gráficos vectoriales, en que estos últimos representan una imagen a través del uso de objetos geométricos, no del simple almacenamiento del color de cada píxel a lo cual se denomina raster.

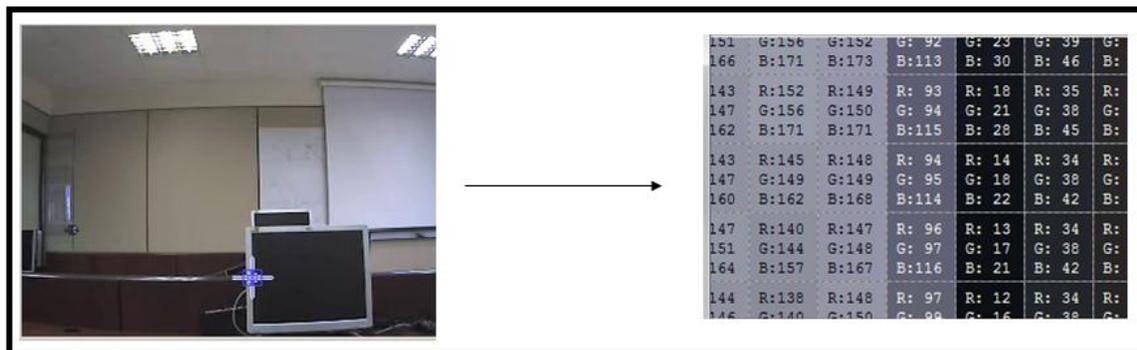


Fig. 2.8 Imagen digital.

Dependiendo del rango de valores que pueda tomar cada pixel podemos distinguir diferentes tipos de imágenes:

- Imágenes binarias: El rango está formado por los valores negro y blanco [0 1] únicamente. Ver Fig. 2.9.

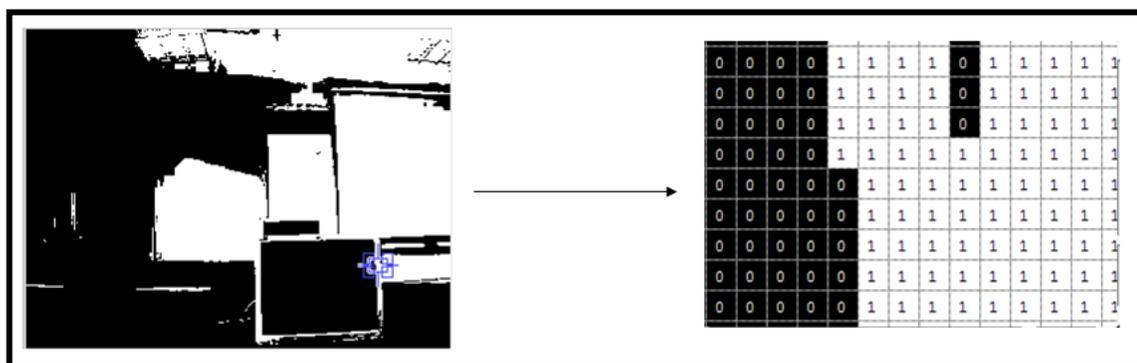


Fig. 2.9 Imagen Binaria.

- Imágenes de intensidad: Conocidas también como imágenes en escala de grises, existen hasta 256 niveles de grises, por lo que su rango está entre [0, 255]. Observar Fig. 2.10.

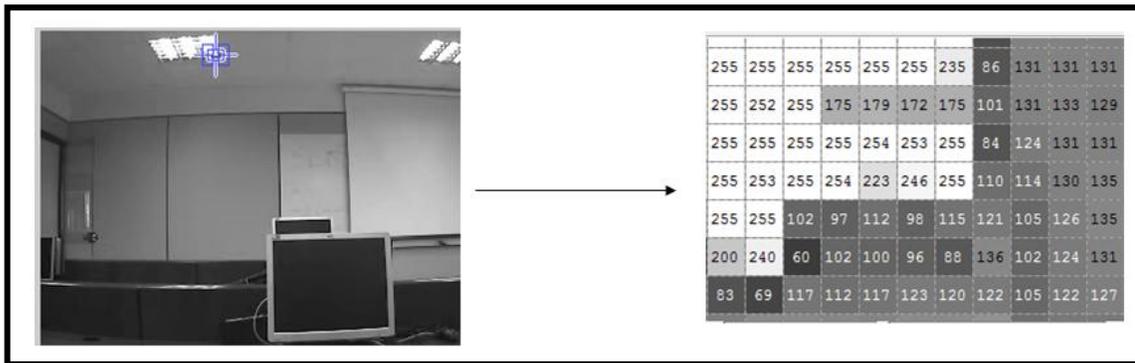


Fig. 2.10 Imagen de Intensidad.

- Imágenes de color: Las imágenes a color tiene tres componentes básicos RGB, por tanto cada píxel tiene un valor para cada componente básico del color. Observar Fig. 2.11.



Fig. 2.11 Imagen a color.

2.1.2.3 Filtro.

El propósito de realizar el proceso de filtrado conlleva ciertas técnicas que permiten extraer características para mejorar y detallar cierta información de la imagen obtenida en el proceso de captura, o inclusive minimizar efectos de ruido.

Las técnicas realizadas son transformaciones de cada uno de los píxeles, los mismos que no dependen únicamente del nivel de gris de un determinado píxel, sino también del valor de los píxeles colindantes. De esta manera, mediante la utilización de



matrices denominadas máscaras, las mismas que son aplicadas en la imagen, dan como resultado una imagen nueva.

El tipo de técnica a realizar es un factor importante ya que dependerá de la aplicación hacia la cual está orientada.

2.1.2.3.1 Filtro de suavizado gaussiano.

Un filtro de suavizado gaussiano produce imágenes donde se reduce el ruido y cambios bruscos de intensidad, lo que da como resultado que las imágenes se tornen algo borrosas. Como desventaja se tiene que posiblemente se pierdan los bordes de las imágenes, debido al filtrado (10).

La distribución gaussiana con desviación típica σ y media μ viene dado por:

$$g_{\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (2.2)$$

La desviación típica $g_{\sigma}(x)$ es un factor de gran importancia en lo que respecta al suavizado de la imagen, ya que si es mayor σ , considera los píxeles más lejanos provocando así un filtrado suavizante mayor. Caso contrario a lo que ocurriría si la desviación típica es menor dando como resultado un cambio mínimo (15).

Es por ello que la desviación típica debe ser la adecuada al ser un valor bajo al indicado, lo que provocará un mínimo filtraje. Por el contrario, si es demasiado grande, la imagen se presentará difuminada.

A continuación se presenta en Fig. 2.12, el resultado de aplicar un filtro de suavizado gaussiano en una imagen digital.

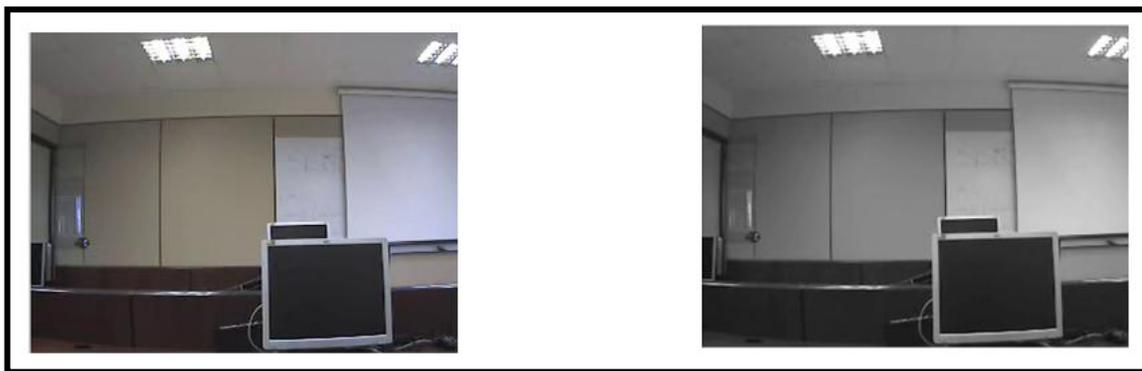


Fig. 2.12 a) Imagen original. b) Resultado de la imagen al aplicar filtro suavizado gaussiano.

2.1.2.4 Detección de movimiento.

La detección de movimiento es una de las funciones fundamentales de un sistema típico de video vigilancia. Cuando la aplicación está destinada a vigilar la actividad dentro de un espacio determinado y el sistema de sensores de captación de imágenes es estático, es decir, está montado sobre estructuras inmóviles, bajo un fondo fijo. La actividad que puede violar las condiciones de seguridad de la zona son los elementos que están en movimiento.

Por tanto, la zona de no interés para el sistema forma la parte inmóvil. En cambio, los elementos nuevos dentro de la escena, que serán normalmente los de la región de interés, están en movimiento. Esto permite reconocer objetos móviles, en este caso, personas.

Es así que el sistema puede determinar en todo momento si la actividad dentro de la escena vigilada es de interés o no.

El método más simple para la detección de movimiento se basa en la diferencia entre imágenes. Por un lado esta técnica resulta poco precisa pero vale recalcar que el tiempo de respuesta es más rápido, lo que es importante en sistemas de tiempo real.

El resultado de esta técnica es una secuencia de imágenes que se obtienen a partir de la diferenciación realizada en el proceso de captura. Esto se puede observar en la Fig. 2.13.



Si bien es cierto en este método los cambios de iluminación no son tan pronunciados. Sin embargo, se puede obtener una imagen del fondo como referencia de esta manera se obtiene el objeto en movimiento y no las variaciones del fondo de la imagen con respecto al objeto. Esta técnica se denomina sustracción de fondo.

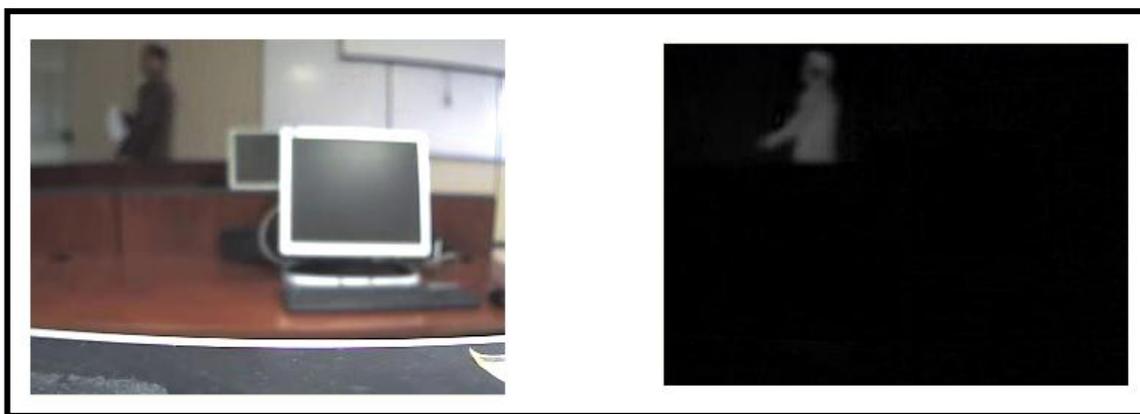


Fig. 2.13 Detección movimiento

2.1.2.5 Sustracción del background.³

La técnica de sustracción de background es una técnica más elaborada y presenta mejores resultados cuando el sistema trabaja en condiciones adversas. Como puede ser imágenes con ruido o movimiento de objetos del fondo.

