



INSTITUTO POLITÉCNICO NACIONAL

CENTRO DE INVESTIGACIÓN EN COMPUTACIÓN

LABORATORIO DE TIEMPO REAL Y AUTOMATIZACIÓN

PROYECTO DE INVESTIGACIÓN:

**“INSTRUMENTOS VIRTUALES PARA
VIDEO-VIGILANCIA Y SEGURIDAD”**

T E S I S

QUE PARA OBTENER EL GRADO DE:

**MAESTRO EN CIENCIAS EN INGENIERÍA DE CÓMPUTO
CON OPCIÓN EN SISTEMAS DIGITALES**

PRESENTA

ING. PAMELA ARACELI RANGEL TIRADO

DIRECTORES

**DR. LUIS PASTOR SÁNCHEZ FERNÁNDEZ
DR. OLEKSIY POGREBNIYAK**

MÉXICO D.F., 2012





INSTITUTO POLITÉCNICO NACIONAL

SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

ACTA DE REVISIÓN DE TESIS

En la Ciudad de México, D.F. siendo las 12:00 horas del día 22 del mes de noviembre de 2012 se reunieron los miembros de la Comisión Revisora de la Tesis, designada por el Colegio de Profesores de Estudios de Posgrado e Investigación del:

Centro de Investigación en Computación

para examinar la tesis titulada:

“Instrumentos virtuales para video-vigilancia y seguridad”

Presentada por el alumno:

RANGEL

Apellido paterno

TIRADO

Apellido materno

PAMELA ARACELI

Nombre(s)

Con registro:

B	1	0	1	8	8	3
----------	----------	----------	----------	----------	----------	----------

aspirante de: **MAESTRÍA EN CIENCIAS EN INGENIERÍA DE CÓMPUTO CON OPCIÓN EN SISTEMAS DIGITALES**

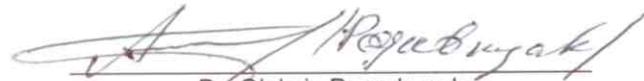
Después de intercambiar opiniones los miembros de la Comisión manifestaron **APROBAR LA TESIS**, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.

LA COMISIÓN REVISORA

Directores de Tesis



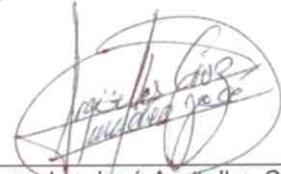
Dr. Luis Pastor Sánchez Fernández



Dr. Oleksiy Pogrebnyak



Dr. Sergio Suárez Guerra



Dr. Amadeo José Argüelles Cruz



Dr. José Juan Carbajal Hernández



Dr. Marco Antonio Moreno Armendáriz

PRESIDENTE DEL COLEGIO DE PROFESORES



Dr. Luis Alfonso Villa Vargas
INSTITUTO POLITÉCNICO NACIONAL
CENTRO DE INVESTIGACIÓN
EN COMPUTACIÓN
DIRECCIÓN



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

CARTA DE CESIÓN DE DERECHOS

En la ciudad de México, el día 22 del mes noviembre del año 2012, la que suscribe **Pamela Araceli Rangel Tirado** alumna del Programa de Maestría en Ciencias en Ingeniería de Cómputo, opción Sistemas Digitales, con número de registro B1018883, adscrita al Centro de Investigación en Computación, manifiesta que es autora intelectual del presente trabajo de Tesis bajo la dirección del Dr. Luis Pastor Sánchez Fernández y del Dr. Oleksiy Pogrebnyak y cede los derechos del trabajo titulado “*Instrumentos virtuales para video-vigilancia y seguridad*”, al Instituto Politécnico Nacional para su difusión, con fines académicos y de investigación.

Los usuarios de la información no deben reproducir el contenido textual, gráficas o datos del trabajo sin el permiso expreso del autor y/o director del trabajo. Este puede ser obtenido escribiendo a las siguientes direcciones pamela.rangel.t@gmail.com, lsanchez@cic.ipn.mx o olek@cic.ipn.mx. Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.

Pamela Araceli Rangel Tirado

AGRADECIMIENTOS

A mis directores de tesis, el Dr. Luis Pastor Sánchez Fernández y el Dr. Oleksiy Pogrebnyak por todo el apoyo, paciencia y guía que me brindaron durante mi estancia en el Centro.

Al Comité Tutorial y Sínodo, integrado por el Dr. Sergio Suárez Guerra, el Dr. Amadeo José Argüelles Cruz, el Dr. José Juan Carbajal Hernández, el Dr. Marco Antonio Moreno Armendáriz y el M. en C. Pablo Manrique Ramírez, además del Coordinador de Programa, el Dr. José Luis Oropeza Rodríguez por sus recomendaciones y consejos que me proporcionaron durante el desarrollo del proyecto.

Al Instituto Politécnico Nacional, al Centro de Investigación en Computación (CIC-IPN), al Consejo Nacional de Ciencia y Tecnología (CONACyT) y al Programa Institucional de Formación de Investigadores (PIFI) por el soporte académico y económico que me otorgaron.

Por último pero no menos importante, a mi familia por el apoyo en cada momento de la vida. A todos mis compañeros, amigos y profesores con los que compartí vivencias, trabajo y ayuda.

RESUMEN

En este trabajo se propone un modelo computacional para automatizar la obtención de imágenes con mayor detalle, a través de un sistema de video-vigilancia dual que está compuesto por X_i cámaras fijas y una móvil, para potenciar los sistemas comerciales de bajas y medias prestaciones a través de la incorporación de instrumentos virtuales que llevan a cabo la detección y procesamiento de eventos relevantes.

La inspección se realiza en todas las cámaras fijas, aún cuando el domo esté atendiendo el procesamiento de un evento relevante. El video es obtenido a través de comunicación TCP/IP, empleando las Herramientas de Desarrollo de *Software* (SDK, por sus siglas en inglés) de los dispositivos.

El evento relevante es detectado en las regiones de interés definidas manualmente en el campo de visión de alguna de las cámaras fijas para localizar al objeto que lo generó. El procesamiento comienza con una resta de imágenes; a continuación se ejecutan operaciones morfológicas y filtrado que antecede al análisis de partículas. Posteriormente, se discriminan los objetos debido a su cercanía con las coordenadas del punto donde se detectó la presencia. De esta forma se puede obtener la imagen que contiene sólo la partícula que causó el evento.

Dos tareas muy importantes antes de buscar la ubicación del objeto de interés son: primero, el preposicionamiento, para asemejar su campo de visión (FOV, por sus siglas en inglés) con el de la cámara fija que generó el evento y así poder buscar el objeto por reconocimiento de patrones; además para comenzar con nuevas localizaciones. Y la segunda, es la calibración intercámara que debido a que la imagen inicial después del procesamiento proviene de una cámara fija y que se desea encontrar el objeto de interés en el FOV de la cámara móvil, es necesario compensar los niveles de iluminación.

Una vez que las condiciones anteriores se han cumplido, se ejecuta la búsqueda de patrón para encontrar el objeto que ha entrado a escena, localizándolo continuamente en el FOV del domo hasta que esté centrado; esto se logra a través de la programación del protocolo de comunicación Pelco-D y el uso de RS-485, además de la lógica aplicada sobre la definición de áreas de vigilancia correspondientes al movimiento y de una relación de aspecto para el acercamiento.

De esta forma, se obtiene una imagen con mayor detalle, misma que es usada en la generación de reportes a usuario en formato PDF. Finalmente, éste se envía por correo electrónico al usuario designado por el administrador.

Como parte de los resultados, se presentan imágenes finales con mayor detalle y de los reportes a usuario derivadas de pruebas realizadas en un estacionamiento exterior restringido.



ABSTRACT

This work proposes a computational model to automate imaging in more detail through a video surveillance system that consists of dual X_i fixed cameras and a mobile camera to promote systems of low and medium commercial performance through incorporating virtual instruments that perform the detection and processing of relevant events.

The inspection is performed on all fixed cameras, even when the dome is processing another relevant event. The video is obtained through TCP / IP channel using the Software Development Kit (SDK) of the devices.

A relevant event is detected defining manually regions of interest in the field of vision of one of the fixed cameras to locate the object that is generated that event. Processing beginning with subtraction of images and particle analysis; after that running mathematical morphology operations and filtering Afterwards, the objects are discriminated because of their proximity to the coordinates where the presence was detected. This way, the image containing only the particle that caused the event is obtained.

Two important tasks before seeking the location of the object of interest are; first, pre-positioning in order to move the dome to an area known to resemble its Fields of Vision (FOV) with fixed camera that generated the event and thus be able find the object by pattern matching; also starting with new locations. The second activity is the inter-camera calibration: because the original image is from a fixed camera and the system searches for the object of interest in the FOV of the moving camera, it is necessary to compensate lighting levels.

Once the conditions are right, the system starts finding the search pattern that match to the object entered the scene, continuously locating it on FOV dome until it is focused. This is achieved through programming the communication protocol Pelco-D and using RS-485, plus the logic applied to the definition of surveillance areas for the movement and an aspect ratio for the zoom in.

This way, an image in more detail is obtained. Then, it is used to generate user reports in PDF. Finally, the report is sent by e-mail to the user designated by the administrator.

As a part of the results, final images with more detail and user reports are obtained by testing in an outdoor parking restricted.

Índice General

RESUMEN	I
ABSTRACT	II
ÍNDICE DE FIGURAS	VI
ÍNDICE DE TABLAS	VIII
GLOSARIO	IX
• ABREVIATURAS.....	IX
• TÉRMINOS	X
CAPÍTULO 1. INTRODUCCIÓN	1
1.1 MOTIVACIÓN Y ANTECEDENTES.....	1
1.2 PLANTEAMIENTO DE PROBLEMAS	2
1.3 JUSTIFICACIÓN	2
1.4 HIPÓTESIS.....	3
1.5 ALCANCE.....	3
1.6 LÍMITES.....	4
1.7 OBJETIVOS	4
1.7.1 <i>Objetivo general</i>	4
1.7.2 <i>Objetivos particulares</i>	4
1.8 CONTRIBUCIONES.....	4
1.9 METODOLOGÍA Y DESARROLLO DE LA INVESTIGACIÓN	5
1.9.1 <i>Propuesta de solución</i>	5
1.9.2 <i>Diagrama</i>	6
1.10 ESTRUCTURA DE LA TESIS.....	7
CAPÍTULO 2. ESTADO DEL ARTE	9
2.1 DOMÓTICA	9
2.2 VIDEO-VIGILANCIA	9
2.2.1 <i>Diseño de sistemas de video-vigilancia</i>	10
2.2.2 <i>Beneficios de la video-vigilancia</i>	10
2.2.3 <i>Sistemas basados en arquitectura cliente/servidor</i>	11
2.2.4 <i>Sistemas de video-vigilancia avanzados</i>	12
2.3 ELEMENTOS DE UN SISTEMA DE VIDEO-VIGILANCIA.....	12
2.3.1 <i>Unidades de captura de video</i>	13
2.3.2 <i>Sensores de movimiento</i>	14
2.3.3 <i>Manipulación de las cámaras: zoom óptico y digital</i>	14
2.3.4 <i>Grabador de video digital</i>	15
2.4 FUNDAMENTOS DE VIDEO	16
2.4.1 <i>Video analógico</i>	17
2.4.2 <i>Video digital</i>	17
2.4.3 <i>Codificación</i>	18
2.4.4 <i>Color</i>	19
2.4.5 <i>Formatos de video</i>	19
2.5 SEGURIDAD.....	20



2.6	CONFIGURACIÓN DE SISTEMAS COMERCIALES	21
2.7	CONCLUSIONES.....	22
CAPÍTULO 3.	MARCO TEÓRICO	23
3.1	COMUNICACIÓN ENTRE DISPOSITIVOS.....	23
3.1.1	<i>Comunicación serial como estándar</i>	<i>23</i>
3.1.2	<i>Comunicación serial en LabVIEW: VISA</i>	<i>24</i>
3.1.3	<i>Comunicación por TCP/IP.....</i>	<i>25</i>
3.2	PROCESAMIENTO DE IMÁGENES	25
3.2.1	<i>Conectividad</i>	<i>26</i>
3.2.2	<i>Regiones de interés.....</i>	<i>27</i>
3.2.3	<i>Histograma</i>	<i>27</i>
3.2.4	<i>Umbralado como herramienta de segmentación.....</i>	<i>28</i>
3.2.5	<i>La resta de imágenes como operador matemático básico.....</i>	<i>29</i>
3.2.6	<i>Morfología matemática.....</i>	<i>30</i>
3.2.7	<i>Detección de bordes</i>	<i>32</i>
3.3	ANÁLISIS DE PARTÍCULAS	32
3.4	MEJORAMIENTO DE LAS CONDICIONES EN LAS IMÁGENES	33
3.4.1	<i>Ajuste de histograma.....</i>	<i>34</i>
3.4.2	<i>Reajuste de los niveles de gris.....</i>	<i>34</i>
3.4.3	<i>Filtrado espacial</i>	<i>35</i>
3.4.4	<i>Especificación del histograma.....</i>	<i>35</i>
3.5	BÚSQUEDA DE PATRONES.....	36
3.5.1	<i>Plantilla.....</i>	<i>36</i>
3.5.2	<i>La búsqueda.....</i>	<i>37</i>
3.6	CONCLUSIONES.....	39
CAPÍTULO 4.	DISEÑO Y DESARROLLO DE LOS INSTRUMENTOS VIRTUALES PARA VIDEO-VIGILANCIA Y SEGURIDAD.....	40
4.1.	DISEÑO DE PROPUESTA.....	41
4.1.1	<i>Requisitos del sistema base.....</i>	<i>42</i>
4.1.2	<i>Conexión y montaje</i>	<i>43</i>
4.2.	OBTENCIÓN DE VIDEO EN TIEMPO REAL	44
4.3.	SELECCIÓN DE LAS REGIONES DE INTERÉS.....	46
4.4.	PREPOSICIONAMIENTO	47
4.5.	GENERACIÓN Y ENVÍO DE REPORTES A USUARIO	48
4.6.	CONCLUSIONES.....	50
CAPÍTULO 5.	PROCESAMIENTO DE EVENTOS RELEVANTES	51
5.1	PROCESAMIENTO DIGITAL DE IMÁGENES PARA DETERMINAR PRESENCIA DE OBJETOS.....	51
5.1.1	<i>Detección de bordes</i>	<i>51</i>
5.1.2	<i>Presencia de objeto.....</i>	<i>53</i>
5.1.3	<i>Discriminación entre partículas.....</i>	<i>54</i>
5.2	CALIBRACIÓN INTERCÁMARA.....	55
5.3	BÚSQUEDA DE PATRÓN	56
5.3.1	<i>Aprendizaje</i>	<i>57</i>

5.3.2	<i>Búsqueda</i>	57
5.4	LÓGICA DE MANDO DE LA CÁMARA TIPO DOMO.....	58
5.4.1	<i>Áreas de vigilancia</i>	58
5.4.2	<i>Control del domo</i>	59
5.4.3	<i>Protocolo de comunicación Pelco-D</i>	62
5.5	DIAGRAMA DE JERARQUÍA.....	63
5.6	CONCLUSIONES	64
CAPÍTULO 6. RESULTADOS EXPERIMENTALES.....		65
6.1	INTEGRACIÓN DEL SISTEMA.....	65
6.2	SISTEMA CLIENTE/SERVIDOR PARA LA TRANSMISIÓN Y RECEPCIÓN DE IMÁGENES.....	65
6.3	DETECCIÓN DE EVENTOS RELEVANTES EN ESTACIONAMIENTOS RESTRINGIDOS	66
6.4	EVALUACIÓN DE TIEMPOS DE EJECUCIÓN DE LOS MÉTODOS Y ALGORITMOS.....	68
6.5	TRABAJO ADICIONAL: DETECCIÓN DE MOVIMIENTO PARA SEGUIMIENTO.....	69
6.6	CONCLUSIONES	73
CONCLUSIONES.....		74
TRABAJOS FUTUROS.....		75
REFERENCIAS		76
ANEXO A. GLOSARIO DE FUNCIONES DE LABVIEW.....		80
ANEXO B. PROTOCOLO PELCO-D		84
B.1	PROTOCOLO “D”	84
B.2	FORMATO DEL CONJUNTO DE COMANDOS.....	84
B.3	FORMATO DEL CONJUNTO DE DATOS	84
B.4	CHECK SUM	85
B.5	RUTINAS PRECONFIGURADAS	85
ANEXO C. SDK IDIS WATCH CLIENT.....		86
C.1	GENERALIDADES.....	86
C.2	MODELOS DE DVR APLICABLES.....	86
C.3	COMPONENTES EMPLEADAS DEL SDK.....	87
ANEXO D. SDK VIVOTEK.....		88
D.1	ALTO PERFIL.....	88
D.2	CONTROL VITAMIN	88
D.3	COMANDOS EMPLEADOS.....	89
ANEXO E. PRECONFIGURACIÓN DE POSICIONES EN EL DOMO		91
E.1	CONFIGURACIÓN	91
E.2	OPERACIÓN.....	91
ANEXO F. DIAGRAMA GENERAL DE SOFTWARE		92
ANEXO G. ESPECIFICACIÓN DE LOS CÁMARAS ANALÓGICAS		93
ANEXO H. PUBLICACIONES		94



Índice de Figuras

Figura 1.1	Esquema propuesto para video-vigilancia	4
Figura 1.2	Diagrama general de etapas.....	6
Figura 1.3	Áreas de vigilancia	6
Figura 1.3	Diagrama de caso de estudio	7
Figura 2.1	Patrón de Bayer	15
Figura 2.2	Escaneo en video analógico	17
Figura 2.3	Elementos del codificador	18
Figura 2.3	Configuración de un sistema de video-vigilancia comercial	22
Figura 3.1	Comunicación serial	24
Figura 3.2	Referencia de píxel	25
Figura 3.3	Píxeles de una imagen	26
Figura 3.4	4-Conectado.....	26
Figura 3.5	6-Conectado.....	26
Figura 3.6	8-Conectado.....	27
Figura 3.7	Imagen con regiones de interés	27
Figura 3.8	Imagen e histograma	28
Figura 3.9	Imagen A	29
Figura 3.10	Imagen B.....	29
Figura 3.11	Imagen destino	30
Figura 3.12	Dilatación	31
Figura 3.13	Erosión	31
Figura 3.14	Detección de bordes	32
Figura 3.15	Análisis de partículas	33
Figura 3.16	Plantilla para búsqueda de patrones	36
Figura 3.17	Correlación cruzada normalizada.....	37
Figura 3.18	Muestreo de plantilla en el método entendimiento de imagen	38
Figura 3.19	Muestreo no uniforme	39
Figura 4.1	Diagrama de bloques de operación general	40
Figura 4.2	Conectividad del sistema básico de video-vigilancia.....	41
Figura 4.3	Esquema de conexión del sistema de video-vigilancia.....	43
Figura 4.4	Configuración general para la obtención del video en tiempo real	44
Figura 4.5.	Diagrama de jerarquía para obtención de video en tiempo real	45
Figura 4.6	Definición de ROI.....	46
Figura 4.7	Mensaje para preposicionamiento.....	47
Figura 4.8	Plano de preposicionamiento.....	48
Figura 4.9	Generación y envío de reportes a usuario	49
Figura 4.10	Diagrama de jerarquía para la generación de reportes a usuario.....	50
Figura 5.1	Diagrama de bloques para el procesamiento de eventos relevantes	51
Figura 5.2	Ajustar parámetros en la detección de objetos.....	52
Figura 5.3	Estado de ejes de detección	52
Figura 5.4	Imagen en tiempo real con detección de bordes.....	53
Figura 5.5	Región de interés global	53

Figura 5.6	Discriminación de partículas	55
Figura 5.7	Histogramas antes de calibración	56
Figura 5.8	Resultado en imágenes después de calibración	56
Figura 5.9	Búsqueda de patrones	57
Figura 5.10	División del FOV en áreas de vigilancia	58
Figura 5.11	Diagrama de bloques del control del domo	60
Figura 5.12	Dimensiones de área 40 y objeto	61
Figura 5.13	Diagrama de bloques para convertir trama	63
Figura 5.14	Diagrama de jerarquía para el procesamiento de eventos relevantes.....	63
Figura 6.1	Sistema de videovigilancia instalado	65
Figura 6.2	Panel Frontal para la visualización del video en LabVIEW	66
Figura 6.3	Imagen resultado 1	66
Figura 6.4	Imagen resultado 2	67
Figura 6.5	Imagen resultado 3	67
Figura 6.6	Reporte a usuario.....	68
Figura 6.7	Gráfica de tiempo de ejecución para la aplicación	69
Figura 6.8	Panel frontal para detección de movimiento	70
Figura 6.9	Cuadro base para discriminación de objetos	71
Figura 6.10	Diagrama de bloques sobre la detección de movimiento	72
Figura B.1	Trama de protocolo Pelco-D	84
Figura B.2	Ejemplo de trama para cálculo de <i>check sum</i>	85
Figura C.1	Partes empleadas del SKD IDIS Watch.....	87
Figura E.1	Pantalla principal del domo	91
Figura E.2	Pantalla de configuración del domo.....	91
Figura F.1	Diagrama de jerarquía total	92



Índice de Tablas

Tabla 2.1	Características de formatos de video MPGE-4	20
Tabla 4.1	Requisitos para el sistema cliente-servidor.....	42
Tabla 4.2	Equipo empleado en el sistema.....	42
Tabla 5.1	Parámetros de filtrado de partículas	54
Tabla 5.2	Parámetros para calibración intercámara.....	56
Tabla 5.3	Datos de configuración para comunicación serial.....	62
Tabla 6.1	Tiempos de ejecución en toda la aplicación.....	69
Tabla B.1	Comandos Pelco-D.....	84
Tabla B.2	Comando 1 Pelco-D.....	84
Tabla B.3	Datos Pelco-D	85
Tabla B.4	Rutinas preconfiguradas Pelco-D.....	85
Tabla C.1	Modelos DVR aplicables	86
Tabla G.1	Cámaras analógicas.....	93

GLOSARIO

- **Abreviaturas**

ID	Unidimensional.
2D	Bidimensional.
3D	Tridimensional.
API	Aplicación de Interfaz de Programación (<i>Application Programming Interface</i>).
CCD	Dispositivo Acoplado por Carga (<i>Charge-Coupled Device</i>).
CMOS	Semiconductor Complementario de Metal Óxido (<i>Complementary Metal-Oxide-Semiconductor</i>).
DNS	Servicio de Nombres de Dominio (<i>Domain Name Service</i>).
DVR	Grabador de Video Digital (<i>Digital Video Recorder</i>).
FOV	Campo de Visión (<i>Field of Vision</i>).
fps	Marcos por segundo (<i>Frames per Second</i>)
IP	Protocolo de Internet (<i>Internet Protocol</i>).
PDF	Formato de Documento Portable (<i>Portable Document Format</i>).
PTZ	Movimiento horizontal (<i>Pan</i>), vertical (<i>Tilt</i>) y acercamiento/alejamiento (<i>Zoom</i>).
PUI	Píxel bajo inspección (<i>Píxel Under Inspection</i>).
ROI	Región de Interés (<i>Region of Interest</i>).
SDK	Herramientas de Desarrollo de <i>Software</i> (<i>Software Development Kits</i>).
SMTP	Protocolo de Transferencia de Correo Simple (<i>Simple Mail Transfer Protocol</i>).
SO	Sistema Operativo.
TCP	Protocolo de Control de Transmisión (<i>Transmission Control Protocol</i>).
UDP	Protocolo de Datagramas de Usuario (<i>User Datagram Protocol</i>).
USB	Bus Serial Universal (<i>Universal Serial Bus</i>).
VISA	Instrumento Virtual de Arquitectura de Software (<i>Virtual Instrument Software Architecture</i>).



- **Términos**

Algoritmo	Conjunto ordenado y finito de operaciones que permite hallar la solución de un problema.
Brillo	Es el atributo de una sensación visual según la cual un área parece emitir más o menos luz. [51]
Color	Es el atributo de una sensación visual según la cual un área parece ser similar a uno de los colores percibidos, rojo, amarillo, verde y azul, o una combinación de dos de ellos. [51]
Contraste	Es la diferencia en las propiedades visual que hace que un objeto distinguible de otros objetos y el fondo. El contraste es determinado por la diferencia en el color y el brillo del objeto y otros objetos dentro del mismo campo de visión.
Corrección Gamma	Se trata de una transformación donde se busca expandir los niveles altos de gris en una imagen mientras suprime los niveles bajos de gris.
Decolar	(<i>Dequeue</i>) Remueve un elemento del inicio de la cola.
Ecuilización	Consiste en encontrar una transformación con la cual el histograma tenga una representación uniforme, es decir, lo más cercano al histograma ideal. [57]
Encolar	(<i>Enqueue</i>) Agrega un elemento en el final de la cola.
Iluminancia	Magnitud que expresa el flujo luminoso que incide sobre la unidad de superficie. Su unidad en el Sistema Internacional es el lux.
Imagen	Representación del mundo a través de una función bidimensional con coordenadas x y y , donde cada $f(x, y)$ es la intensidad de luz para un píxel dado.
Imagen binaria	Fotograma en el cual todos sus píxeles tienen valores solo de dos clases, 0 o 1.
Imagen digitalizada	Fotograma que ha sido discretizada tanto en coordenadas espaciales como en brillo. Es una matriz cuyos índices de filas y columnas identifican un punto en la imagen y su valor de gris en ese punto.
Imagen en escala de grises	Fotograma con coordenadas espaciales x y y con sus respectivos valores de intensidad. Este tipo de imágenes son las más sencillas y su uso es muy frecuente en el análisis. [53]
Lab VIEW	Lenguaje de programación de alto nivel, de tipo gráfico, y enfocado al uso en instrumentación. Los programas desarrollados con Lab VIEW se llaman Instrumentos Virtuales (VI, <i>Virtual Instruments</i>).[51]
Luminancia	Es proporcional a la energía de la luz emitida por unidad de área de la fuente, pero esta energía se pondera de acuerdo a la sensibilidad espectral del ojo. [51]
Lux	Unidad de iluminancia del Sistema Internacional, que equivale a la iluminancia de una superficie que recibe un flujo luminoso de un lumen por metro cuadrado. (Símbolo lx).

<i>Método</i>	Procedimiento que se sigue en las ciencias para hallar la verdad y enseñarla.
<i>Píxel</i>	Unidades básicas de imagen en arreglos de dos dimensiones para formar una imagen o marco. Acrónimo formado con las palabras inglesas ' <i>Picture</i> ' y ' <i>element</i> '.
<i>Saturación</i>	Es el colorido de un área juzgado en proporción a su brillo. [51]
<i>Vecindario de píxeles</i>	Conjunto de píxeles alrededor de un punto; la fortaleza de su unión depende del tipo de conexión entre ellos.
<i>Visión artificial</i>	Subdisciplina de la inteligencia artificial que busca analizar las imágenes para extraer información valiosa para controlar un proceso o actividad. [53]
<i>Vision Lab VIEW</i>	Conjunto de herramientas para el procesamiento de imágenes en LabVIEW. [50]

INSTRUMENTOS VIRTUALES PARA VIDEO-VIGILANCIA Y SEGURIDAD

CAPÍTULO 1. INTRODUCCIÓN

De los cinco sentidos que posee el ser humano, la vista es uno de los más importantes, pues éste proporciona el 80% de la información que es tomada del exterior durante el día y casi tres cuartas partes de las células receptoras sensoriales del cuerpo están ubicados en la parte posterior de los ojos: las retinas. Muchas decisiones son tomadas a través de lo que se puede ver y cómo se interpretan esos datos. Esto es la base del desarrollo científico y tecnológico realizado en el campo de la visión por computadora y la aplicación en áreas tan importantes en la actualidad, como son la video-vigilancia y seguridad.

1.1 Motivación y antecedentes

La palabra domótica proviene de la unión de la palabra “domo” y el sufijo “tica”. La primera, etimológicamente viene del latín *domus* que significa casa, y la segunda de la palabra automática. Aunque algunos autores también diferencian entre “tic” de tecnologías de la información y de la comunicación y “a” de automatización. En general, este término se deriva de la palabra francesa *domotique* que es definida como el concepto de vivienda que integra toda la tecnología permitiendo gestionar de forma energéticamente eficiente, segura y confortable para el usuario, los distintos aparatos e instalaciones domésticas tradicionales que los conforman. [1]

La domótica busca el aprovechamiento al máximo de la energía, adecuando su comportamiento a las necesidades, al mismo tiempo que mejora la seguridad personal y patrimonial en las edificaciones. Es precisamente este último punto donde la domótica juega un papel trascendental, pues la seguridad es una de las funciones más importantes debido a que disminuye el riesgo de daños a las personas y a los bienes materiales. Dentro de este rubro, la video-vigilancia es una de las herramientas básicas, pues además de frenar a los intrusos, permite detectar sospechosos antes de que ocurran los hechos delictivos o proporcionar pruebas de delitos.

Sin embargo, la video-vigilancia puede utilizarse en otras funciones, tales como:

- Monitoreo y control de tráfico terrestre, urbano y rural.
- Caracterización del tipo de tráfico terrestre y las horas críticas.
- Obtención de características del tráfico terrestre tales como velocidad, aceleración, entre otras.
- Mejora de procesos industriales.
- Gestión empresarial.
- Actividades médicas.



Además, gracias a la video-vigilancia se soluciona uno de los principales objetivos de la prevención de riesgos, la supervisión ininterrumpida de las eventuales alarmas y averías en los diferentes sistemas de seguridad instalados, descargando a los seres humanos de ciertas responsabilidades. En este sentido, la video-vigilancia puede complementar otros sistemas de monitoreo y control, para de esta forma obtener herramientas que mejoren las directrices para tomar mejores decisiones, e incluso adquirir reportes e históricos sobre los eventos relevantes.

Relacionado con lo anterior, un ejemplo con posibilidad de complemento con video-vigilancia es el monitoreo del ruido ambiental en áreas urbanas, donde el tráfico juega un papel fundamental, siendo un fuerte generador de contaminación acústica [2] [3] [4] [5] [6] [7] [8] [9] [10] [11] la cual se produce, entre otras causas, por las siguientes:

- Intenso tráfico.
- Hacer sonar el claxon, innecesariamente.
- Aceleraciones innecesarias.
- Sirenas con elevado nivel sonoro.
- Ruido de silbatos producido por agentes encargados de controlar el tráfico en situaciones críticas.

Para este caso, se han creado métodos computacionales novedosos para el reconocimiento de patrones de ruido. Dichos trabajos se han presentado en eventos científicos internacionales de alto impacto [12] [13] [14] [15] [16] [17].

1.2 Planteamiento de problemas

- I. La localización automática de un objeto en tiempo real, la detección y procesamiento de eventos relevantes, son temas complejos que requieren de algoritmos y métodos avanzados para su realización, así como imágenes que permitan la extracción de rasgos característicos, siendo estas actividades ejecutadas por equipos sofisticados que suelen tener precios elevados
- II. A fin de disminuir los costos, se emplean equipos de video-vigilancia de baja y media resolución, lo que conlleva a la pérdida de detalle en las imágenes.
- III. Los sistemas comerciales de medias y bajas prestaciones no realizan tareas inteligentes ni comunicación entre cámaras, además de ser generalmente operados de forma manual.
- IV. Monitorear un área abierta grande con una cámara implica perder detalles importantes que permitan el análisis de rasgos, el reconocimiento óptico de caracteres (OCR) o la identificación de personas.

1.3 Justificación

Existe una frase muy famosa que dice que “una imagen vale más que mil palabras”, esto ha caracterizado la evolución de los medios de comunicación y tecnologías. Grandes inventos como la televisión, los videograbadores y las cámaras son un claro ejemplo de la voluntad humana para interactuar con las imágenes, así como con audio y texto.

Tanto para el hogar, la seguridad pública, como para cualquier tipo de negocio o empresa, la videovigilancia se convirtió en una herramienta útil que se ha popularizado en los últimos años, empleándose para el monitoreo continuo de las acciones que realizan los seres humanos. Sin embargo, esto no ha sido suficiente, pues existe la importante necesidad de ampliar la video-vigilancia para combatir los actos delictivos en diversas zonas y ciudades, creando sistemas integrales de monitoreo y seguridad que ayuden a definir las directrices para mejorar la vida de las personas.

Las imágenes son de vital importancia en el análisis de delitos, como evidencia de abusos o simplemente para seguimiento de eventos, sin embargo, la mayoría de los sistemas proporcionan grabación continua mientras ocurra cualquier suceso detectado mediante el sensor de movimiento que tienen intrínsecamente las cámaras, además, esto requiere que un administrador verifique los videos y pueda definir si existió alguna anomalía o evento trascendente.

Una solución posible a estos problemas, es potenciar los sistemas comerciales de video-vigilancia de bajas y medias prestaciones con algoritmos avanzados que realicen actividades de mayor complejidad, como la localización automática de objetos o individuos, para minimizar la actividad del administrador y maximizar la utilidad de cámaras que no necesariamente sean de última tecnología o con las prestaciones más avanzadas.

1.4 Hipótesis

Mediante la combinación de una cámara fija (X_i) de bajas o medias prestaciones y una cámara móvil de mediana o alta resolución, se puede detectar un objeto en múltiples regiones de interés (ROI, por sus siglas en inglés) definidas dentro del FOV amplio y obtener detalles del objeto mediante la cámara móvil, sin necesidad de emplear sistemas costosos de altas prestaciones.

1.5 Alcance

El sistema ejecutará las siguientes actividades:

- I. Dada la presencia estacionaria de un objeto en una región de interés (ROI) detectado por una cámara fija X_i , se realiza el mando de la cámara tipo domo para tomar imágenes de mayor detalle.
- II. Los reportes a usuario se realizan con la imagen adquirida a través de la actividad antes descrita.

La manipulación de las cámaras se realiza con un valor mayor a 0.01 lux, esto es, cuando las condiciones de iluminación son equivalentes a la luz de día o con alumbrado público de mediana intensidad.

Otro punto importante a mencionar es la condición que el tiempo de vida o permanencia del evento, debe ser siempre mayor que el tiempo de respuesta del sistema, que está entre 15 y 20 segundos.

1.6 Límites

En la Fig. 1.1 se muestra el esquema del sistema propuesto, en el cual se tienen X_i cámaras fijas de circuito cerrado, donde $i = 1, 2, \dots, n$, y n depende del FOV de la cámara tipo domo, sin embargo, si este número es muy grande, la operación del sistema puede ser más lenta al presentarse muchos eventos de interés de manera simultánea y aumentar así el tiempo de respuesta del domo.

