



**Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Ingeniería Mecánica Eléctrica**

**PROPUESTA DE UN SISTEMA DE VIGILANCIA UTILIZANDO
LA DETECCIÓN DE MOVIMIENTO EN UNA SECUENCIA DE
IMÁGENES CAPTADAS POR CÁMARA DE VIDEO DIGITAL.**

MANUEL ANTONIO ARREDONDO ESCOBAR
Asesorado por Ingeniera Ingrid Rodríguez de Loukota

Guatemala, octubre de 2005.

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**PROPUESTA DE UN SISTEMA DE VIGILANCIA UTILIZANDO
LA DETECCIÓN DE MOVIMIENTO EN UNA SECUENCIA DE
IMÁGENES CAPTADAS POR CÁMARA DE VIDEO DIGITAL.**

TRABAJO DE GRADUACIÓN
PRESENTADO A JUNTA DIRECTIVA DE LA
FACULTAD DE INGENIERÍA
POR

MANUEL ANTONIO ARREDONDO ESCOBAR

ASESORADO POR INGENIERA INGRID RODRÍGUEZ DE LOUKOTA
AL CONFERÍRSELE EL TÍTULO DE
INGENIERO ELECTRÓNICO

GUATEMALA, OCTUBRE DE 2005.

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA



NÓMINA DE JUNTA DIRECTIVA

DECANO	Ing. Murphy Olympo Paiz Recinos
VOCAL I	
VOCAL II	Lic. Amahán Sánchez Alvarez
VOCAL III	Ing. Julio David Galicia Celada
VOCAL IV	Br. Kenneth Issur Estrada Ruiz
VOCAL V	Elisa Yazminda Vides Leiva
SECRETARIA	Inga. Marcia Ivonne Véliz Vargas

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

DECANO	Ing. Sydney Alexander Samuels Milson
EXAMINADOR	Ing. Enrique Edmundo Ruiz Carballo
EXAMINADOR	Ing. Luis Eduardo Durán Córdoba
EXAMINADOR	Ing. Marvin Marino Hernández Fernández
SECRETARIO	Ing. Pedro Antonio Aguilar Polanco

HONORABLE TRIBUNAL EXAMINADOR

Cumpliendo con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

PROPUESTA DE UN SISTEMA DE VIGILANCIA UTILIZANDO LA DETECCIÓN DE MOVIMIENTO EN UNA SECUENCIA DE IMÁGENES CAPTADAS POR CÁMARA DE VIDEO DIGITAL,

tema que me fuera asignado por la Dirección de la Escuela de Ingeniería Mecánica Eléctrica con fecha 30 de mayo de 2005.

Manuel Antonio Arredondo Escobar

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES	V
GLOSARIO	IX
OBJETIVOS	XV
RESUMEN	XVII
INTRODUCCIÓN	XIX
1. SISTEMAS DE VIGILANCIA	1
1.1. Medios técnicos pasivos	2
1.1.1. Protección perimetral	3
1.1.2. Protección periférica	3
1.1.3. Protección del bien	4
1.2. Medios técnicos activos	5
1.2.1. Detectores	7
1.2.2. Señalizadores	12
1.3. Sistemas de vigilancia modernos	13
2. CARACTERÍSTICAS DE VIDEO DIGITAL	21
2.1. Video análogo	21
2.1.1. Proceso de exploración de la imagen	22
2.1.2. Frecuencias de exploración horizontal y vertical	24
2.1.3. Las señales de color	25
2.2. Video digital	29
2.2.1. Muestreo	29
2.2.2. Cuantización	33
2.2.3. Codificación	34
2.2.3.1. Codificación de las señales compuestas	35
2.2.3.2. Codificación en componentes	37