Consiste en generar una estimación del fondo a partir de las imágenes anteriores, es decir procesamiento de video en base a la separación de la parte variable de la imagen de la parte fija. Por lo tanto, permite adquirir la parte de información que emite la parte variable para lograr una respuesta a dicho evento, lo que se conoce como alarmas.

De esta forma, el sistema puede aprender el comportamiento de la escena en condiciones normales y determinar la presencia de objetos de interés mediante decisiones más complejas. A estas técnicas se las denominan como algoritmos de

³ BACKGROUND: en relación a una imagen constituye lo que representa el segundo plano, o fondo.



fondo. A continuación en la Fig. 2.14 se muestra el resultado de aplicar un algoritmo mediante sustracción de background.



Fig. 2.14 Sustracción de background.

2.1.2.5.1 Método de imagen de referencia.

Este tipo de método probablemente es uno de los más sencillos de implementar debido a que su función radica en la detección de movimiento en base a la comparación de la imagen actual con la imagen de referencia (1).

La intensidad es un factor importante. En ambientes controlados se puede solucionar de manera casi imperceptible diferenciando un fondo de intensidad homogénea con uno de intensidad similar y de las mismas dimensiones. Para lograr esto se requiere una captura previa del fondo donde no existen objetos de primer plano.

Bajo este concepto este tipo de algoritmos presentan un problema considerable ya que al momento de efectuarse la umbralización, el umbral obtenido de la diferencia del valor absoluto de la imagen de referencia con la actual será el mismo. Por consiguiente, no es muy útil en una secuencia de video en el que se encuentra presente el ruido y por ende no es el mismo a lo largo de toda la imagen.

De manera que es aplicable en entornos donde los objetos son fijos, o donde las variaciones del nivel de grises en los píxeles son básicamente debidas al sensor de captación de imagen. Sin embargo, en zonas donde existan objetos móviles, la variación de intensidad puede ser debido al ambiente, es decir, la hora del día, lugares



donde existen presencia de espejos, vidrios e inclusive el agua, por nombrar algunas; las cuales serán mayores que las introducidas por el sensor de captura.

2.1.2.5.2 Método estadístico.

Los métodos estadísticos sugieren un umbral dependiente del nivel de variación de cada pixel de la imagen, por ello, este método con relación al de imagen de referencia, dispone de mayor cantidad de información del fondo.

El módulo de detección de movimiento sería más complejo, debido a que el inconveniente de este tipo de sistemas es el tiempo de cómputo para obtener los datos estadísticos necesarios, caso contrario a lo que ocurre al tener una imagen de referencia fija pero que no permite trabajar en tiempo real (1).

Por otro lado, para poder calcular los datos estadísticos de forma precisa sería necesario que la escena estuviera libre de objetos de primer plano, al menos durante los primeros instantes de funcionamiento del sistema.

Además, en otros casos no resulta suficiente contar con un fondo fijo, es por ello que se requiere actualizar el fondo a lo largo del tiempo. Para ello se utiliza la mediana como función estadística, lo que permite adquirir un fondo en base a los pixeles que han cambiado en el proceso de adquisición de imagen a través de la captura de fotogramas fotograma por segundo. De esta manera se lograr tener la información de la parte variable de la imagen en la secuencia de video.

2.1.3 Segmentación.

La segmentación es el proceso de dividir una imagen digital en varias regiones homogéneas en base a una determinada característica. De esta manera cada uno de los pixeles que forman parte de elementos visuales similares serán parte de una región, logrando así que la imagen resultante sea más interesante para el análisis.

En segmentación de imágenes se deben tomar en cuenta tres conceptos básicos:



- **Similitud:** Los píxeles que formarán parte de un grupo deberán tener características similares bajo una determinada propiedad (nivel de gris, borde, textura, etc)
- **Conectividad:** Los objetos forman parte de un área de píxel con conectividad, las regiones son continuas de píxeles.
- **Discontinuidad:** Permite delimitar objetos por lo que se diferencian unos objetos de otros.

2.1.3.1 Vecindad entre píxeles.

Un píxel p en las coordenadas (x, y) tiene cuatro vecinos horizontales y cuatro verticales, cuyas coordenadas están dadas por: (10)

$$(x + 1, y), (x - 1, y), (x, y + 1), (x, y - 1) \quad (2.3)$$

Este grupo de píxeles se nota como $N_4(p)$.

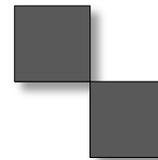
Por otro lado, las vecindades diagonales con el punto (x, y) vienen dadas por $ND(p)$, y sus coordenadas son:

$$(x + 1, y + 1), (x + 1, y - 1), (x - 1, y + 1), (x - 1, y - 1) \quad (2.4)$$

Para definir de forma adecuada el concepto de vecindad es necesario revisar el de adyacencia. Dos píxeles son adyacentes si, y solo si, tienen en común una de sus fronteras, o al menos una de sus esquinas. La Fig. 2.15, muestra píxeles adyacentes.



(a) Adyacentes por frontera.



(b) Adyacentes por esquina.

Fig. 2.15 Píxeles Adyacentes.

Dos píxeles son vecinos si cumplen con la definición de adyacencia. Si los píxeles comparten una de sus fronteras, se dice que los mismos son vecinos directos; si sólo se tocan en una de sus esquinas, se denominan vecinos indirectos.



Para ilustrar las definiciones anteriores, se puede observar, en la Fig. 2.16, vecindades de 4 y vecindades de 8. La primera formada por pixeles que son vecinos directos, mientras que la vecindad de 8 está formada tanto por vecinos directos como por indirectos.



Fig. 2.16 Vecindades

2.1.3.2 Conectividad.

La conectividad entre pixeles es un concepto utilizado para establecer los límites en objetos y regiones de componentes en una imagen. Para establecer la conectividad entre dos pixeles es necesario determinar si son adyacentes en sentido específico (si tiene 4 vecindades) y si su nivel de gris satisface un criterio especificado de similitud (si son iguales) (10).

Por ejemplo, en una imagen binaria con valores 0 y 1, dos pixeles pueden tener vecindad de 4, pero sólo se consideran conectados si tienen el mismo valor.

De tal manera en una matriz V se almacena los niveles de grises bajo un rango específico o predeterminado, que serán utilizados para definir la conectividad en la imagen:

A continuación se mencionará dos tipos de conectividad:

- a) **Conectividad 4.** Dos pixeles, p y q , con valores de V , están conectados si q pertenece a $N_4(p)$.
- b) **Conectividad 8.** Dos pixeles, p y q , con valores de V , están conectados si q pertenece a $N_8(p)$.



2.1.3.3 Detección de Contornos.

En las técnicas de segmentación aplicando detección de bordes en una imagen digital se pueden definir como transiciones entre dos regiones de niveles de gris significativamente distintos, detectando así las regiones donde existe cambio en la intensidad o bordes. Por lo que es un aporte de información sobre las fronteras de los objetos y puede ser utilizada para segmentar la imagen, reconocer objetos, etc (16).

Existen diferentes técnicas para detectar bordes, la mayoría de ellas emplean métodos con el uso de operadores.

2.1.3.3.1 Operadores basados en la primera derivada (Gradiente).

La derivada de una señal continua proporciona las variaciones con respecto a la variable de forma que el valor de la derivada es mayor cuanto más rápidas son las variaciones.

En el caso de funciones bidimensionales $f(x,y)$, la derivada es un vector que apunta en la dirección de la máxima variación de $f(x,y)$ y cuyo módulo es proporcional a dicha variación. Este vector se denomina gradiente y se define:

$$\nabla f(x,y) = \begin{bmatrix} \frac{\partial f(x,y)}{\partial x} \\ \frac{\partial f(x,y)}{\partial y} \end{bmatrix} \quad (2.5)$$

$$\text{Mag} [\nabla f(x,y)] = \sqrt{\left(\frac{\partial f(x,y)}{\partial x}\right)^2 + \left(\frac{\partial f(x,y)}{\partial y}\right)^2} \quad (2.6)$$

$$\theta = \text{arctag} \frac{\frac{\partial f(x,y)}{\partial x}}{\frac{\partial f(x,y)}{\partial y}} \quad (2.6)$$

En el caso bidimensional discreto, las distintas aproximaciones del operador gradiente se basan en diferencias entre los niveles de grises de la imagen. La derivada parcial $f_x(x,y)$ gradiente de fila $G_f(i,j)$ puede aproximarse por la diferencia de pixeles adyacentes de la misma fila.

$$\frac{\partial f(x,y)}{\partial x} \approx \nabla_x f(x,y) = f(x,y) - f(x-1,y), \quad \{-1 \ 1\} \quad (2.7)$$

La discretización del vector gradiente en el eje $G_c(i,j)$, será para columna:



$$\frac{\partial f(x,y)}{\partial y} \approx \nabla_y f(x,y) = f(x,y) - f(x,y-1), \quad \{-1 \ 1\} \quad (2.8)$$

El gradiente de la fila G_f y de la columna G_c en cada punto se obtiene la convolución de la imagen con las máscaras H_f y H_c

$$G_f(i,j) = F(i,j) \otimes H_f(i,j) \quad (2.9)$$

$$G_c(i,j) = F(i,j) \otimes H_c(i,j) \quad (2.10)$$

La magnitud y orientación del vector gradiente suele aproximarse por la expresión:

$$|G_c(i,j)| = \sqrt{G_f^2 + G_c^2} \approx |G_f(i,j)| + |G_c(i,j)| \quad (2.11)$$

2.1.3.3.1.1 Operador de Roberts.

La principal característica de estos operadores radica en su buena respuesta en bordes diagonales (17), y a su localización, por otro lado presenta inconvenientes de extrema sensibilidad al ruido y por ende detecta deficientemente.

Los resultados de este operador se puede observar en la Fig. 2.17.



Fig. 2.17 Detección de contornos mediante algoritmo de Roberts.

2.1.3.3.1.2 Operadores de Prewitt y Sobel.

En el operador Prewitt se involucran a los vecinos de filas y columnas adyacentes para proporcionar mayor inmunidad al ruido. (17). Se muestra en la Fig. 2.18 resultados del operador de Prewitt.



Fig. 2.18 Detección de contornos mediante algoritmo de Prewitt.

El operador Sobel se supone que es más sensible a los bordes diagonales que el de Prewitt, y más inmune al ruido además le otorga más importancia espacial a píxeles cercanos al centro de la máscara. (17). En la Fig. 2.19 se muestran los resultados tras evaluar el operador Sobel.



Fig. 2.19 Detección de contornos mediante algoritmo de Sobel.

2.1.4 Morfología.

La morfología matemática es una herramienta muy utilizada en el procesamiento de imágenes como una herramienta para extraer componentes de una imagen, así como también para resaltar características importantes y eliminar aspectos irrelevantes (10), como se observa en la Fig. 2.20.

De esta manera la morfología matemática es útil en procesamiento de imágenes no solo para la reducción de ruido, sino también para extraer elementos de la imagen que son útiles en la representación y descripción de formas, los que llamaremos funciones estructurales.



2.1.4.1 Elementos del procesado morfológico.

El procesado morfológico y su análisis se basa fundamentalmente en el álgebra de conjuntos. Se destacan tres elementos principales:

- Conjuntos.
- Elementos estructurales.
- Operadores morfológicos (dilatación, erosión, apertura y cierre).