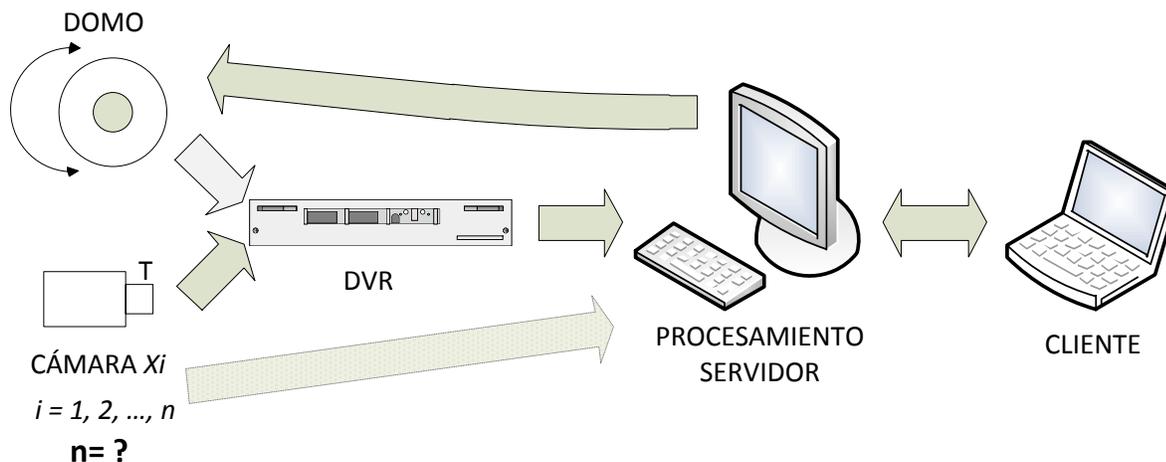


Figura 1.1 Esquema propuesto para video-vigilancia

1.7 Objetivos

1.7.1 Objetivo general

Desarrollar instrumentos virtuales para video-vigilancia y seguridad, con operación local y remota.

1.7.2 Objetivos particulares

- I. Generar métodos, algoritmos y *software* que permitan potenciar sistemas comerciales de video-vigilancia de bajas y medias prestaciones.
- II. Integrar un sistema de video-vigilancia local y remoto, mediante cámaras analógicas y digitales (cámaras IP).
- III. Estructurar un sistema cliente/servidor para la transmisión y recepción de imágenes y avisos sobre eventos relevantes obtenidos mediante un sistema de video-vigilancia.
- IV. Detectar eventos relevantes relacionados con la seguridad en edificaciones, tales como los generados en estacionamientos en zonas restringidas, adquiriendo imágenes y videos con mayores detalles y generando reportes a usuarios.

1.8 Contribuciones

Con el presente trabajo se logrará:

- I. Incorporar desarrollo tecnológico de costo reducido en los sistemas de video-vigilancia comerciales de bajas y medias prestaciones, mejorando su desempeño.

- II. Potenciar los sistemas comerciales de bajas y medias prestaciones, debido al complemento de la detección, procesamiento y localización de eventos relevantes.
- III. Un sistema de video-vigilancia accesible para Pequeñas y Medianas Empresas (PyME) con aplicación en grandes área de vigilancia para edificaciones, vía pública y con posibilidad de ser complemento para un sistema distribuido de monitoreo permanente de ruidos ambientales urbanos.

1.9 Metodología y desarrollo de la investigación

En la Fig. 1.2 se puede observar un diagrama que enuncia todas las etapas que el sistema desarrolla. Cada una forma parte de la metodología que da solución a los problemas expuestos. Con el fin de llevar a la práctica la teoría estudiada, se diseñarán, programarán e implementarán un conjunto de algoritmos que procesan la secuencia de fotogramas comparando el actual con el inmediato anterior para obtener parámetros que faciliten la detección y localización de objetos.

1.9.1 Propuesta de solución

A partir del esquema de un sistema automático como el mostrado en la Fig. 1.1, se obtienen imágenes continuas en tiempo real a través del par por una cámara fija y una móvil.

Las etapas son mostradas en la Fig. 1.2. La primera consiste en la adquisición de video del grabador digital por medio del Protocolo de Control de Transmisión (TCP, por sus siglas en inglés), esto se realiza a través de un conjunto de herramientas de desarrollo propio del grabador y de la cámara IP, con la resolución original estándar del grabador, que es de 336 x 224 píxeles y de la cámara IP, 320 x 240 píxeles.

La presencia de objetos se hace con la evaluación de una ROI para verificar cambios en los ejes de una línea de referencia.

Por otro lado, se realiza un segmentado en el área de video-vigilancia para tener una correspondencia de áreas de vigilancia, como se muestra en la Fig. 1.3, con cada una de las acciones de movimiento que se pueden realizar con la cámara tipo domo y con ello realizar la localización de objetos.

Una vez que se tiene la imagen extraída y un arreglo de áreas, se realiza la comparación: si el objeto de interés se encuentra en cualquier área diferente a la cero (A0), entonces se tiene que realizar el movimiento del domo y posiblemente un acercamiento; de otra forma se detiene el movimiento y comparación, se toma una fotografía; y se genera el reporte a usuario.



Figura 1.2 Diagrama general de etapas

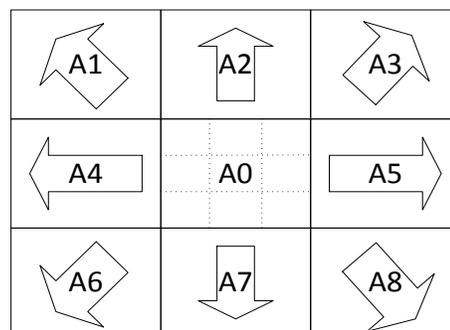


Figura 1.3 Áreas de vigilancia

Las características específicas del sistema son:

- a) La obtención de imágenes a través de X_i cámaras de circuito cerrado que envían el video a un grabador digital (DVR, por sus siglas en inglés) o directamente tomado de una cámara IP, donde $i = 1, 2, \dots, n$ y n depende del campo de visión global de la cámara tipo domo.
- b) Los fotogramas obtenidos serán empleados en los algoritmos programados en la computadora central que también tendrá el rol de servidor, éste proporcionará la información procesada en reportes a usuario a través de una computadora cliente, además de modificar la posición de la cámara tipo domo para realizar la localización de objetos y la toma de fotografías con mayor detalle.
- c) La cámara móvil localiza continuamente el objeto y lo posiciona en el centro de su FOV, modificando las condiciones PTZ (*Pan, Tilt, Zoom*). Es necesario que esta cámara posea enfoque y manejo de diafragma automáticos. Además, es recomendable que el domo tenga un zoom óptico mayor o igual a 22x.

1.9.2 Diagrama

A continuación se muestra la Fig. 1.4, la cual contiene el diagrama sobre la metodología de la investigación.

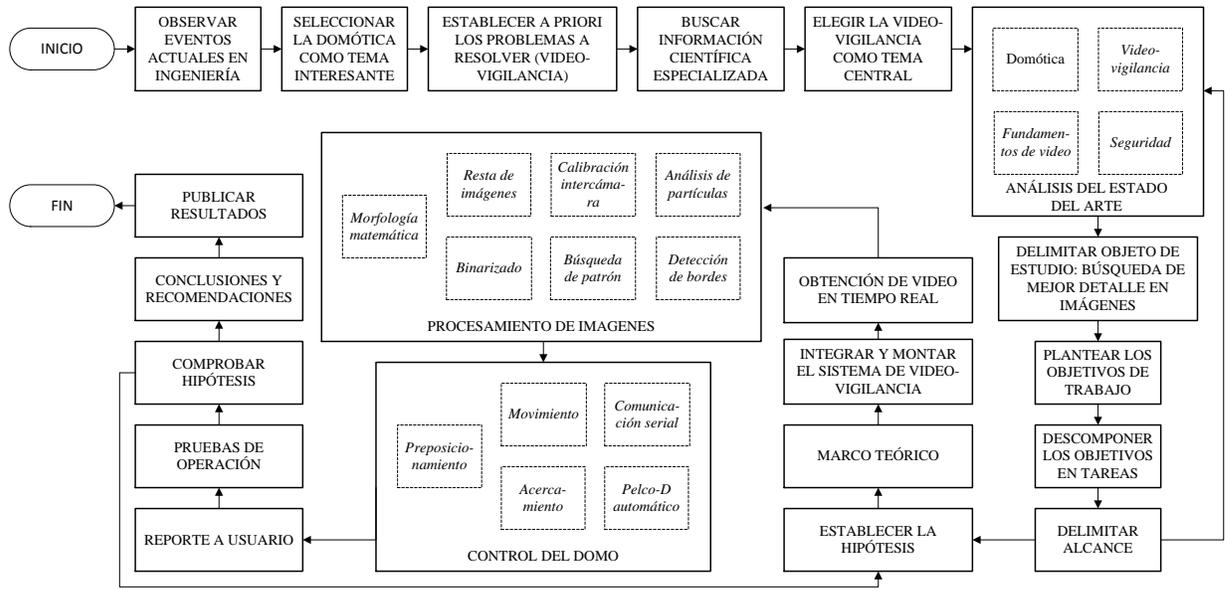


Figura 1.3 Diagrama de caso de estudio

1.10 Estructura de la tesis

Este documento se encuentra dividido en 6 capítulos, bibliografía y anexos. Los primeros dos capítulos explican de forma detallada el procedimiento utilizado para dar solución al problema de estudio; los capítulos restantes muestran la metodología empleada, los resultados, las conclusiones y las propuestas de trabajos futuros.

El capítulo 2 presenta un panorama sobre los trabajos y técnicas desarrolladas en otras investigaciones de la misma área.

El capítulo 3 muestra la base teórica que sustenta la presente investigación, presentando los conceptos y definiciones necesarias.

El capítulo 4 explica el diseño y montaje del sistema propuesto, además la metodología sobre la obtención de video en tiempo real, la selección de regiones de interés, el preposicionamiento y la generación de reportes a usuario.

El capítulo 5 se trata los temas de visión por computadora para realizar la detección y procesamiento de eventos relevantes. Estos temas abarcan la detección de bordes, el procesamiento de imágenes a través de la operación de resta, binarizado y análisis de partículas, así como la calibración intercámara y el reconocimiento de patrones, componente fundamental para lograr la ubicación y acercamiento al objeto de interés, además de la lógica de mando de la cámara tipo Domo

El capítulo 6 se presenta la integración del sistema y la obtención del video en tiempo real, así como imágenes de los resultados de las pruebas finales y los tiempos de ejecución para todas las secciones de la aplicación. Adicionalmente, se muestra un bloque de programación que sienta las bases para realizar el seguimiento de objetos en movimiento.



En los anexos se presenta la recopilación de las funciones más importantes de LabVIEW empleadas en la aplicación; la explicación del protocolo de comunicación y mando de las cámaras que emplean control PTZ; los SDK para la obtención del video en tiempo real; la preconfiguración de posiciones del domo; el diagrama general de jerarquía de software; las condiciones de resolución para los dispositivos; y por último las publicaciones derivadas del presente trabajo.

CAPÍTULO 2. ESTADO DEL ARTE

En este capítulo se muestra la perspectiva general de los temas que fundamentan el presente trabajo, para tener una visión sobre la situación actual de este tipo de desarrollos a nivel mundial.

2.1 Domótica

La evolución de la tecnología en los hogares fue marcada por el uso de la electricidad y el desarrollo de la electrónica y computación, todo lo cual ha permitido elevar el nivel de confort y ha dado paso a la entrada de los electrodomésticos, máquinas capaces de realizar tareas cotidianas de forma casi autónoma y que emplea la electrónica, permitiendo realizar rutinas. El siguiente paso en la evolución es la domótica, que se encarga de la integración y regulación de sistemas eléctricos y electrónicos, de tal manera que ante una solicitud prevista, da la respuesta adecuada dentro de una gama acotada y ordenada al mecanismo correspondiente para que actúe en consecuencia. Incluye tres áreas: confort, ahorro energético y seguridad. Además el desarrollo de la electrónica permitió un control descentralizado de estos procesos; y finalmente la informática permitió una gestión del edificio en su control y centralización. [1]

Actualmente existen sistemas [18] que integran nuevas herramientas computacionales para unificar en una sola aplicación los contenidos en el hogar de servidores en casa y los compartidos por otros usuarios, a través de servicios como Flickr, YouTube, iTunes y Facebook para comunicar contenidos en forma segura y acceder a ellos desde cualquier lugar.

Según algunas fuentes [19], los últimos 20 años de desarrollo no han logrado los resultados previstos. La demanda y el suministro han sido bajos. Para diseñar estos sistemas es necesario considerar la construcción, infraestructura informática y aspectos de prestación de servicios; las necesidades del usuario e intenciones deben ser estudiadas con más detalle y respetadas siempre que sea posible; e investigar los problemas legales y éticos, el usuario y la aceptación del proveedor.

Se cree que el futuro de las redes domésticas dependerá en gran medida de la utilidad de los servicios para el hogar y en su facilidad de uso, donde el objetivo principal será proporcionar una buena experiencia para usuarios que puedan acceder a su "vida personal digital" en un una manera segura en cualquier lugar.

2.2 Video-vigilancia

En el aspecto social, el aumento de delitos es uno de los principales factores que han influido en la demanda de sistemas de video-vigilancia remota [20], pues busca disuadir, prevenir y combatir ilícitos. Sin embargo, pueden aplicarse en otros campos que incluyen la comunicación y mejora de relaciones entre proveedores y clientes que no se encuentran cerca. [21] Es común que grupos de trabajo sean monitoreados para tener registro o consulta en línea de una correcta ejecución de funciones, horarios y hábitos.

En la industria existen procesos elaborados por máquinas o lugares de alto riesgo en los cuales es prácticamente imposible contar con supervisión presencial [22], por lo que la video-vigilancia es útil para tener un monitoreo masivo de procesos. [23]



2.2.1 *Diseño de sistemas de video-vigilancia*

La visión artificial para sistemas de video-vigilancia realiza varias funciones comunes necesarias para el procesamiento de bajo nivel de datos de video. Estos algoritmos incluyen [24]:

- Adquisición de imágenes desde una cámara.
- Calibración de la cámara.
- Segmentación de una imagen en píxeles de primer plano y de fondo.
- Ubicación de los objetos que poseen componentes conectadas.
- Obtención de las descripciones de objetos.
- Seguimiento de objetos dentro del campo de visión de una cámara.

Además de las funciones anteriores, la video-vigilancia también puede realizar el seguimiento de las trayectorias de objetos, almacenamiento y recuperación de recorridos para el análisis a largo plazo.

De manera general, lo indicado por algunos autores [20] [25], un sistema de video-vigilancia avanzado debe contener las siguientes funcionalidades:

- Obtención de imágenes de las cámaras a través de una red privada.
- Transmisión del servidor a los clientes por comunicación basada en HTTP.
- Acumulación temporal de datos como imágenes en un buffer interno del servidor.
- Reducción de tamaño de los datos en el servidor para la comunicación con dispositivos móviles.
- Almacenamiento de imágenes en su base de datos interna para la recuperación remota de clientes.
- Escalabilidad del sistema.
- Operación en tiempo real.
- Bajo consumo de recursos.
- Control de la transmisión del servidor con los clientes.
- Reconfiguración en tiempo de ejecución.

Adicionalmente, San Miguel et al. [25] proponen una clasificación general unificada con otros autores:

- Según el propósito o fin para el cual fue diseñado.
- Según la distribución del sistema.
- Según la centralización del control del sistema.

2.2.2 *Beneficios de la video-vigilancia*

Tecnologías como la video-vigilancia permiten observar, responder, intervenir y mitigar riesgos que se puedan presentar. [26] Los sistemas de este tipo poseen las siguientes bondades:

- Disponibilidad. En los últimos años, la tecnología relacionada con la video-vigilancia se ha expandido y popularizado, existiendo cada vez más compañías dedicadas a desarrollar hardware y software para este fin, por lo que es más fácil adquirir sistemas convencionales de monitoreo y video, generalmente de bajas y medias prestaciones.
- Monitoreo en tiempo real. Gracias al internet, es posible realizar un monitoreo simple a través de una página web. En el pasado, los eventos capturados en video eran usados para recabar información importante y emplearlos como evidencia posteriormente. Sin embargo, las nuevas tecnologías permiten a los usuarios monitorear y responder a las alarmas en tiempo real.

- Acceso remoto. Una de las ventajas más fuertes de la video-vigilancia es la posibilidad de monitorear remotamente desde cualquier lugar mediante dispositivos como computadoras o celulares.

2.2.3 *Sistemas basados en arquitectura cliente/servidor*

Boyd et al. [24] describen un sistema de video-vigilancia que asimila la información de múltiples cámaras en un único modelo de una escena. En una aplicación en tiempo real con procesadores separados para realizar las operaciones de bajo nivel (segmentación de movimiento y reconocimiento de objetos), y luego compartir, a través de una red de área local, la información con otras partes del sistema de vigilancia. Otra propuesta la presenta Sayles et al. [27], quienes describen el CAML (*CAmera Markup Language*), un sistema de cámaras de red, que extrae el procesamiento de bajo nivel referido por el sistema de Boyd et al. [24] para crear un servidor de información de video. Este sistema ofrece, a través de conexiones de red, una descripción de funciones de la cámara, una secuencia de video y un flujo de documentos XML que describen objetos detectados y sus trayectorias en coordenadas de la imagen. Para detectar el movimiento, el sistema calcula el flujo de CAML por óptica de Lucas-Kanade [28], umbrala e identifica grandes componentes conectados de píxeles en movimiento. El sistema está implementado en el lenguaje de programación Python, y hace uso de la librería de desempeño primitivo integrada por Intel [29] para optimizar la velocidad de ejecución.

Así, Sayles et al. [27], presentaron su sistema de servidor CAML con una aplicación cliente que combina las trayectorias de la imagen y los datos de calibración de la cámara (realizada por los servidores) para calcular las trayectorias en un modelo tridimensional de una zona sometida a video-vigilancia. Sin embargo, Boyd et al. [30] mejoran este sistema con innovaciones tales como que los servidores, los cuales son dinámicamente configurables por determinados clientes, proporcionan la descripción del objeto y su trayectoria de manera separada, con comunicación bidireccional, además los clientes poseen nombres de usuario y contraseñas para solicitar información al servidor con lo cual se evitan conflictos de mando a una misma cámara por diferentes clientes. Este sistema tiene ciertas similitudes con las nuevas cámaras MPEG-7 [31] pues la información en los servidores también se maneja con metadatos a través de documentos XML para la comunicación con el cliente.

Por otro lado, Imai et al. [20] proponen un sistema de seguridad para monitoreo con acceso remoto a través de internet operando a través de plataformas como Open PLANET y Java que facilitan las operación de aplicaciones pensadas para teléfonos celulares. Otra alternativa empleando estos dispositivos, es la presentada por Lien et al. [32], donde emplean el *software* Microsoft MSN, el cual permite mandar mensajes instantáneos o mostrar el video en tiempo real al cliente, el cual puede iniciar una sesión regular en MSN desde un equipo remoto o PDA y enviar comandos de texto simple como una solicitud para monitorear las instalaciones.

Muchos sistemas como los anteriormente descritos, emplean tecnologías TCP/IP como protocolo de transmisión y la arquitectura es en su mayoría cliente/servidor (C/S), o navegador/servidor (B/S) si el cliente es un PC. [33] Aún con la ventaja de alta velocidad de datos y la reutilización de *software*, esta solución no está disponible para todas las situaciones, debido a su diseño de *hardware* complicado y alto consumo de energía [34]. Para eliminar estos inconvenientes, Yang et al. [35] proponen un sistema basado en Linux con el protocolo de comunicación Bluetooth con arquitectura cliente/servidor, pues, exponen, que en comparación con otras soluciones, ese protocolo tiene una menor tasa de datos, pero más



fácil de implementar en *hardware*. Combinado con MPEG-4, que es una tecnología de codificación de video efectiva, Bluetooth puede ser una valiosa exploración en el sistema de video-vigilancia.

2.2.4 *Sistemas de video-vigilancia avanzados*

En los sistemas de video-vigilancia, existen algunos que realizan el seguimiento de objetos de interés, ya sea con una única cámara o con sistemas más avanzados. En referencia al primer caso, Varcheie y Bilodeau [36] realizaron un proyecto donde buscan identificar y reconocer objetos, así como detectar y seguirlos, empleando una cámara con movimiento (PTZ) con comunicación vía Internet; este método se basa en la comparación de muestras elípticas con el objetivo, evaluando la probabilidad de semejanza para estimar la localización del objeto, a través de un estudio difuso sobre las áreas sin movimiento y con movimiento, empleando el algoritmo de flujo óptico de Lucas Kanade. Así mismo se realiza la predicción de posición del objeto para ayudar al seguimiento. Este sistema tiene la peculiaridad que es necesario tomar la muestra manualmente, lo que lo hace útil para identificación de personas previamente elegidas.

Otro sistema innovador es el propuesto por Chu-Sing et. Al [37], que además de emplear algoritmos de seguimiento para interiores, también adiciona la localización geográfica de personas.

Existe otro sistema propuesto por Hsien-Chou y Wei-Yi [38], el cual propone un sistema de cámara dual, empleando cámaras con movimiento PTZ para asemejar la identificación y seguimiento de un objeto como lo hacen las águilas, empleando la primer cámara para ver el panorama general y la otra para el detalle en una misma área.

2.3 Elementos de un sistema de video-vigilancia

Un Circuito Cerrado de Televisión (CCTV) es indispensable para el sistema de seguridad inteligente, los cuales hoy en día son operados de forma manual en los centros de control [39]. En general los sistemas digitales de video-vigilancia están compuestos de tres módulos:

- Unidades de captura de video. Son un conjunto de cámaras digitales o analógicas con el apoyo de un dispositivo codificador de video que es capaz de realizar la transición analógico-digital. En este módulo se captura el video, comprime los datos en bruto y la codifica en un formato popular estándar (MPEG, JPEG, H261, H263, entre otros).
- Red transmisión. Proporciona el flujo de video codificado a través de una red que puede ser de área local (LAN) o incluso de Internet.
- Módulo de control central. Puede visualizar y grabar cada canal de video. También controla las acciones de las cámaras mediante el envío de comandos de control

Por otro lado, dos tipos de datos fluyen entre los módulos:

- Control de datos. El transmisor de control y mando puede ser una computadora, un teclado de control o una matriz de conmutación. El destinatario puede ser una cámara IP PTZ (*pan-tilt-zoom*), un conmutador de matriz, o un DVR.
- Secuencia de video. El video digital requiere hasta 4 Mbps de ancho de banda de red.

El sistema propuesto por Hintermaier y Steinbach [40] presenta un sistema novedoso para aplicaciones de asistencia al conductor con una cámara donde el funcionamiento ejemplifica la arquitectura general de un sistema de video-vigilancia:

La óptica está determinada por el campo de aplicación y caracterizada por el ángulo de abertura. La conversión de una imagen óptica en una señal digital se realiza por un sensor de imagen. Las técnicas de compresión más adecuadas a la perspectiva de hoy son el estándar de JPEG y H.264. Y por último el protocolo de red para esta aplicación es IP / Ethernet.

2.3.1 Unidades de captura de video

Las cámaras de seguridad juegan un papel importante de los sistemas de video-vigilancia y varias empresas han diseñado y producido muchos tipos [20]. A continuación se describe el manejo de algunas de ellas:

a. Analógica

Los sistemas de video-vigilancia se comenzaron a desarrollar en la década de 1970, incorporando formatos de video de tipo NTSC (Comisión Nacional del Sistema de Televisión) y PAL (Fase de Línea Alternada) obtenido de las cámaras analógicas convencionales que son conectadas por medio de un cable coaxial, inicialmente a grabadoras de VHS y más recientemente a grabadores de video digital, en una estación de monitoreo. Algunos sistemas analógicos tienen la desventaja de proporcionar el video en blanco y negro, además de mantener bajos niveles de calidad en las imágenes, con un pobre procesamiento de la señal y la necesidad de grabar la información en una cinta, todo lo cual puede no ser una prueba de delito en un tribunal de justicia.

b. Domo

Debido a la capacidad de *pan-tilt-zoom* (PTZ) para cubrir al mismo tiempo una zona panorámica además de mantener imágenes de alta resolución y con el fin de ampliar la gama de video-vigilancia, las investigaciones en este tipo de sistemas automáticos han empleado cámaras tipo domo, pero su movimiento aumenta la dificultad de la segmentación del objeto en movimiento para aplicaciones avanzadas.

Según Liu y Xiao [39], hay muchos algoritmos de reconocimiento en la video-vigilancia, pero los primeros pasos de ellos son comunes: la segmentación de objetos en movimiento. Ellos proponen un algoritmo donde toman como referencia una imagen y después procesan la información del movimiento empatando las demás; esto se aplica para calcular la distancia el movimiento horizontal de la secuencia de imágenes en escala de grises.

Chen et al. [35] presentan un interesante diseño para la manipulación de múltiples domos a través de parámetros como los ángulos de *pan*, *tilt* y posiciones objetivo en x, y, z para el *zoom*, en el cual dividen



las fases de operación en dos: la de adquisición, cuyo propósito es recoger información deseada para relacionar los parámetros; y la de montaje, que emplea un modelo de aproximación polinomial unificado entre los valores de giro (*pan*) e inclinación (*tilt*) de las cámaras PTZ con parámetros intrínsecos desconocidos y configuraciones del sistema en la escena.

c. Con Protocolo de Internet (IP)

La tecnología de internet hace que la vigilancia por video sea mucho más fácil de aplicar y más rentable que el cableado tradicional en los CCTV. Si bien hay varias ventajas de la tecnología IP para sistemas de video-vigilancia, los sistemas también están expuestos a nuevas amenazas de seguridad, lo cual es un factor crítico al hacer que estos sistemas sean prácticos y utilizables. [41]

Combinando cámaras tradicionales y la tecnología de video en red, la cámara IP puede comprimir y entregar los clips de video en directo a Internet sin utilizar una computadora para que el usuario pueda ver de forma remota las áreas protegidas. Por lo tanto, las cámaras IP son dispositivos clave de los sistemas de vigilancia actuales. El rendimiento de las cámaras de vigilancia IP es generalmente evaluado en velocidad de codificación, resolución de video, velocidades de transmisión de fotogramas y distorsión, dispersión de potencia, entre otros. [42]

2.3.2 *Sensores de movimiento*

Existen dos tipos principales de sensores de imagen:

- CMOS (*Complementary Metal-Oxide-Semiconductor*) es el más empleado. Detecta miles de imágenes por segundo. Posee un conjunto de elementos electrónicos para cada píxel, lo cual da la posibilidad de elegir cierta área de visión y con ello aumentar la velocidad. La principal desventaja es que es difícil su uso en exteriores.
- CCD (*Charge-Coupled Device*), es más barato que el CMOS. Controla millones de píxeles y proporciona una mejor calidad de imagen.

Una nueva propuesta surgió en 2004 [43] cuyo objetivo era ampliar las capacidades de los sistemas de sensor de corriente y garantizar la seguridad de vulnerabilidades que de otro modo no se habían tratado. Este tipo de sensor no sólo detecta movimiento tridimensional de video, sino también proporcionan información de alto nivel que no estaba disponible con otras tecnologías de sensores volumétricos, cuyos datos de la evaluación están disponibles de manera inmediata al operador, lo que permite tomar decisiones críticas en el momento oportuno. Este sistema emplea una red de sensores independientes que opera en una plataforma PC. El inconveniente de esta innovación es que si alguna parte del objeto bajo observación no es vista por dos cámaras al menos, entonces no es parte de la detección.

2.3.3 *Manipulación de las cámaras: zoom óptico y digital*

Para entender el funcionamiento del *zoom* óptico es necesario conocer la base física de la fotografía: la luz atraviesa una lente y concentra la información en una zona, donde se encuentra el sensor. En función de la distancia que haya entre la lente y el sensor, la imagen resultante será mayor o menor, a esto se le conoce como distancia focal, a partir de la cual las fotos que obtenemos se ven aumentadas. La calidad de este aumento se debe a la propia calidad de la lente y a la información que pueda recoger el sensor.

La intención de emplear el *zoom* en una cámara de video es la obtención de información visual más detallada que el punto de vista de fotograma completo. Los mecanismos de *zoom* óptico proporcionan esta información adicional, sin embargo, cuando el *zoom* es óptico, la cámara deja de detectar el punto de vista de fotograma completo. En aplicaciones tales como la video-vigilancia, es necesario detectar el punto de vista de imagen máxima en todo momento, sin embargo, la cámara debe ser capaz de proporcionar información visual adicional cuando se acerca a la región de interés (ROI, por sus siglas en inglés). Los lentes de *zoom* óptico son caros y difíciles de mantener. La región de conocimiento (ROA, por sus siglas en inglés) [44] representa el campo que es constantemente monitoreado por la cámara.

Por lo general, la ROA se espera que sea área mucho más grande en comparación con la ROI. Sin embargo, con una cámara de *zoom* óptico, la ROA puede ser igual a la ROI.

Para supervisar estas áreas de estudio es necesario tener un sistema para manipulación PTZ.

Al emplear el *zoom*, el ROA se mantiene constante, mientras que el ROI es ajustada a la salida de video al tamaño del marco, en este caso, esta ampliación puede un causar un desenfoque de video y el deterioro de la calidad de imagen; en comparación con la solución de *zoom* óptico producido por algoritmos avanzados, que producen la misma resolución y calidad de imagen en todos los niveles de *zoom*. Un método para mejorar la resolución de captura electrónica es el uso de un sensor de imagen de gran tamaño [44] o desplegar varias cámaras para capturar el ROA [43].

Un método popular de *zoom* electrónico es producir el marco de color de la imagen completa de la ROA total y seleccionar la ventana de la ROI para ajustarse al marco de salida. Sin embargo, este método incurre en cálculos innecesarios en la reproducción del color en las áreas de la ROI y lleva a cabo la ampliación en el color interpolado del fotograma de video de entrada [45]. En los sistemas de un solo sensor de imagen electrónica, el color de la escena se adquiere por el submuestreo en los planos de tres colores para captar los datos de color de imágenes de forma simultánea para rojo, verde y componentes de color azul. Por lo general esto se logra mediante un conjunto de filtros de color (CFA) en un conjunto de sensores 2D. Un tipo de CFA llama el patrón de Bayer [45] se muestra en la Fig. 2.1:

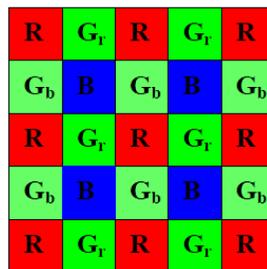


Figura 2.1 Patrón de Bayer

2.3.4 Grabador de video digital

Un grabador de video digital (DVR, por sus siglas en inglés) emplea una unidad de disco duro para almacenar video. Tiene varias ventajas sobre los grabadores antiguos, incluyendo una mayor fidelidad, acceso aleatorio y la capacidad de grabar y reproducir simultáneamente. Hasta hace poco, este sistema era



prohibitivamente caro; las reducciones de circuitos integrados y los costos de la unidad de disco han hecho de éste un producto viable.

Los elementos fundamentales de un grabador de video digital son: [46]

- Codificador.
- Método de compresión.
- Formato de datos.
- Procesador y sistema operativo para control.
- Interconexión de componentes periféricos (PCI, por sus siglas en inglés).
- Procesador de video.

Actualmente existen grabadores de video digital con funciones muy avanzadas de video análisis, tales como conteo de personas y objetos, intrusión y seguimiento, detección de objetos removidos y olvidados, detección de multitudes y hasta detección de rostros. Además proporciona otras funcionalidades tales como notificación de eventos por correo electrónico, acceso por un explorador de internet, control remoto de parámetros y de cámaras y compatibilidad con teléfonos celulares inteligentes.

2.4 Fundamentos de video

Una imagen fija es una distribución espacial de la intensidad que es constante con respecto a tiempo, donde la intensidad se entiende por una medida en un intervalo del espectro electromagnético de la corriente de energía que se irradia desde la superficie. Por lo general se mide en Watts por metro cuadrado.[47]

El video, por el contrario, es un patrón espacial de la intensidad que cambia con el tiempo, es decir, es una secuencia de imágenes.

Existen algunos parámetros importantes para considerar en las señales de video que a continuación se enumeran:

- Relación de aspecto: es la relación entre el ancho a la altura de un marco.
- Resolución vertical: es el número de líneas a escanear por marco.
- Velocidad de marcos: es el número de fotogramas escaneados por segundo. El efecto de movimiento suave se logra utilizando una velocidad de 25 a 30 marcos por segundo. Sin embargo, con esta tasa, el ojo humano capta el parpadeo producido por la actualización de la pantalla entre los marcos. Para evitar esto, la tasa de actualización de la pantalla debe estar por encima de 50 Hz.

La industria emplea diferentes combinaciones de parámetros de video. Por ejemplo, las computadoras utilizan la exploración progresiva, con una velocidad de fotogramas de 72 marcos por segundo. La industria de la televisión utiliza de 25 o 30 marcos por segundo para mantener el movimiento suave.

2.4.1 Video analógico

El término de video analógico se refiere a la señal eléctrica unidimensional que se obtiene con el muestreo del patrón de intensidad de video en las coordenadas vertical y temporal y convirtiéndola en una representación eléctrica. Este proceso se conoce como escaneo.[47]

La exploración de la trama de comienza en la esquina superior izquierda y avanza en sentido horizontal con una ligera pendiente vertical, a través de la imagen, como se ve en la Fig. 2.2. Cuando se llega al extremo derecho, regresa a la esquina superior izquierda para comenzar con una nueva línea. Al llegar a la esquina inferior derecha, un fotograma completo ha sido escaneado y la digitalización se ajusta a la esquina superior izquierda para comenzar un nuevo marco.

Hay tres principales sistemas de video analógico, en la mayoría de Europa occidental, el sistema PAL 625/50 se utiliza; en Rusia, Francia, oriente medio y Europa del este, el sistema 625/50 SECAM se emplea; en América del norte y Japón, el sistema de 525/60 NTSC se utiliza.