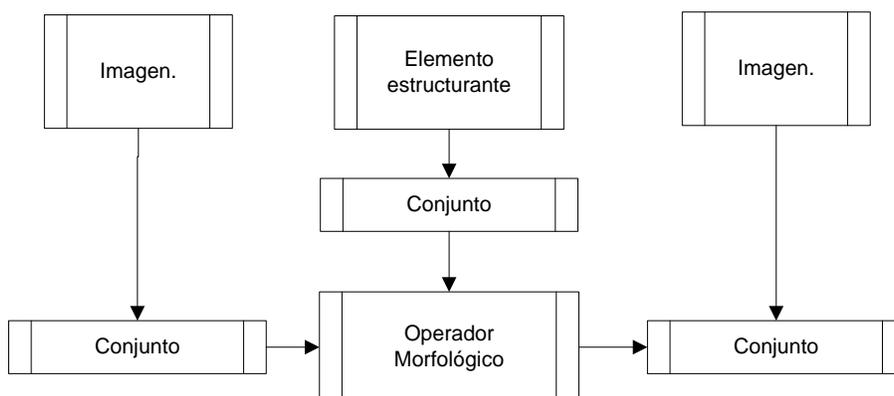


Fig. 2.20 Esquema del procesado morfológico.

2.1.4.1.1 Conjuntos.

Cada conjunto representa las formas de los objetos en una imagen.

En las imágenes binarias cada conjunto de todos los píxeles negros y, blancos son una descripción completa de la imagen: Si son blancos su valor es 1 y si son negros su valor es 0.

En las imágenes monocromáticas, una imagen de niveles de grises puede ser representada como conjuntos cuyas componentes se encuentran en un espacio 3D. En este caso, dos componentes de cada elemento de un conjunto se refiere a las coordenadas del píxel, y la tercera componente está relacionada con la intensidad.

2.1.4.1.2 Elementos estructurantes.

Elemento estructurante es un patrón que permite sondear la estructura de una imagen. La forma del elemento estructurante como el tamaño permite determinar regiones bajo



ciertas características, es decir el EE (elemento estructurante), puede ser de cualquier forma. (10)

2.1.4.1.3 Operadores morfológicos.

Operadores morfológicos: Permiten realizar cualquier transformación a la imagen cuyo resultado es otra imagen con características deseables (10).

2.1.4.1.3.1 Dilatación.

La dilatación permite añadir o rellenar los puntos del fondo que a su vez son parte del borde de un determinado objeto. Por lo que hace un barrido tomando en cuenta el fondo y los objetos o estructuras geométricas dentro de la imagen.

2.1.4.1.3.2 Erosión.

La erosión reduce el tamaño del objeto. La principal ventaja es que elimina todos los grupos de píxeles que no son parte del elemento estructurante, de tal manera que la erosión es principalmente importante en aplicaciones donde se requiere la eliminación de detalles irrelevantes. Esto se consigue de tal manera que el elemento estructurante sea de un tamaño menor al tamaño de los objetos que queremos extraer.

2.1.4.1.3.3 Apertura.

La apertura al igual que la erosión elimina los píxeles que no son parte del elemento estructurante. Además, permite alisar los contornos de los objetos geométricos dando lugar al realce de la imagen y a la separación entre objetos donde los puntos son angostos.

2.1.4.1.3.4 Cierre.

Entre las principales características de este operador morfológico se encuentra el relleno de los contornos del objeto dentro de la imagen y en relación al fondo. Además, elimina las minucias y permite unir objetos con otros. Estas funciones en ciertas aplicaciones suele causar ciertos inconvenientes en los resultados.



CAPÍTULO III

DESCRIPCIÓN DEL SOFTWARE Y RESULTADOS.

3.1 Modelo de desarrollo del software.

El modelo es de tipo secuencial el cual está formado por una serie de iteraciones las cuales son una respuesta para cada una de las etapas. Esto permitirá determinar el bloque funcional completo de la aplicación, como podemos observar en la Fig. 3.1

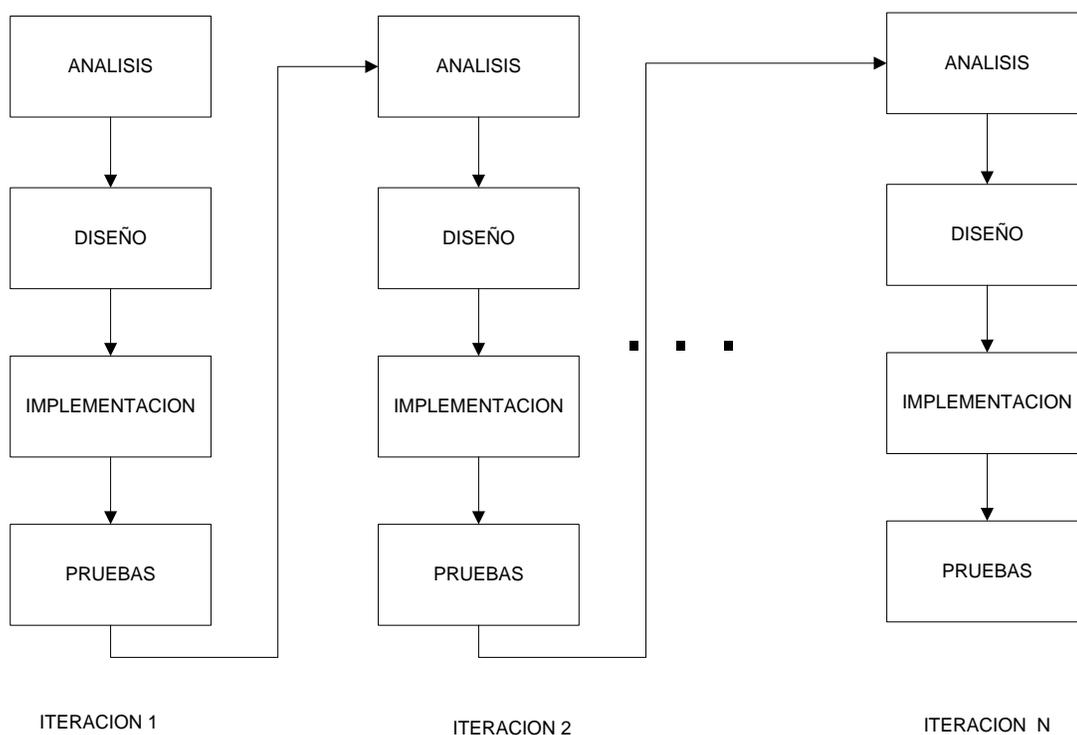


Fig. 3.1 Esquema del modelo de software.

3.1.1 Requisitos funcionales.

Los requisitos funcionales son características requeridas del sistema y determinan la funcionalidad. A continuación se presentan los requisitos funcionales para el proyecto:



- El sistema debe hacer un conteo de personas.
- El sistema debe detectar movimiento.
- El sistema debe detectar objetos abandonados y emitir señales de alarma en caso de cumplirse el estado de abandono.

3.1.2 Requisitos no funcionales.

Los requisitos no funcionales son características requeridas del sistema, las mismas que determinan una restricción, pero no una funcionalidad. A continuación se detallan los requisitos no funcionales.

- Es necesario tener una iluminación moderada.
- Es necesario que la cámara se encuentre en estado fijo.
- Todas las herramientas para la implementación del programa deben ser compatibles.

3.2 Resultados obtenidos y análisis.

Esta sección está dedicada a describir y presentar los resultados obtenidos en las pruebas realizadas sobre la aplicación.

3.2.1 Algoritmo de detección de personas.

Para obtener el requerimiento de conteo de personas se propone el siguiente diagrama de flujo:

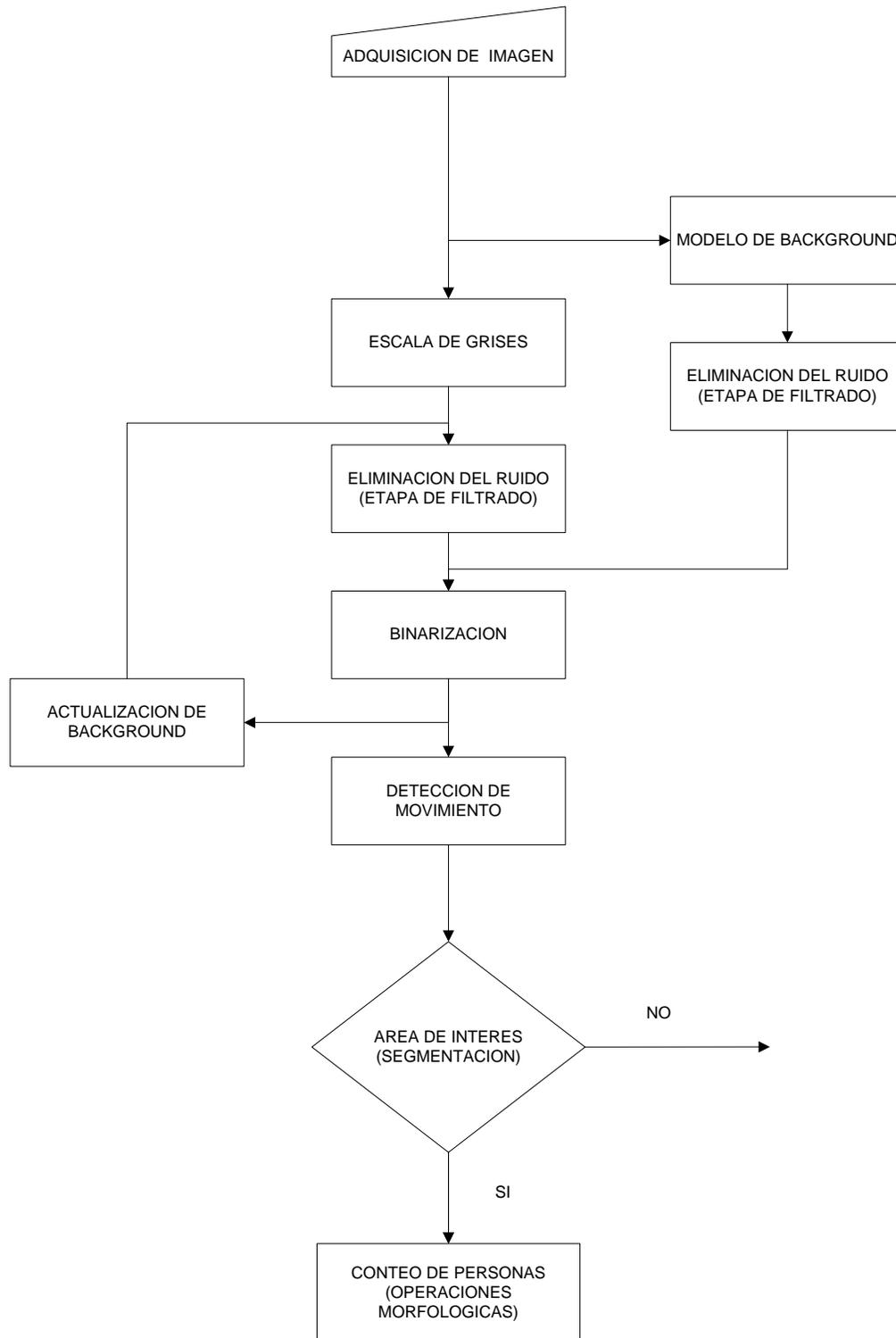


Fig. 3.2 Diagrama de flujo del algoritmo de detección de personas.