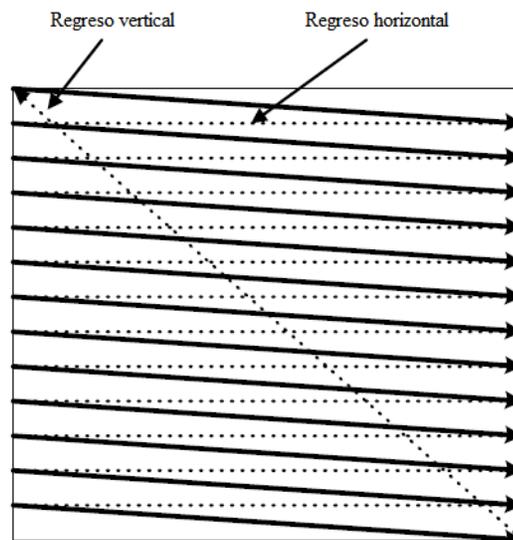


Figura 2.2 Escaneo en video analógico

2.4.2 Video digital

El video digital tiene las siguientes ventajas sobre el video analógico:

- Facilidad de editar, mejorar y crear efectos especiales.
- Conversión de *software* fácil de un estándar a otro.
- Robustez al ruido y la facilidad de cifrado.
- Facilidad de escalabilidad (espacial, temporal, o relación señal-ruido, SNR, por sus siglas en inglés). Esto facilita la prestación del mismo servicio en una amplia gama de redes y plataformas de hardware.
- Interactividad.
- Facilidad de búsqueda y recuperación.

Para convertir el video analógico a digital es necesario realizar la digitalización, que consiste en el filtrado, muestreo y cuantificación: la operación de filtrado se utiliza para limitar la banda de la señal de

entrada y acondicionarlo para la operación de muestreo siguiente; la amplitud de la señal analógica filtrada es luego muestreada en los instantes específicos para generar una señal discreta; la tasa mínima de muestreo se conoce como la tasa de Nyquist y es igual a dos veces el ancho de banda de la señal. La operación de cuantificación se utiliza para asignar esos valores en un conjunto finito de amplitudes discretas que pueden ser representadas por un número finito de bits.

Cada tiempo discreto y amplitud discreta de la muestra es llamada un elemento de imagen y es generalmente abreviado a una pel o un píxel. El video digital se compone de una secuencia de tales marcos. Para video en color, las operaciones anteriores se repiten para cada componente, generalmente 3 básicas de dos dimensiones (rojo, verde y azul).

2.4.3 Codificación

La codificación de video o compresión es una tecnología clave para compactar la información contenida en un video, de manera que se pueda almacenar y reproducir de una forma más conveniente al uso y transportación de los datos. Así, el objetivo de la codificación de video es el de reducir el número de bits utilizados, eliminando las redundancias presentes en la señal de video, las cuales son:

- a) Estadística: está presente porque ciertos patrones de datos son más probables que otras. Esto se debe principalmente a la alta resolución espacial (*intraframe*) y temporal (*interframe*) que correlaciona a los píxeles vecinos.
- b) Psicovisual: se debe al hecho de que el sistema visual humano (HVS) es menos sensible a cierta información visual que a otra.
- c) De código: esto ocurre cuando el video está codificado en una forma que usa más símbolos de los que son absolutamente necesarios.

Para realizar esta reducción de bits son necesarios tres elementos indispensables que tiene el codificador:

- a) El mapeador transforma los datos de entrada en bruto en una representación que reduce los datos más susceptibles a la compresión en etapas posteriores. La transformación es uno a uno y el mapeo, por tanto, reversible.
- b) El cuantificador reduce la precisión de la salida del mapeador, de acuerdo con algún criterio de fidelidad, para reducir la redundancia psicovisual. Este es un mapeo de muchos a uno, por tanto, irreversible.
- c) El codificador de símbolos asigna un código de acceso, una cadena de bits binarios de cada símbolo en la salida del cuantificador. El código disminuye la redundancia de codificación. Esta operación es reversible.

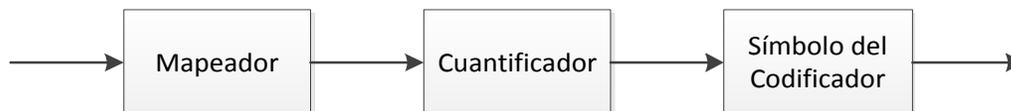


Figura 2.3 Elementos del codificador

En general, los métodos de compresión se pueden clasificar en métodos sin pérdidas y con pérdida. En los primeros, la reconstrucción de datos es idéntica a los datos originales. Esto significa que tales métodos no utilizan un cuantificador. En los métodos con pérdida, la reconstrucción no es idéntica a los datos

originales debido a la cuantificación del proceso. Tales métodos son por lo tanto, irreversibles, y por lo general logran una mayor compresión de los métodos sin pérdidas.

2.4.4 Color

Según la teoría tricromática de la visión del color [48], éste se percibe a través de tres tipos de conos o foto receptores en el ojo. En consecuencia, un video en color puede ser producido por la superposición de tres señales. Cada señal representa uno de los tres colores primarios: rojo, verde y azul (RGB, por sus siglas en inglés), que es un sistema de color aditivo. Esto significa que cuando todas las componentes primarias son añadidas en cantidades máximas iguales, el color blanco es percibido.

Los sistemas de video normalmente convierten esta representación RGB a un espacio de color diferente de la luminancia (que está estrechamente relacionada con la percepción de la luminosidad) y crominancia (que se relaciona con la percepción de la tonalidad de color y saturación). Esta representación tiene dos propósitos. En primer lugar, la luminosidad asegura la compatibilidad con video monocromo. En segundo lugar, esta representación se presta más fácilmente a la compresión de video.

Hay tres principales sistemas de codificación de color: PAL (*Phase Alternation Line*), SECAM (*Sequential Couleur Avec Memoire*) and NTSC (*National Television System Committee*). Se diferencian principalmente en la forma de calcular los componentes de luminancia y crominancia de los componentes RGB.

2.4.5 Formatos de video

Los datos de compresión y descompresión han sido desde hace mucho tiempo la respuesta a las preocupaciones de seguridad y autenticación. En los últimos años, varios algoritmos y técnicas se han utilizado para asegurar las comunicaciones de datos. [49]

Los estándares internacionales permiten que las imágenes y video puedan ser procesados en diferentes plataformas de hardware, ser almacenadas y ser transmitidas en redes de comunicación. Esta interoperabilidad abre un gran mercado para equipos de video y al mismo tiempo ofrece a los consumidores una amplia gama de servicios.

A continuación se describe brevemente el formato de video comercial más empleado en la actualidad: [47]
[48]



Tabla 2.1 Características de formatos de video MPGE-4

ESTÁNDAR	FECHA DE CREACIÓN Y DESARROLLADOR	APLICACIONES	SUBTIPOS, SEMEJANZAS CON OTROS ESTÁNDARES	TÉCNICAS DE PROCESAMIENTO
MPEG-4	1993, ISO/IEC JTC1/SC29/WG11 MPEG	Empleado en comunicaciones, informática y entretenimiento (TV/cine)	Agrega contenido interactividad, la compresión y el acceso universal	Utiliza un modelo de representación basado en objetos, en donde se representa una escena, codificada, y manipulada como objetos audiovisuales individuales

2.5 Seguridad

El enfoque desarrollado para resguardar la seguridad en la construcción y automatización de los sistemas en las edificaciones (BACS, por sus siglas en inglés) trata de armonizar dos enfoques distintos sobre la seguridad, para reducir el riesgo en las construcciones. [50] Aunque en español la palabra seguridad tiene un uso general, en el idioma inglés pueden verse más claramente las dos variantes: *security* y *safety*.

Safety, está definido bajo la norma IEC 61508 y se refiere a la protección de las personas, entendiendo esto como procurar la ausencia de riesgos inaceptables de daño físico o daño a la salud de las personas [50], teniendo presente que este concepto atiende a la idea sobre si en un lugar los habitantes están seguros o no.

Por otro lado, *Security*, esta normalizado por el proyecto de Criterio Común (CC) desarrollado en Estados Unidos en 1993, específicamente dentro de ISO/IEC 15408 y trata sobre la protección al abuso de los recursos y la información [51], así como buscar patrones de eventos delictivos previamente planeados a través de procedimientos para encontrar pruebas y preservarlas.

En relación con lo anterior, la domótica busca mantener la seguridad integrando tres campos que normalmente están controlados por sistemas distintos: [1]

Seguridad de las personas (*safety*)

1. Teleasistencia y telemedicina para las personas mayores, enfermos o discapacitados.
2. Acceso a los servicios de vigilancia sanitaria, policía, etc.

Incidencia y averías (*safety*)

1. La detección y control de todo tipo de averías de agua, gas, etc.
2. Detección de incendios y alarmas
3. Detectar averías en los accesos, en los ascensores, o cualquier otro sistema.

Seguridad de los bienes (*security*)

1. Gestión del control de acceso con reconocimiento o identificación de los usuarios.
2. Control de presencia y detección de intrusismo y de la posterior persuasión.
3. Detección de rotura de cristales y forzado de puertas.
4. Simulación de presencia, memorizando acciones cotidianas para su repetición.
5. Video-vigilancia a través de cámaras

2.6 Configuración de sistemas comerciales

Usar una IP estática en una red tipo LAN/WAN, requiere de altos gastos de administración de la red y costos para los usuarios. Por tanto, los sistemas comerciales soportan direcciones IP dinámicas, para redes usando DHCP (*Dynamic Host Configuration Protocol*). Esto reduce los gastos para los usuarios y costos de los sistemas. [52]

Los sistemas comerciales de video-vigilancia están conformados, en general, por:

- Grabador de video, que entre otras características, posee entradas para una o más cámaras de video, salida de video para monitor, codificación de diversos formatos de video, modos de grabación (por evento y/o por tiempo) y conexiones para alarmas.
- *Software* de usuario.
- Servidor de sistema.

El proceso de operación de estos sistemas para monitoreo remoto es el siguiente:

1. Se registra el nombre de la unidad y la dirección IP del grabador en el servidor.
2. En el *software* de usuario, se busca la dirección IP con que fue registrado el grabador de video para comunicarse con el servidor.
3. El servidor notifica al *software* de usuario que la dirección IP es válida y se ha dado de alta.
4. Se conecta el grabador de video usando la dirección IP a través del *software* de usuario para entablar comunicación continua con el servidor.

Para direcciones dinámicas, es necesario realizar los pasos b) al d) cada vez que se conecte el grabador de video, con la finalidad de tener siempre una comunicación exitosa.

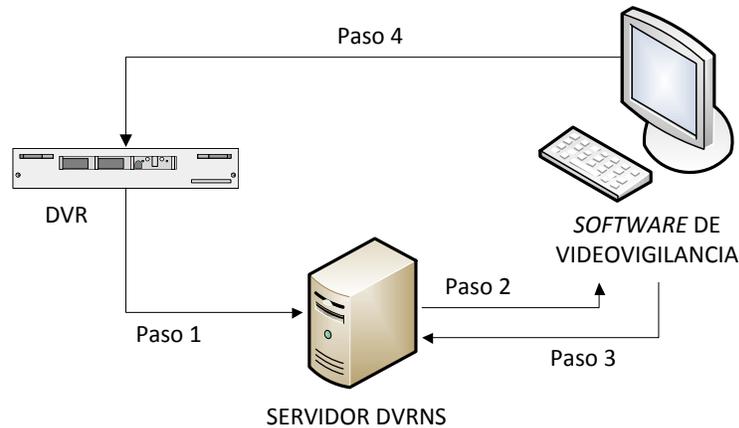


Figura 2.3 Configuración de un sistema de video-vigilancia comercial

2.7 Conclusiones

La domótica es un conjunto de sistemas y técnicas que buscan satisfacer tres objetivos principales: mejorar el confort, emplear de manera eficiente la energía y mejorar la seguridad. Es precisamente en este último punto donde existe una herramienta que ha tenido auge en la última década: la video-vigilancia; ésta busca desestimular a los intrusos y detectar sospechosos antes de que ocurran los eventos, así como proporcionar pruebas sobre abusos o delitos con imágenes y video.

El equipo que integra este tipo de sistemas se define de acuerdo al diseño o requerimientos que se deseen cubrir. Sin embargo, básicamente se deben considerar las unidades de captura de video, conociendo el tipo de sensor de detección, la red de transmisión y el modulo de control central que almacena la información no siempre valiosa, pues los sistemas comerciales de video-vigilancia de bajas y medias prestaciones no realizan operaciones inteligentes tales como el procesamiento de video o la comunicación intercámara.

Debido a lo anterior, se han desarrollado algoritmos para el procesamiento cuadro por cuadro en un video para desarrollar tareas más complejas. Entre los sistemas de este tipo, destacan los llamados “sistemas de cámaras dual”, donde se combinan distintos tipos de cámaras que tienen una interrelación para detectar, localizar y tomar una imagen final con más detalle del objeto de interés.

CAPÍTULO 3. MARCO TEÓRICO

En este capítulo se definirán los términos importantes para el presente trabajo, así como la justificación sobre su empleo para el desarrollo de la aplicación.

3.1 Comunicación entre dispositivos

La comunicación es un conjunto de métodos que buscan compartir información entre dispositivos. Sin embargo, no es suficiente que los equipos “hablen” el mismo lenguaje, también es necesario que las aplicaciones posean cierta estructura.

Para que una aplicación computacional esté bien organizada, se tienen tres niveles de desarrollo, lo que permite maximizar el reuso del código; el tiempo de mantenimiento y modificación es minimizado; y el código gana abstracción. Estas capas son: [53]

- Principal: proporciona las funciones de interfaz de usuario y prueba la ejecución de las aplicaciones.
- Lógico: este nivel es llamado por el principal; contiene la parte de prueba y medidas de subrutinas.
- De controlador: este nivel entabla las comunicaciones necesarias para que los instrumentos y dispositivos sean usados. Los controladores pueden ser clasificados en: los empleados para medición; los de configuración; y los controladores de estado que verifican el estado de un instrumento por medio de la lectura de sus registros.

Cuando la comunicación con un dispositivo externo o programa se está realizando, las operaciones de entrada y salida se ejecutan en el nivel lógico o inclusive en el nivel principal, en lugar de utilizar un controlador.

3.1.1 Comunicación serial como estándar

Como se dijo en la sección anterior, los controladores o *drivers*, como son comúnmente conocidos, se emplean para mejorar la legibilidad, reutilizar el código y mejorar la velocidad de la aplicación. El propósito principal de un controlador es abstraer el código de bajo nivel.

Actualmente, hay muchas maneras de entablar comunicación entre dispositivos. La comunicación serial es una de las más empleadas. Una de sus ventajas con respecto a otros métodos es su disponibilidad, pues casi todas las computadoras poseen un puerto serial. Otro beneficio es la habilidad que presenta al controlar dispositivos a grandes distancias.

Existen muchos equipos que emplean la comunicación serial incluyendo RS-232C y RS-422A que soportan la llamada comunicación *full duplex*, como se ve en la fig. 3.1 (a). Mientras que en la fig. 3.1 (b) se observa el formato de RS-485 que trabaja bajo la comunicación *half duplex*. La primera tiene muchas ventajas sobre la segunda, sin embargo, debido a las propiedades eléctricas, la distancia de comunicación para RS-232C es solo de 15 m, en cambio, para RS-485 alcanza los 1200 m. Así, tanto el RS-422A, como RS-485 tienen casi las mismas propiedades eléctricas, con dos extremos TXD y RXD. [54]

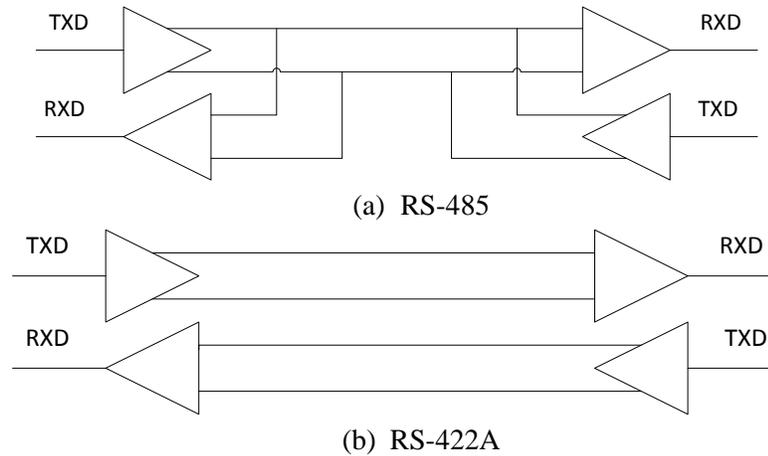


Figura 3.1 Comunicación serial

Además de lo anterior, otras ventajas de emplear RS-485 son las siguientes:

- Estable a interferencias (anti-jamming).
- Amplio uso industrial.
- Emplea el mismo cableado para conectar dos dispositivos.

3.1.2 Comunicación serial en LabVIEW: VISA

Para tener comunicaciones exitosas entre el puerto serial y un dispositivo, la configuración en ambos debe ser igual, de manera que todos los parámetros sean correspondientes; éstos son:

- Tamaño del buffer.
- Número de puerto.
- Velocidad de transmisión.
- Número de bits de datos.
- Número de bits de paro.
- Datos de paridad.
- Control de flujo de datos.

Desde la versión LabVIEW 7, la comunicación serial tradicional fue cambiada por el conjunto de herramientas llamadas VISA (*Virtual Instrument Software Architecture*); es una API para entablar comunicación con instrumentos de entrada y salida.

Una de las ventajas más notables de VISA es que es una plataforma independiente, pero esto no quiere decir que existan problemas con otras aplicaciones, pues VISA realiza las conversiones pertinentes. Es importante recordar que en la comunicación serial es necesario que, una vez que se termina la comunicación, el puerto sea liberado, pues no se podrá utilizar nuevamente hasta que LabVIEW sea cerrado completamente.

Además de VISA, otra variante en la comunicación serial ha sido ampliamente usada USB (*Universal Serial Bus*), que permite conectar dispositivos en un puerto común, y proporciona la habilidad de sustituir los elementos sin necesidad de detener o alterar la operación normal de la computadora donde se encuentran; además de poder conectar más de un elemento USB al mismo puerto gracias a los concentradores. [53] USB permite una velocidad de transferencia de 1.5 Mbps a 5 Gbps.

3.1.3 Comunicación por TCP/IP

Hay tres principales protocolos de comunicación a través de redes: Protocolo de Control de Transmisión (TCP), Protocolo de Internet (IP) y Protocolo de Datagramas de Usuario (UDP). TCP está encima de la capa de IP; divide los datos en paquetes para enviar a IP. TCP también realiza la comprobación de los datos para asegurar que los datos lleguen a su destino de forma completa. Los datos TCP/IP constan de 20 bytes de información de IP, seguido de 20 bytes de información de TCP y finalmente los datos que se envían.

Cada computadora en una red IP tiene una dirección de Internet exclusiva. Esta dirección es un número entero de 32 bits que se divide en cuatro enteros de 8 bits separados por puntos decimales. El Servicio de Nombres de Dominio (DNS) del sistema es una base de datos de las direcciones IP asociadas cada una a nombres únicos.

Hay una serie de estándares que utilizan TCP/IP que se puede implementar utilizando LabVIEW. Por ejemplo, SMTP se utiliza para aplicaciones de correo.[53]

En cuanto a operación, al igual que en la comunicación serial, es necesario cerrar la sesión TCP, pues de otra forma ninguna otra aplicación podrá entablar contacto por este medio.

3.2 Procesamiento de imágenes

Las imágenes que han sido adquiridas por las unidades de captura no tienen el formato y características adecuadas para analizarlas, por lo que es necesario tratarlas de manera adecuada, a esto se le conoce como procesamiento.

Por convención, la referencia espacial del píxel con coordenadas (0,0) está localizada en la esquina superior izquierda, por tanto el incremento en horizontal es a la derecha y el vertical hacia abajo como se observa en la Fig. 3.2, donde $f(x,y)$ es un píxel de la imagen digital.[55]

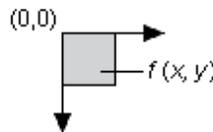


Figura 3.2 Referencia de píxel

3.2.1 Conectividad

El procesamiento de imágenes altera el valor de cada píxel y de su vecindario en el fotograma; a esto se le conoce como Conectividad.

Los arreglos de Conectividad más usados son 4 y 8, aunque también existen algunas ocasiones en que se emplea la configuración 6-Conectado. A continuación la explicación de cada uno. Considérese una imagen como la vista en la Fig. 3.3. [53]

00	01	02	03	04
10	11	12	13	14
20	21	22	23	24
30	31	32	33	34
40	41	42	43	44

Figura 3.3 Píxeles de una imagen

- 4-Conectado

Se refiere a la rutina de procesamiento de imágenes en forma cardinal. Supóngase que el píxel actual es el número 22, para analizar su vecindario, se toman los píxeles 12, 21, 23 y 32 como se muestra a continuación:

00	01	02	03	04
10	11	12	13	14
20	21	22	23	24
30	31	32	33	34
40	41	42	43	44

Figura 3.4 4-Conectado

- 6-Conectado

Representa una conectividad hexagonal, que se puede ver más fácilmente si se traslapa media posición alternadamente las filas de píxeles, de manera que el procesamiento es como sigue:

00	01	02	03	04
10	11	12	13	14
20	21	22	23	24
30	31	32	33	34
40	41	42	43	44

Figura 3.5 6-Conectado

- 8-Conectado

Es la conectividad más usada. Se realiza el procesamiento en el vecindario del píxel actual de manera cardinal y subcardinal. De manera que el vecindario se ve en la Fig. 3.6.

00	01	02	03	04
10	11	12	13	14
20	21	22	23	24
30	31	32	33	34
40	41	42	43	44

Figura 3.6 8-Conectado

3.2.2 Regiones de interés

También llamadas subimagen, una ROI (*Region of Interest*) es una estructura de píxeles que permite la definición solo de un área de la imagen dada, siendo ésta un marcador de posición que recuerda una ubicación dentro de una imagen, para realizar el procesamiento y análisis.[53] [55]

Las propiedades de una ROI son:

- Puede tener forma irregular, no necesariamente rectangular o cuadrada; inclusive rotada.
- Al tener varias ROI de una misma imagen, se pueden procesar por separado, evitando hacer cálculos para todo el marco y perder eficiencia computacional.
- Se pueden modificar las ROI dinámicamente durante la ejecución.

En la Fig. 3.7 se muestra una imagen que contiene diversas formas de regiones de interés.

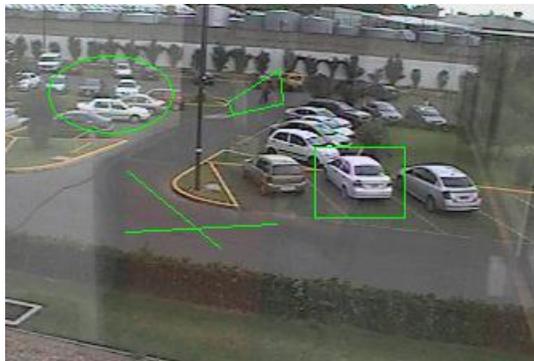


Figura 3.7 Imagen con regiones de interés

3.2.3 Histograma

Un histograma es una herramienta muy útil para determinar e interpretar de forma objetiva las características de una imagen, tales como fondo, objetos y ruido, sin necesidad de observarla en forma directa. La definición formal es:

El trazado de frecuencias en que se repite cada uno de los niveles de grises de una imagen. En el eje x se coloca cada uno de los niveles de grises l , ($0 \leq l < L$) de la imagen, mientras que en el eje y se coloca la frecuencia de aparición $h(l)$ de cada uno de estos niveles. [56]

Se pueden detectar fundamentalmente dos criterios a partir del histograma:

- Saturación. Demasiada luz causa una sobre exposición del sensor. Imágenes con saturación no contienen toda la información que se requiere para inspeccionar la escena que se observa, de hecho, tienen muchos píxeles con valores muy bajos de la escala de grises; esto aparece como un pico en extremo izquierdo del histograma. [55]
- Bajo contraste. Este es un problema muy común que se presenta en el análisis de imágenes. En la sección 3.5 se plantean procedimientos para mejorar las condiciones de luz.

En la Fig. 3.1 (b) se muestra el histograma $h(l)$ de la imagen de Fig. 3.1 (b):

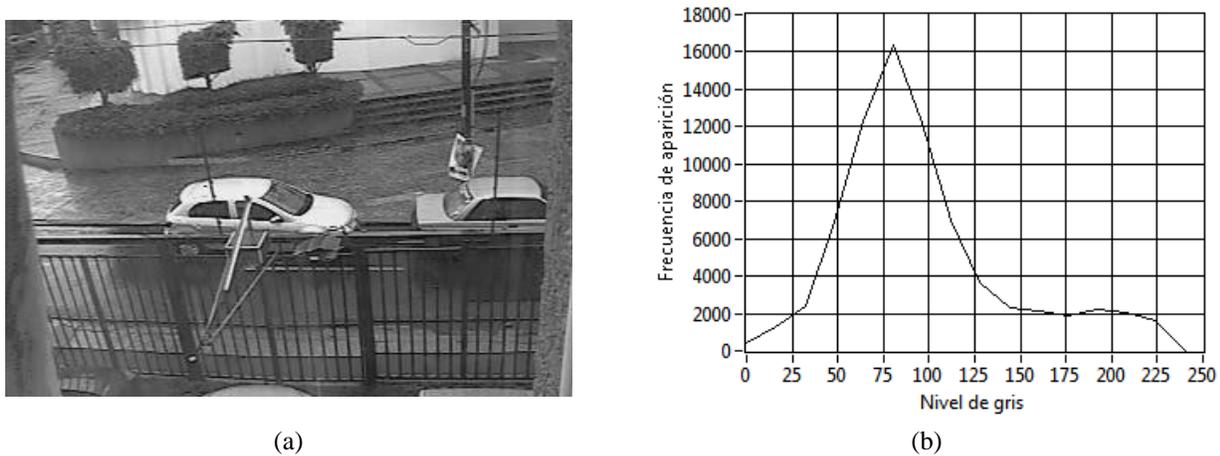


Figura 3.8 Imagen e histograma

3.2.4 Umbralado como herramienta de segmentación.

La segmentación es una de las etapas cruciales de cualquier sistema de análisis de imágenes, tanto por las dificultades que conlleva como por la importancia de sus resultados. El objetivo fundamental de esta operación es el de separar los objetos de interés del resto no relevante el cual es considerado como fondo. [56]

Realizar la segmentación no es una tarea trivial; su uso depende del problema a resolver y de la imagen de entrada. Es por esta razón que se han originado diferentes técnicas de segmentación; uno de los más antiguos y clásicos es el método basado en el umbralado a partir del histograma de la imagen, el cual busca transformar una imagen en escala de grises o color a binaria.

El umbralado básico corresponde a la siguiente definición:

Dado un conjunto N de números naturales sean (x, y) las coordenadas espaciales de una imagen y sea $G = \{0, 1, \dots, 2^B\}$ un conjunto de números positivos que representan los niveles de grises de cada píxel $f(x, y)$, donde B es el número de bits. Entonces se puede definir una función de imagen como $f: N \times N \rightarrow G$. Sea $\tau \in G$ un umbral y sea $C = \{c_0, c_1\}$ un par de niveles de gris binarios ($c_0, c_1 \in G$). El resultado de segmentar una función de imagen respecto a un umbral es la función como se representa a continuación:

$$g(x, y) = \begin{cases} c_0 & \text{si } f(x, y) < \tau \\ c_1 & \text{si } f(x, y) \geq \tau \end{cases} \quad \text{Ec. 3.1 Umbralado}$$

Así, se tienen dos clases, donde cada uno corresponde al fondo y al objeto de interés respectivamente. En este caso se dice que la imagen ha quedado binarizada. El éxito de este método depende de la distribución de los niveles de gris de la imagen y, por tanto de la forma del histograma.

3.2.5 La resta de imágenes como operador matemático básico

Los operadores matemáticos básicos ejecutan cálculos píxel por píxel con los valores de intensidad de cada uno a lo largo de la imagen fuente y una función en dos modos: con otra imagen, restando las intensidades entre ellas; o solo con la imagen fuente y una constante.

Existen distintas operaciones aritméticas entre imágenes, tales como la suma, la multiplicación, la división, el módulo y la resta. Además otras lógicas como AND, OR y XOR, entre otras.

Específicamente, la resta es calculada entre matrices, de modo que:[53]

$$I_{Dest}(x, y) = I_A(x, y) - I_B(x, y) \quad \text{Ec. 3.2 Resta entre imágenes}$$

Es importante recordar que si se trata de imágenes en escala de grises, la operación solo se realiza en el único canal de color; además, si el resultado es un número negativo, simplemente se pondrá con una intensidad igual a cero (negro) en dicho píxel. Sin embargo, si se trata de cualquier otra configuración, será necesario hacer la resta por componente, es decir, si se trabaja en RGB, se tendrá que hacer la resta en tercera dimensión. Indiscutiblemente esto representa mayor dificultad en los cálculos.

Las Fig. 3.9 y 3.10 muestran las imágenes fuente para calcular la resta:



Figura 3.9 Imagen A



Figura 3.10 Imagen B

La imagen A contiene mayor información que la imagen B, de manera que lo que se busca con ello es aislar el objeto que ha entrado en escena, despreciando todo el fondo y objetos ya existentes. De esta forma se tiene La Fig. 3.11 muestra la imagen resultante al aplicar la resta.



Figura 3.11 Imagen destino

3.2.6 Morfología matemática

La morfología matemática se refiere a la alteración de una imagen usando rutinas computacionales.[53] Es un marco teórico muy poderoso para el procesamiento digital no lineal de imágenes, cuando éstas presentan ruido u objetos con orificios, efectos que no son deseables en las aplicaciones de visión.

Las operaciones morfológicas están basadas en la modificación del PUI (por sus siglas en inglés, píxel bajo inspección) con respecto a su vecindario. Posee dos operaciones fundamentales, la erosión y la dilatación. A partir de éstas, se derivan las demás. A continuación la demostración matemática tomada de la fuente [56].

El conjunto B es conocido como “elemento de estructura”, pues define la forma en que la erosión o dilatación operará.

Sea $B \in P(Z^2)$ un conjunto arbitrario. El conjunto traspuesto B , denotado como \tilde{B} es el conjunto:

$$\tilde{B} = \{-x, x \in B\} \quad \text{Ec. 3.3 Conjunto traspuesto de } B$$

Además, considérese que para cualquier $a \in Z^2$, la traslación de B , denotada como B_a por el vector a , es el conjunto:

$$B_a = \{a + d, d \in B\} \quad \text{Ec. 3.4 Traslación de } B \text{ por el vector } a$$

La dilatación de X por B es el conjunto de los puntos x en Z^2 tal que la intersección entre X y B_x es no vacía. Se refiere a la expansión de un objeto para aumentar su tamaño, rellenar huecos y conectar píxeles vecinos.[53] El conjunto resultante se denota por:

$$X \oplus \tilde{B} = \{x \in Z^2, X \cap B_x \neq \emptyset\} \quad \text{Ec. 3.5 Expresión matemática de dilatación}$$

En la Fig. 3.12 se muestra la operación de dilatación.[53]

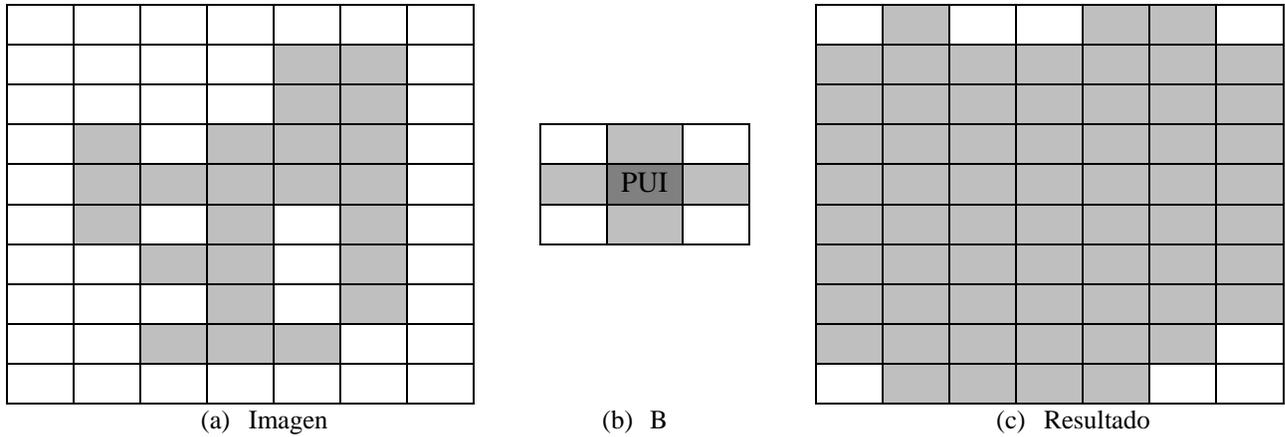


Figura 3.12 Dilatación

La erosión desgasta los bordes de la imagen, disminuyendo su tamaño y abriendo agujeros en el objeto. Se define de forma similar a la dilatación, como se expresa a continuación

$$X \ominus \tilde{B} = \{x \in Z^2, B_x \in X\} \quad \text{Ec. 3.6 Expresión matemática de erosión}$$

En la Fig. 3.13 se muestra la operación de erosión.[53]

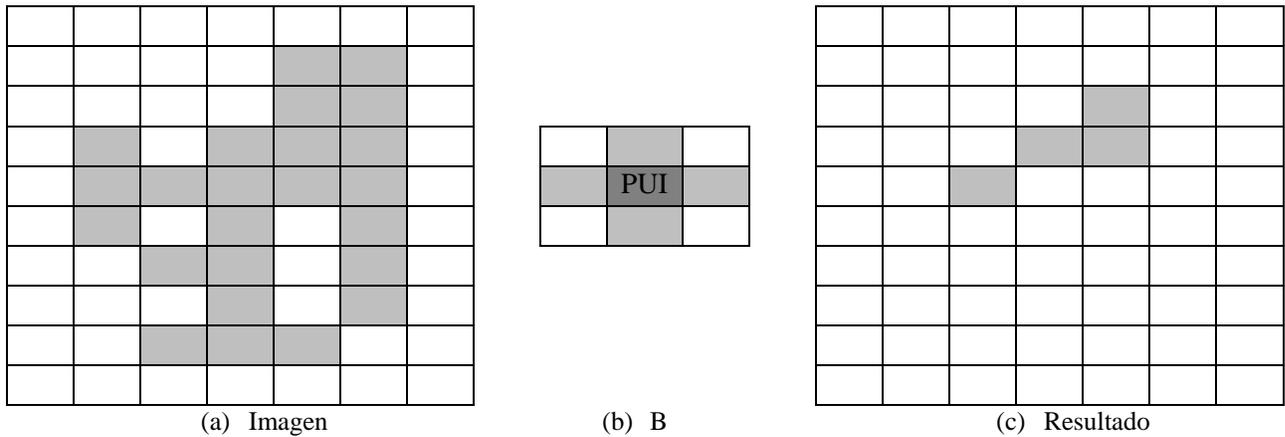


Figura 3.13 Erosión

Estas operaciones son duales, lo que quiere decir que se complementa una a la otra, es decir:

$$X \oplus \tilde{B} = (X \ominus \tilde{B})^c \quad \text{Ec. 3.7 Dualidad erosión-dilatación}$$

3.2.7 Detección de bordes

Un eje en una imagen es generalmente definida como una característica del objeto con alto contraste; si la intensidad de una imagen cambia drásticamente dentro de un intervalo pequeño, entonces se puede considerar como un borde. [53]

La inspección se realiza analizando la intensidad de gris de un conjunto de píxeles que se encuentran dentro de una ROI, que puede ser una línea o las aristas de cualquier figura geométrica.