3.2.1.1 Escena 1.- Conteo de personas.

Las pruebas se realizaron tomando en cuenta la distancia de las personas. Como se puede observar en la Fig. 3.3 una persona es detectada aproximadamente a una distancia de 2 metros de la cámara, dando una detección de movimiento de manera que el sistema determina la región donde se produce el cambio dentro de la escena, haciendo detección de bordes de la zona de interés.

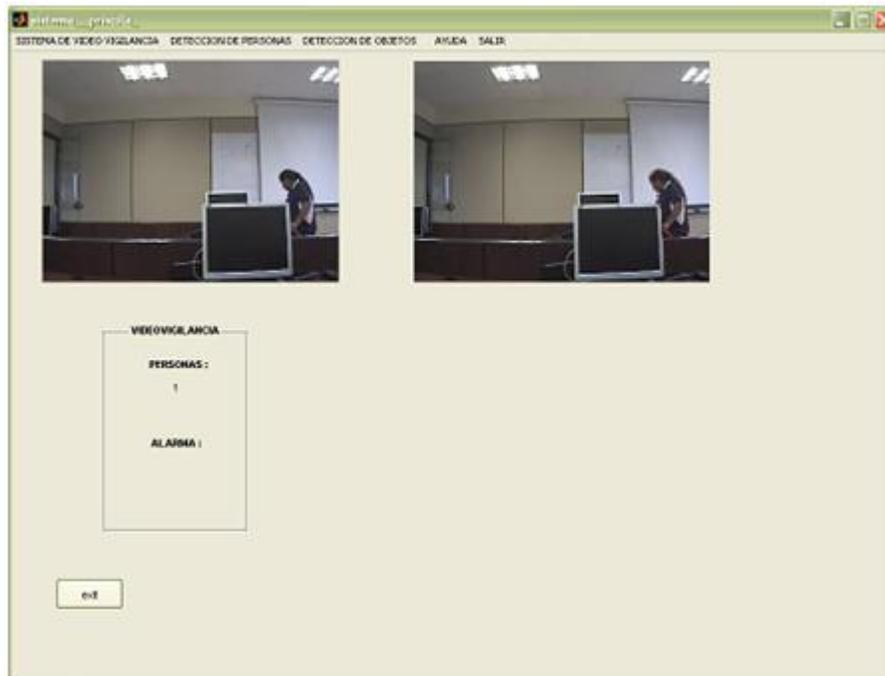


Fig. 3.3 Individuo a 3 metros de la cámara.

A continuación se puede observar en la la Fig. 3.4 que el sistema realiza el conteo de personas a través de la detección de movimiento a una distancia aproximada de 1 m.

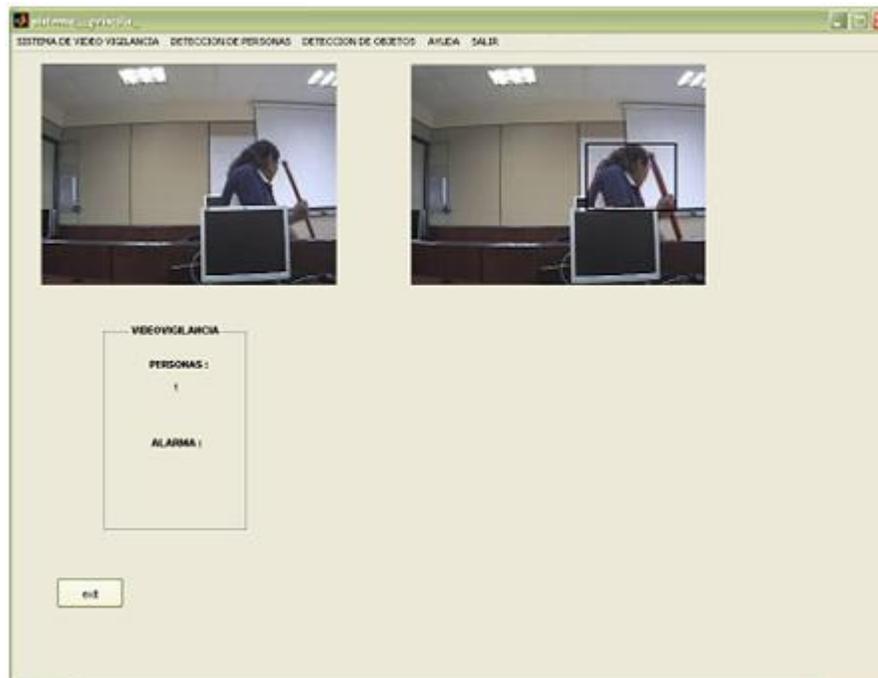


Fig. 3.4 Individuo a 1 metro de la cámara.

Se puede observar en la Fig. 3.5, que el sistema cubre todo el fotograma de manera que no existe pérdida de información en los bordes de la imagen, por lo tanto, esta procesando todo el evento y detectando movimiento en la parte derecha e izquierda de la imagen como se observa en la Fig. 3.5 (a) y Fig. 3.5 (b), respectivamente de manera que realiza el conteo.

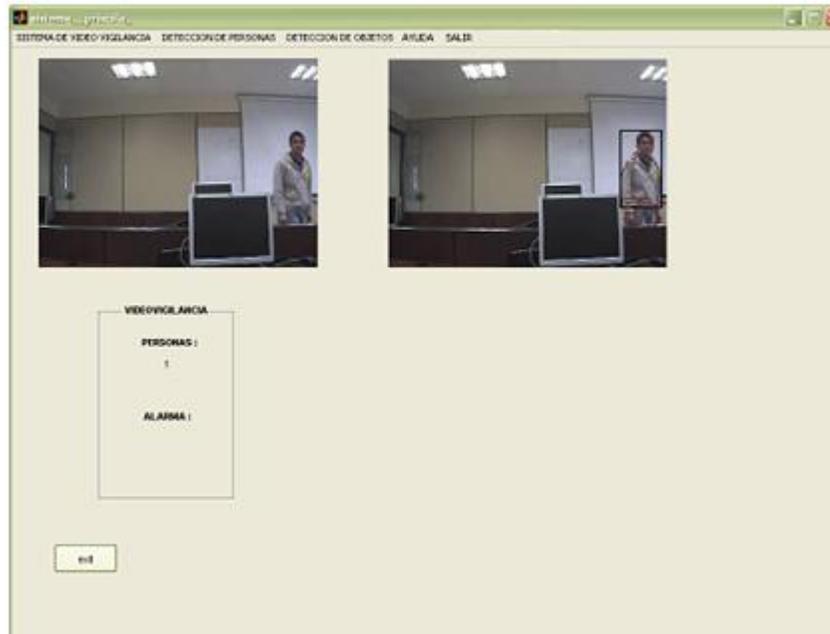


Fig. 3.5 (a) Individuos captados en el borde derecho de la imagen.

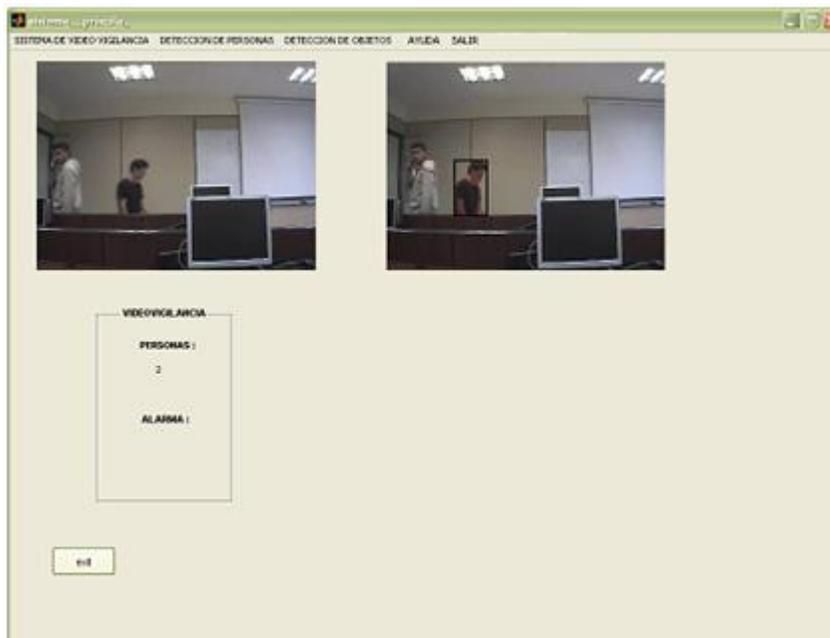


Fig. 3.5 (b) Individuos captados en borde izquierdo de la imagen.



El sistema detecta en varias regiones de la imagen los cambios que se producen, de manera que procede a realizar el conteo de personas. Como se observa en la Fig. 3.6. El sistema se encuentra realizando la detección a una distancia de 2 metros de la cámara IP. Sin afectarle en este caso la iluminación, de manera que en esta escena los cambios de luminosidad son casi imperceptibles para el sistema.

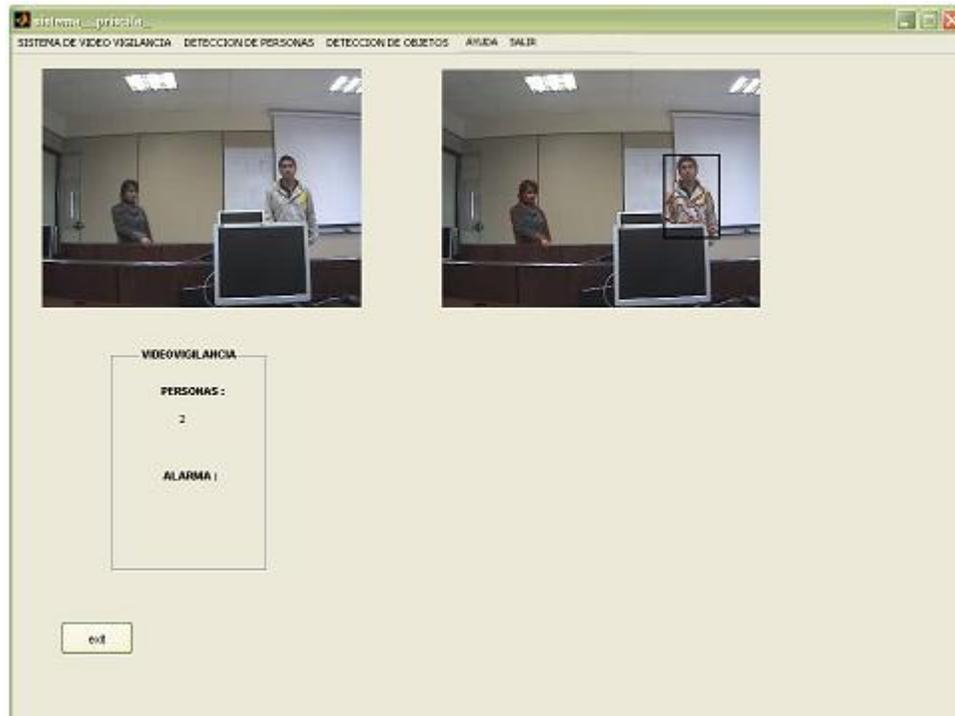


Fig. 3.6 Individuos a 3 metros de la cámara.