En la Fig. 3.14 se muestra la gráfica que define el borde.

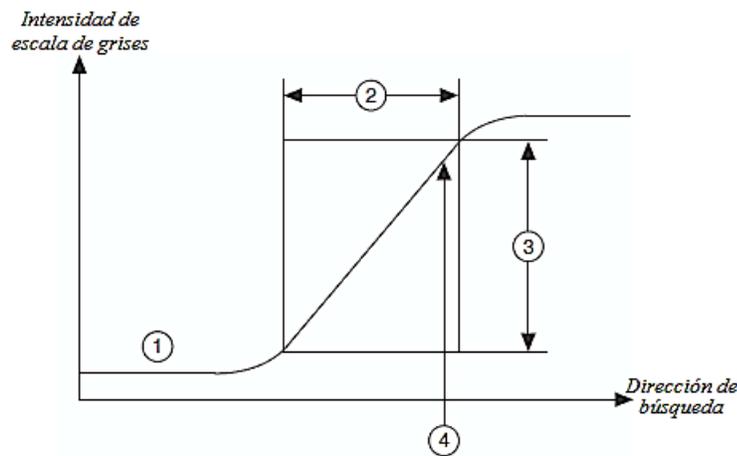


Figura 3.14 Detección de bordes

Donde:

1. Perfil de la escala de grises.
2. Distancia de evaluación de píxeles dentro de la ROI para detectar cambios en la escala de grises.
3. Contraste de la imagen. Diferencia entre el nivel de gris del fondo y del objeto. Para mejores resultados, este intervalo debe ser pronunciado, de lo contrario podrían ocurrir errores en la detección.
4. Punto de evaluación $B_d(x, y)$ donde se detectó el borde.

Existe otro concepto importante, que es la polaridad. Éste se refiere a la dirección del cambio de iluminación. Para el punto $B_d(x, y)$ la polaridad es positiva pues el flanco ascendente se caracteriza por un aumento de la escala de grises valores a medida que cruzan el borde, a diferencia de la polaridad negativa donde ocurre lo contrario. Este parámetro está vinculado con la dirección de búsqueda.

3.3 Análisis de partículas

El análisis de partículas se emplea para detectar las regiones conectadas de una imagen y luego hacer mediciones seleccionadas de esos grupos de píxeles. Estas regiones se denominan comúnmente como partículas, por tanto son porciones de imagen contiguas de píxeles distintos de cero.

El analizar las imágenes segmentando partículas, da la posibilidad de extraer solo las porciones que son significativas. También es útil filtrar las partículas que no cumplen con ciertos parámetros preestablecidos, de esta forma se realiza la elección de material útil.

Un proceso de análisis de partículas típico se describe en la Fig. 3.15, donde primero se debe crear una imagen binaria mediante un proceso de umbralado. A continuación, puede mejorar la imagen binaria utilizando transformaciones morfológicas por ejemplo, la dilatación y erosión. Otra actividad importante es el filtrado, que busca eliminar espurios que provocan ruido en la imagen. Después se explora toda la imagen para detectar las partículas. Se hacen mediciones sobre las partículas en la imagen. Finalmente se construye un informe detallado sobre cada partícula, para su posterior uso en el procesamiento de imágenes. [55]

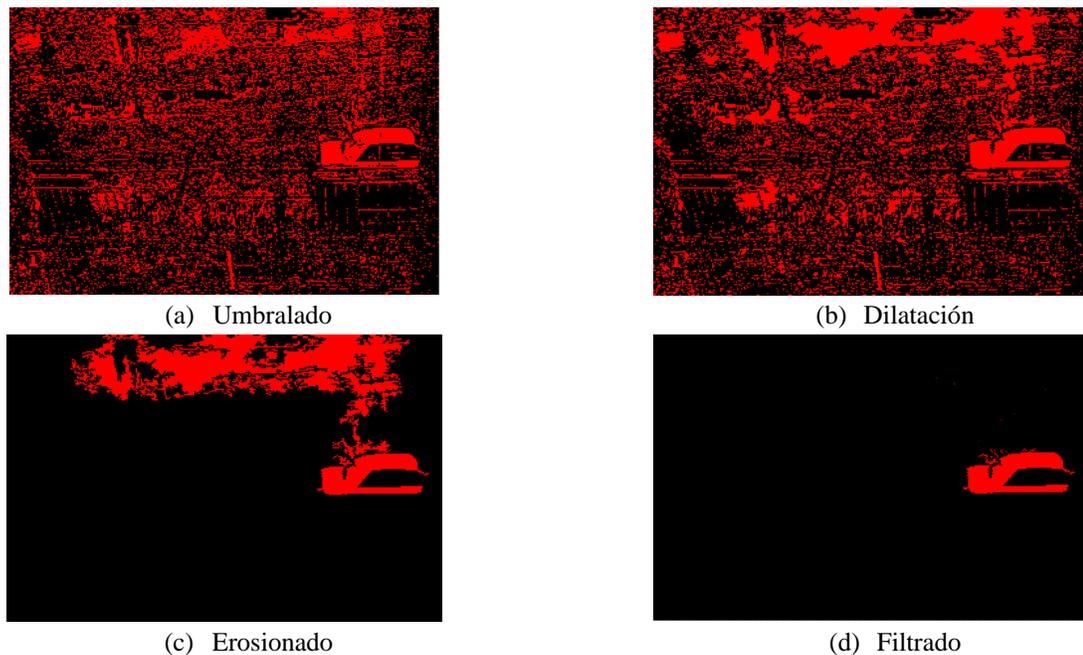


Figura 3.15 Análisis de partículas

3.4 Mejoramiento de las condiciones en las imágenes

El objetivo fundamental del mejoramiento digital de imágenes es destacar información acerca de un objeto no claramente observable en la imagen original. [56] Los aspectos más importantes son:

- Mejorar el contraste
- Eliminación o atenuación del ruido.
- Calibración entre imágenes cuya fuente es distinta.

Las operaciones a realizar en el mejoramiento digital de imágenes se pueden enmarcar en tres grupos:

- Operaciones puntuales: su resultado depende únicamente del valor del píxel de entrada en una posición dada.



- Operaciones locales: depende tanto del valor del píxel de entrada, como de los valores de los píxeles de su vecindad.
- Operaciones globales: cuando un punto de la imagen de salida depende en gran parte de los valores de los píxeles de la imagen de entrada.

3.4.1 Ajuste de histograma

El uso de operaciones puntuales tiene muchas ventajas, entre las cuales se encuentran alcanzar una alta velocidad en su ejecución y expresiones matemáticas más sencillas. El ajuste del histograma es una de las operaciones puntuales más conocidas para mejorar el contraste, donde su aumento es un paso previo a llevar a cabo procedimientos más complejos. [56]

Para realizar el ajuste de histograma es necesario calcular el histograma acumulativo de una imagen, que es la suma de todas las frecuencias de los niveles de gris en el histograma original $h(i)$ de la imagen en la resolución de B bits de la imagen.

$$A = \sum_{i=0}^{2^B} h(i) \quad \text{Ec. 3.8 Histograma acumulativo de una imagen}$$

Si $f(x, y)$ es el valor de un píxel con sus coordenadas y $h_0(f(x, y))$ el histograma de la imagen de entrada, se tiene una transformación T , cuya función es monótona, que se aplica a esa imagen, entonces el valor del píxel de salida está dado por:

$$g(x, y) = T(f(x, y)) \quad \text{Ec. 3.9 Píxel de salida después de transformación}$$

Esto origina un nuevo histograma $h_1(g(x, y))$.

3.4.2 Reajuste de los niveles de gris

Cuando los objetos aparecen concentrados solo en una parte de la escala de los $2^B - 1$ niveles de gris, se pueden tener problemas al realizar ciertas transformaciones. El nivel del fondo (Z_1) puede ser fijado a cero y el valor de gris de uno de los objetos (Z_2) se fija en el máximo de la escala, el intervalo entre Z_1 y Z_2 puede ser reajustado linealmente entre valores de nivel de gris de todo el histograma.

El procedimiento propuesto en [56] busca elevar el contraste por medio de la siguiente ecuación:

$$g(x, y) = \begin{cases} \frac{(f(x, y) - Z_1)(2^B - 1)}{Z_2 - Z_1} & Z_1 < f(x, y) < Z_2 \\ 2^B - 1 & f(x, y) \geq Z_2 \\ 0 & f(x, y) \leq Z_1 \end{cases} \quad \text{Ec. 3.10 Transformación de la escala de niveles de gris}$$

Donde $g(x,y)$ es la imagen transformada y $f(x,y)$ el valor de gris de la imagen original en las coordenadas dadas.

3.4.3 Filtrado espacial

Los métodos de realce de imágenes se dividen en dos categorías: los métodos de realce en el dominio espacial, que manipulan directamente los píxeles de la imagen; y los métodos de realce en el dominio de la frecuencia, que corresponden a técnicas basadas en la representación de los píxeles, a través de una transformación hacia el dominio frecuencial.

En general, los filtros se dividen en lineales (también llamados convolución) y los no lineales; su operación se basa en alterar el valor de los píxeles con respecto a variaciones de intensidad de luz en su vecindario, el cual está definido por una matriz, donde del píxel central se modificará su nivel de gris. Los filtros espaciales se dividen en dos categorías: [55]

- El filtro pasa baja atenúa las variaciones de la intensidad de luz, tienen además, la tendencia a eliminar detalles y contornos borrosos.
- El filtro pasa alto enfatiza las variaciones de intensidad en una imagen, detecta bordes, elimina ruido y mejora detalles finos de una imagen.

3.4.4 Especificación del histograma

Si se desea ajustar una imagen por medio del histograma de otra imagen para tener una distribución semejante y así ajustar el brillo, se necesita definir funciones de densidad de probabilidad $P_r(r)$, que puede estimarse de la imagen de entrada y $P_z(z)$ que es la función de densidad de probabilidad especificada, considerándose ambas continuas; donde r es el nivel de gris y z la imagen procesada. [57]

Si se toma como base la expresión para linealizar un histograma (ecualización), se tiene que:

$$s = \int_0^r P_r(w)dw = T(r) \qquad \text{Ec. 3.11 Transformación de ecualización para } r$$

Además si la variable z se define de la siguiente forma:

$$v = G(z) = \int_0^z P_z(t)dt \qquad \text{Ec. 3.12 Transformación de ecualización para } z$$

Para igualar los histogramas, es necesario que tenga la forma $T(r) = G(z)$, entonces z debe satisfacer la condición:

$$z = G^{-1}(s) = G^{-1}[T(r)] \qquad \text{Ec. 3. 13 Especificación del histograma}$$

Una vez que se tiene definida z , se aplica esa transformación a todos los píxeles de la imagen. El resultado de este procedimiento es una imagen cuyos niveles de gris z tendrán la función de densidad de probabilidad especificada, $P_z(z)$.

3.5 Búsqueda de patrones

El reconocimiento de objetos es una de las actividades más importantes y complejas del análisis de imágenes. Después de la segmentación, la porción extraída será la imagen de entrada, con la finalidad de obtener una serie de parámetros descriptores que facilitarán la tarea de buscar, detectar y reconocer patrones; de esta forma, se pueden emplear también clasificadores, que cataloguen la información en un conjunto de hipótesis indicando al usuario qué objetos se pueden encontrar en la imagen de prueba.[58]

La búsqueda de patrones tiene tres usos:[55]

1. Alineación: determina la posición y orientación de un objeto.
2. Medición: localiza el objeto y mide ciertas características predefinidas.
3. Inspección: detecta fallas y presencia/ausencia de objetos.

Para comenzar la búsqueda, se requiere definir el objeto que se desea encontrar; se trata de una imagen que es llamada plantilla, así como una lista de grados de libertad permitidos y una rutina de búsqueda en el fotograma deseado.

3.5.1 Plantilla

La plantilla requiere tener un buen nivel de contraste. En Lab VIEW, el algoritmo que más adelante se explicará no se limita a buscar las coincidencias exactas en los niveles de gris de las imágenes, por lo que puede encontrar objetos bajo diferentes condiciones de luz, ya que la información de fondo es un auxiliar en la detección. [53]

Generalmente, esta plantilla es tomada como una ROI de la misma imagen donde se desean encontrar todos los objetos de ese tipo. Sin embargo, existen otras formas; una de ellas es que el usuario de la aplicación proporcione la plantilla desde una fuente externa, es decir, otro programa con la que se guarde la porción de imagen que se desea buscar, aunque la información adicional valiosa podría perderse. Otra opción es encontrar la plantilla en un imagen y buscar el patrón en otra, en cuyo caso se deberán considerar opciones de nivelación de contraste y brillo interimagen para hacer equivalentes las imágenes y que el algoritmo pueda funcionar adecuadamente.

Otro punto importante es elegir una imagen que tenga un fondo significativo, de manera contraria podría causar ruido y la búsqueda no sería exitosa.



Figura 3.16 Plantilla para búsqueda de patrones

3.5.2 La búsqueda

Para encontrar la plantilla en la imagen se requiere poner en práctica un algoritmo que sea capaz de hallarlo de manera eficiente y bajo ciertas características que el usuario determinará, tales como si el objeto sufre cambios en escala y/o rotación.

El algoritmo base para la búsqueda del patrón es la “correlación cruzada normalizada”, sin embargo éste tiene algunas desventajas, pues debido a que su operación se trata de una serie de operaciones de multiplicación, el proceso puede consumir mucho tiempo. Para atacar ese problema se han desarrollado tecnologías que realizan los cálculos de manera paralela. Otra solución que se ha probado es reducir el tamaño de la imagen y restringir la región de la imagen en la que la coincidencia se produce.

- Correlación cruzada normalizada

En esta parte del algoritmo, la plantilla se coloca sobre la imagen buscando la mayor coincidencia a través de multiplicaciones píxel a píxel, después todos esos resultados se suman para producir un valor de correlación simple. Ver Fig. 3.17

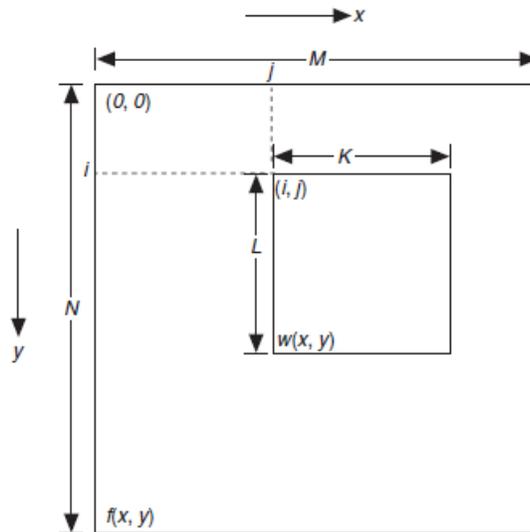


Figura 3.17 Correlación cruzada normalizada

Donde:

x	Desplazamiento horizontal	
y	Desplazamiento vertical	
M	Tamaño horizontal de imagen	
N	Tamaño vertical de imagen	
K	Tamaño horizontal de plantilla	
L	Tamaño vertical de plantilla	
j	Coordenada horizontal	$i = 0, 1, \dots, M - 1$
i	Coordenada vertical	$j = 0, 1, \dots, N - 1$
$f(x, y)$	Imagen	
$w(x, y)$	Plantilla	

La plantilla se mueve entonces una casilla y el proceso se repite formando una matriz:

$$C(i, j) = \sum_{x=0}^{L-1} \sum_{y=0}^{K-1} w(x, y) f(x + i, y + j) \quad \text{Ec. 3.14 Matriz de Correlación}$$

Finalmente se elige el valor máximo y así localizar el grupo de píxeles donde existe la mayor probabilidad de encontrar el objeto buscado. Sin embargo, este método es sensible a los cambios de amplitud de la imagen y de la plantilla, por lo que es necesario calcular el coeficiente de correlación normalizado como se muestra en la Ec. 3.15.

$$t(i, j) = \frac{\sum_{x=0}^{L-1} \sum_{y=0}^{K-1} (w(x, y) - \bar{w}) (f(x + i, y + j) - \bar{f}(i, j))}{\sqrt{\sum_{x=0}^{L-1} \sum_{y=0}^{K-1} (w(x, y) - \bar{w})^2 \left[\sum_{x=0}^{L-1} \sum_{y=0}^{K-1} (f(x + i, y + j) - \bar{f}(i, j))^2 \right]}} \quad \text{Ec. 3.15 Coeficiente de correlación}$$

Uno de los defectos más grandes de correlación cruzada es su incapacidad para que coincida con los objetos de una imagen de origen que son o bien de un tamaño diferente a la plantilla, o han sido rotados. Para ello se emplean otros métodos para elegir las características más valiosas de la plantilla y con esto incrementar la velocidad de procesamiento y reducir los cálculos. Ejemplos de estos algoritmos son la búsqueda piramidal, que reescala tanto la imagen como la plantilla para encontrar el mejor candidato; y el llamado entendimiento de la imagen, procedimiento que reduce la información redundante y enfatiza la característica para permitir una eficiente, pero robusta, la correlación cruzada.

- *Image understanding*

El entendimiento de la imagen (*image understanding*) utiliza una técnica de muestreo no uniforme que incorpora comprensión de imágenes para describir a fondo y de manera eficiente la plantilla. Esta técnica de muestreo inteligente incluye específicamente una combinación de píxeles del borde y los píxeles de la región como se muestra en la Fig. 3.18.

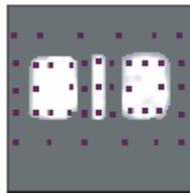


Figura 3.18 Muestreo de plantilla en el método entendimiento de imagen

Para observar a mayor detalle el muestreo no uniforme, considérese la Fig. 3.19 (a) como una imagen binaria con cierta forma. En la Fig. 3.19 (b) se ve cómo se toman secciones de píxeles con las mismas intensidades que sus vecinos.[59]

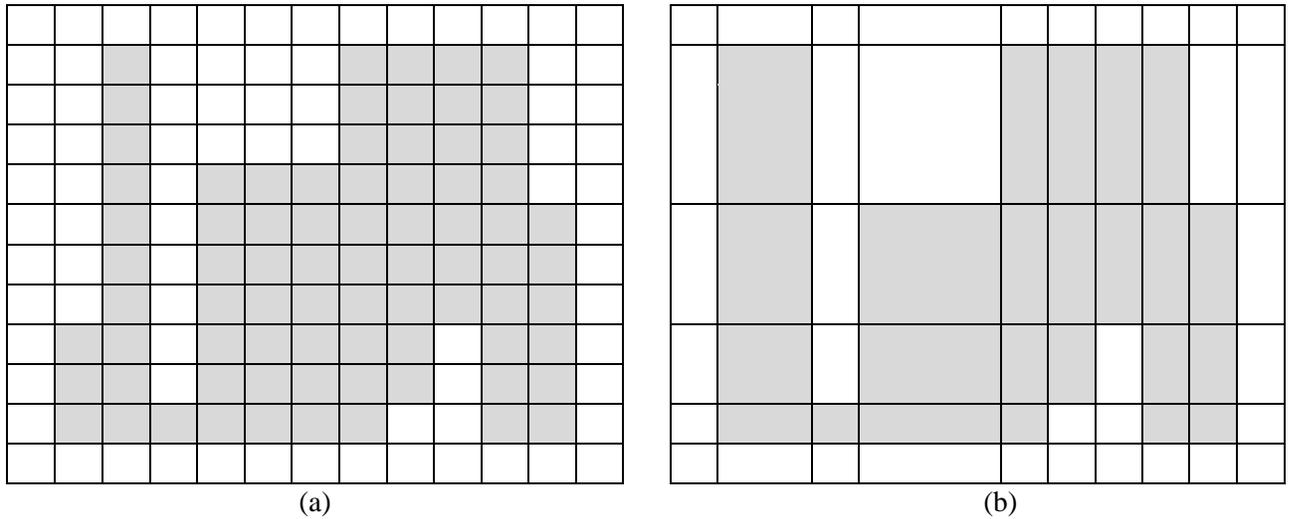


Figura 3.19 Muestreo no uniforme

3.6 Conclusiones

En este capítulo se enumeró la teoría que sustenta la aplicación que se explicará en los capítulos subsecuentes. La organización fue de acuerdo a la lógica de operación del sistema, yendo desde un panorama sobre las comunicaciones, específicamente del tipo serial, pasando por los conceptos fundamentales sobre la visión artificial. Posteriormente el análisis de partículas para rellenar huecos y remover pequeñas porciones de imagen binarizada que no son útiles en este tratamiento, siendo esto una alternativa al filtrado. Después se revisaron técnicas para mejorar el contraste, esto con la finalidad de hacer una calibración intercámara y preparar las imágenes para la “búsqueda de patrones” y encontrar el objeto deseado.

CAPÍTULO 4. DISEÑO Y DESARROLLO DE LOS INSTRUMENTOS VIRTUALES PARA VIDEO-VIGILANCIA Y SEGURIDAD

En este capítulo se describe el diseño y montaje del sistema, así como el desarrollo de los Instrumentos Virtuales utilizados para la visualización del video en tiempo real, la comunicación entre cámaras, el preposicionamiento de la cámara móvil, la elaboración y envío de los reportes a usuario.

En la Fig. 4.1 (a) se muestra el diagrama de bloques donde comienza la aplicación con el escaneo de todas las cámaras fijas X_i , para determinar si existe o no evento en las regiones de interés que el usuario configuró previamente. Cuando la presencia de un objeto ha sido detectada, tanto el identificador de la cámara que generó el evento, como las imágenes limpia y con el objeto, son puestos en un elemento de la cola y enviados a la siguiente parte de la aplicación, mientras se continúa con la verificación de las cámaras, haciendo uso del procesamiento paralelo o pseudoparalelo. La línea de retroalimentación se refiere a que aún cuando los datos se envían, la inspección de las cámaras continúa para verificar si existe un evento en otra cámara.

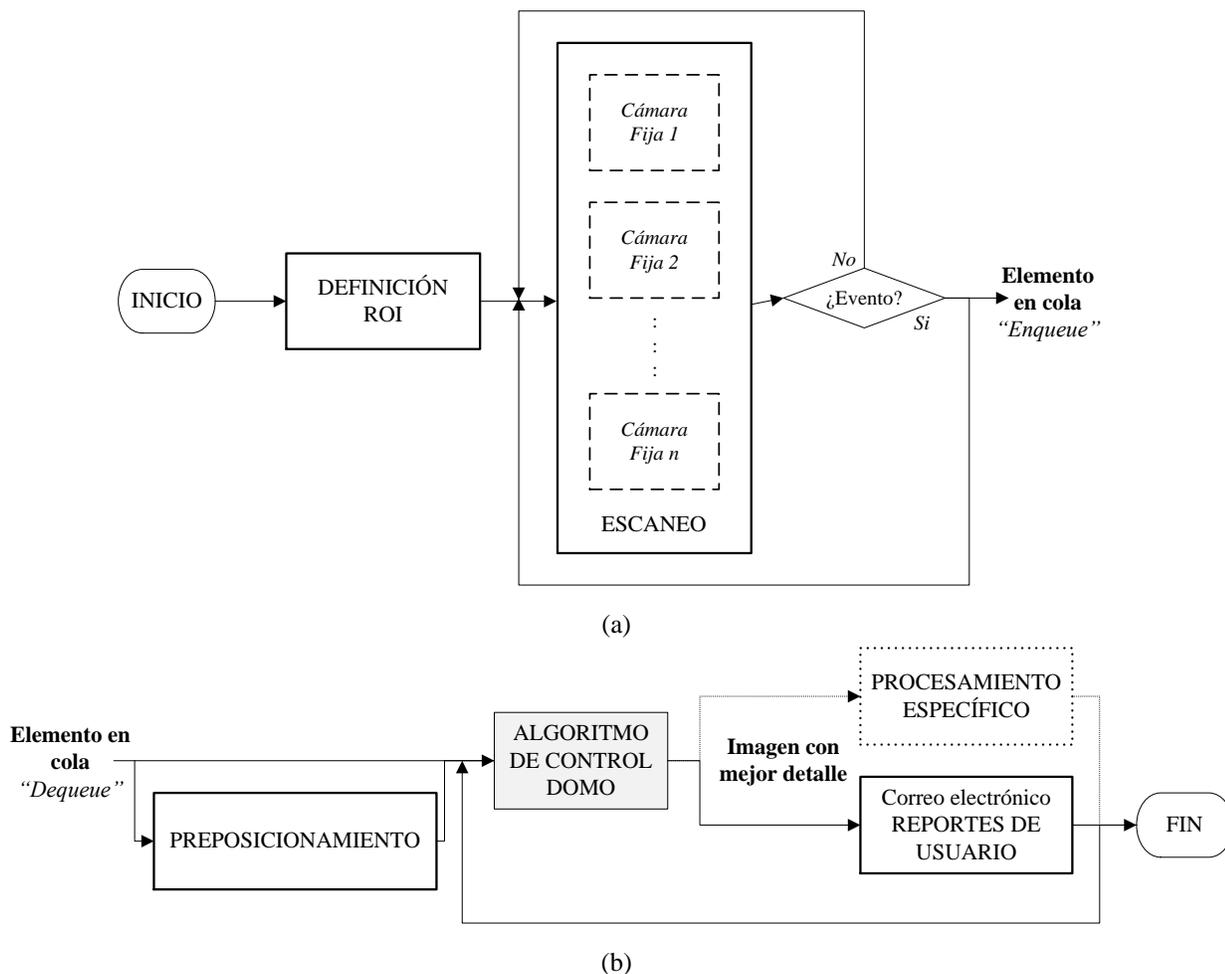


Figura 4.1 Diagrama de bloques de operación general

En la Fig. 4.1 (b) se indica que cuando la cola tiene elementos, se ejecuta el algoritmo de control del domo, que se verá en el siguiente capítulo. Al obtener una imagen con mejor detalle y con un acercamiento aceptable se genera el reporte a usuario, mismo que es enviado a una cuenta de correo electrónico. El bloque de “procesamiento específico” se refiere a la utilidad que puede tener la imagen final para otro tipo de análisis no atacado en el presente trabajo, tal como la determinación de placas de automóviles o rasgos como texturas, color e incluso modelo.

4.1. Diseño de propuesta

La idea básica del sistema de video-vigilancia se muestra en la Figura 4.2. Se tienen $n = 2$ cámaras fijas, una de es tipo IP y la otra analógica, además de un domo.

En el Centro de Investigación en Computación se tiene todo el equipo y servidor del sistema. La cámara IP se conecta a la PC ubicada en el Laboratorio de Tiempo Real y Automatización a través de Internet; el el mismo sitio se encuentran la cámara analógica fija, el domo y el DVR; las dos primeras están conectadas a éste último en comunicación serial RS-485. Tanto el servidor como el DVR tienen las funciones mostradas en la Fig. 4.2 y ambos están en comunicación a través de internet gracias al *software* comercial del grabador y la obtención del video programado para esta aplicación.

Por otro lado, el cliente representa, en principio, cualquier computadora que cumpla con ciertas características que se mencionan más adelante. En este dispositivo se mostrarán los reportes a usuario y se podrá tener acceso al sistema por medio del *software* comercial.

La configuración de todo el sistema la realizará el administrador en la computadora servidor.

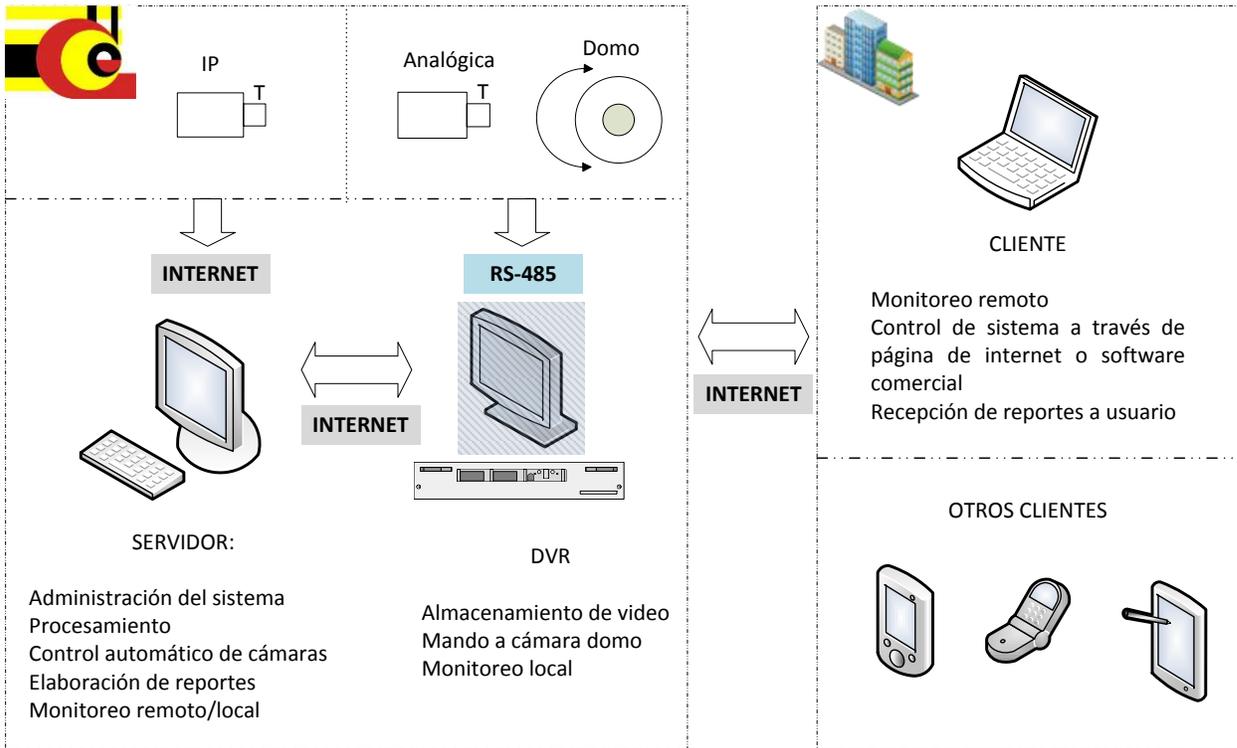


Figura 4.2 Conectividad del sistema básico de video-vigilancia

4.1.1 Requisitos del sistema base

A continuación se enlistan los requisitos de *software* y *hardware* para el cliente y para el servidor:

Tabla 4.1 Requisitos para el sistema cliente-servidor

	CLIENTE	SERVIDOR
Software	SO Microsoft Windows. Adobe Reader o similar. Acceso a Internet. Correo electrónico google (gmail)	SO Microsoft Windows (preferentemente Windows 7). Lab VIEW 2010. Toolkit Lab VIEW Visión. Toolkit Lab VIEW generación de reportes para Microsoft Office. Bull zip PDF Printer. SDK Vivotek, Vitamin Decoder. SDK IDISWatch.
Hardware	-	Tarjeta RS-485/USB NI

Como se mencionó anteriormente, el equipo empleado en este sistema es:

Tabla 4.2 Equipo empleado en el sistema

CÁMARAS	Minitrax Dome Camera, Syscom 	3DNR Super Night Vision Outdoor Camera, Epcor 	IP Camera Vivotek 
DVR	4-channel digital video recorder, Syscom 		
CABLEADO	Cable UDP	Cable coaxial	Cable ethernet
	Tarjeta RS-485/USB NI 		
PROTOCOLO	Protocolo Pelco D (ver Anexo C)		

4.1.2 Conexión y montaje

Particularmente para el diseño del sistema de pruebas, se emplean dos cámaras analógicas (domo y bala) que son conectadas por cableado UDP a través de sus conexiones seriales RS-485 al DVR de cuatro canales, éste digitaliza el video. En la Fig 4.3 se muestra el esquema de conexión física empleada.

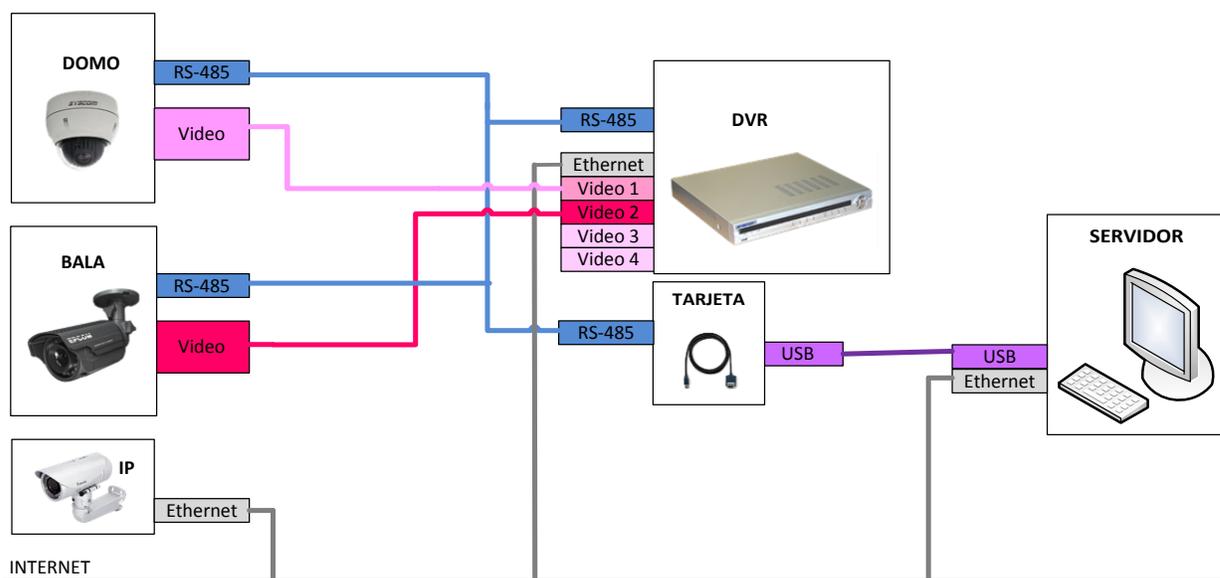


Figura 4.3 Esquema de conexión del sistema de video-vigilancia

Para la comunicación serial RS-485 en el domo solo se emplea el canal de recepción, pues el domo no mandará ningún mensaje al DVR, por lo que no será necesario emplear el canal Tx. También, se usa conexión para enviar el video analógico al grabador por medio de cable coaxial.