En la Fig. 3.7 se da un falso positivo de tal forma que al juntarse en este caso dos personas, el sistema la detecta como una sola debido a la mínima distancia que existe entre ellas.

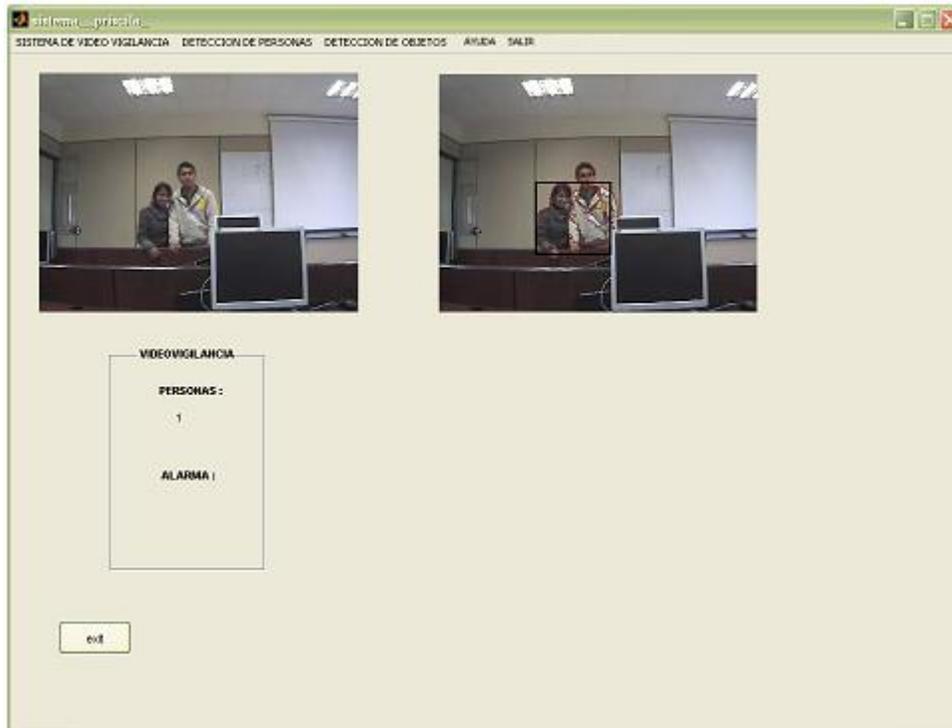


Fig. 3.7 Individuos a mínima distancia entre píxeles.

En la Fig. 3.8 se muestra el comportamiento del sistema al producirse un cambio en los eventos. Por lo tanto, el sistema aplica el módulo de actualización de background, detección de movimiento, operaciones de segmentación y morfología dando como resultado el conteo de personas.

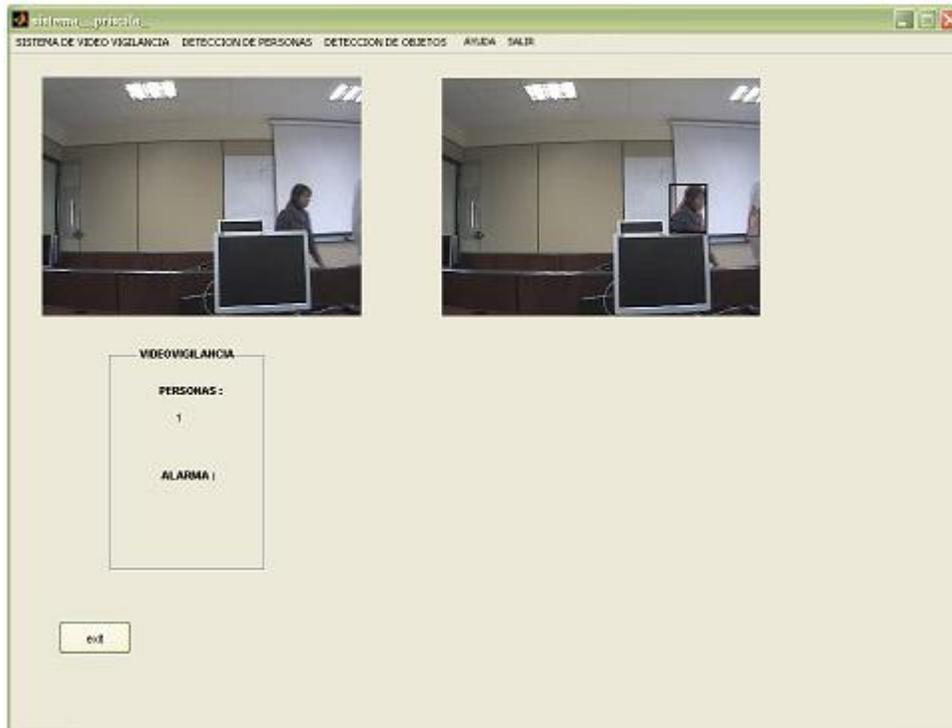


Fig. 3.8 Actualización de background, detección de movimiento, segmentación y técnicas de morfología.

En la siguiente escena se puede observar cómo, al cambiar la iluminación, el sistema se ve afectado debido al umbral. Estos cambios pueden ser debido a la presencia de espejos, vidrios etc. En este caso se procedió a cambiar la iluminación en modo apagado, por lo cual el sistema sufrió cambios al realizar el procesamiento como observamos en la Fig. 3.9.

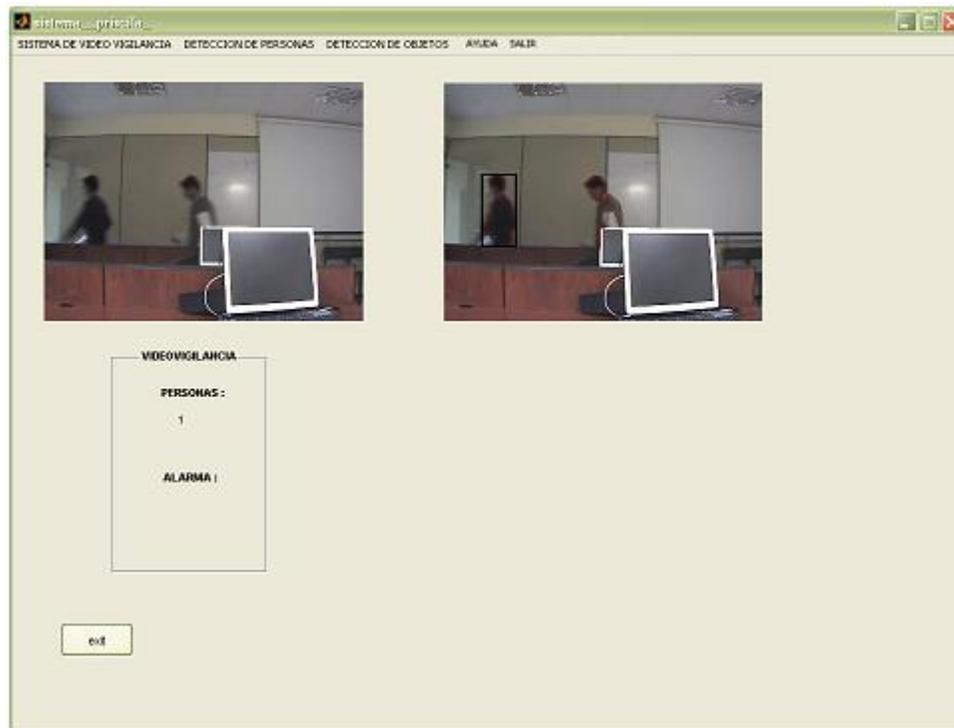


Fig. 3.9 Cambio de iluminación en la escena.

Al cambiar el modo de encendido, el sistema sufre cambios debido a la iluminación como se observa en la Fig. 3.10. Por lo tanto, el sistema debe permanecer en reposo al menos un tiempo mínimo de 10 segundos antes de empezar el procesamiento.

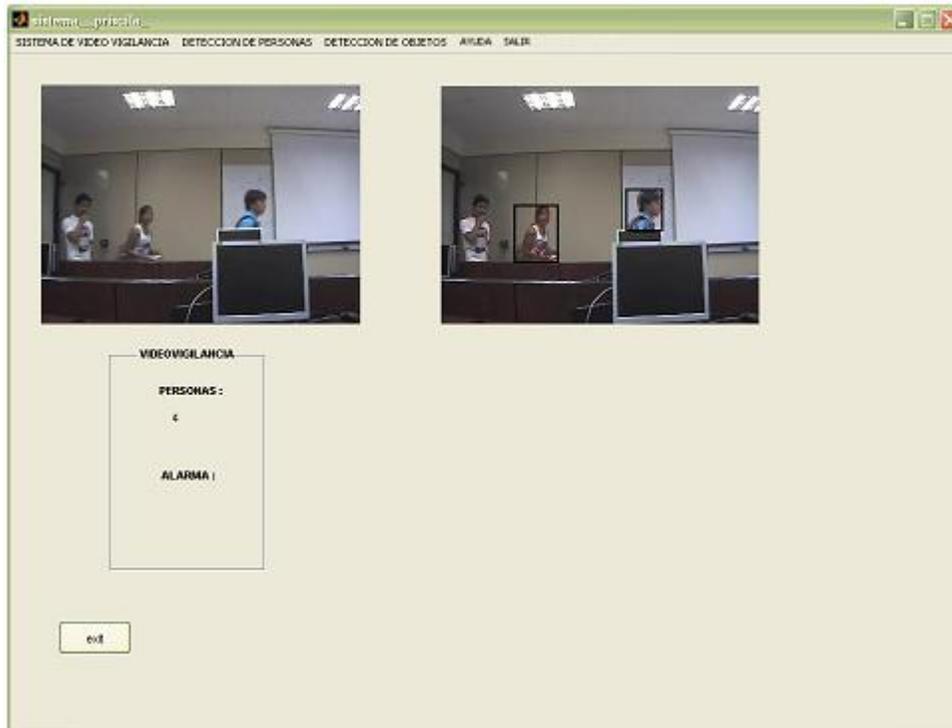


Fig. 3.10 Cambio de iluminación en el entorno.

3.2.2 Algoritmo para la detección de objetos abandonados.

Para obtener el requerimiento de objetos abandonados se plantea el siguiente diagrama de flujo:

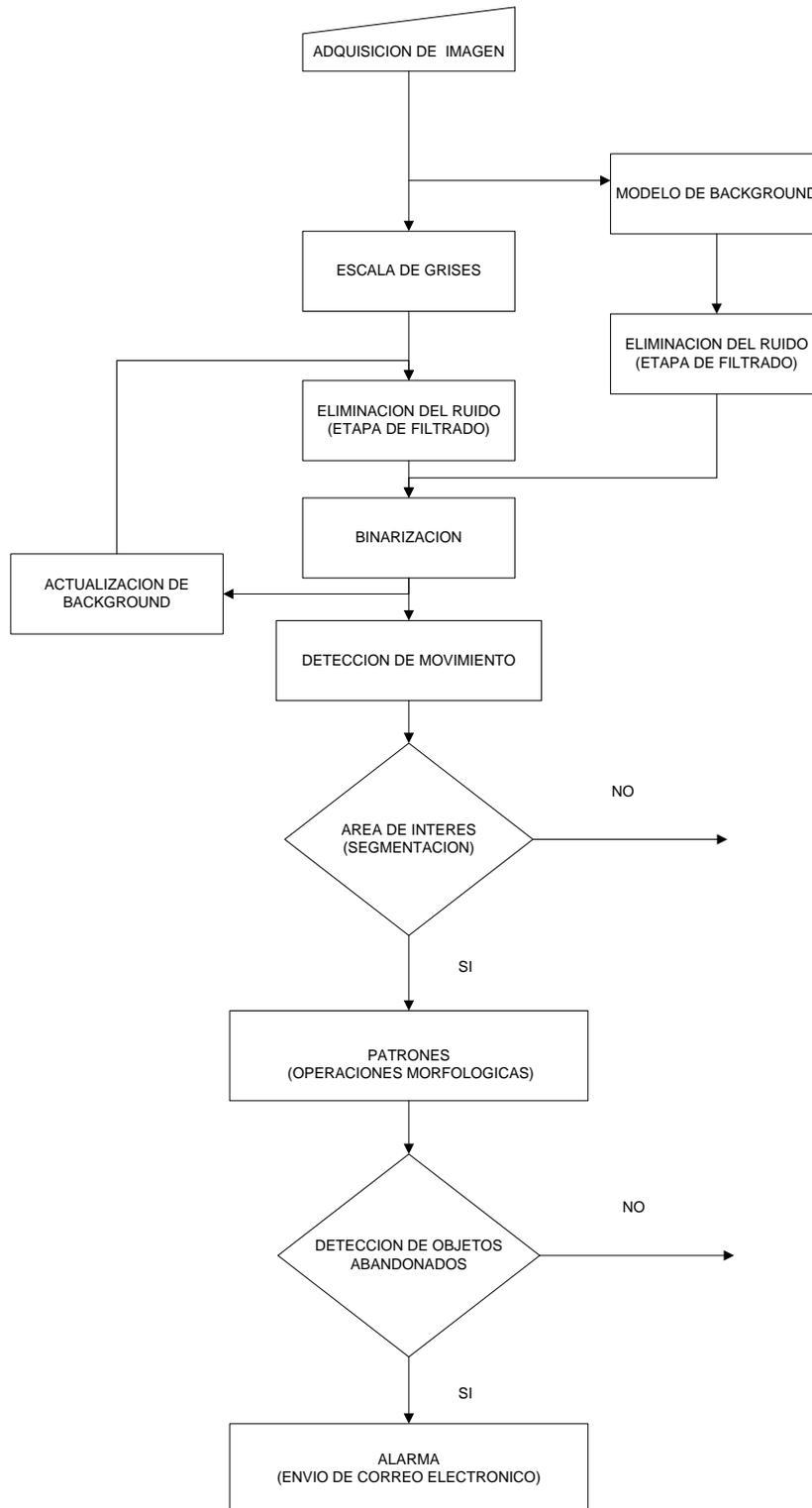


Fig. 3.11 Diagrama de flujo del algoritmo de detección de objetos abandonados.



3.2.2.1 Escena 2.- Detección de objetos y activación de alarma.

Como se observa en la Fig. 3.12. Una persona abandona un objeto, al existir el estado de reposo del objeto, la alarma comienza a evaluar el estado incrementándose hasta llegar a un límite en el cual se activará y enviará un correo electrónico al administrador de la red.

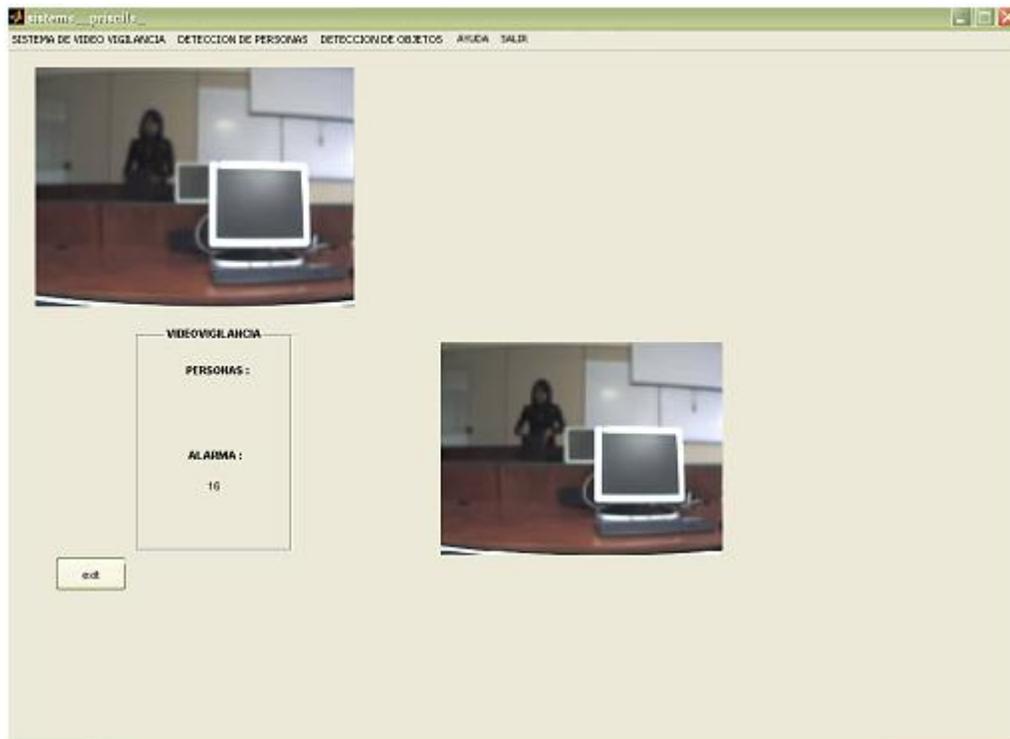


Fig. 3.12 Individuo abandonando objeto.

En la Fig. 3.13. Se muestra lo que sucede al quitar el objeto de la escena. Cambia de estado del objeto lo que determina que no se encuentra en estado de abandono, por lo tanto la alarma toma su valor inicial fijándose en cero [0] lo cual no dará lugar a la activación del correo electrónico.

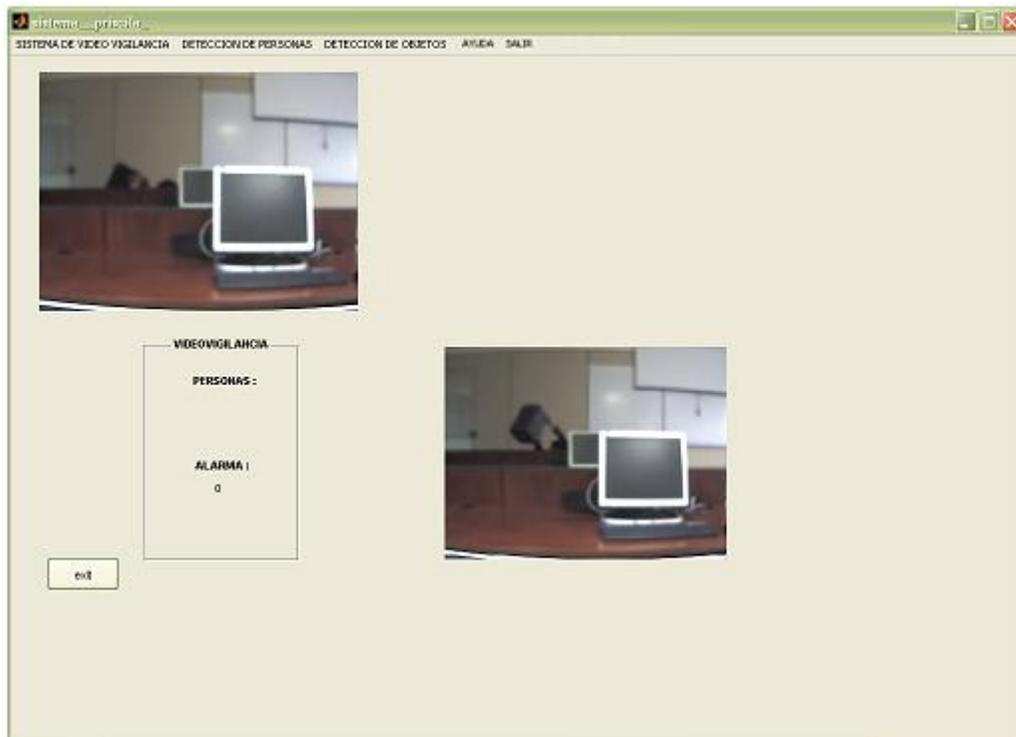


Fig. 3.13 Individuo retira el objeto.

En este caso de la Fig. 3.14, el objeto se encuentra a una distancia más próxima a la cámara. De igual manera el objeto es detectado como abandonado por lo cual al cumplirse el tiempo determinado se enviará la señal de alarma.

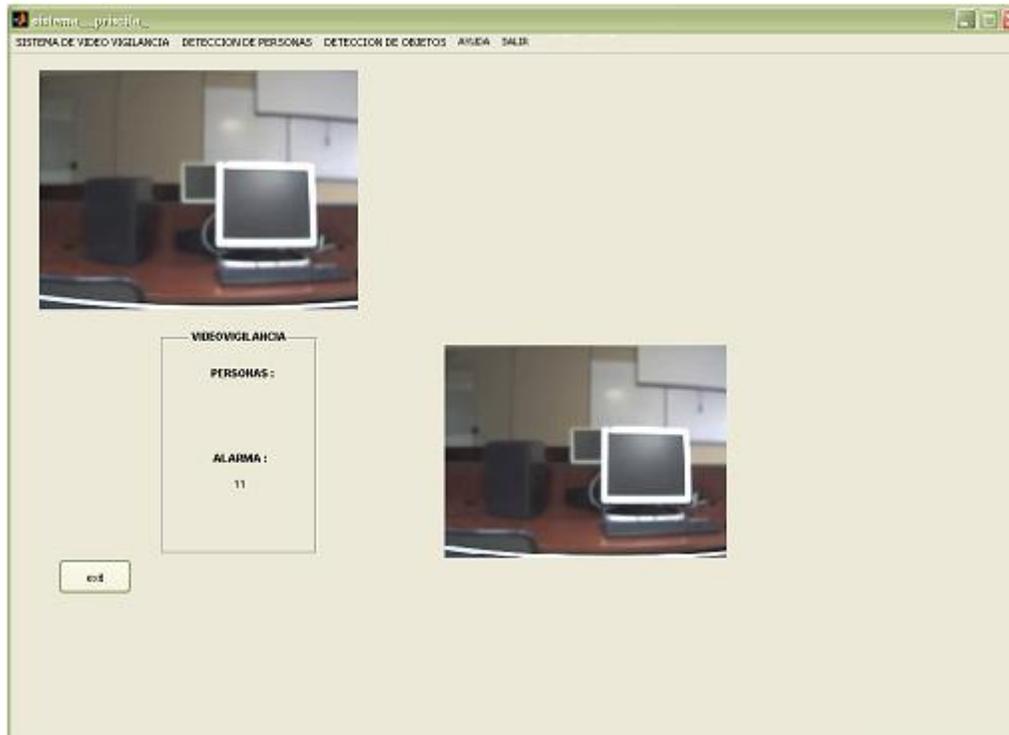


Fig. 3.14 Objeto a 1 metro de la cámara.