De manera semejante al anterior, la cámara tipo bala emplea el cableado UTP, pero con la diferencia que el video lo envía directamente por cable coaxial y que, aunque la comunicación serial está conectada, no se tiene manipulación de las condiciones PTZ a través de la interfaz comercial, únicamente por el menú propio de la cámara. También emplea un transformador para proporcionarle los niveles correctos de energía.

Las imágenes de las cámaras anteriores, son enviadas al DVR. Este dispositivo está conectado a Internet, lo que dará la posibilidad de visualizar el video en tiempo real desde cualquier computadora cliente, además de ser el medio de almacenamiento de video, siempre y cuando posea un disco duro en su interior.

La tarjeta que adapta la señal de RS-485 a USB es conectada a la computadora donde se realiza todo el procesamiento. Este adaptador tiene el objetivo de entablar la comunicación para manipular las condiciones PTZ de la cámara tipo domo a través de aplicaciones diferentes al *software* comercial, por ejemplo LabVIEW, pues al ser parte del equipo de National Instruments, la comunicación es transparente.

Por otro lado, la cámara IP solo debe conectarse a la alimentación y a Internet para que pueda funcionar. El video es almacenado en la ubicación que sea elegida dentro de la computadora principal de usuario, pero éste mismo puede ser visualizado en cualquier computadora cliente a través de su dirección IP, nombre de usuario y contraseña.

Para la aplicación del presente trabajo, el video del DVR y de la cámara IP puede visualizarse por medio del *software* comercial por Internet, o empleando el SDK del grabador en LabVIEW. Esta última opción fue la que se eligió. Para mayor información, consultar Anexos C y D.

4.2. Obtención de video en tiempo real

Existen dos formas para obtener el video en tiempo real a partir del grabador digital, la primera es tener una tarjeta para conectarla en directo al DVR, con entrada coaxial y salida con puerto compatible a la computadora (USB, HDMI u otros), o bien, a través de comunicación TCP/IP, entrando al servidor del DVR.

Aunque la primer opción resulta tener mayor fidelidad en la transmisión de información, las tarjetas con entrada coaxial han sido descontinuadas desde mucho tiempo atrás, debido a la nueva generación de cámaras digitales, por lo que resulta muy difícil conseguir una tarjeta que podría no tener una respuesta confiable. Por esta razón, se emplea la segunda solución al problema: comunicación TCP/IP.

En la Fig. 4.4 se muestra la configuración general para obtener el video en tiempo real a través de la comunicación con el DVR y la cámara IP.

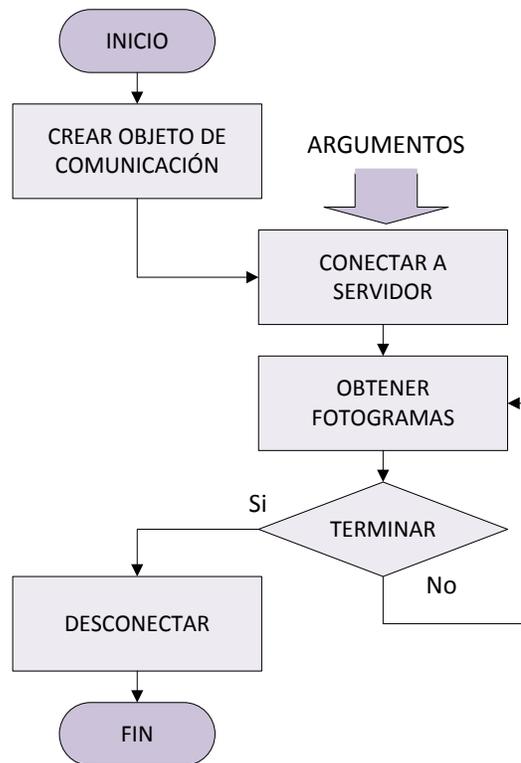
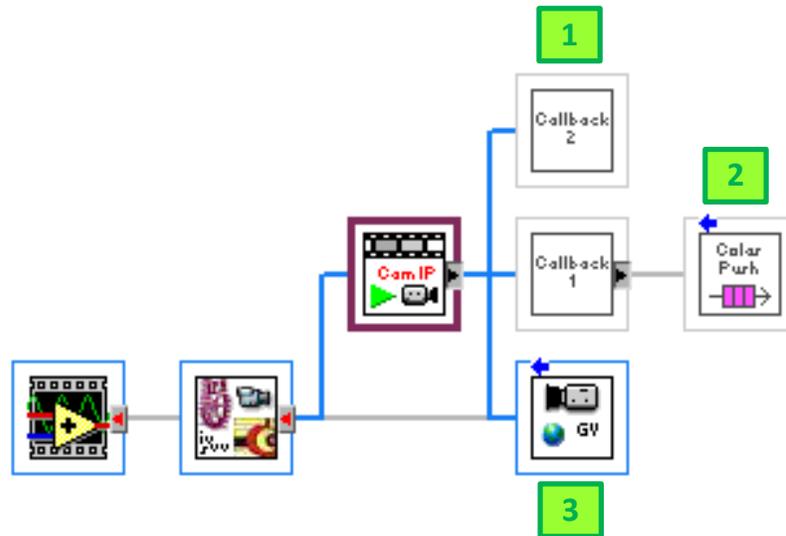


Figura 4.4 Configuración general para la obtención del video en tiempo real

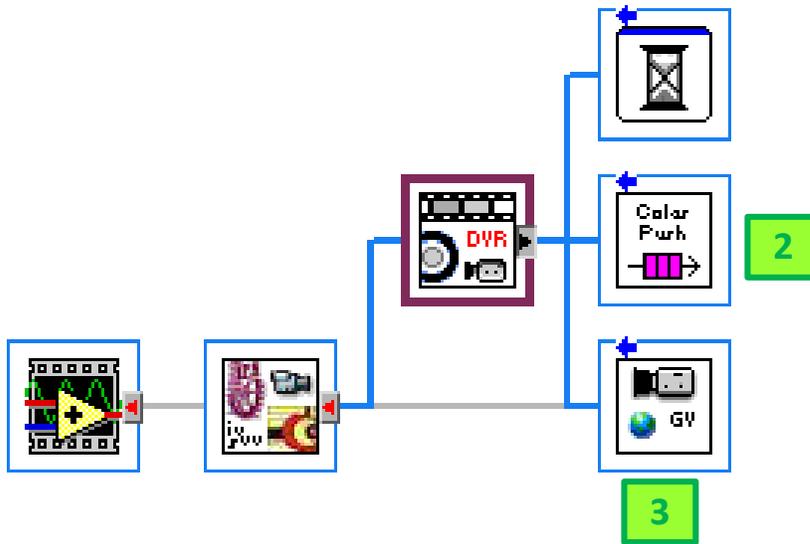
Debido a que se emplean los SDK (Anexos C y D) tanto del grabador como de la cámara IP, es necesario crear un objeto de comunicación. Para conectar con el servidor de cada uno de los dispositivos, se definen algunos argumentos, tales como nombre del sitio, dirección IP, usuario, contraseña, identificador y puerto de comunicación. Después se obtienen los fotogramas que conforman el video, ubicándolos en colas para

la visualización del mismo y su posterior uso en el procesamiento. El programa espera hasta que le sea indicada la desconexión del servicio.

En la Fig. 4.5 (a) contiene el diagrama de jerarquía de la aplicación solo para la obtención del video en tiempo real de la cámara IP. El esquema del mismo tipo para el DVR se muestra en la Fig. 4.5 (b).



(a)



(b)

Figura 4.5. Diagrama de jerarquía para obtención de video en tiempo real

Los diagramas muestran que para ambos dispositivos se tiene la misma lógica de operación, con la diferencia que en el caso de la cámara IP se emplean funciones de llamado (*callback 1* y *2*) necesarias para solicitar las imágenes a través de su SDK, como se observa en la Fig. 4.5 (b), punto 1.

Los fotogramas son almacenados en una cola (punto 2) que guarda cada uno de ellos para evitar el conflicto que se crea al no tener una tasa constante de marcos por segundo, lo cual es un problema característico de los sistemas de video-vigilancia.

Al mantener comunicación constante con el resto de la programación a través de las variables globales (punto 3), se garantiza que la aplicación está sincronizada para el paro del sistema o desconexión, como se mencionó en la explicación de la Fig. 4.4.

4.3. Selección de las regiones de interés

Las ROI se seleccionan para tener un punto de comparación en la detección de bordes y verificar el cambio de iluminación. Bajo ese esquema, se puede emplear cualquier figura geométrica; la recomendada es la línea, pues puede ponerse en cualquier dirección.

Como se guardan los descriptores de estas ROI en un arreglo, pueden elegirse hasta $2^{31} - 1$ elementos. Sin embargo, esto no es necesario. Dependiendo de la amplitud de la zona que se desea vigilar, es suficiente poner hasta cinco líneas, pudiendo cada una de ellas abarcar tanto espacio como el usuario desee.

Es importante mencionar que se debe tener cuidado al colocar estas ROI sobre la imagen, pues con el criterio de evaluar el cambio de iluminación, pueden tenerse medidas erróneas al sobreponer la líneas en bordes a primera vista observados. En la Fig. 4.6 se muestran un ejemplo de una deficiente selección de regiones de interés (a) y (b) el uso recomendado en el presente trabajo para una detección de bordes adecuada.



Figura 4.6 Definición de ROI

Aunque en el primer caso se elige solo una línea como ROI, esta selección induce un error en la detección temprana de bordes que evidentemente existen, pero que no proporcionan ninguna información valiosa. Si se consideran dos puntos como mínimo para generar evento, en esta imagen se haría un falso llamado a las funciones de procesamiento. A diferencia de lo anterior, en el segundo caso se tiene una mejor distribución de ROI en la imagen, pudiendo evaluar la existencia de eventos sin objetos que produzcan ruido y con una mejor distribución para monitorear ciertas zonas del FOV.

4.4.Preposicionamiento

El preposicionamiento emplea rutinas antes configuradas directamente en la memoria del domo. Esto se hace ingresando al menú principal de esta cámara por medio del *software* comercial, ahí se establecen puntos con valores PTZ (procedimiento en Anexo E). De esta forma se puede hacer uso de las subrutinas con solo elegir el punto deseado (parámetro “número”, según la Fig. 4.7) y el domo modificará sus condiciones de acuerdo a lo programado en el mensaje que se envía. (Ver Anexo B, sobre protocolo Pelco-D).

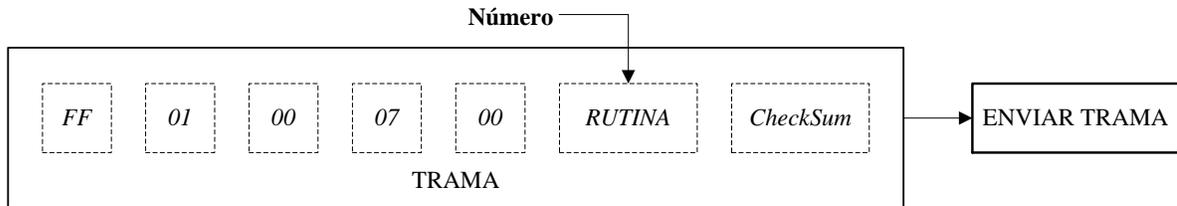


Figura 4.7 Mensaje para preposicionamiento

Este bloque tiene la finalidad de mover el domo hacia una zona conocida para asemejar su FOV con el de la cámara fija X_i que generó el evento y así poder buscar el objeto por reconocimiento de patrones; esto es esencial, de otra forma el domo no tendrá una imagen adecuada de referencia y por lo tanto no habrá éxito en la localización. La preconfiguración de los puntos se muestra en la Fig 4.8, donde las posiciones 3, 6 y 7 lograrían el propósito antes mencionado.

Otro punto en donde se emplea el preposicionamiento es después de adquirir la imagen con mayor detalle cuando es deseable poner el domo en un punto para comenzar con nuevas localizaciones y no mantener su FOV cubierto por el objeto de interés que ya se encontró. Ejemplos de estas ubicaciones serían las posiciones 1, 2 y 8 en la Fig. 4.8 según el FOV que se desee cubrir mientras se está libre de tareas asignadas.

Probablemente para este caso no sería necesario configurar las posiciones 4 y 5, a menos que se tuvieran más cámaras fijas o que se deseara observar ese FOV.

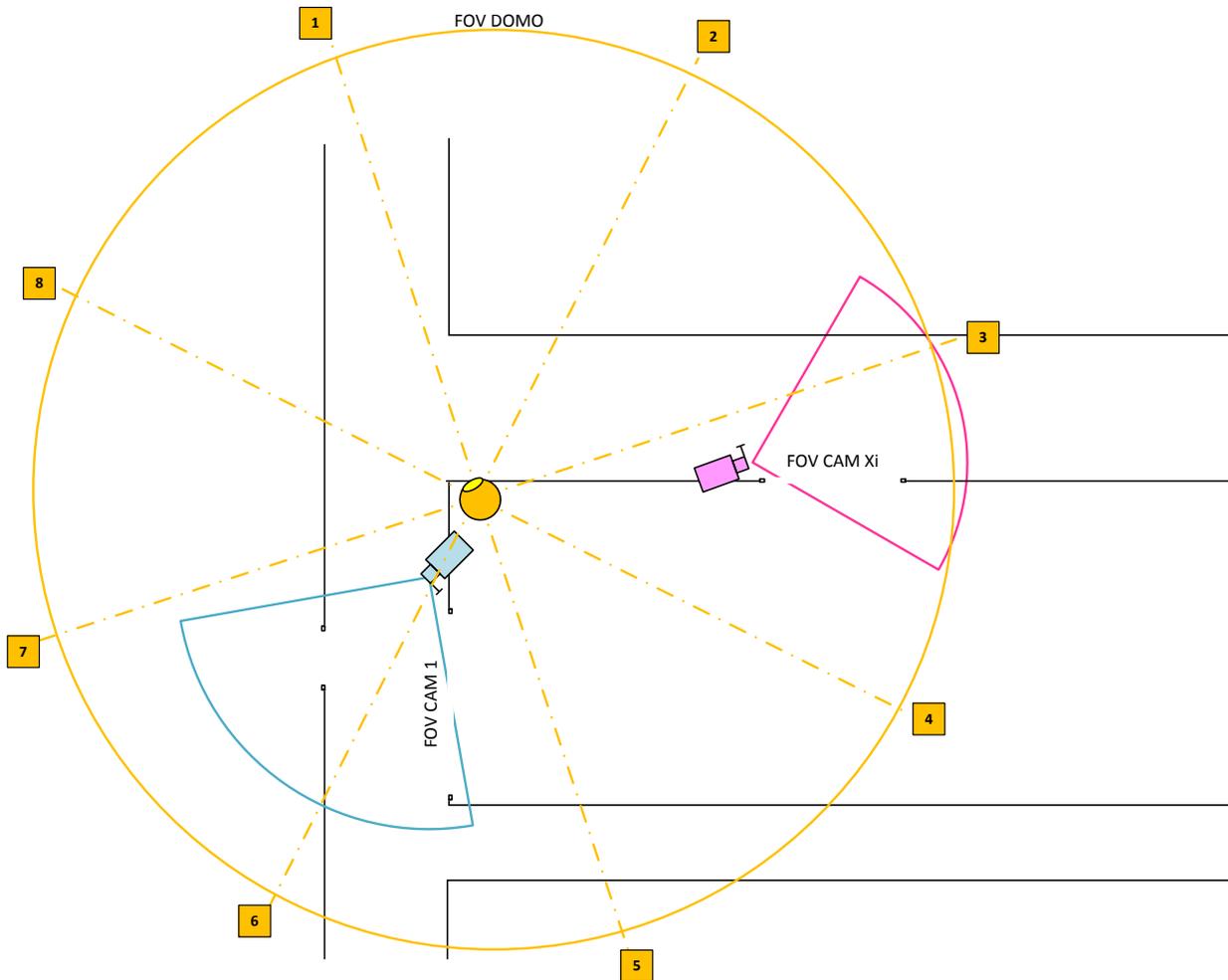


Figura 4.8 Plano de posicionamiento

Se observa que se pueden elegir tantos puntos como se requiera y permita la memoria del domo, siempre cercanos al FOV de las otras cámaras y al menos uno que sea el lugar a donde vuelva el domo después de la localización.

4.5. Generación y envío de reportes a usuario

Cuando se detecta que un objeto existe en una región de interés, se hace el procesamiento de video para tener una imagen que contenga detalles mejor definidos. Una vez terminada esta secuencia se complementa generando un reporte a usuario con la fecha, la hora y la fotografía del evento.

En la Fig. 4.9 se muestra la lógica para generar el reporte a usuario.



Figura 4.9 Generación y envío de reportes a usuario

Con las funciones de LabVIEW para Microsoft Office es posible llenar un reporte basado en una plantilla previamente definida. Una de las formas más sencillas es tomando una tabla en el documento y rellenar las celdas de manera programada. Después se guarda y se libera la memoria.

Para tener un documento general y de fácil acceso, se transformó por medio de LabVIEW y el SDK de Bull Zip Writer, el archivo de Office Word a PDF bajo la siguiente lógica: [60]

1. Cargar la configuración de Bullzip.PdfWriter.
2. Establecer el nombre de la impresora.
3. Crear un archivo temporal para comprobar su estado de impresión.
4. Establecer todos los parámetros de nombre de archivo.
5. Realizar los ajustes de PDF para ejecutarse sólo una vez.
6. Cargar las utilidades de Bullzip.PdfWriter.
7. Imprimir el archivo de origen.
8. Esperar el tiempo en que el archivo se crea.
9. Comprobar si el archivo existe.
10. Eliminar el archivo temporal, liberar memoria.

Una vez que se ha guardado el archivo PDF, éste se puede adjuntar en un correo electrónico. LabVIEW tiene funciones para envío a través de SMTP. Sin embargo, estas funciones no soportan la autenticación que casi todos los servicios comerciales tienen, por lo que fue necesario emplear un programa para envío de correos con archivos, con la limitante que está desarrollado para gmail de Google Inc., aunque puede ser modificado para otros servidores. [61]

En la Fig. 4.10 se muestra el diagrama de jerarquía para la elaboración de reportes a usuario correspondiente a la lógica antes explicada. Una vez más, la variable global proporciona una sincronización y constante comunicación con el resto de la aplicación.

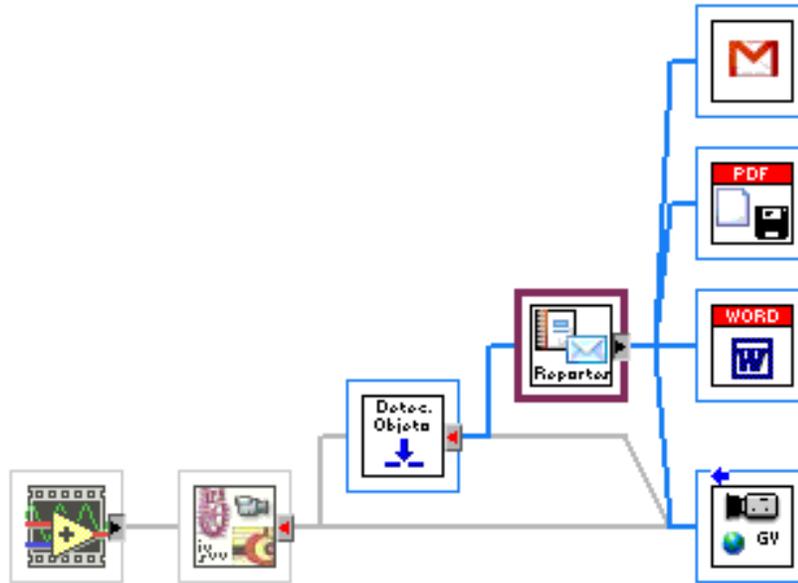


Figura 4.10 Diagrama de jerarquía para la generación de reportes a usuario

4.6. Conclusiones

En este capítulo se explicó el diseño de los instrumentos virtuales para video-vigilancia y seguridad, recordando que se trata de un sistema base con dos cámaras fijas y una móvil, siendo la lógica X_i cámaras fijas para una móvil, trabajando un par a la vez. También se explicó la manera en que se obtiene el video en tiempo real a través de los SDK de los fabricantes (para mayor información, consultar Anexos C y D).

También se dijo que la selección de las ROI se hace de acuerdo a la amplitud de la zona que se desea vigilar, teniendo un número máximo de $2^{31} - 1$ de ellas. Aunque puede ser utilizada cualquier forma geométrica, se sugiere el uso de la línea, debido a que puede colocarse en cualquier dirección.

El preposicionamiento del domo hacia el área vigilada por la cámara que genera el evento, es muy importante para la posterior búsqueda de patrones. Esto se logra a través de establecer puntos en la memoria del propio domo a través de su menú principal.

Después de todo el procesamiento y control de movimiento, podría ser suficiente tener la imagen deseada, pero el usuario no podría visualizarla fácilmente. Para atacar este problema, se añadió la generación de reportes a usuario realizándolo primero en un archivo de Office Word para después cambiar el formato a PDF y tener una mayor disponibilidad. Finalmente se envía por correo electrónico de Gmail al usuario que desee el administrador.

CAPÍTULO 5. PROCESAMIENTO DE EVENTOS RELEVANTES

En este capítulo se explican los temas relacionados a la visión artificial empleados en el algoritmo del presente trabajo. En la Fig. 5.1 es mostrado el diagrama de bloques para el procesamiento de eventos relevantes comenzando con la obtención de imágenes de la cola para hacer una resta entre ellas, análisis de partículas y calibración intercámara en el “procesamiento de imágenes”. Después, el “control del domo” hace una comparación entre las coordenadas proporcionadas por la “búsqueda de patrón” y la división de las “áreas de vigilancia”. Una vez que se ha determinado la ubicación y además se ha hecho el acercamiento al objeto se guarda una imagen con mejor detalle.

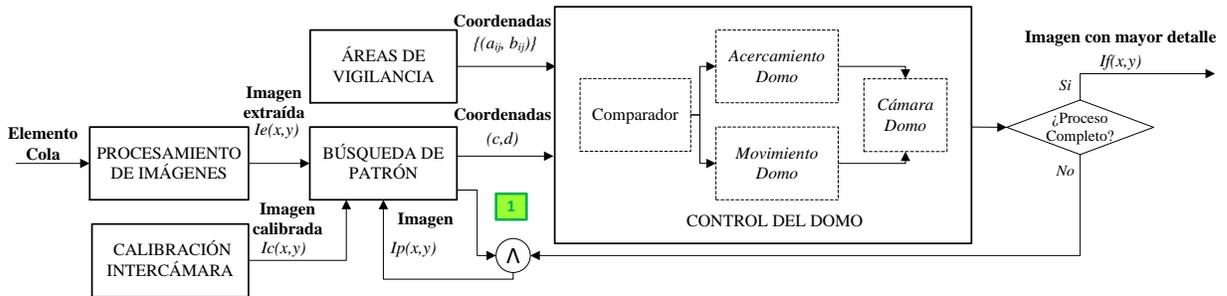


Figura 5.1 Diagrama de bloques para el procesamiento de eventos relevantes

5.1 Procesamiento digital de imágenes para determinar presencia de objetos

Esta parte de la aplicación tiene la finalidad de detectar la presencia de nuevos objetos en una imagen de alguna de las cámaras fijas, sin poseer ninguna información extra. Como se ve en la Fig. 5.1, la entrada es un elemento de cola que contiene dos imágenes que se usarán en la subsección 5.1.2, y el identificador de la cámara de generó el evento, parámetro empleado con anterioridad en el preposicionamiento.

La lógica del procesamiento se divide en tres secciones.

5.1.1 Detección de bordes

Este algoritmo es empleado para determinar que un evento ha ocurrido, a través de la detección de cambios de iluminación a lo largo de las regiones de interés.

A continuación se explica la lógica acorde a la estructura de secuencia que se emplea en la programación:

0. Inicialización de variables; abrir puerto serial.
1. Limpiar las regiones de interés que puedan existir; inicializar en cero arreglos correspondientes.
2. Agregar regiones de interés con el botón “limitar”, e indicar el número mínimo de cambios de iluminación que se requiere para tener un disparo de evento, llamado “puntos de detección”. En la Fig. 5.2 se observa esta parte en la preconfiguración:

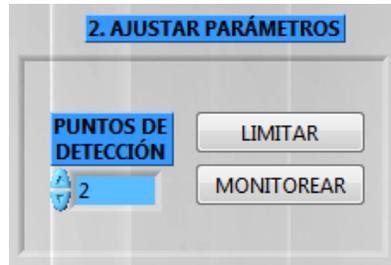


Figura 5.2 Ajustar parámetros en la detección de objetos

3. Una vez determinadas todas las ROI, se guarda una imagen inicial limpia. Ésta será el punto de referencia en el procesamiento de imágenes explicado en la siguiente sección 5.1.2.
4. Se agrupan todos los descriptores de ROI para obtener un rectángulo global que abarque todas ellas.
5. Procesamiento principal:
 - 5.0 Con cada uno de los descriptores de ROI y el número de puntos de detección como argumentos, se evalúa si existe un cambio de iluminación en todos los ejes. Esto cambiará el estado de los ejes de detección y activará la bandera de objeto detectado. En la Fig. 5.3 se observa el panel de detección en la interfaz de usuario.



Figura 5.3 Estado de ejes de detección

- 5.1 Cuando se ha detectado un cambio de iluminación, los puntos son sobrepuestos, ver Fig. 5.4. Esto no solamente genera una indicación para el usuario, sino también generar un arreglo que guarde todas las coordenadas de dichos puntos. Esto será de gran importancia para la “discriminación de partículas” explicado en la sección 5.1.3.



Figura 5.4 Imagen en tiempo real con detección de bordes

5.2 Si uno o más objetos se han detectado, se ejecuta el bloque de “presencia del objeto” (sección 5.1.2) para finalmente mover la cámara como al objetivo.

6. Cerrar puerto serial.

5.1.2 Presencia de objeto

En esta sección se obtienen los nuevos objetos que aparecen en la imagen. Esto se realiza teniendo dos imágenes, una es limpia, después de seleccionar las regiones de interés; y la otra cuando el objeto ha entrado a escena.

Las regiones de interés empleadas son líneas, pero éstas tienen un rectángulo que las encierra. En la Fig 5.5 se muestra un conjunto de ROI. Si se toma las esquinas superior izquierda mayor (punto 1) y la que esté más a la derecha abajo de entre todas ellas (punto 2), se obtiene el rectángulo exterior que abarca a todas. Esta será la región de interés global.

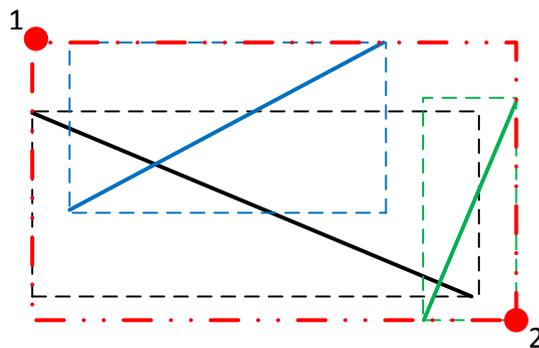


Figura 5.5 Región de interés global

Con esas coordenadas, y con la finalidad de establecer de manera más fuerte los límites de procesamiento, además de eficientar el proceso, las imágenes completas son cortadas para solo analizar esas porciones de imágenes. Se tienen, por tanto, dos fotogramas cortados en la misma proporción y ubicación.



Entonces se procede a realizar una resta de imágenes. Esta operación matemática básica tiene la finalidad, al igual que en la aritmética, de señalar en escala de color o grises los píxeles que no existen en la imagen limpia, convirtiendo todos los demás en cero (negro), como la Fig. 3.11 mostrada en el marco teórico.

Esto es de gran utilidad, pues se tiene una imagen en escala de grises con un gran contraste, razón por la cual es conveniente emplear un binarizado sencillo. Para ello se toman los datos de su histograma, tomando el valor medio como umbral, de manera que se divida en dos clases la distribución de la imagen.

Ahora comienza el análisis de partículas. Primero se rellenan los huecos internos en todas los elementos que se hayan binarizado. Después se erosiona 5 veces con una ventana de 3x3. Esto garantiza que la mayor parte del ruido se elimine tras las erosiones. Una vez realizado lo anterior se procede a filtrar las partículas, eliminando las que no se cumplan con los siguientes parámetros:

Tabla 5.1 Parámetros de filtrado de partículas

PARÁMETRO	VALOR	EXPLICACIÓN
Tipo de filtrado	Por área	Este concepto de filtrado permite quitar las partículas más pequeñas y/o más grandes.
Límite menor	400	Estos límites están dados acorde a la resolución de las imágenes en píxeles.
Límite mayor	800	
Filtrado fuera de límites	F	Esto presupone el filtrado dentro de los límites establecidos
Medidas en unidades reales	F	La medición se hace en píxeles, si este caso fuera verdadero, se tendría que realizar una calibración previa

Finalmente se obtienen los datos de localización de las partículas, guardando en un arreglo las coordenadas del rectángulo circunscrito, lo que será de muy importante para la siguiente sección.

5.1.3 Discriminación entre partículas

De manera ideal, solo debería detectarse una partícula en la imagen, sin embargo esto puede no suceder. Para evitar confusión en el sistema, se discriminan los objetos debido a su cercanía con las coordenadas del punto donde se detectó presencia en las regiones de interés.

Supóngase que se tienen dos coordenadas (X_1, Y_1) y (X_2, Y_2) que corresponden a partículas obtenidas después del análisis donde se eliminaron espurios. Además considérese que el punto donde que disparó el evento en la región de interés es (X_a, Y_a) , pero se tiene un margen de 10 píxeles para definir si la partícula corresponde al objeto. Ver Fig. 5.6.

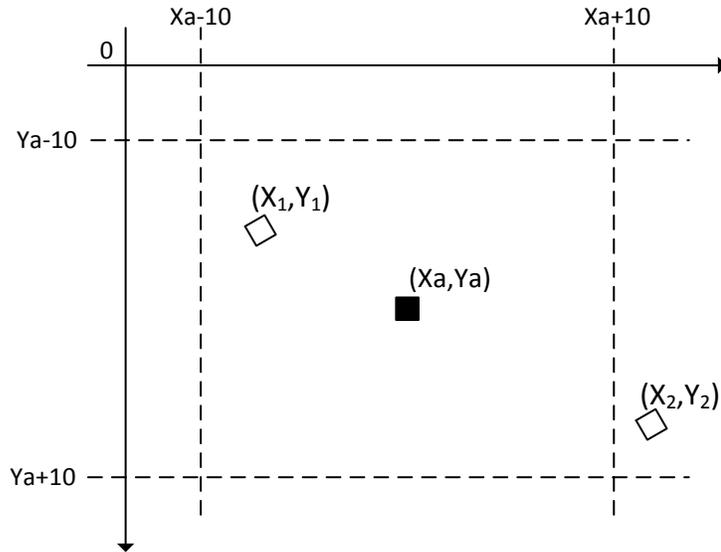


Figura 5.6 Discriminación de partículas

De manera matemática, la lógica de operación es de la siguiente forma:

$$P_i = O_i \leftrightarrow [X_a - 10 \leq X_i < X_a + 10] \wedge [Y_a - 10 \leq Y_i < Y_a + 10] \quad \text{Ec. 5.1 Objeto de interés}$$

Esto es, la partícula P_i (con coordenadas iguales a su subíndice) es el objeto de interés O_i , si y solo si se encuentra dentro de los límites establecidos por la tolerancia ± 10 píxeles en ambas direcciones.

De esta forma se puede obtener la imagen extraída $I_e(x, y)$ que contiene solo la partícula que causó el evento, eliminando todas las demás posibilidades; para la Fig. 5.6, solo la partícula P_1 se considerará. Es importante mencionar que en la práctica no solamente un punto en la ROI es detectado, por lo cual se tiene un arreglo que almacena todas las coordenadas que generaron evento y cada una de las partículas tendrá esta evaluación para encontrar la adecuada.

5.2 Calibración intercámara

Debido a que la imagen inicial después del procesamiento y extracción proviene de una cámara fija y se desea encontrar el objeto de interés en el FOV de la cámara móvil, es necesario compensar los niveles iluminación antes de buscarlo por reconocimiento de patrones.

Lo anterior se realiza midiendo la intensidad de luz de la imagen del domo para tomar algunos parámetros y efectuar la transformación en el histograma de la imagen extraída de la cámara fija.

En la Fig. 5.7 (a) se muestra el histograma original de la imagen fija y en el inciso (b) el histograma de la imagen del domo.

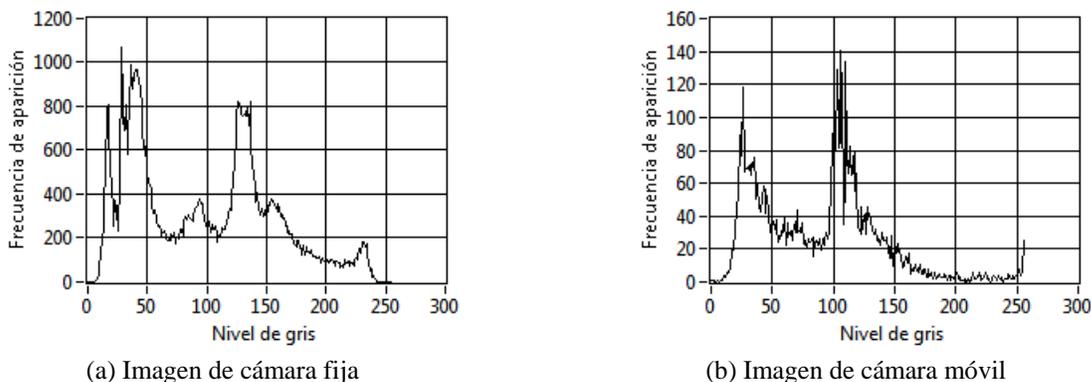


Figura 5.7 Histogramas antes de calibración

Los parámetros se muestran en la Tabla 5.2. En la segunda columna se encuentran los valores del histograma del domo y en la tercera el factor que afecta el parámetro para ajustarse a las condiciones buscadas.