Sin embargo, como se observa en la Fig. 3.15 antes de cumplir con el tiempo para determinar el abandono del objeto, éste ha sido retirado de la escena por tanto la alarma vuelve a su valor inicial.

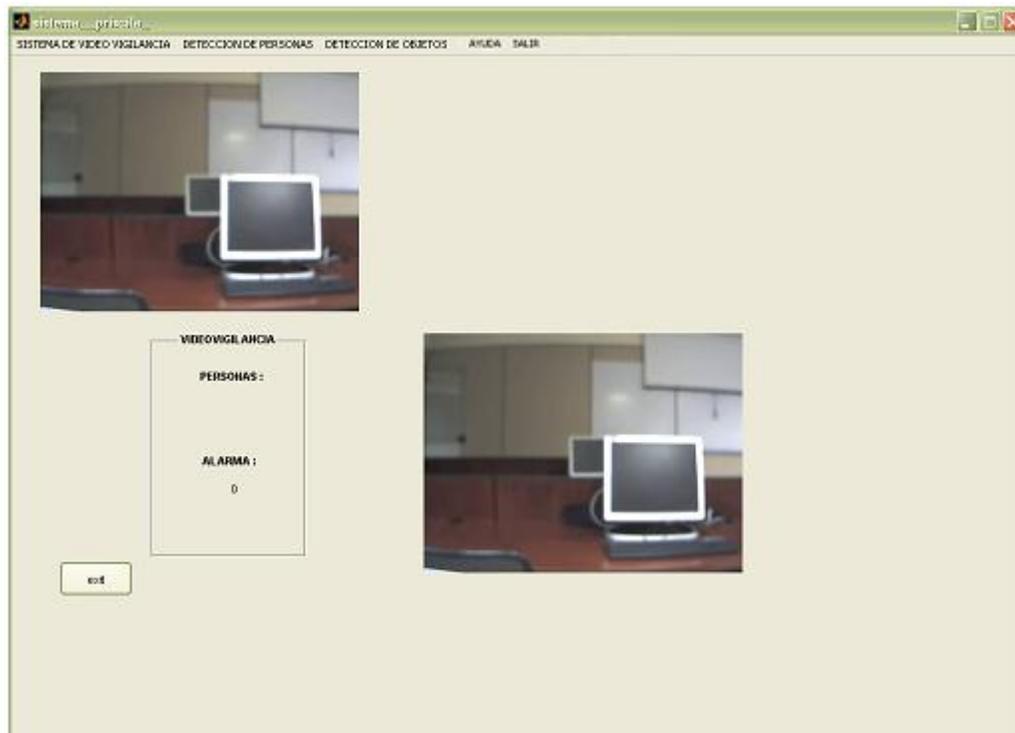


Fig. 3.15 Objeto retirado de la escena

A continuación en la Fig. 3.16. se presenta la evaluación con objetos de mayor tamaño, por ende, con respecto al número de píxeles será mayor. Primero se lo hace con respecto a una distancia considerable a la del sensor de visión. Como se observa el objeto es detectado.

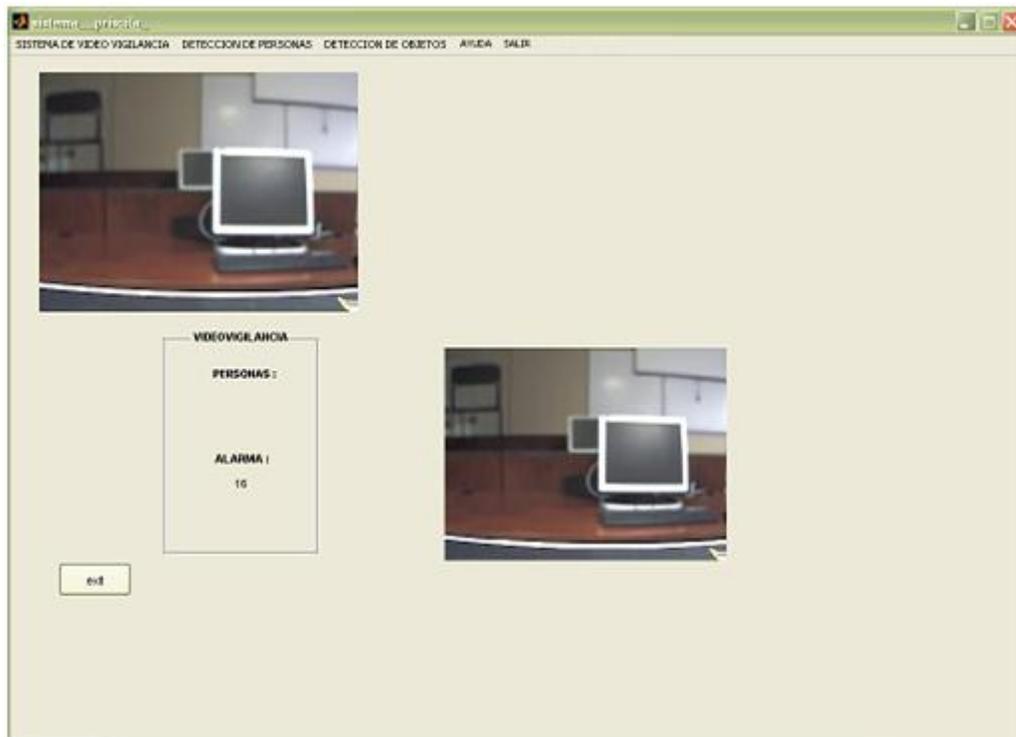


Fig. 3.16 Objeto de mayor tamaño a 2 metros de la cámara.

Como se observa en la Fig. 3.17. El objeto ha sido retirado y de esta manera la alarma vuelve a su valor inicial



Fig. 3.17 Objeto a 2 metros de la cámara es retirado de la escena.

En el caso presentado en la Fig. 3.18. Se evalúa el objeto de mayor tamaño en una posición cercana a la cámara IP.

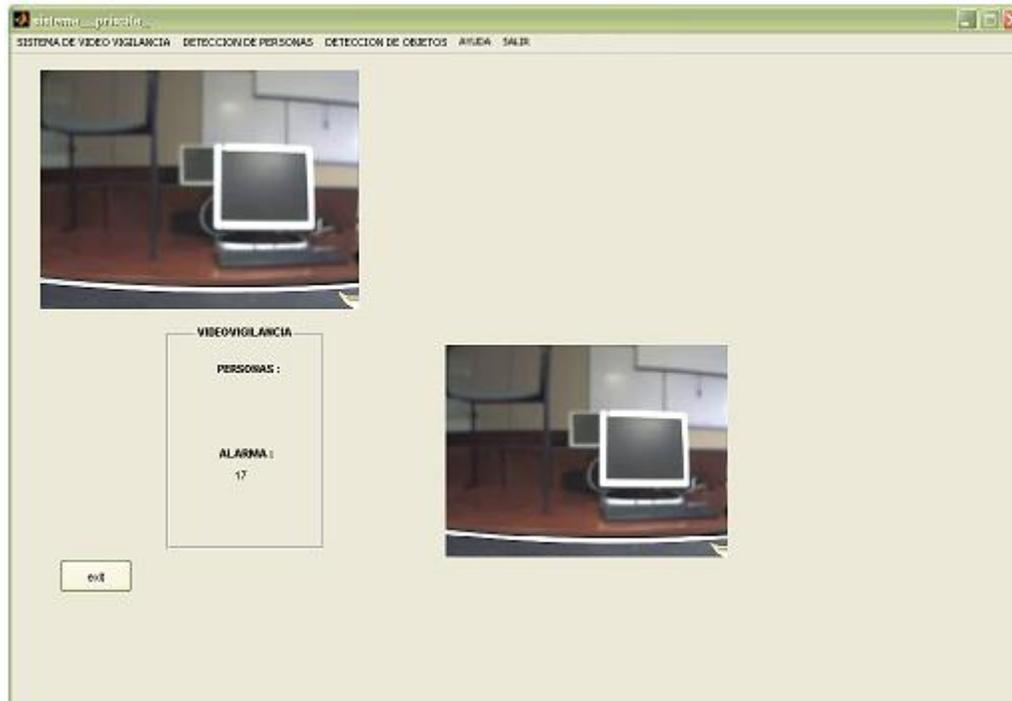


Fig. 3.18 Objeto de mayor tamaño a 1 metro de la cámara

Como se observa en la Fig. 3.19. El objeto ha sido retirado, por lo tanto, la alarma vuelve a su valor inicial.

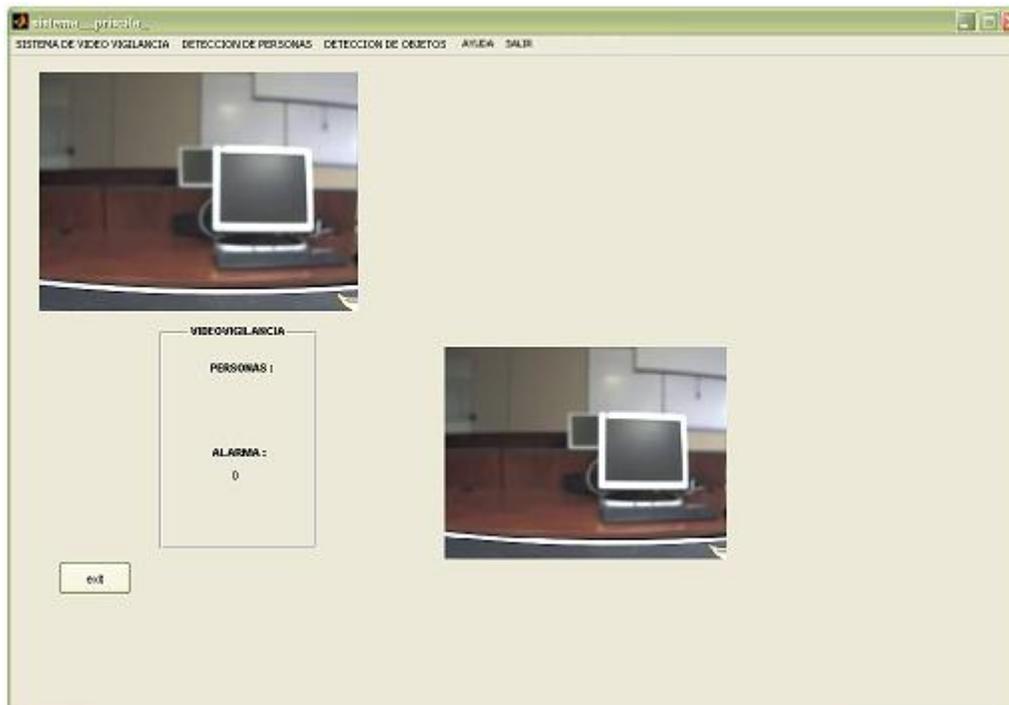


Fig. 3.19 Objeto a 1 metro de la cámara es retirado de la escena.

3.3 Rendimiento del sistema de video vigilancia.

El rendimiento del sistema está en relación directa con las características del computador donde se ejecuta el software. A continuación se detallan las características del computador donde se realizaron las pruebas como se observa en la Fig. 3.20:



Fig. 3.20 Características técnicas del computador.

El tiempo de cómputo del software se realizó con las funciones de Matlab 'tic' y 'toc'. La función 'tic' toma el valor del tiempo actual mientras que la función 'toc' retorna entre el tiempo transcurrido. Se han realizado algunas mediciones para obtener datos reales, para el algoritmo de detección de personas y detección de objetos abandonados, como se observa en la Fig. 3.21, Fig. 3.22 respectivamente.

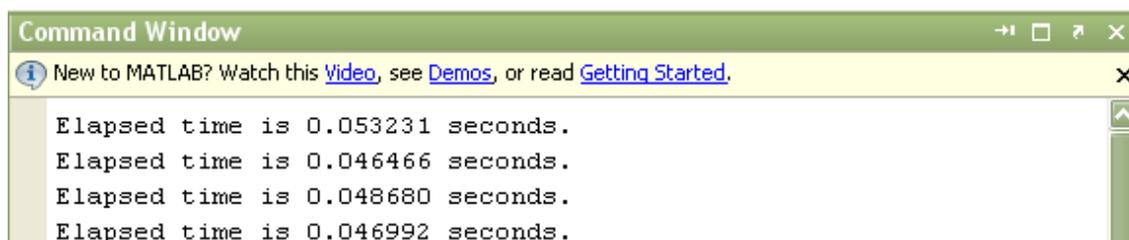


Fig. 3.21 Tiempo de respuesta obtenido para el algoritmo de detección de personas.



```
Command Window
New to MATLAB? Watch this Video, see Demos, or read Getting Started.
Elapsed time is 0.057698 seconds.
Elapsed time is 0.064488 seconds.
Elapsed time is 0.064200 seconds.
Elapsed time is 0.070578 seconds.
```

Fig. 3.22 Tiempo de respuesta obtenido para el algoritmo de detección de objetos abandonados.



CONCLUSIONES

- Los sistemas de vídeo vigilancia de segunda generación son una alternativa óptima con respecto a los de primera generación porque son basados en técnicas de procesado, lo que permite obtener información relevante y de esta manera descartar información que no es necesaria. Además, el administrador de la red mejorará su rendimiento de trabajo por la eficiencia del sistema de vigilancia.
- Un sistema de vídeo vigilancia IP puede tener accesibilidad al área vigilada, desde cualquier terminal conectado a internet de manera que se accede a los datos desde cualquier momento y lugar. Además, éste sistema mediante la configuración de red alámbrica como inalámbrica se puede configurar listas de acceso y brindar seguridad mediante clave respectivamente.
- En el sistema de video vigilancia se realizaron una serie de funciones para el preprocesamiento, las cuales son elementales para obtener una imagen mejorada que nos permita destacar la información necesaria para la eficiencia del sistema de video vigilancia.
- La implementación de una etapa de filtrado desempeña una fase primordial en el sistema porque minimiza los efectos de ruido, por lo que no depende únicamente del nivel de gris de un determinado pixel, sino que también del valor de los pixeles colindantes. Por lo tanto, la desviación es un factor importante ya que determinará el nivel de filtrado alrededor de toda la imagen.
- La detección de movimiento se realizó en base a la diferencia y umbralizado de las imágenes capturadas en el tiempo. Para detectar la parte móvil en foreground⁴ de la parte pasiva en background, etiquetando el cambio producido en la escena evaluada como movimiento.

⁴ FOREGROUND: en relación a una imagen todo lo que se encuentra en primer plano con respecto al background.



- El algoritmo utilizado para la detección de personas estima un tiempo de procesamiento de 0.048 segundos. Para el algoritmo de detección de objetos abandonados presenta un procesamiento de aproximadamente 0.064 segundos. El tiempo de procesamiento es importante ya que se evalúa la escena en tiempo real.
- La etapa de sustracción de background representa una de las etapas más importantes debido al método estadístico. Ya que no siempre es el mismo background a lo largo de toda la secuencia del video. Por lo que es indispensable los detalles estadísticos de cada pixel a lo largo del tiempo y con ello conseguir la actualización de la imagen de fondo.
- Las características comunes entre los pixeles permite realizar la segmentación. De esta manera se determina las regiones de interés para que la imagen resultante esté lista al análisis en base a elementos estructurantes.
- La etapa de morfología o reconocimiento de patrones permite que el sistema logre reconocer un objeto abandonado si, en un determinado tiempo no ha sido retirado del escenario evaluado; en caso de que el objeto sea etiquetado como en estado de abandono, se activa la alarma enviando una señal visual en el software y un correo electrónico al administrador del sistema de video vigilancia.
- El algoritmo empleado tanto en la fase de detección de movimiento, sustracción de background, detección de contornos y morfología, permitieron el conteo de personas en la escena evaluada.
- Con los resultados obtenidos a lo largo de este trabajo, se pudo observar que existen varios factores que influyen para el correcto funcionamiento del sistema. Uno de los principales es la iluminación debido a que es un factor externo que interviene directamente en la percepción de las imágenes y su procesado. Es



por ello que el sistema debe comportarse mediante diferentes eventos a través de la actualización de background.

- La red debe proporcionar un ancho de banda flexible que debe ser administrado correctamente para que no provoque fallas en el sistema ni se vea afectada, debido a que el proceso de adquisición de imágenes (fotogramas por segundo) es en tiempo real y el tamaño de la imagen es variable porque depende de la resolución, el códec y el escenario.
- Todo el conjunto de etapas que forman parte del sistema son fases fundamentales para el correcto funcionamiento del sistema de vídeo vigilancia tomando en cuenta de igual manera los requisitos funcionales y no funcionales del mismo.



TRABAJOS FUTUROS

Los estudios realizados, junto con los algoritmos desarrollados en este proyecto de fin de carrera sientan las bases de líneas de investigación futuras, entre las que cabe citar:

- Respecto a la detección de personas, actualmente el sistema realiza el conteo de las mismas; sería interesante basarse en diferentes tipos de patrones que debe cumplir una persona y, de esta manera, relacionarlo bajo cierta conducta que nos permita determinar el tipo de emoción de una persona generando diversas alarmas.
- En base a las funciones de detección de movimiento es necesario llevar un seguimiento de la trayectoria realizada por la persona, permitiendo de esta forma tener conocimiento del recorrido que realizó, frecuentemente esta técnica es utilizada para determinar en base al comportamiento que cumplen determinadas personas para provocar ataques es por tanto que en base a estos patrones se puede prevenir accidentes bajo ciertos eventos comunes en diferentes lugares.
- En cuanto a la detección de objetos abandonados, mediante la cantidad de pixeles detectados que el sistema realiza; se podría determinar la forma del objeto, obteniendo como resultado información específica del objeto.
- El sistema de vídeo-vigilancia actual se basa en una sola cámara que cubre un área específica; sería importante tener un arreglo de cámaras que nos permitan vigilar diferentes áreas o escenarios desde un lugar determinado.



REFERENCIA

1. **Jiménez, Pedro Gil.** [En línea] 21 de Octubre de 2009.
<http://dspace.uah.es/jspui/bitstream/10017/6597/1/TesisPedroGil.pdf>.
2. **Jiménez, Javier Alonso Albusac.** [En línea] 9 de Julio de 2008.
<http://www.esi.uclm.es/www/jalbusac/doc/masteralbusac.pdf>.
3. Escuela Politécnica Nacional. [En línea] 2009.
<http://bieec.epn.edu.ec:8180/dspace/bitstream/123456789/942/2/T10705CAP3.pdf>.
4. Grupo de investigacion ORETO. [En línea]
<http://oreto.esi.uclm.es:9673/oreto/Proyectos/Hesperia>.
5. Grupo de investigación ORETO. [En línea]
<http://oreto.esi.uclm.es:9673/oreto/Proyectos/ARGOS>.
6. Grupo de investigacion GAVAB. [En línea] <http://www.gavab.es/reactivos/index.html>.
7. Grupo de investigacion GAVAB. [En línea] <http://www.gavab.es/capo/vatrap/>.
8. Grupo de investigacion GAVAB-CAPO. [En línea]
<http://www.madrimasd.org/informacionidi/noticias/noticia.asp?id=42633>.
9. BBC- Mundo. [En línea]
http://www.bbc.co.uk/mundo/ciencia_tecnologia/2010/03/100311_video_camara_seguridad_inteligente_reino_unido_np.shtml?s.
10. **Rafael C. González, Richard E. Woods.** *Tratamiento digital de imágenes*. Massachusetts : ADDISON-WESLEY IBEROAMERICANA, S.A.
11. VIVOTEK. [En línea]
http://www.vivotek.com/products/model.php?network_camera=ip7137&tab=over.
12. VIVOTEK. [En línea] http://www.vivotek.com/products/network_cameras.php.
13. VIVOTEK. [En línea]
http://www.vivotek.com/products/model.php?network_camera=ip7137.
14. **José Ramón Alcalá, Guillermo Navarro.** Universidad de Castilla la Mancha. [En línea]
http://publicaciones.uclm.es/files/ebook_chapter_pdf/ebook_chapter_pdf_00185_MCd1.pdf.



15. **Jimenez, Pedro Gil.** [En línea] 15 de 02 de 2009.

<http://www.escet.urjc.es/~jjpantrigo/PFCs/PFCobjetosAbandonados.pdf>.

16. **Carrera, Antonio Collazos.** Escuela Tecnica Superior de Ingenieria Informatica. *Modelado de Escenarios para Aplicaciones de Videovigilancia.* [En línea] 23 de 06 de 2010.

<http://eciencia.urjc.es/dspace/bitstream/10115/4117/1/pfcModelado.pdf>.

17. **Rivas., Melva Ramos.** Universidad de las Américas Puebla. [En línea] 18 de 12 de 2003.

http://catarina.udlap.mx/u_dl_a/tales/documentos/lis/ramos_r_m/capitulo3.pdf.

18. **Franco, Ana Carcedo Y.** Universidad de las Américas Puebla (UDLAP). [En línea] 7 de 1 de 2004. http://catarina.udlap.mx/u_dl_a/tales/documentos/lem/carcedo_y_a/capitulo1.pdf.



ANEXOS

ANEXO 1 Especificaciones de la Cámara IP137

Cámara IP 137 VIVOTEK	
ILUMINACIÓN MÍNIMA	1.5 Lux / F2.0 (typical)
VIDEO	Compression: MPEG-4 Streaming: MPEG-4 streaming over UDP, TCP, HTTP or HTTPS Supports 3GPP mobile surveillance Frame rates: MPEG-4: Up to 30/25 fps at 640x480
IMAGEN	Adjustable image size, quality, and bit rate Time stamp and text caption overlay Flip & mirror Configurable brightness, contrast, saturation, sharpness, white balance and exposure AGC AWB AES
CREACIÓN DE REDES	10/100 Mbps Ethernet, RJ-45 Built-in 802.11b/g WLAN Protocols: IPv4, TCP/IP, HTTP, UPnP, RTSP/RTP/RTCP, SMTP, FTP, DHCP, NTP, DNS, DDNS
SEGURIDAD	Multi-level user access with password protection IP address filtering
USUARIOS	10 clients on-line monitoring at the same time
POWER	5V DC
HERRAMIENTAS DE INSTALACIÓN Y MANTENIMIENTO	Installation Wizard 2 32-CH central management software Supports firmware upgrade
APLICACIONES	SDK available for application development and system integration