Tabla 5.2 Parámetros para calibración intercámara

PARÁMETRO HISTOGRAMA	VALOR	FACTOR	PARÁMETRO TRANSFORMADO
Intensidad mínima	0~7	(× 2)	Brillo
Desviación estándar	48.80~72	(÷ 100)	Gamma

El resultado de las imágenes se muestra en la Fig. 5.8, en el inciso (a) se tiene una porción de la imagen de referencia tomada del domo, en el (b) la imagen de la cámara fija original y en (c) la imagen de la cámara fija transformada.



Figura 5.8 Resultado en imágenes después de calibración

Como se puede ver, la imagen calibrada $I_c(x, y)$ es más oscura que el FOV del domo, esto podría corregirse modificando también el contraste, que tiene el nivel neutro (45). Sin embargo, para este caso es suficiente con tener la imagen de este modo para que el reconocimiento de patrones pueda encontrarlo.

5.3 Búsqueda de patrón

El reconocimiento de patrones tiene muchas utilidades. Para este caso se emplea para encontrar el objeto que ha entrado a escena, obteniendo su ubicación y pudiendo con ello, localizarlo continuamente en el FOV del domo hasta que esté centrado.

En la Fig. 5.1 se encuentra un rectángulo con el número 1, la lógica se explica a continuación: una vez que se ha aprendido la imagen plantilla a buscar $I_p(x, y)$, se espera a que se haga el control del domo, esto para obtener un FOV actualizado del domo. Después de se realiza el movimiento del domo pero el proceso no se ha completado, entonces se toma la misma plantilla $I_p(x, y)$ para buscarlo nuevamente. Esto asegura que el objetivo sea igual y así evitar un error en la selección del objeto de interés.

La búsqueda de patrón se compone de dos etapas: el aprendizaje y la búsqueda.

5.3.1 Aprendizaje

En esta etapa, se extraen los rasgos descriptivos de los objetos de interés contenidos en la plantilla. Para este caso es necesario extraer la información de partículas que estén rotadas, escaladas y/o trasladadas.

Con los datos anteriores se hace una descripción del objeto, escribiendo en una cadena la información necesaria para realizar la búsqueda. Todo lo cual se guarda en la misma imagen de plantilla.

5.3.2 Búsqueda

Durante la fase de búsqueda, el descriptor se extrae de la imagen de la plantilla y se utiliza para buscar el objeto en la imagen de inspección. Para esto es necesario especificar el contraste siendo éste la diferencia mínima esperada entre el valor de píxel máximo y el valor mínimo de píxel en la imagen de inspección. Además definir si los objetos que se buscan pueden estar rotados y finalmente la información que se guardó en la imagen durante la etapa de aprendizaje.

En la salida se tiene un conjunto de datos del objeto encontrado. Las más importantes para la localización del mismo son las coordenadas del centro del objeto encontrado en la imagen de inspección y la dimensión del rectángulo donde está inscrito. Otros datos son el porcentaje de éxito en la búsqueda y el número de patrones encontrados. Para este caso solo se desea encontrar un objeto. En la Fig. 5.9 (a) se muestra la imagen con el patrón encontrado y en el inciso (b) los parámetros para cada objeto encontrado.



(a)



(b)

Figura 5.9 Búsqueda de patrones

5.4 Lógica de mando de la cámara tipo domo

Esta sección es una de las más importantes de la aplicación, pues gracias a esta cámara se tendrá el movimiento para localizar al objeto de interés, haciendo un acercamiento que proporcionará una imagen con mayor detalle, siendo capaz la observación de características importantes. Esto se logra a través de la programación del protocolo de comunicación Pelco-D para dar el mando PTZ, además de la lógica aplicada sobre la definición de áreas de vigilancia correspondientes al movimiento y de una relación de aspecto para el acercamiento. Por último, emplear el preposicionamiento a una localidad conocida facilita la espera en el escaneo para una nueva búsqueda.

5.4.1 Áreas de vigilancia

Según la Fig. 5.1, este bloque de programación es la base para el control del domo, pues fracciona su FOV de 2 dimensiones (2D) en 81 áreas de vigilancia, esto con la finalidad de correlacionar cada una de las porciones con las direcciones de acciones de mando. De esta forma, el campo de visión queda de la siguiente forma:

0	1	2	3	4	5	6	7	8
9	10	11	12	13	14	15	16	17
18	19	20	21	22	23	24	25	26
27	28	29	30	31	32	33	34	35
36	37	38	39	40	41	42	43	44
45	46	47	48	49	50	51	52	53
54	55	56	57	58	59	60	61	62
63	64	65	66	67	68	69	70	71
72	73	74	75	76	77	78	79	80

	A0	Centro, se detiene.
	A1	Noreste
	A2	Norte
	A3	Noroeste
	A4	Este
	A5	Oeste
	A6	Sureste
	A7	Sur
	A8	Suroeste

Figura 5.10 División del FOV en áreas de vigilancia

En el cuadro central se tienen áreas de cada color para hacer un movimiento fino y poder centrar de mejor manera al objeto.

Las coordenadas son guardadas en una matriz de acuerdo a las dimensiones de cada área de vigilancia teniendo como referencia la esquina superior izquierda y la inferior derecha para su comparación. Esto se expresa matemáticamente en la Ec. 5.2

$$(x, y) = \{(a, b)_{i,j} | a \in x \wedge b \in y\} \quad \text{Ec. 5.2 Conjunto de pares ordenados de áreas de vigilancia}$$

Donde:

$$i, j = 0, 1, 2, \dots, (m^2 - 1) \quad m: \text{Es el número de áreas de vigilancia por lado, asumiendo que tendrá una representación de matriz cuadrada.}$$

5.4.2 Control del domo

El control del domo se logra a través del protocolo de comunicación Pelco-D y RS-485, por medio de una cadena de comando en hexadecimal que es enviada para su manipulación.

En la Fig. 5.11 se muestra el diagrama de bloques de la lógica de mando del domo, comenzando con el cálculo de la condición inicial a partir de las coordenadas de las áreas de vigilancia y del objeto de interés (punto 1). Tomando como referencia la Ec. 5.2, se comparan por pares de coordenadas para conocer el área de vigilancia donde el objeto se encuentra, realizando m^2 cálculos bajo la Ec. 5.3, comenzando con el índice cero.

$$(x_2 \geq P_x > x_1) \wedge (y_2 \geq P_y > y_1) \quad \text{Ec. 5.3 Cálculo de condición para asignación de área de vigilancia}$$

Donde:

- (x_1, y_1) Coordenada área de vigilancia, esquina superior izquierda.
- (x_2, y_2) Coordenada área de vigilancia, esquina inferior derecha.
- (P_x, P_y) Coordenada de objeto de interés derivada de la búsqueda de patrones.

Dependiendo del número de cálculo que resultó verdadero, el selector de área de vigilancia (Fig. 5.11, punto 2) dará el identificador para definir la dirección del movimiento que el domo realizará.

Una vez que se encuentra en la zona 40, el mando del domo se detiene para tomar hacer el acercamiento al objetivo.

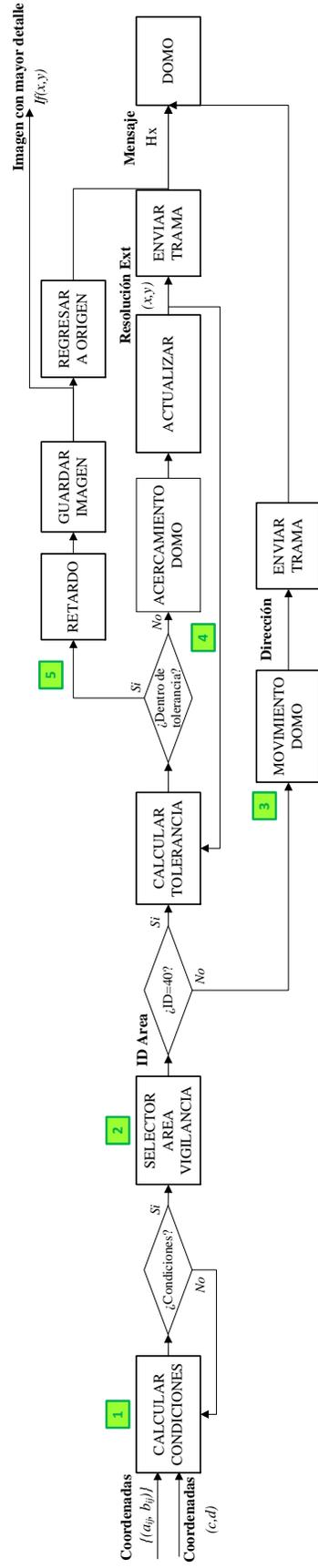


Figura 5.11 Diagrama de bloques del control del domo

- *Movimiento*

Según la Fig. 5.11 (punto 3), si el identificador es diferente a 40, entonces entra a la rutina del movimiento. Una vez que se sabe cuál es la ubicación del objeto, se manda un único pulso para trasladar el domo en esa dirección a través de un mensaje que es enviado vía RS-485. Se envía la trama por medio de la comunicación al domo para realizar el movimiento como se vió en la Fig. 5.10. Después de esto, se regresa a la Fig. 5.1 donde el proceso no está completo, por lo que vuelve a la búsqueda de patrón para que el ciclo se repita hasta que el objeto se encuentre en el centro del FOV del domo.

- *Acercamiento*

De acuerdo a la Fig 5.11 (punto 4), cuando el selector de área de vigilancia proporciona el identificador 40, quiere decir que la imagen está lista para comenzar con el acercamiento. Las dimensiones del FOV del domo se muestran en la Fig. 5.12 (a); para el acercamiento, lo primero que se debe calcular son las tolerancias establecidas en la Ec. 5.4.

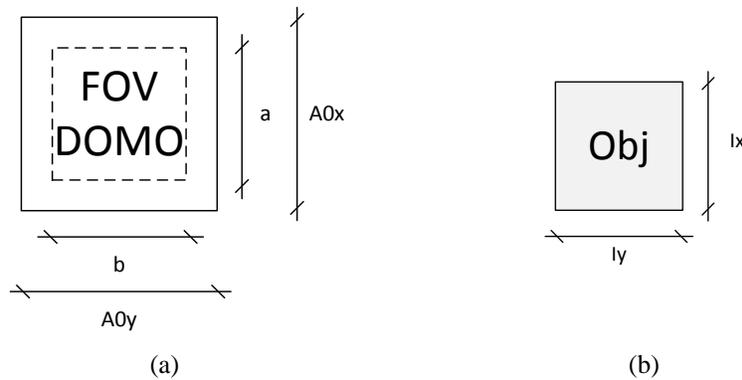


Figura 5.12 Dimensiones de área 40 y objeto

Donde:

$$a = A0_x - 10$$

$$b = A0_y - 10$$

Ec. 5.4 Tolerancia para acercamiento

Después de realizar pruebas de ajuste, se encontró que el factor 1.1 es adecuado para hacer una relación de aspecto, este parámetro se ocupa para redimensionar la imagen extraída solo como referencia para los incrementos que se hacen en el acercamiento, pero no afecta completamente a la imagen.

La lógica se detendrá hasta que el tamaño del objeto sea menor o igual a la resolución en alguno de los lados del FOV del domo, como se muestra matemáticamente en la Ec. 5.5.

$$(I_x \geq a) \vee (I_y \geq b)$$

Ec. 5.5 Expresión matemática para detener acercamiento

El pseudocódigo se presenta a continuación:

```

While  $((I_x \geq a) \wedge (I_y \geq b))$ 
{
     $I_{Nx} = (1.1)(I_x)$ ;
     $I_{Ny} = (1.1)(I_y)$ ;
     $I_P(x, y) = \text{resize}(I_{Nx}, I_{Ny})$ ;
     $I_x = \text{size}.x(I_P(x, y))$ ;
     $I_y = \text{size}.y(I_P(x, y))$ ;
}

```

- *Salida*

Una vez finalizado el cálculo y operación de acercamiento, Fig. 5.11 (punto 5), se procede a guardar la imagen final $I_f(x, y)$ anteponiendo un retardo de 7 segundos, tiempo que necesita la comunicación con el SDK para actualizar la imagen mostrada en LabVIEW. Una vez que se tiene la imagen en una localidad conocida, se manda el domo a la posición preestablecida que se elige, bajo la misma lógica que en el preposicionamiento.

5.4.3 Protocolo de comunicación Pelco-D

La manipulación de la posición del domo se realiza por medio de una tarjeta de comunicación serial RS-485 con conexión USB para unir la computadora Servidor con el DVR; de manera que la programación del protocolo Pelco-D desarrollado en LabVIEW permite tener características tales como movimiento manual que provee el *software* comercial, con la diferencia que éste proporciona un control automático del domo.

La programación del protocolo Pelco-D (Anexo B) comienza con la función de “abrir serial” y así entablar la comunicación con la tarjeta RS-485/USB, abriendo el puerto con los parámetros mostrados en la Tabla 5.3.

Tabla 5.3 Datos de configuración para comunicación serial

Nombre de dispositivo	COM 4, es necesario verificar en dispositivos
Velocidad de transmisión	9600
Paquete de bits	8
Paridad	Ninguno
Bits de parada	1
Control de flujo	Ninguno

Dependiendo del cambio que se haga de dirección o acercamiento, las posiciones de los arreglos binarios serán cambiadas. Además, otros parámetros como velocidad para movimiento y número identificador del domo (siempre 1).

Después de ello, se debe hacer una transformación a para enviar la trama en el formato adecuado (Fig. 5.13). Consiste en la inicialización de variables; la conversión de formatos en los datos; el cálculo de

la suma de verificación (*checksum*), la cual es una función que tienen como propósito principal detectar cambios accidentales en una secuencia de datos para proteger la integridad de datos, verificando que no haya discrepancias. Finalmente todos los datos se empaquetan en una cadena que es enviada al bloque “escribir serial”.

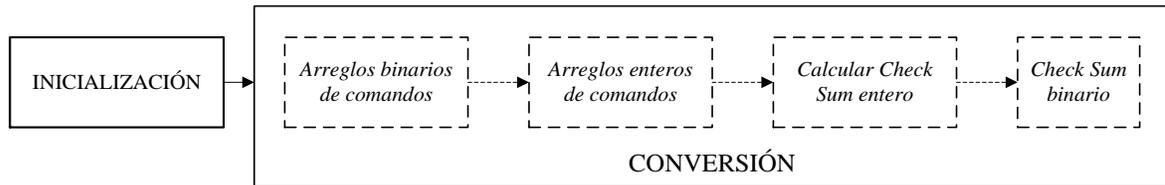


Figura 5.13 Diagrama de bloques para convertir trama

La comunicación serial concluye cerrando el puerto. Esto es muy importante, pues como se vio en el capítulo 3, de lo contrario podría causar problemas operativos.

5.5 Diagrama de jerarquía

En la Fig. 5.14 se muestra el esquema sobre los bloques de programación descritos en este capítulo; está numerado de acuerdo a la sección en donde fue explicado. El punto 5.1 trata el procesamiento digital de imágenes; el 5.2 la calibración intercámara; en el 5.3 la búsqueda de patrón; y el recuadro de línea discontinua es el punto 5.4 la lógica de mando del domo, que incluye la división en áreas de vigilancia, el control de movimiento/acercamiento y el manejo del protocolo Pelco-D que hace un llamado al preposicionamiento para regresar a un origen dado, además del envío del mensaje por comunicación serial.

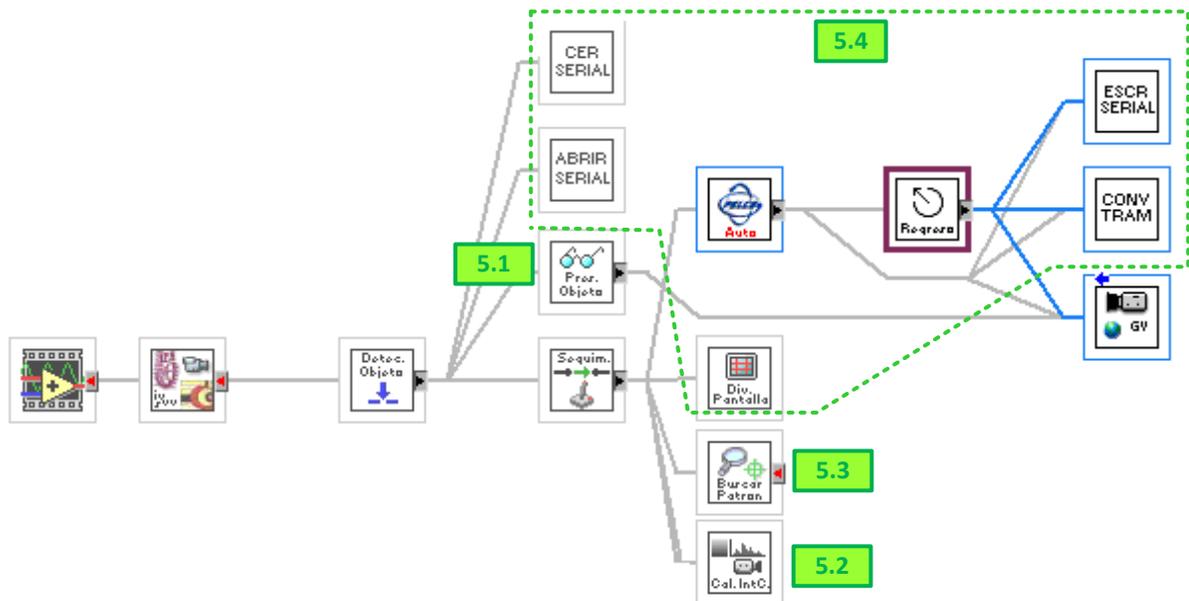


Figura 5.14 Diagrama de jerarquía para el procesamiento de eventos relevantes



5.6 Conclusiones

La visión por computadora es una rama de la inteligencia artificial que permite alcanzar varios objetivos, entre los cuales se encuentran analizar imágenes a través de su procesamiento, esto solo como operación preparatoria para tareas más complejas como la detección de eventos relevantes, empleando búsqueda de patrones para ello.

En este capítulo se explicó el desarrollo de los algoritmos para analizar los fotogramas de video en tiempo real. El primer paso consiste en descubrir que un objeto ha entrado en escena a través de la detección de bordes, que observa el cambio de contraste en la imagen a lo largo de una línea (ROI). El siguiente paso es procesar las imágenes por medio de una resta entre ellas, binarizado y posteriormente con un análisis de partículas para determinar cuántos nuevos objetos existen. Finalmente, la discriminación entre partículas complementa al análisis previo de éstas, pues evalúa la cercanía que existe entre el centro del objeto y las coordenadas del punto donde el evento se desencadenó.

Por otro lado, la calibración intercámara es una parte fundamental de los sistemas duales de video-vigilancia, los cuales combinan imágenes de diferentes fuentes para encontrar un objetivo en común. Debido a la diferencia de condiciones entre estos captores, es necesario asemejar los niveles de iluminación como actividad previa a la búsqueda de patrones, que finalmente determinará la ubicación de los objetos para su localización por medio de la cámara móvil.

Una de las partes más importantes del presente trabajo es la lógica de mando del movimiento del domo, pues sirve para ubicar el objeto de interés dentro de las áreas de vigilancia, hacer un acercamiento y con ello conseguir una imagen de mejor calidad.

CAPÍTULO 6. RESULTADOS EXPERIMENTALES

En este capítulo se presenta la integración del sistema y la obtención del video en tiempo real. Los resultados de las pruebas finales que se realizan únicamente durante el día no lluvioso. Se muestran los tiempos de ejecución para todas las secciones de la aplicación, así como imágenes con mayor detalle de los resultados experimentales. Adicionalmente, se muestra la programación de la “detección de movimiento con vectores de movimiento” como trabajo adicional, sienta las bases para realizar el seguimiento de objetos.

6.1 Integración del sistema

En la Fig. 6.1 se muestra la imagen del montaje final del sistema. Como se mencionó en la sección 4.1.1, se tienen dos cámaras analógicas, una de ellas fija y la otra domo, conectadas a un DVR para el almacenamiento y visualización del video. Además se tiene la cámara digital tipo IP y una tarjeta de comunicación RS-485/USB para el mando del domo.



Figura 6.1 Sistema de videovigilancia instalado

Ambas cámaras analógicas (puntos 1 y 2) se encuentran montadas en techo y pared en el Centro de Investigación en Computación con vista hacia el estacionamiento exterior. La IP (punto 3) está sobre un tripie para mayor movilidad al hacer las pruebas. La tarjeta de comunicación (punto 4) está conectada a la computadora central donde se encuentra el software desarrollado.

6.2 Sistema cliente/servidor para la transmisión y recepción de imágenes

Como se explicó en la sección 4.2, el video se obtiene a través de los SDK de cada dispositivo. Estas imágenes tienen un retraso con la realidad de aproximadamente 0.5 s en el software comercial. Cuando se

conecta con LabVIEW solo para la visualización, se tiene un retraso con una diferencia despreciable con respecto al caso anterior. Sin embargo, cuando se realiza el procesamiento, ese retardo es mayor, de manera que al final de la rutina, es necesario adicionar 7 segundos en la espera de la actualización de la imagen final en LabVIEW.

En la Fig. 6.2 se muestra la pantalla del panel frontal para la visualización del video.

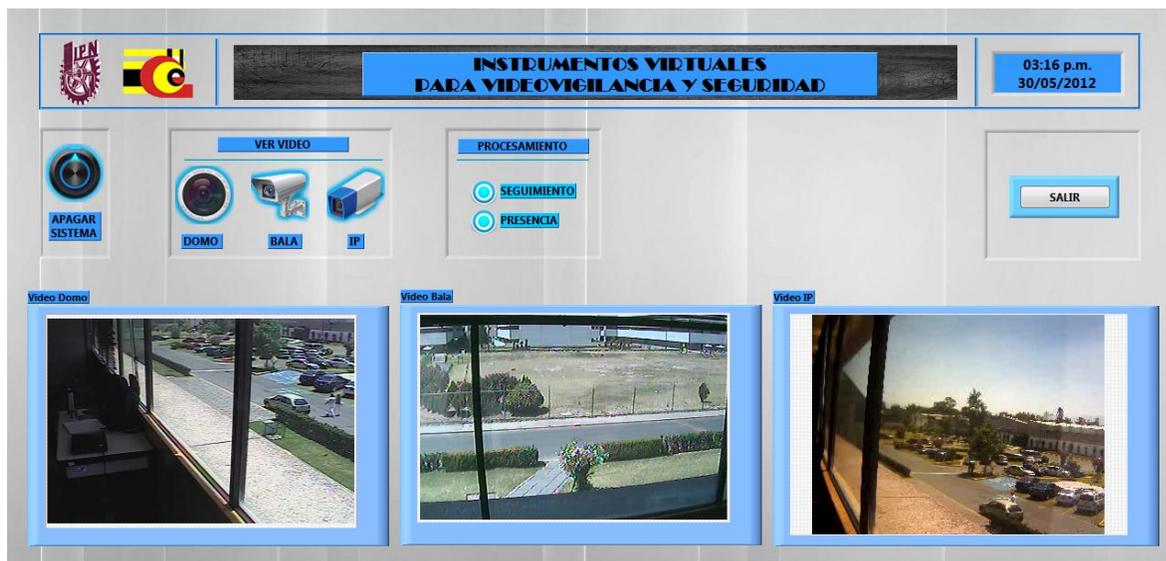


Figura 6.2 Panel Frontal para la visualización del video en LabVIEW

6.3 Detección de eventos relevantes en estacionamientos restringidos

A continuación se muestran imágenes como resultado del proceso que en este trabajo se reporta. Las pruebas fueron realizadas en un estacionamiento restringido, obteniendo acercamientos de diversos automóviles.

En las Fig. 6.3 se muestra claramente las placas, modelo y marca de un automóvil.



Figura 6.3 Imagen resultado 1

Al igual que el caso anterior, en la Fig. 6.4 se muestra otro auto; en esta fotografía se muestra un mayor detalle y una perspectiva más recta, lo que da una mayor posibilidad de procesamiento OCR, por ejemplo



Figura 6.4 Imagen resultado 2

En la tercer imagen resultado (Fig. 6.5) se muestra otro automóvil, sin embargo, en esta ocasión no se alcanzan a visualizar la placa, pero se me mejor el tipo y estructura de auto.



Figura 6.5 Imagen resultado 3

Por otro lado, en la Fig. 6.6 se presenta el reporte a usuario que se envía al usuario. En este documento se presenta la imagen del objeto que generó el evento, así como la fecha y la hora.



Instituto Politécnico Nacional

LA TÉCNICA AL SERVICIO DE LA PATRIA



Centro de Investigación en Computación

LABORATORIO TIEMPO REAL Y AUTOMATIZACIÓN

INSTRUMENTOS VIRTUALES PARA VIDEOVIGILANCIA Y SEGURIDAD
REPORTE A USUARIO

IMAGEN



DATOS DEL EVENTO

FECHA	20/11/2012
HORA	1:49 pm

Figura 6.6 Reporte a usuario

El sistema falla bajo las siguientes condiciones:

1. Cuando las ROI se colocan de manera inadecuada, como se explicó en la sección 4.3.
2. En el momento que un objeto en movimiento cubre alguna ROI y se envía una señal falsa, no arrojando ninguna partícula como resultado después del procesamiento.
3. Cuando los FOV de la cámara fija que genera el evento y el domo no son similares.

6.4 Evaluación de tiempos de ejecución de los métodos y algoritmos

En esta sección se muestran la gráfica de tiempos de ejecución una vez que el usuario ha seleccionado las ROI y ha configurado la opciones del envío de reportes a usuario.

La Tabla 6.1 reporta los tiempos de ejecución para todas las partes de la aplicación antes mencionadas, además del total de segundos que el sistema demora en realizar el escaneo, procesamiento, generación y envío de reportes al usuario en una exhibición.

Tabla 6.1 Tiempos de ejecución en toda la aplicación

ACTIVIDAD		DURACIÓN [s]
1	Obtención del video	2
2	Escaneo cámaras	
3	Preposicionamiento	
4	Detección de presencia	15
5	Mando del domo	
6	Generación y envío reportes a usuario	
TOTAL		17

En la Fig. 6.7 se pueden visualizar los tiempos empleados en la ejecución del programa. El resultado final es que el proceso se realiza en un promedio de 17 segundos, debido a las operaciones que se realizan en paralelo o pseudoparalelo.

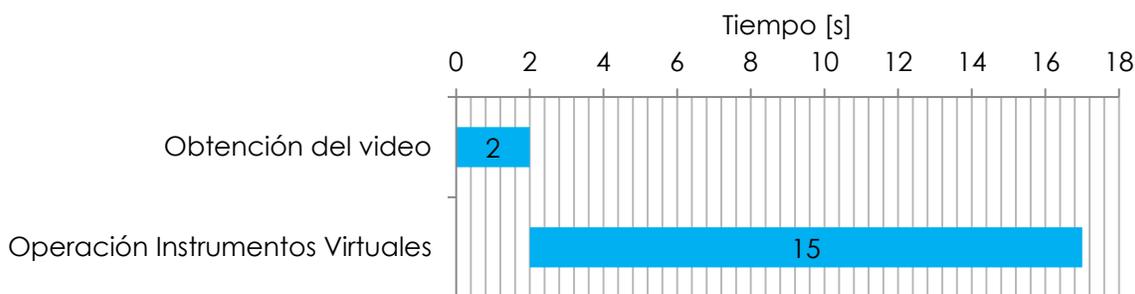


Figura 6.7 Gráfica de tiempo de ejecución para la aplicación

6.5 Trabajo adicional: detección de movimiento para seguimiento

Para detectar el movimiento, el sistema calcula el flujo óptico por el algoritmo de Horn-Shunk para obtener los vectores de movimiento; umbrala e identifica grandes componentes conectados de píxeles en movimiento. Este procedimiento es la base del seguimiento de objetos en tiempo real, pues conociendo su dirección y orientación es posible estimar la trayectoria que podría seguir el objeto.

En la Fig. 6.10 se observa el panel frontal donde se localiza la imagen binaria después de un tratamiento de imágenes y análisis de partículas a partir de la imagen que proporciona la función de “estimación de movimiento”. En la parte superior izquierda se tiene el video original en tiempo real para monitoreo; a su derecha, el marco binarizado del objeto en movimiento; abajo izquierda, la imagen del video original en escala de grises con los vectores de movimiento, el rectángulo y centroide superpuestos; abajo derecha; la imagen extraída.



Figura 6.8 Panel frontal para detección de movimiento

A continuación se describe la Fig. 6.10, la cual muestra la programación de la detección de movimiento:

1. Estimación de flujo por el algoritmo Horn y Schunk. A la salida de la función se emplea la imagen correspondiente a la magnitud de la velocidad calculada.
2. Dibujado de vectores de movimiento sobre la imagen en escala de grises.
3. Creación del histograma.
4. Obtención de parámetros a partir del histograma.
5. Aumento de contraste, obteniendo el valor absoluto de la imagen y a la salida, se ecualiza.
6. Se binariza la imagen, tomando el valor del umbral como mayor al promedio obtenido del histograma.

7. En este paso se realiza la operación morfológica de dilatación, para unir las pequeñas partes de las partículas y formar el objeto en movimiento. Después se emplea un filtro gaussiano. Se hacen dos nuevas operaciones morfológicas, de erosión, para eliminar las partículas muy pequeñas que no muestren información sustancial; y nuevamente la dilatación para volver a unir las que hayan quedado un poco separadas del objeto principal.
8. Para este caso se realiza un análisis de partículas, donde se busca rellenar los hoyos que puedan existir en el objeto principal; nuevamente se eliminan los pequeños grupos de píxeles que no estén dentro del grupo principal; se erosionan los últimos detalles de la imagen.
9. Devuelve el número final de partículas, las posiciones de cada lado para un rectángulo que delimita al objeto en movimiento, su centro de masa en dos coordenadas y el área del objeto.
10. Aquí se toman los parámetros obtenidos anteriormente, para obtener las medidas del rectángulo que encierra al objeto. Además se obtiene una relación para discriminar entre los objetos que sean de interés y los que no, tomando únicamente los más grandes y realizando el paso 11 solo en caso verdadero; lo anterior bajo la siguiente lógica:

Teniendo la siguiente figura:

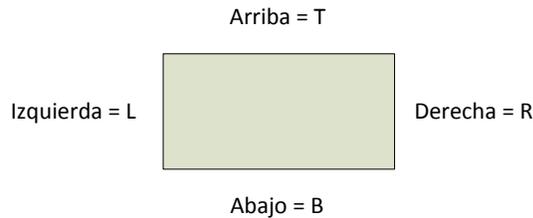


Figura 6.9 Cuadro base para discriminación de objetos

Se hace una relación de área para tomar los objetos más grandes a una referencia:

$$\left. \frac{(R \times B)}{A}, si \right\} \begin{array}{l} > 0.1, Verdadero \\ En otro caso, Falso \end{array} \quad Ec (6.1)$$

11. Este punto tiene la finalidad de completar la visualización del objeto de interés. Se extrae la porción de la imagen que contiene al objeto en movimiento para colocarla en un arreglo; se dibuja el rectángulo y su centroide en la imagen en escala de grises.

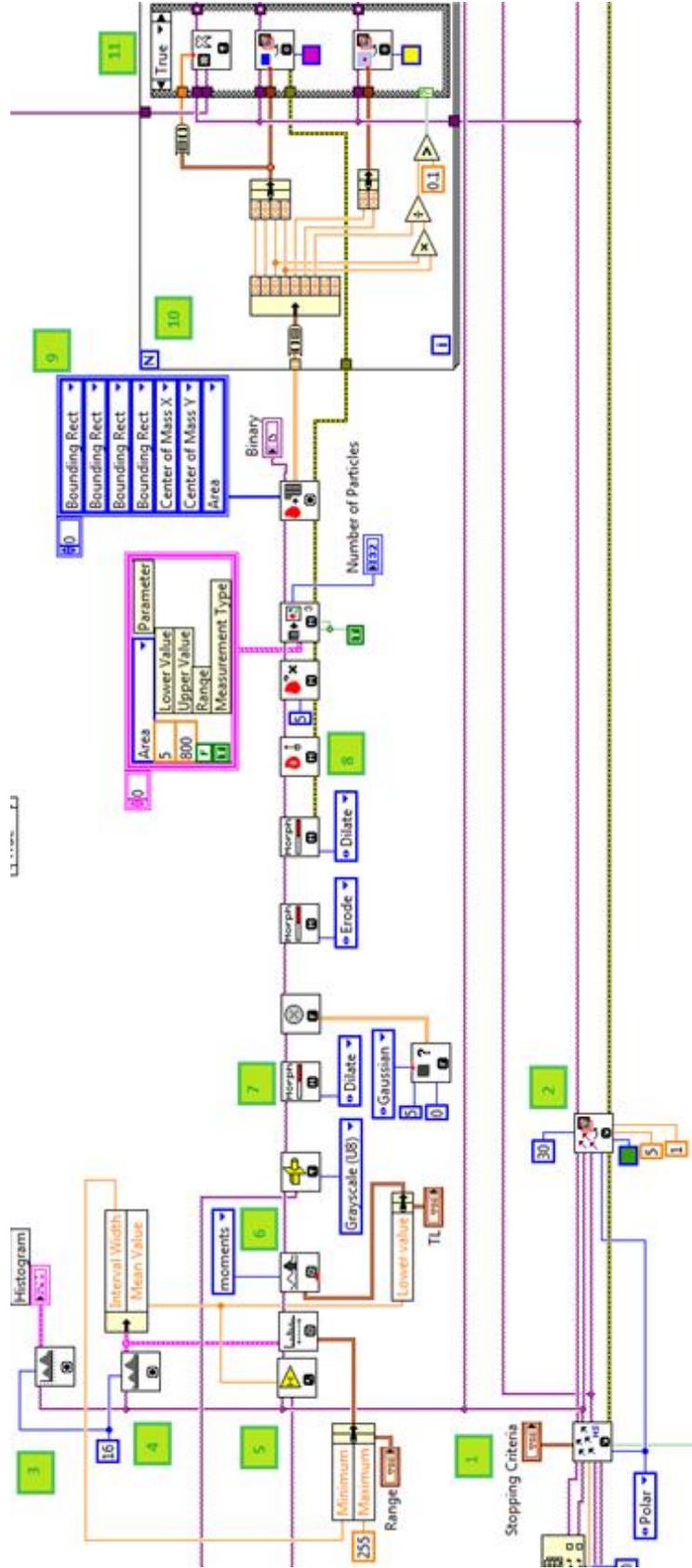


Figura 6.10 Diagrama de bloques sobre la detección de movimiento

6.6 Conclusiones

En este capítulo se mostró la imagen del sistema instalado y la evaluación de los tiempos de ejecución del sistema, obteniéndose como resultado que se consume en promedio 17 segundos en realizar todas las tareas, desde la obtención del video hasta la generación y envío de reportes a usuario.

Por otro lado, como parte de los resultados se muestran también imágenes de las pruebas realizadas en un estacionamiento restringido, donde se observan detalles característicos de autos que generaron el evento: en dos de las fotografías presentadas es posible observar la placa y nombre de modelo y marca, siendo éstas útiles en otro tipo de análisis; en la tercer imagen aunque no se puede observar la placa, se ve con mayor facilidad el modelo del automóvil.

El último resultado presentado es el reporte a usuario que es envía vía correo electrónico al usuario que el administrador determine. En este documento se presenta la imagen final con mayor detalle, así como la fecha y la hora en que el evento se produjo.

En la sección final se incluye la explicación de la detección de movimiento como trabajo adicional. Este bloque de programación hace uso de los vectores de movimiento, pudiendo obtener información valiosa a partir de éstos, tal como determinar la ubicación y dirección que el objeto de interés sigue. Esto forma parte del trabajo a futuro que puede seguirse.



CONCLUSIONES

Se desarrolló el modelo computacional con instrumentos virtuales, empleando equipo de video-vigilancia, con operación local y remota.

Se generaron métodos, algoritmos y *software* para potenciar sistemas comerciales de video-vigilancia de bajas y medias prestaciones, de manera que cualquier usuario puede tener acceso y/o acondicionar la operación de equipos de circuito cerrado.

Se integró un sistema de video-vigilancia, que trabaja con cámaras analógicas y digitales, formando un par con una de las X_i del tipo fijo de bajas/medias prestaciones y otra móvil de mediana/alta resolución, para detectar, localizar y realizar un acercamiento al objeto que generó el evento de interés, obteniendo así una imagen con mayor detalle que es colocada en un reporte a usuario enviado por correo electrónico.

Se estructuró un sistema cliente/servidor para la obtención del video en tiempo real y avisos de eventos relevantes de la siguiente manera:

Se detectó un objeto de interés que aparece en las ROI del fotograma de alguna de las cámaras fijas y se procesó la imagen para extraer el objeto de interés. Después, por búsqueda de patrones se localizó continuamente la partícula para posicionarla en el centro del FOV del domo y hacer el acercamiento adecuado, modificando las condiciones PTZ, para obtener el resultado deseado. Para comunicar la dupla de cámaras mediante *software*, se hizo una calibración intercámara antes de la búsqueda de patrón para asemejar los niveles de iluminación de los dos FOV, modificando los parámetros de brillo y gamma, teniendo como efecto una imagen muy oscura pero capaz de proporcionar la información necesaria para ser encontrada como patrón.

Los resultados presentados se obtuvieron haciendo pruebas durante el día en un estacionamiento restringido, logrando mostrar algunas características importantes como placas y modelo de los automóviles. Se comprobó que el sistema falla bajo tres condiciones que contemplan la selección inadecuada de las ROI, objetos en movimiento que causan ruido y cuando los FOV de las cámaras de interés no son similares, siendo en este último punto donde radica la importancia del preposicionamiento.

Se evaluó el tiempo de ejecución total de la aplicación, siendo éste de aproximadamente 17 segundos, empleando una cámara móvil con enfoque y manejo de diafragma automáticos.

Con el modelo computacional propuesto (ver Fig. 1.1), se prepararon las condiciones e imágenes para lograr otras tareas avanzadas, analizando los rasgos característicos de los objetos en eventos relevantes.

También se presentó la metodología sobre la detección de movimiento que sienta las bases para el seguimiento de objetos en tiempo real a través de cámaras de circuito cerrado, mediante el uso de vectores de movimiento y análisis de partículas aplicado en el video.

De esta forma, se obtuvieron tres aportaciones: desarrollo tecnológico de costo reducido; el reforzamiento de sistemas de video-vigilancia comerciales de bajas y medianas prestaciones, mediante algoritmos, métodos y *software* propios; y el desarrollo de un sistema accesible para PyME con aplicación en grandes área de vigilancia tales como edificaciones, vía pública y como complemento para otros sistemas.

Se verificó la hipótesis y se cumplieron los objetivos, general y particulares, del trabajo.

TRABAJOS FUTUROS

Para trabajos a futuro se dan las siguientes propuestas:

- a) Facilitar la posibilidad que el usuario establezca prioridades de atención en los eventos generados por las cámaras fijas.
- b) Mejorar la discriminación de partículas de acuerdo a rasgos característicos u orden de aparición en el FOV de la cámara que se esté evaluando, cuando existan más de una partícula de interés.
- c) Hacer más eficiente el algoritmo para la localización y acercamiento al objeto de interés, pudiendo intercalar estas acciones para obtener mayor velocidad y una ubicación más precisa.
- d) Con el documento aquí expuesto, se preparan las condiciones para el análisis de los eventos relevantes, pudiendo emplearse otras técnicas de visión artificial como el reconocimiento óptico de caracteres (OCR) para lectura de placas o herramientas biométricas más avanzadas para inspeccionar rostros a través de sistemas de video-vigilancia.
- e) A partir de los rasgos característicos que se pueden obtener de las imágenes, es posible implementar topologías de redes neuronales o memorias asociativas que sean fácilmente adaptables a la incorporación de nuevas categorías con el mínimo o nulo retrenamiento para la clasificación de automotores.
- f) Implementar el seguimiento de objetos en tiempo real a partir de los vectores de movimiento calculados en la sección 6.5, los cuales pueden proporcionar información sobre la dirección y sentido que la partícula sigue de acuerdo al elemento del conjunto de vectores que posea la mayor magnitud, con lo que se puede incluir un modelo matemático que tenga la capacidad de predecir la ruta que el objeto puede llevar.
- g) Adicionar algoritmos para modificar el enfoque y diafragma automáticamente y con ello poder emplear el sistema cuando la iluminación es baja (noches), dado que las luces externas o de los automóviles provocan ruido y efectos no deseados haciendo que las imágenes parezcan borrosas.
- h) Probar el sistema ampliado con la aplicación de los algoritmos y métodos presentados para otra configuración que guarde el carácter dual entre los dispositivos, colocando X_i cámaras fijas y una móvil, trabajando por pares.

REFERENCIAS

- [1] R., Cristóbal, V., Francisco y C. Carlos, *Domótica e Inmótica, viviendas y edificios inteligentes*, Tercera edición, Editorial Alfaomega Ra-Ma, México, Junio de 2011.
- [2] K., Karl D. *The effects of Noise in Man*. Segunda Edición, Academic Press, 1985.
- [3] H., Cyril. *Manual de Medidas Acústicas y Control de Ruido*. Mc. Graw Hill. 1995.
- [4] Rojo, A. y Sánchez Fernández, L.P. *Sistema de monitoreo de ruidos ambientales producidos por aviones en el AICM (Aeropuerto Internacional de la Ciudad de México)*, Tesis de Maestría terminada, CIC-IPN, junio 2008.
- [5] Valencia, M. y Sánchez Fernández, L.P. *Sistema distribuido para análisis de ruidos ambientales*, Tesis de Maestría terminada, CIC-IPN, diciembre, 2007.
- [6] Domínguez Urbán María Guadalupe, Sánchez Fernández, L.P. *Medición y procesamiento avanzado de indicadores de ruido, en zonas críticas localizadas dentro del Distrito Federal*. Tesis de Maestría, 2009.
- [7] Villafuerte Ramírez, Miguel Santiago y Sánchez Fernández, L.P. *Desarrollo de un Módulo de Adquisición de Señales Acústicas de Alto Rango Dinámico*. Tesis de Maestría terminada, 2010.
- [8] López Pacheco María Guadalupe, Molina Lozano Herón y Sánchez Fernández L.P. *Clasificador difuso de señales acústicas ambientales basado en análisis de componentes independientes*. Tesis de Maestría en desarrollo, 2010.
- [9] Díaz Gudiño Luis Jorge, Sánchez Fernández L.P. y Oleksiy Pogrebnyak. *Modelo computacional para el análisis avanzado de señales de ruidos ambientales*. Tesis de Maestría en desarrollo, 2010.
- [10] Sánchez Fernández, L.P. *Proyecto CONACYT: Sistema Avanzado de Monitoreo Ambiental de Sonidos y Vibraciones, Clave: 51283-Y, Vigencia 2007-2010*.
- [11] *Convenio de desarrollo y transferencia de tecnología: "Sistema Distribuido de Monitoreo Permanente Experimental, de Ruidos Ambientales en Puntos Críticos del Centro Histórico de La Ciudad de México (SIMAR-CH)". 2008-2009. IPN-ICYTDF*
- [12] Sánchez Fernández, L.P., Oleksiy Pogrebnyak, José Luis Oropeza Rodríguez and Sergio Suárez Guerra: *Noise pattern recognition of airplanes taking off: task for a monitoring system*. *Lecture Notes in Computer Science*, Springer-Verlag, Germany. Vol. 4756, pp. 831-840, 2007.
- [13] Sánchez F. L.P., L. Pérez and L.A. Sánchez. *Environmental noise monitoring. 1st stage: Aircrafts noise patterns recognition*. *Research in Computing Science*, Vol. 24, pp. 95-104, 2006. LatinIndex Data base. ISSN 1870-4069
- [14] Sánchez Fernández, L.P. R. Herrera Charles, Oleksiy Pogrebnyak. *Spectral patterns for the generation of unidirectional irregular waves*. *Lecture Notes in Computer Science*, Vol. 3773. 2005. ISSN 0302-9743, Springer-Verlag, Germany. ISI Data base.
- [15] Sanchez Fernandez, L.P., Oleksiy Pogrebnyak and Cornelio Yañez. *Neural Network and Trend Prediction for Technological Processes Monitoring*. *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, Vol. 3789. 2005. ISSN 0302-9743, Springer-Verlag, Germany. ISI Data base.
- [16] Sánchez Fernandez, L.P., L. Gorostiaga, Oleksiy Pogrebnyak, *Method for Alarm Prediction*. *Proceedings Second International Conference on Informatics in Control, Automation and Robotics*, Vol. IV, Barcelona, Spain, 2005. ISBN 972-8865-31-7, Indexed in Data Base BDLP, Germany
- [17] Sánchez, L.P., *Error en las mediciones computarizadas de oleaje en laboratorios de investigación, Ingeniería Hidráulica en México*, Vol. XIX, No. 2, abril-junio de 2004. ISSN-0186-4076 (ISI Data base: Science Citation Index Expanded)

- [18] Roman, R., López J., Dugeon, O., Lacoste, M., Plaza, P. Y., Bel, M., Advanced secure multimedia services for digital homes. En *Springer Science+Business Media*, 2010.
- [19] Chan, M., Estève, D., Escriba, C., Campo, E., A review of smart homes-Present state and future challenges. En *Computer methods and programs in biomedicine 91*, 2008.
- [20] Imai, Y., Sugiue, Y., Hori, Y. y Masuda, S., Application of a Remote Surveillance System with Mobile Phone-Enhanced User Interface.
- [21] G. Forman et. al., *The challenges of mobile computing*, IEEE Computer, April 1994, pp. 38-47.
- [22] Sánchez, L.P, Patrones de sistemas de primero y segundo orden, en un ambiente de instrumentación virtual, *Computación y Sistemas*, Vol. VII, No. 2, 2003 (registrada en el Índice de Revistas Mexicanas de Investigación Científica y Tecnológica del CONACYT).
- [23] Sistemasmgs.com, Video-vigilancia IP, Mantenga el control, Sistemasmgs.com, revisado 19 Noviembre 2010, <<http://sistemasmgs.com/video-vigilancia.html>>.
- [24] J. E. Boyd, E. Hunter, P. Kelly, L. Tai, C. Phillips, y R. Jain. Mpi-video infrastructure for dynamic environments. En *IEEE Conference on Multimedia Systems 98*, Austin, TX, Enero 1998.
- [25] San Miguel, J. C., Bescós, J., Martínez, J. M. y García, A., DiVA: a Distributed Video Analysis framework applied to video-surveillance systems. En *Ninth International Workshop on Image Analysis for Multimedia Interactive Services*.
- [26] Jenkins, W., Top Benefits of video surveillance, Articlesbase, Free Online Articles Directory, revisado 19 Noviembre 2010, <<http://www.articlesbase.com/customer-service-articles/top-benefits-of-video-surveillance-1233905.html>>.
- [27] M. Sayles, X. Wu, y J. E. Boyd. Caml: Camera markup language for network interaction. En *SPIE Internet Imaging IV*, volume 5018, pp. 248–256, Santa Clara, CA, Enero 2003.
- [28] B. D. Lucas y T. Kanade. An iterative image registration technique with an application to stereo vision. En *International Joint Conference on Artificial Intelligence*, pp. 674–679, 1981.
- [29] Intel Corporation. *Intel Integrated Performance Primitives for Intel Architecture: Reference Manual*, 2003.
- [30] Boyd, J. E., Sayles, M., Olsen, L. y Tarjan, P., Content Description Servers for Networked Video Surveillance. En *Proceedings of the International Conference on Information Technology: Coding and Computing*, 2004.
- [31] T. Ebrahimi, Y. Abdeljaoued, R. Figueras, I. Ventura y O. Divorra Escoda. Mpeg-7 camera. En *ICIP01, pages Multimedia Indexing, Browsing, y Retrieval*, 2001.
- [32] Lien, C., Chen, P., Ying-Wen, B. y Ming-Bo y L., Monitoring System with Moving Object Detection Based on MSN Messenger. En *IEEE International Instrumentation and Measurement Technology Conference*, Victoria, Vancouver Island, Canada, 2008.
- [33] ZhengbingZhang, ZhongnianLi, ZhenhuaXia, GuangxiZhu, "IP-based video lab-monitor system", *Proceedings of SPIE - The International Society for Optical Engineering*, Electronic Imaging and Multimedia Technology IV, pp. 372-376, Junio 2005.
- [34] Jingwei Yang, *The Design and Implementation of Embedded Wireless IP Camera*, pp. 5-6
- [35] Chen, C., Yi, Y., Drira, A., Koschan, A. y Abidi, M., Cooperative Mapping of Multiple PTZ Cameras in Automated Surveillance Systems. ISBN 978-1-4244-3991-1/09, 2009, IEEE.
- [36] Varcheie, P.D.Z.; Bilodeau, G.-A. "Active people tracking by a PTZ camera in IP surveillance system," *Robotic and Sensors Environments*, 2009. *ROSE 2009. IEEE International Workshop on*, vol., no., pp.98-103, 6-7 Nov. 2009.



- [37] Chu-Sing Yang; Ren-Hao Chen; Chao-Yang Lee; Shou-Jen Lin; , "PTZ camera based position tracking in IP-surveillance system," *Sensing Technology*, 2008. ICST 2008. 3rd International Conference on , vol., no., pp.142-146, Nov. 30 2008-Dec. 3 2008
- [38] Hsien-Chou Liao; Wei-Yi Chen. "A dual-PTZ-camera system for visual tracking of a moving target in an open area", *Advanced Communication Technology*, 2009. ICACT 2009. 11th International Conference on, vol.01, no., pp.440-443, 15-18 Feb. 2009.
- [39] Liu, Y. y Xiao, M., A Motion Object Detecting Algorithm Applied for Constant Speed Dome Camera in Surveillance Camera System. En *2010 International Conference on Computer Application and System Modeling (ICCAASM 2010)*, pp. V6-361-V6-365.
- [40] Hintermaier, W. y Steinbach, E., A System Architecture for IP-camera based Driver Assistance Applications. En *2010 IEEE Intelligent Vehicles Symposium*, San Diego, California, Estados Unidos de América, 2010.
- [41] Liu, Z., Dichao, P. y Zheng, Y., Communication Protection in IP-based Video Surveillance Systems. En *Proceedings of the Seventh IEEE International Symposium on Multimedia (ISM'05)*, ISBN 0-7695-2489-3/05, 2005.
- [42] Yang, M., Tham, Y., Wu, D. y Huang Goh, K., Cost Effective IP Camera for Video Surveillance. En *ICIEA 2009*, 2009.
- [43] Wang R., Huang T., "Fast camera motion analysis in MPEG domain", *International Conference on Image Processing 1999*, pp. 691 -694 vol.3
- [44] Jong-Il Park, Inoue S., Iwadata Y., Estimating Camera Parameters From Motion Vectors of Digital Video", *IEEE Workshop Multimedia Signal Processing 1998*pp.105-110
- [45] A Probabilistic Model for Camera Zoom Detection Rong Jin, Yanjun Qi, Alexander Hauptmann *Proceedings of the 16th International Conference on Pattern Recognition (ICPR'02)* 1051-4651/02
- [46] Solari, S., Architecture for a digital video recorder. En *ITUAM 3.4*, 2000
- [47] Ebrahim, M., Canagarajh, C. N. y Bull, D. R. Video Coding for Mobile Communications Efficiency, Complexity, and Resilience. Editorial Academic Press, San Diego, California, Estados Unidos, ISBN-0-12-053079-1, 2002.
- [48] Anatassios, G. y Bradley, T., Security for Real-Time MPEG Compressed Video in Distributed Multimedia Applications. En *IEEE*, 1996.
- [49] Srinivasan, K., Porkumaran, K. y Sainarayanan, G., Improved Background Subtraction Techniques for Security in Video Applications.
- [50] Novak, T., Treytl, A. y Palensky, P., "Common approach to functional safety and system security in building automation and control systems", *Emerging Technologies and Factory Automation, 2007. ETFA. IEEE Conference on*, vol., no., pp.1141-1148, 25-28 Sept. 2007.
- [51] "IEC 15408 – Information technology – Security technique – Evaluation criteria for IT security", 1999.
- [52] Manual de usuario del Sistema de Administración Remota (RAS Plus).
- [53] Bitter, R., Mohiuddin, T. y Nawrocki, M., Lab VIEW Advanced Programming Techniques, Editorial CRC-Press, Estados Unidos de América, 2007. ISBN 0-8493-3325-3.
- [54] Xunwen, S., Shaoping, W., Dongmei Z. y Zhu, Q., "RS-485 serial port pseudo-full-duplex communication research and application," *Prognostics and Health Management Conference*, 2010. PHM '10. , vol., no., pp.1-5, 12-14 Jan. 2010.
- [55] National Instruments Corporation, NI Vision Concepts Help, National Instruments documentation. August, 2011.

- [56] Rodríguez, R. y Sossa, J. *Procesamiento y Análisis Digital de Imágenes*, Editorial Ra-Ma, España, 2011. ISBN 978-84-9964-077-8.
- [57] Escalante, B., *Apuntes de curso de Procesamiento Digital de Imágenes*, 2006. <http://verona.fi-p.unam.mx/boris/teachingnotes/Capitulo4.pdf>. Fecha de recuperación 30 de agosto de 2012.
- [58] Relf, C., *Image Acquisition and Processing with Lab VIEW*, Editorial CRC Press, Estados Unidos de América, 2004. ISBN 0-8493-1480-1.
- [59] Thomas, K., *Image processing with Lab VIEW and IMAQ visión*, Editorial Prentice Hall Professional, Estados Unidos de América, 2003. ISBN 0-13-047415-0.
- [60] NI Developer Zone, NI Community, “Lab VIEW-save a document in PDF file using Freeware Bull Zip PDF Write”, 2012. <https://decibel.ni.com/content/docs/DOC-22555>. Fecha de recuperación 4 de septiembre de 2012.
- [61] NI Developer Zone, “GMAIL.vi”, 2009. <http://zone.ni.com/devzone/cda/epd/p/id/5992>. Fecha de recuperación 4 de septiembre de 2012.

ANEXO A. GLOSARIO DE FUNCIONES DE LABVIEW

ICONO	NOMBRE	DESCRIPCION
VISION		
	IMAQ Create	Crea una locación de memoria temporal de una imagen
	IMAQ Cast Image	Convierte el tipo actual a otro tipo de imagen, especificando el tipo de imagen.
	IMAQ Dispose	Destruye la imagen y libera el espacio que ocupaba en la memoria. Este VI se requiere para cada imagen creada en una aplicación.
	IMAQ Optical Flow (HS)	Calcula el flujo óptico (velocidad de flujo) de información en dos imágenes sucesivas utilizando el algoritmo de Horn y Schunck.
	IMAQ Overlay Motion Vectors	Sobreescribe información de velocidad en la imagen.
	IMAQ Get Kernel	Lee un kernel predeterminado. Este código se compone de tres unidades separadas: la familia del kernel, tamaño de kernel y el número del kernel.
	IMAQ Histogram	Calcula el histograma de una imagen.
	IMAQ Absolute Difference	Resta una imagen de otra o una constante de una imagen y devuelve el valor absoluto de la diferencia.
	IMAQ Equalize	Produce una ecualización del histograma de la imagen. Este VI redistribuye los valores de píxel de una imagen para linealizar el histograma acumulado.
	IMAQ Convolute	Filtra una imagen usando un algoritmo lineal. Los cálculos se realizan con enteros o de punto flotante, dependiendo del tipo de imagen y el contenido del kernel.
	IMAQ AutoBThreshold 2	Calcula el valor umbral óptimo para una imagen o una región de interés y se aplica el umbral calculado.
	IMAQ Morphology	Realiza operaciones de morfología matemática en las imágenes.

ICONO	NOMBRE	DESCRIPCION
	IMAQ FillHole	Rellena los huecos presentes en una partícula. Los agujeros que se encuentran en contacto con el borde de la imagen nunca se llenan, porque es imposible determinar si estos agujeros son parte de una partícula.
	IMAQ RemoveParticle VI	Elimina o mantiene las partículas resistentes a una ventana de 3 x 3 con erosión. Las partículas que se mantienen son exactamente la misma forma que los que se encuentran en la imagen de la fuente original.
	IMAQ Particle Filter	Filtra (mantiene o elimina) cada partícula en una imagen de acuerdo a sus medidas.
	IMAQ Particle Analysis	Devuelve el número de partículas detectadas en una imagen binaria y una matriz 2D de las mediciones solicitadas acerca de las partículas.
	IMAQ Extract	Extrae (reduce) una imagen o parte de una imagen con el ajuste de la resolución horizontal y vertical.
	IMAQ Overlay Rectangle	Sobreescribe un rectángulo en la imagen.
	IMAQ Overlay Points	Sobreescribe un punto o arreglo de puntos en una imagen.
	IMAQ Copy	Copias de las especificaciones y los píxeles de una imagen en otra imagen del mismo tipo. Si la imagen de origen contiene información adicional, ésta también se copia en la imagen de destino
	IMAQ Get Image Size	Proporciona información sobre el tamaño (resolución) de la imagen.
	IMAQ Convert Rectangle to ROI	Convierte un rectángulo en un descriptor de región de interés (ROI).
	IMAQ Overlay ROI	Superpone una región de interés (ROI) en una imagen.
	IMAQ Simmetry	Transforma una imagen a través de su simetría.

ICONO	NOMBRE	DESCRIPCION
	IMAQ ArrayToColorImage	Crea una imagen en color a partir de una matriz 2D.
	IMAQ Light Meter (Rectangle) VI	Mide la intensidad de iluminación de cada pixel dentro de un rectángulo.
	IMAQ BCGLookup VI	Se aplica una corrección de brillo, contraste y gamma a una imagen.
DIALOGO		
	General Error Handler	Indica si hubo un error, si es así, este VI devuelve su descripción y, opcionalmente, muestra un cuadro de diálogo.
CADENAS		
	Search/Split String	Divide una cadena simple en dos subcadenas de determinada longitud.
	Search and Replace Pattern	Busca las cadenas de subcadenas que coincidan con la expresión regular y las reemplaza. Una expresión regular requiere una combinación específica de caracteres.
	Decimal String to Number	Convierte los caracteres numéricos de la cadena en un número.
	Decimate 1D Array	Divide los elementos del arreglo sucesivamente en los arreglos de salida.
COMUNICACIÓN TCP		
	TCP Open Connection	Abre una conexión de red TCP con la dirección y el puerto remoto o nombre de servicio.
	TCP Write	Escribe los datos en una conexión de red TCP.
	TCP Read	Lee un número de bytes de una conexión de red TCP, devolviendo los resultados de los datos a cabo.
	TCP Close Connection	Cierra una conexión de red TCP con la dirección y el puerto remoto o nombre de servicio.

ICONO	NOMBRE	DESCRIPCION
COMUNICACIÓN SERIAL		
	VISA Configure Serial Port VI	Inicializa el puerto serie especificado por el nombre del recurso VISA con la configuración
	VISA Close Function	Cierra una sesión de dispositivo u objeto evento especificado por VISA
REPORTES MICROSOFT OFFICE		
	New Report	Crea un nuevo reporte
	Word Edit Cell	Edita la celda especificada por fila y columna
	Save Report to File	Guarda un informe HTML en el archivo especificado en la ruta del archivo del informe
	Dispose Report	Cierra el informe y libera la memoria.



ANEXO B. PROTOCOLO PELCO-D

Fuente: <http://cvs.ru/files/pelco-d.pdf>

Este protocolo es usado entre los sistemas de conmutación de matriz y los receptores/controladores. Es popular en el control PTZ de cámaras de vigilancia.

B.1 Protocolo "D"

El formato del mensaje o trama es:

Byte 1	Byte 2	Byte 3	Byte 4	Byte 5	Byte 6	Byte 7
Sincronización	Dirección	Comando 1	Comando 2	Datos 1	Datos 2	Check Sum

Figura B.1 Trama de protocolo Pelco-D

Donde:

- Todos los valores están mostrados en hexadecimal.
- El byte 1 de sincronización es siempre \$FF.
- La dirección es la dirección lógica del receptor/controlador al que se manipula.

B.2 Formato del conjunto de Comandos

Los comandos 1 y 2, vistos en los bytes 3 y 4 respectivamente, tienen el siguiente formato:

Tabla B.1 Comandos Pelco-D

	Bit 7	Bit 6	Bit 5	Bit 4	Bit 3	Bit 2	Bit 1	Bit 0
Comando 1	Sensitividad	Reservado en 0	Reservado en 0	Escaneo Auto/Man	Cámara enc/apag	Cerrar diafragma	Abrir diafragma	Enfoque cerca
Comando 2	Enfoque lejos	Zoom ancho	Zoom alto	Arriba	Abajo	Izquierda	Derecha	Siempre 0

Donde, en el comando 1, el bit de sensibilidad, indica la aceptación de los bit 3 y 4, bajo la siguiente lógica:

Tabla B.2 Comando 1 Pelco-D

Sensibilidad	Bit 3		Bit 4		El comando...
	1	0	1	0	
1	✓		✓		Permitirá el auto escaneo y encender la cámara
		✓		✓	No existe acción
0	✓		✓		Permitirá el escaneo manual y apagar la cámara
		✓		✓	No existe acción

B.3 Formato del conjunto de Datos

Los datos 1 y 2, para los byte 5 y 6 respectivamente, tienen la siguiente funcionalidad:

Tabla B.3 Datos Pelco-D

	Función	Detenido	Alta	Turbo
Datos 1	Velocidad horizontal (<i>pan</i>)	\$00	\$3F	\$FF
Datos 2	Velocidad vertical (<i>tilt</i>)	\$00	\$3F	-

B.4 Check Sum

Es un medio de evaluación que emplea la suma de los comandos en hexadecimal para verificar que el mensaje haya sido bien enviado y bien recibido por los dispositivos. El byte 7 de la trama contiene la suma de byte 6 al byte 1, que se calcula de la siguiente forma: se supone que el mensaje enviado es dirigido al receptor 10, con cámara encendida, enfoque lejos activado, movimiento a la izquierda con turbo velocidad horizontal, es decir:

FF	0A	88	90	00	40	62
Sincronización	Dirección	Comando 1	Comando 2	Datos 1	Datos 2	Check sum

Figura B.2 Ejemplo de trama para cálculo de *check sum*

El procedimiento comienza transformando los bytes de hexadecimal a binario para entonces poder realizar la suma normalmente; después se vuelve a transformar a hexadecimal.

0A		00001010		
88		10001000		
	Subtotal	10010010	92	
90		10010000		
	Subtotal	00100010	22	Módulo 256 permite que el bit alto se ponga a cero
00		00000000		
	Subtotal	00100010	22	
40		01000000		
		01100010	62	Valor final de <i>check sum</i>

B.5 Rutinas preconfiguradas

Además de los comandos PTZ, hay comandos de control que permiten el acceso a las funciones más avanzadas de algunos equipos. Esto se logra modificando los Comandos 1 y 2, así como los Datos 1 y 2 de la trama principal. Algunos ejemplos se muestran en la siguiente tabla:

Tabla B.4 Rutinas preconfiguradas Pelco-D

Comando	Byte 3	Byte 4	Byte 5	Byte 6
Ir al <i>preset</i>	00	07	00	01 a 20
Girar 180°	00	07	00	21
Ir al origen	00	07	00	22
Correr patrón	00	23	00	00
Escanear zona	00	1B	00	00
No escanear zona	00	1D	00	00
Ajustar velocidad de zoom	00	25	00	00 al 03



ANEXO C. SDK IDIS WATCH CLIENT

Fuente: <http://www.avlogicsfileserver.com>.

Fecha de recuperación: 30 de septiembre de 2011

Un kit de desarrollo de software o SDK es un conjunto de herramientas de desarrollo de software que permite crear aplicaciones para un sistema en concreto.

Este SDK es desarrollado por la empresa coreana dedicada a la vigilancia de edificaciones, IDIS Co., Ltd., como una interfaz de programación de aplicaciones o API (del inglés *application programming interface*) creada para permitir el uso del conjunto de funciones que hacen posible la comunicación en tiempo real con algunas marcas y modelos de DVR. A continuación una visión general de este SDK.

C.1 Generalidades

El IDIS Watch SDK está desarrollado en el lenguaje de programación C++. Está compuesto de archivos *.dll (dynamic linking library file), *.lib (static binding library file) y archivos de cabecera (*.h) de C++, que son:

- idiswatch.dll
- idiswatch.lib
- IDIS_WatchSDK.h
- idisdecoder.dll
- idisdecoder.lib
- IDIS_DecoderSDK.h
- IDIS_Define.h

Los prototipos del IDIS WatchClient API definidos en el archivo de cabecera IDIS_WatchSDK.h son el conjunto de funciones proporcionados por el Cliente RASplus WatchClient (*software* propio del DVR empleado en el presente trabajo). Los datos regresados por la función IDIS WatchClient son especificados en el archivo cabecera IDIS_Define.h. La información adicional acerca de la descompresión de video puede ser adquirida por medio de la función Callback FRAMELOADED con algunos parámetros del archivo IDIS_Define.h. Para descomprimir los marcos del video, es necesario emplear la sección de IDIS_DecoderSDK.h Desplegar las imágenes de acuerdo con la cámara seleccionada y mostrar la información del sistema, se realizará sólo por el programador de interfaz de usuario que use el IDIS Watch SDK, ya que este SDK no incluye ningún código para la aplicación visual.

C.2 Modelos de DVR aplicables

Este *software* puede ser empleado para los siguientes modelos:

Tabla C.1 Modelos DVR aplicables

Modelos DVR			
IDR/PCR 6.06.042/6.06.070/6.07.000/6.07.001/6.08.000	HDR1600 1.1.7 / 2.0.0 / 2.0.1	SDR410 1.7.4/ 1.8.2	SDR402 2.1.0/3.0.1/3.1.0/3.2.0
ADRPRO 1.4.2	XDRDLX 2.0.0 / 2.1.0 / 2.3.1	XDR/XDRPRO 3.4.0/ 4.0.0	SDR421 1.0.0
MDR400 1.0.0	XDRPRO2 2.1.0 / 2.3.1	SDR800/1600 1.4.0 / 1.5.1 / 1.7.0	CDR420 2.1.0

C.3 Componentes empleadas del SDK

El SDK contiene numerosas funciones, sin embargo no todas son empleadas para este caso particular. A continuación se muestra en la Fig. D.1 las partes fundamentales de la aplicación para obtener el video en tiempo real a través del DVR.

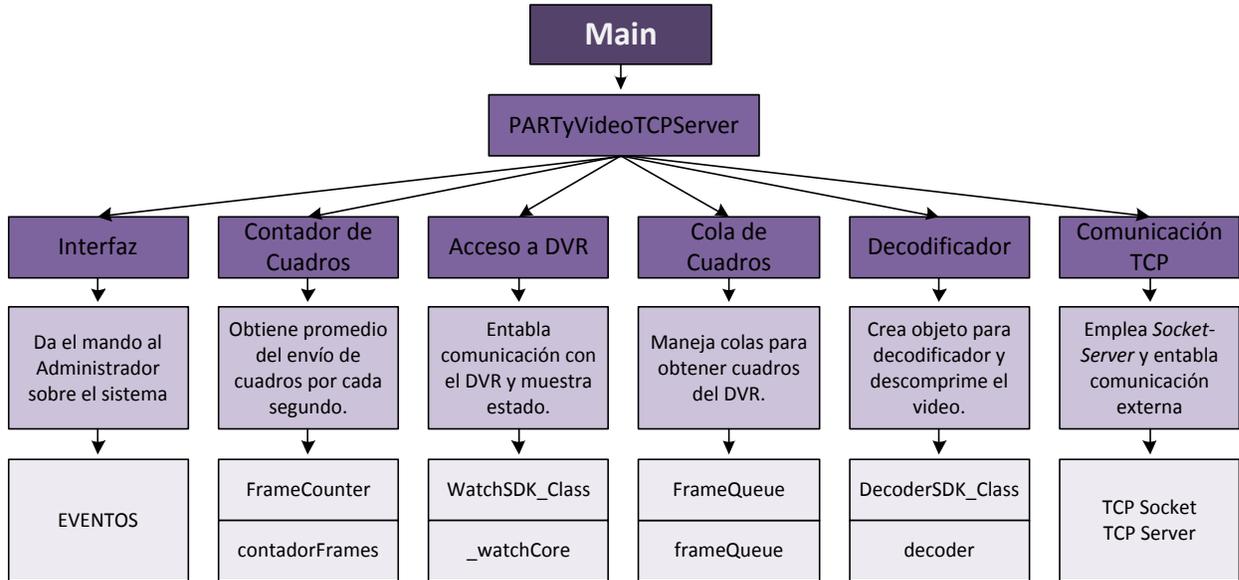


Figura C.1 Partes empleadas del SKD IDIS Watch

La raíz del diagrama representa el programa principal, que hace el llamado a la primera clase, PARTyVideoTCPServer, ésta contiene los constructores de los objetos a emplear. En el siguiente nivel del diagrama, se tienen los elementos que se emplearon para utilizar el SDK; para cada uno corresponde la descripción y los objetos y/o nombres de clases asociados a ellos.



ANEXO D. SDK VIVOTEK

Fuentes: <http://www.vivotek.com/support/sdk.php>,
<ftp://www.direktronik.se/Support/Kameror/Vivotek%20SDK/VitaminControl.pdf>

Fecha de recuperación: 14 de enero de 2012

El SDK de VIVOTEK proporciona herramientas, documentación y ejemplos de código que le permiten sacar el máximo provecho de los avances tecnológicos de esta línea de productos. VIVOTEK ofrece dos perfiles completos para ayudarle a desarrollar las aplicaciones, alto y principal.

Con el SDK de perfil principal (*main profile*), se tienen que desarrollar todas las funciones propias para conectar con el servidor de streaming (dentro de la cámara de red o el servidor de vídeo), manipular audio y video y solicitar los datos desde el servidor de medios de comunicación.

D.1 Alto perfil

Alto perfil (*high profile*), se compone de dos partes. Uno de ellos es el control Vitamin, y el otro es el MediaPlayer. Ambos son un tipo de control ActiveX.

El control Vitamin, desarrollado para conectar a las cámaras de red de VIVOTEK o servidores de video, puede mostrar el audio y video decodificado y/o registrar los datos. Es un componente integrado con funciones completas y contiene los aspectos más genéricos del SDK.

El MediaPlayer, por otra parte, se utiliza para la reproducción del video y de audio, grabados por la ST3402 o el control Vitamin, a partir de la base de datos.

Por lo tanto, cuando se requiere una rápida acumulación de programa para controlar y ver los productos de VIVOTEK, el control Vitamin es el método más fácil para hacerlo. Sin embargo, la flexibilidad y la eficiencia podrán ser relativamente menores.

Los controles ActiveX se utilizan para denotar los componentes reutilizables del software, se basa en el Modelo de Microsoft de componentes objetos (COM, por sus siglas en inglés). En términos generales, el programa sólo puede ser desarrollado con en el sistema operativo Microsoft Windows.

D.2 Control Vitamin

La función principal del control Vitamin es la de proporcionar un rápido desarrollo de la aplicación que se puede utilizar para monitorear, controlar, registrar y actualizar la configuración.

Para instalar el control Vitamin en la PC, se tienen que seguir estos pasos:

1. Descargar VitaminDecoder.zip para la instalación del archivo.
2. Ejecutar el VitaminDecoder.exe.
3. El directorio de instalación contiene dos subdirectorios.

D.3 Comandos empleados

A continuación se enumeran los comandos utilizados en el presente trabajo para obtener el video en tiempo real de la cámara tipo IP Vivotek.

- *Connect*

Conectar con el servidor visual para iniciar la descarga.

- Parámetros: el código de retorno de la función. 0 significa éxito, otros no significa.
- Nota: si la conexión ya está establecida, esta convocatoria no hace nada y devuelve 0.
- Requerimientos: los métodos *RemotelIPAddr*, *HttpPort*, *Url*, *ConnectionProtocol* deben estar listos antes de llamar a *Connect*.

- *CurrentControlCam*

Leer o establecer el índice de control de la cámara actual. Esta propiedad es significativa para el modelo de 4 canales.

- Tipo: long, el índice de la cámara.

- *RemotelIPAddr*

Obtiene o establece la dirección IP remota del servidor visual.

- Tipo: string, el formato debe ser IP con puntos, es decir: 'a.b.c.d'.

- *Password*

Obtener o establecer la contraseña utilizada para la autenticación de la página web.

- Tipo: string.

- *UserName*

Selecciona u obtiene el nombre de usuario para la autenticación web.

- Tipo: string.

- *AutoServerModelType*

Reconoce el tipo de modelo de servidor de forma automática. Con esta configuración, los usuarios ahora pueden especificar sólo el nombre de IP / usuario / contraseña para conectarse al servidor.

- Tipo: boolean, verdadero significa que el model fue encontrado de manera automática.
- Nota: para decidir el tipo de servidor, el control tiene que comunicarse con el éste antes de conectarse. Por lo tanto, se necesitará más tiempo para completar la operación.



- *NotifyVideoData*

Determina si el control debe enviar los datos de nuevo video decodificado por evento.

- Tipo: boolean, verdadero significa notificar a cada pieza.

- *AutoReconnect*

Decidir si el control debe intentar volver a conectarse al servidor si se encuentra la conexión rota.

- Tipo: boolean, verdadero significa que pudo reconectarse automáticamente.

- *ViewStream*

Selecciona u obtiene flujo de video/audio que el control se conectará.

- Tipo: EDualStreamOption.

- *ConnectionProtocol*

Establece el protocolo de conexión utilizado para conectar con el servidor.

- Tipo: EConnProtocol.

- *Display*

Selecciona u obtiene el interruptor para encender o apagar la pantalla de la gráfica en la pantalla.

- Tipo: boolean, verdadero significa que es posible mostrar el gráfico.
- Nota: para computadoras con tarjeta gráfica lenta, este interruptor es muy útil para evitar que el sistema se sature.

- *DecodeAV*

Obtener o establecer el indicador que determina si los datos de audio y vídeo deben ser decodificados después de recibido.

- Tipo: boolean, verdadero significa que el video y audio deben decodificarse.

ANEXO E. PRECONFIGURACIÓN DE POSICIONES EN EL DOMO

Fuente: <http://www.revoamerica.com/Portals/3/support-docs/REVDPTZ10-1.pdf>

Fecha de recuperación: 25 de enero de 2011

Esta sección se extrajo del manual de usuario del domo empleado según los datos de la Tabla 4.2 para establecer puntos que serán empleados en el preposicionamiento.

E.1 Configuración

En el menú principal se tienen varias opciones (Fig. E.1), de las cuales se elige *preset*.

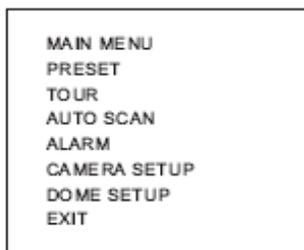


Figura E.1 Pantalla principal del domo

Puede guardar hasta 60 puntos preestablecidos. Cada punto preconfigurado tiene componentes de pan, tilt, zoom según se hayan definido como se ve en la Fig E.2.

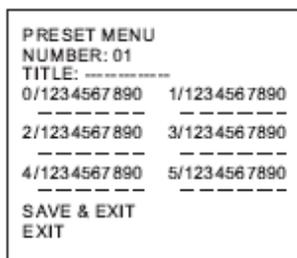


Figura E.2 Pantalla de configuración del domo

E.2 Operación

Para mandar llamar las rutinas será suficiente con cambiar el mensaje Pelco-D de acuerdo al Anexo B según la Tabla B.4 e indicar el número con el que fue guardada la preconfiguración.

ANEXO G. ESPECIFICACIÓN DE LOS CÁMARAS ANALÓGICAS

A continuación se enlistan las características más importantes de las cámara analógicas empleadas para las pruebas realizadas.

Tabla G.1 Cámaras analógicas

EQUIPO	DESCRIPCIÓN
<p>Minitrax Dome Camera, Syscom</p>	<ul style="list-style-type: none"> -Más Zoom 352X (22X óptico/16X digital). -Día/Noche Real (filtro ICR Incorporado). -Compatible con DVR's (Protocolos Fastrax, Pelco D/P). -Detección Automática de protocolos de datos RS485. -CCD 1/4" SONY Exview HAD. -580 líneas de resolución (Color). -0.5 Lux (Color), 0.05 Lux (B/N), 0.001 Lux (Low Shutter). -Enfoque automático. -Lente de 3.9 a 85.8 mm. -120 Presets programables.
<p>3DNR Super Night Vision Outdoor Camera, Epcom</p>	<ul style="list-style-type: none"> -Zoom 352X (22X óptico/16X digital). -Día/Noche Real (filtro ICR Incorporado). -Compatible con DVR's (Protocolos Fastrax, Pelco D/P). -CCD 1/4" SONY Exview HAD. -560 líneas de resolución (Color). -600 líneas de resolución (B/N). -0.2-0.00001 Lux



ANEXO H. PUBLICACIONES

- 12th. International Congress on Computer Science, CORE 2012. *Computational model for video surveillance, dual detection and processing of relevant events*. Presentado el día 28 de noviembre de 2012 en la Ciudad de México. (Autora principal)
- 16° Convención Científica de Ingeniería y Arquitectura, CUJAE. *Environment noise map estimation based on real-time measurements, a case study*. Presentado el día 26 de noviembre de 2012 en la Habana, Cuba. (Co-autora).

Computational model for video surveillance, dual detection and processing of relevant events

Pamela Araceli Rangel Tirado, Luis Pastor Sánchez Fernández, Oleksiy Pogrebnyak

Centro de Investigación en Computación, Instituto Politécnico Nacional,
Av. Juan de Dios Bátiz s/n esq. Miguel Othón de Mendizábal,
Col. Nueva Industrial Vallejo, CP 07738, Mexico City, Mexico
pamela.rangel.t@gmail.com, {lsanchez, olek}@cic.ipn.mx

Abstract. This paper proposes a computational model for automated obtaining images in greater detail, through the combination of multiple fixed cameras and a mobile camera, working a couple at a time. It detects an event of interest and its location to make object pan, tilt, zoom. To achieve this objective, computer vision techniques, image processing, pattern matching and calibration are used and inter-camera Communication Protocol Pelco-D is employed. The application was programmed in LabVIEW 2010. The tests are performed in a restricted parking, resulting in a car zooming in.

Keywords: Surveillance, dual system, LabVIEW, virtual instruments, fixed camera, PTZ, pattern matching, computer vision, image processing.

1 Introduction

Nowadays, the detection and recognition of relevant events is one of the most studied areas of research. Images have a vital importance in the analysis of crime, as evidence of abuse or simply to track events. One of the outstanding problems is the fact that fixed cameras have a wide field of view (FOV), but cannot be used to view details of an object, such as a car plate or a person's face. One possible solution is to leverage commercial video surveillance systems of low and medium performance with advanced algorithms and a combination of fixed and mobile cameras, to perform more complex activities to maximize their utility.

This paper proposes an automated system that employs X_i fixed cameras to detect the presence of an object in a large FOV, processing images to segment them and find valuable information for pattern recognition of the object location. This is done positioning the mobile camera in the center of FOV and making zoom for obtaining an image with more details. The methods can be the basis for achieving more complex activities, such as tracking an object in real time.

The paper is structured as follows. Section II provides an overview of the state of the art. Section III describes the system architecture. Section IV presents the method-

ology to develop this system, with all the algorithms for image processing and control of the dome. Section V provides some results in images. Section VI concludes the paper.

2 Related works

On the social side, the increase in crime is one of the main factors that have influenced the demand for remote video surveillance systems [1], as it seeks to stop, prevent and combat illegal. Technologies such as video surveillance can observe, respond, intervene and mitigate risks that may arise. [2] There are many works in this area of research. Boyd et al. [3] describe a video surveillance system that assimilates information from multiple cameras in a single stage model. In a real time application with separate processors to perform low level operations (motion segmentation and recognition of objects), and then shared through a local area network, information with other parts of the monitoring system.

In such systems, there are some that will track objects of interest, either with a single chamber or more advanced systems. Regarding the first case, and Bilodeau Varcheie [4] conducted a project that seeks to identify and recognize objects and detect and track them using a moving pan, tilt, zoom (PTZ) camera with communication via Internet. This method is based on the comparison of elliptical samples to the objective evaluating the similarity probability when estimating the location of the object, through a study of the diffuse areas movement without employing the optical flow algorithm Lucas-Kanade. Likewise prediction of the object position is made to help track. This system has the peculiarity that the sample must be taken manually, which makes it useful for identifying previously chosen persons.

Another innovative system is proposed by Chu-Sing et. al [5], which in addition to tracking algorithms employed internally also adds geographical location of people. Another system proposed by Hsien-Chou and Wei-Yi [6], which proposes a dual camera system with PTZ using motion sensors to resemble the identification and tracking of an object as eagles do, using the first camera see the wide FOV and the other for the detail in the same area.

3 System Architecture

This paper presents the development of virtual instruments for video surveillance and security system (Figure 1), with getting video through X_i CCTV cameras, which may or may not require conversion to digital video (depending on the type of sensor employed), where $i = 1, 2, \dots, n$ and n depends on the FOV of the dome camera. The obtained frames are used in the algorithms to change the position of the dome camera and get detailed pictures. Once the process is complete, the user reports are generated and sent to a client computer. The method developed for image analysis and PTZ movement will be explained in Section IV.

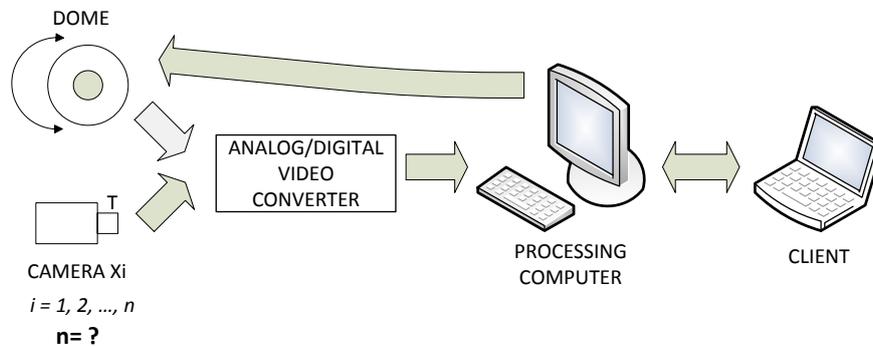


Fig. 1. Method of operation of the proposed system

4 Methodology

The application starts with the scanning of X_i fixed cameras to determine if an event happens in the regions of interest (ROI) that user previously determined, Fig. 2.

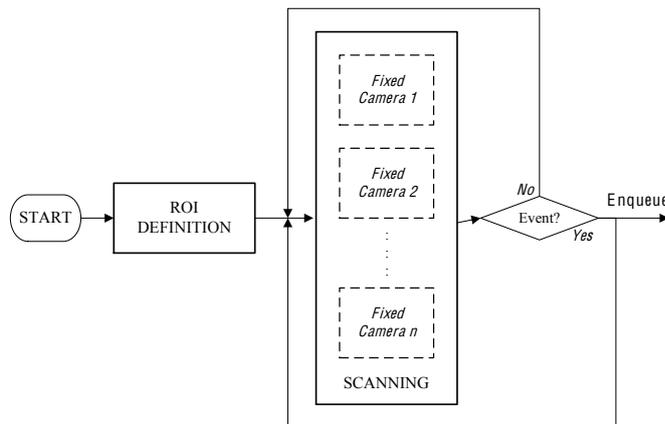


Fig. 2. General block diagram of the system, part 1

When the presence of an object has been detected, the identification number of the camera that generated the event and the images clean and with the object, are placed in a queue and sent to the next part of the application. The feedback line means that even when the data is sent, inspection continues to check for other event.

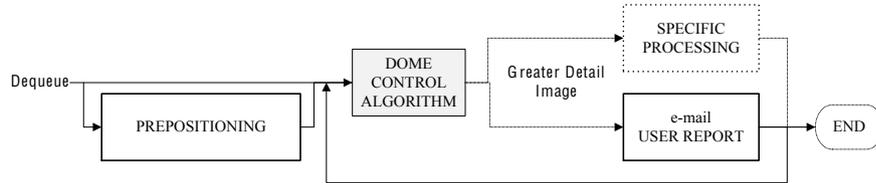


Fig. 3. General block diagram of the system, part 2

Fig. 3 shows that when the queue has elements, system starts running the prepositioning that uses preconfigured routines directly in the memory of the dome. This block is used to move the dome to a known area to resemble its FOV with the fixed camera X_i that generated the event, thus to find the object by pattern recognition.

Moreover, the dome control algorithm is executed as shown in Fig. 4; it will be explained in the following subsections.

Finally, when a focused image is obtained, the user report are generated and sent them to an email account. The specific processing block refers to the utility that can be the final image for other analyzes not attacked in this work, such as the determination of license plates or traits as textures, color and even model.

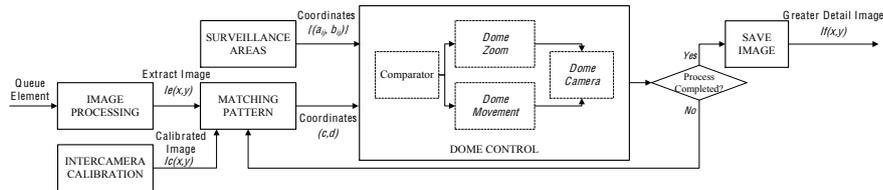


Fig. 4. Block diagram for dome control algorithm

4.1 Image Processing

The following explains the image processing components, which after the object discrimination, extract the object $I_{ext}(x, y)$ that triggered the event. The queue element consists of the identifier of the camera that generated the event, a clean image as a reference and the other with the object of interest.

Edge detection.

This algorithm is used to determine that an event has occurred, through the detection of changes in illumination along the ROI that the user determined before the inspection. Once done, a clean original image is saved; this will be the benchmark for image processing explained in the following subsection. When change lighting is detected, the coordinates of the points are stored in an array. This is very important for correct particle discrimination.

Image subtraction.

This is done taking two images: the reference image $I_A(x, y)$ after selecting the ROI, and the other image $I_B(x, y)$ containing the object, which has entered into the scene. At this step, the subtraction of images pixel by pixel, Eq. (1) is made [7]:

$$I_{Dest}(x, y) = I_A(x, y) - I_B(x, y). \quad (1)$$

This basic mathematical operation has the purpose of indicating in grayscale which pixels are not in the clean image, turning every other to zero (black), having an output image with high contrast. This is the reason that the image is binarized using only the average value of its histogram as the threshold value.

Particle analysis.

Firstly, a dilation Eq. (2) is calculated [7], with a structure element B of 3x3 size to fill internal voids in all non-zero elements:

$$X \oplus \tilde{B} = \{x \in Z^2, X \cap B_x \neq \emptyset\}. \quad (2)$$

Then, 5 erosions with the same structure element B of dilation according to Eq. (3) [7] are performed:

$$X \ominus \tilde{B} = \{x \in Z^2, B_x \in X\}. \quad (3)$$

This ensures that noise decreases its size. Then, particles are filtered per area. Finally, the coordinates of a circumscribed rectangle of their location are stored in an array.

Particle discrimination.

Ideally, only one object should be detected in the image, but in practice the system observes various objects. To prevent the system might be confused, particle are discriminated by their proximity to the coordinates where the object presence was detected in the ROI. Suppose two coordinates (X_1, Y_1) and (X_2, Y_2) correspond to particles obtained after analysis that eliminated spurious particles. Also, consider that the point where they shot the event in the ROI is (X_a, Y_a) , but it has a margin of 100 pixels to define if the particle corresponds to the object (see Fig. 5). Mathematical logic operation is as follows:

$$P_i = O_i \leftrightarrow [X_a - 100 \leq X_i < X_a + 100] \wedge [Y_a - 100 \leq Y_i < Y_a + 100]. \quad (4)$$

That is, the particle P_i (with coordinates are equal to its subscript) is the object of interest O_i , if and only if it is into the tolerance limits set by ± 100 pixels in both directions.

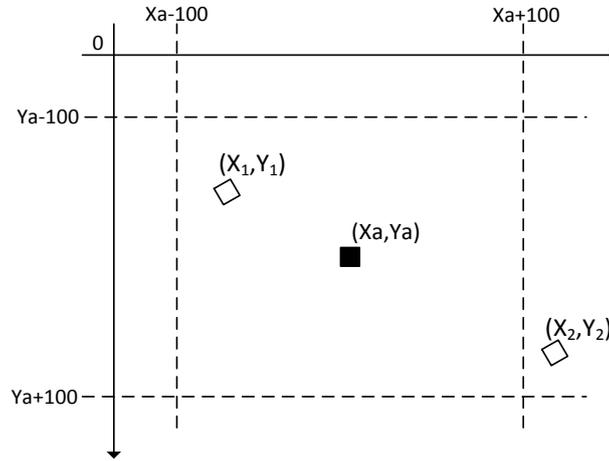


Fig. 5. Particle discrimination

4.2 Inter-camera calibration

The initial image after processing and extraction from a fixed camera is searched only once in the dome FOV, so it is necessary to compensate the difference in lighting levels before pattern matching. This is done by measuring the light intensity of the image of the dome to take some parameters and modified the extracted image histogram of the stationary camera. [8] That is, if the equalization expression, Eq. (5), is taken as a basis, the similarly with the other image is calculated according to Eq. (6):

$$s = \int_0^r P_r(w)dw = T(r) . \tag{5}$$

$$v = G(z) = \int_0^z P_z(t)dt . \tag{6}$$

To equalize the histograms, it is necessary to take the form $T(r) = G(z)$, and then z must satisfy the condition of Eq. (7):

$$z = G^{-1}(s) = G^{-1}[T(r)] . \tag{7}$$

4.3 Pattern matching

Pattern recognition has many uses. In this case it is used to find the object $H(x, y)$ that has entered the scene, giving its location and can thus continuously positioning in the dome FOV until it is centered. [7]

Learning.

At this stage, information is extracted from the objects of interest contained in the template through non-uniform sampling (see Fig. 6), where sections with the same intensities of pixel neighbors are taken to assure that the descriptive features have better quality and to reduce computational consumption. It is also important to define if the particles are rotated, scaled and/or moved.

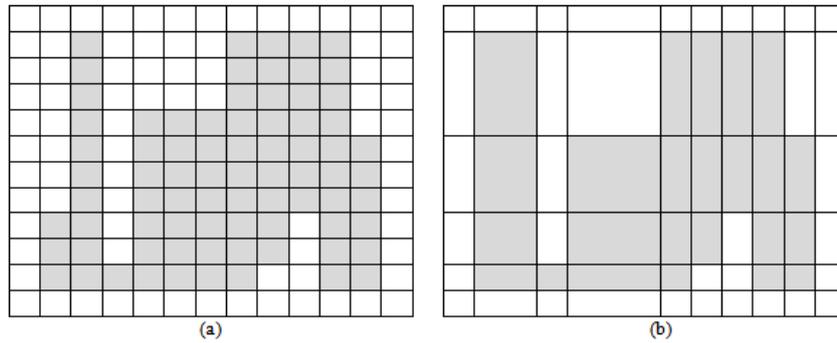


Fig. 6. Non-uniform sampling

Search.

During this phase, the descriptor is used to find the object in the inspected image. As it is shown in Fig. 7, the template is moved pixel by pixel to calculate a cross-correlation matrix according to Eq. (8) [9]:

$$C(i, j) = \sum_{x=0}^{L-1} \sum_{y=0}^{K-1} w(x, y) f(x + i, y + j), \quad (8)$$

where $f(x, y)$ is the image, $w(x, y)$ is the template, x is a horizontal offset, y is a vertical offset, K is a horizontal template size, L is a vertical template size; j is a horizontal coordinate $j = 0, 1, \dots, M - 1$, i is a vertical coordinate $i = 0, 1, \dots, N - 1$, and $M \times N$ is the image size.

Once the cross-correlation matrix C is calculated, its highest value indicates that with a greater probability in the neighborhood of the corresponding pixel is the desired object. Therefore, the output will have the coordinates (c, d) of the template center in the inspected image.

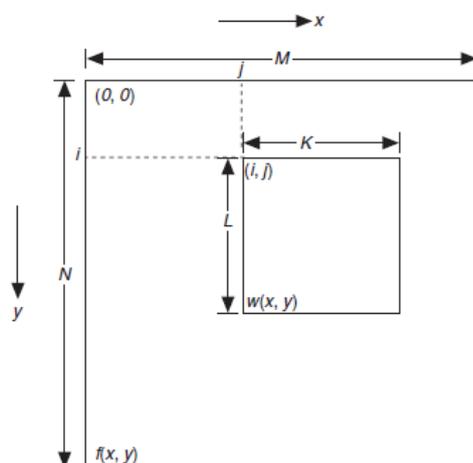


Fig. 7. Cross-correlation calculation

4.4 Surveillance areas

In order to correlate each of the data portions with control actions, the picture block of 2 dimensional (2D) video splits into 81 surveillance areas. Fig. 8 shows the FOV where in the central square there are areas of each color for a fine movement and a better focus on the object. Once located in the zone 40, the dome movement is stopped to zoom in the object.

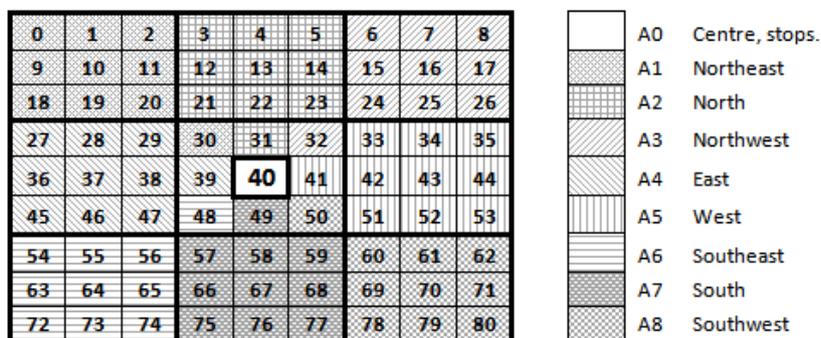


Fig. 8. Division FOV in surveillance areas

4.5 Dome Control

The movement of the dome is one of the most important parts of this application because thanks to this camera that moved to the object of interest location with a

zooming, a better quality image is provided. This is accomplished through communication protocol Pelco-D and RS-485, by sending of a command string in hexadecimal.

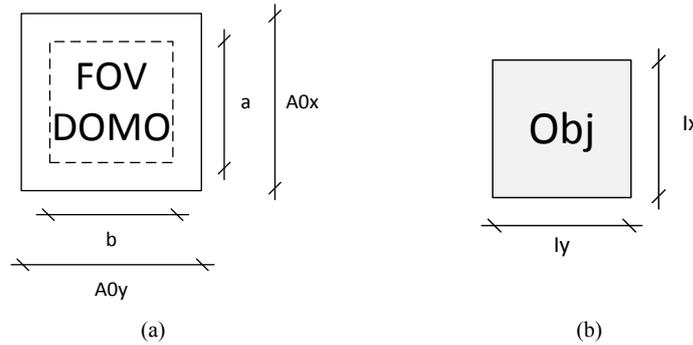


Fig. 9. Logic for zoom in

For zooming, when the surveillance area selector provides the identifier 40, it means that the image is ready to begin zooming. The dimensions of the FOV dome shown in Fig. 9 (a), and for the zoom in, the first step is to calculate the tolerances according to Eq. (9):

$$Tolerances \begin{cases} a = A0_x - 10 \\ b = A0_y - 10 \end{cases} \quad (9)$$

Next, the extracted image is increased by a factor 1.1 preserving the aspect ratio. This value, 1.1, was found experimentally is suitable for resizing after various adjustment tests. This parameter is used only to make increases for zoom in, but not affect completely the image.

The increasing is repeated until the object size is less than or equal to the resolution in either side of the dome FOV, as shown mathematically in Eq. (10):

$$(I_x \geq a) \vee (I_y \geq b), \quad (10)$$

where I_x, I_y are the object size, a, b are the object size in the scale of the dome FOV.

The corresponding pseudocode is shown below:

```
While ((Ix ≥ a) ∧ (Iy ≥ b))
{
    INx = (1.1) (Ix);
    INy = (1.1) (Iy);
    Ip (x, y) = resize (INx, INy);
    Ix = size.x (Ip (x, y));
    Iy = size.y (Ip (x, y));
}
```

5 Experiments and analysis

5.1 Results

The equipment employed for the tests is shown in Table 1. The event was presented for the IP camera.

Table 1. Equipment used

CAMERAS	Minitrax Dome Camera, Syscom	3DNR Super Night Vision Outdoor Camera, Epc-com	IP Camera Vivotek
CONVERTER	4-Channel Digital Video Recorder (DVR), Syscom		

Fig. 10 shows the detection axes of an automobile because of two of the four ROI lines. Red dots indicate where lighting change was detected.



Fig. 10. Edge Detection

When the event happen, the image processing (see Fig. 8) is responsible for searching for the global coordinates of the ROI, which indicates where the subtraction is performed. Fig. 11 illustrates that after thresholding, particle analysis and discrimination, the object of interest is delimited, showing its centroid.



Fig. 11. Presence of objects

Then, when the dome has a FOV similar to having the camera that detected the event (see Fig. 12), the extracted image is compared with the FOV to calibrate its light intensity to find the object.

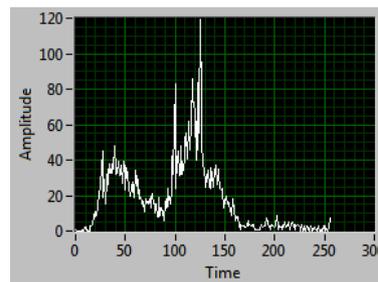


Fig. 12. Dome FOV

Fig. 13 (a) shows the extracted image and Fig. 13 (b) shows its histogram. Furthermore, after the measurement of illumination, the parameters brightness and gamma are modified, maintaining the contrast in the neutral value (45) (see Fig. 14 (a)).



(a)



(b)

Fig. 13. Image search

As shown in Fig. 14 (b), the image is very dark, because it did not change the contrast value. However, this is not a limitation, because the pattern matching algorithm succeeds in finding the object. After transformations, the histogram is modified (see Fig. 14 (c)).

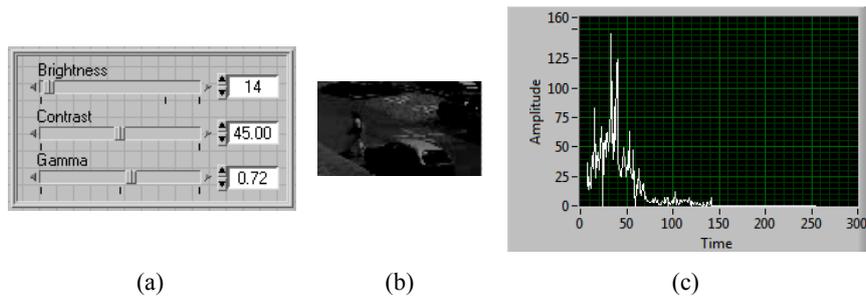


Fig. 14. Modified image search

Pattern matching continually monitors the target location and gives the image coordinates, which are compared by the dome control to locate the relevant surveillance area. Fig. 15 shows how pattern is found. This is repeated until the target is in the center of the FOV of the dome.



Fig. 15. Pattern Matching

After positioning in the center and then making zoom in as necessary, this steps will stop when some of the lengths of the pattern (vertical or horizontal) is greater than or equal to the stated tolerance.

Finally, the result is shown in Fig. 16. This image, $I_f(x, y)$, according to the block diagram of Fig. 1, can be used to perform other advanced processing or for display it in a report that is sent to the user email account.



Fig. 16. Result

5.2 Discussion

The design of the system corresponds to a dual system that employs a pair formed by a fixed camera and a mobile one that work together, as proposed in the paper of Hsien-Chou y Wei-Yi [6].

The test was conducted in a restricted parking that complicates the algorithm having uncontrolled lighting, unlike Chu-Sing et. al work [5] which is done outside.

One advantage of the proposed system over other developed by Varcheie y Bilodeau [4], is that the present work employs RGB color to grayscale transformations, which can be a benefit for the low-performance commercial systems that provide monochromatic images. Thus, matching pattern is not dependent on color, the object can be found through inter-camera calibration.

6 Conclusion

In this paper a computational model is proposed and applied for the security of buildings, in this case parkings in restricted areas. The presence of an object is detected by a fixed camera then the command is send to the dome camera to move it to that point and zoom in to get a picture with a greater detail. This is achieved using advanced algorithms of image processing, pattern matching and Pelco-D communication protocol. The application was programmed in LabVIEW 2010. The final image can be used for other purposes that computer vision allows, such as feature analysis and optical character recognition (OCR). In this paper, user reports are generated and sent via email. Future work will be related to sharing information in a local network, and integrating other image processing techniques with the purpose of object tracking in real time for a more robust and complete system.

7 References

1. Imai, Y.; Sugiue, Y.; Hori, Y.; Masuda, S.; , "Application of A Remote Surveillance System with Mobile Phone-Enhanced User Interface," Mobile Business, 2006. ICMB '06. International Conference on , vol., no., pp.27, 26-27 June 2006.
2. Jenkins, W., Top Benefits of video surveillance, Articlesbase, Free Online Articles Directory, <http://www.articlesbase.com/customer-service-articles/top-benefits-of-video-surveillance-1233905.html>
3. Boyd, J.E.; Hunter, E.; Kelly, P.H.; Li-Cheng Tai; Phillips, C.B.; Jain, R.C.; , "MPI-Video infrastructure for dynamic environments," Multimedia Computing and Systems, 1998. Proceedings. IEEE International Conference on , vol., no., pp.249-254, 28 Jun-1 Jul 1998.
4. Varcheie, P.D.Z.; Bilodeau, G.-A.; , "Active people tracking by a PTZ camera in IP surveillance system," Robotic and Sensors Environments, 2009. ROSE 2009. IEEE International Workshop on , vol., no., pp.98-103, 6-7 Nov. 2009.
5. Chu-Sing Y.; Ren-Hao C.; Chao-Yang L.; Shou-Jen L.; , "PTZ camera based position tracking in IP-surveillance system," Sensing Technology, 2008. ICST 2008. 3rd International Conference on , vol., no., pp.142-146, Nov. 30 2008-Dec. 3 2008.
6. Hsien-Chou L.; Wei-Yi C.; , "A dual-PTZ-camera system for visual tracking of a moving target in an open area", Advanced Communication Technology, 2009. ICACT 2009. 11th International Conference on , vol.01, no., pp.440-443, 15-18 Feb. 2009.
7. Relf, C., Image Acquisition and Processing with LabVIEW, Editorial CRC Press, Estados Unidos de América, 2004.
8. Escalante, B., Apuntes de curso de Procesamiento Digital de Imágenes, 2006. <http://verona.fi-p.unam.mx/boris/teachingnotes/Capitulo4.pdf>.
9. National Instruments Corporation, NI Vision Concepts Help, National Instruments documentation. August, 2011.