2.2.5. Ventajas del video digital	40
3. PROCESAMIENTO DE SECUENCIAS DE IMÁGENES	43
3.1. Fundamentos	43
3.1.1. Ruido en imágenes	46
3.1.2. Relaciones básicas entre píxeles	48
3.1.3. Pasos para el procesamiento de imágenes	50
3.2. Mejoramiento de imágenes	52
3.2.1. Mejora mediante procesamiento puntual	52
3.2.2. Filtrado espacial	55
3.2.3. Mejora en el dominio de las frecuencias	56
3.3. Segmentación de imágenes	58
3.3.1. Regiones y bordes	58
3.3.2. Segmentación basada en umbralización	60
3.3.3. Segmentación basada en bordes	61
3.4. Morfología matemática	68
3.4.1. Operaciones básicas sobre conjuntos	69
3.4.2. Operaciones morfológicas	69
4. DETERMINACIÓN DE MOVIMIENTO	73
4.1. Flujo óptico	76
4.2. Factores que influyen en la detección del movimiento	77
4.2.1. Factores fotométricos	78
4.2.2. Factores geométricos	79
4.3. Métodos de determinación del movimiento	80
4.3.1. Método diferencial	80
4.3.2. Método de correlación	84
4.3.3. Método de error	86
5. PROPUESTA DE DISEÑO	89
5.1. Elementos del sistema de vigilancia	89
5.1.1. Fuente de alimentación	90

5.1.2.	Equipo de seguridad	90
5.1.3.	Detectores	91
5.1.4.	Señalizadores	92
5.1.5.	Red	92
5.1.6.	Diagrama del sistema propuesto	93
5.2.	Descripción del diseño	94
5.2.1.	Descripción del programa	95
5.2.2.	Factores que causan distorsión en el sistema	98
5.3.	Costo del sistema de vigilancia propuesto	99
5.4.	Ventajas y desventajas del sistema de vigilancia propuesto	100
CONCLUSIONES		101
RECOMENDACIONES		103
BIBLIOGRAFÍA		105
APÉNDICE		109

ÍNDICE DE ILUSTRACIONES

FIGURAS

1.	Alarma de infrarrojos LAS3255	14
2.	Alarma de seguridad LAS5201	15
3.	Detector de movimiento LAS1618	16
4.	Asegurador de puertas LAS5200	16
5.	Sistema de niebla activa LAS3209	17
6.	Alarmas de tele-asistencia LAS1495	18
7.	Microcámaras	19
8.	Variación infinita de un parámetro continuo en función del tiempo	21
9.	Representación simplificada del barrido progresivo	23
10.	Barrido entrelazado 2:1. Cuadro completo	23
11.	Los campos de un entrelazado 2:1	24
12.	Obtención de la señal Y	26
13.	Valores de luminancia relativa	27
14.	Representación de los colores en NTSC	28
15.	Muestreo de una señal análoga $e(t)$ por un tren de impulsos $u(t)$	30
16.	Banda base de la señal de video	31
17.	Espectro de una señal de video muestreada a la frecuencia f_o	31
18.	Cuando la frecuencia de muestreo es $f_o > 2f_s$	32

19.	Cuando la frecuencia de corte del filtro PB es superior a $f_o - f_s$	32
20.	Error de cuantización	33
21.	Codificación de la señal compuesta	35
22.	Codificación de componentes	35
23.	Transición de análogo a digital de las señales compuestas	36
24.	Cuantización de la señal de luminancia	38
25.	Cuantización de la señal de crominancia	38
26.	Angulo de la irradiación y reflectancia de la imagen	44
27.	Diagrama principal de una cámara digital	45
28.	CCD	46
29.	Célula Peltier	47
30.	Pasos del procesamiento de imágenes	51
31.	Ejemplos de histograma	54
32.	Imagen con dos modas dominantes	60
33.	Seguimiento de un contorno	62
34.	Transformada de Hough. (a) Línea recta y (b) su transformada	63
35.	Perfiles de bordes	64
36.	Dilatación. (a) elemento estructural, (b) imagen, (c) resultado	70
37.	Erosión. (a) elemento estructural, (b) imagen, (c) resultado	71
38.	Mecanismos de medida del movimiento	75
39.	Cubo de convolución tridimensional	82
40.	Arquitectura para detección de movimiento por gradientes	83

41.	Arquitectura para detección de movimiento por gradientes	86
42.	Diagrama del sistema de vigilancia propuesto	93
43.	Pantalla de interacción del sistema de vigilancia	94

TABLAS

I.	Parámetros de la norma 4:2:2 CCIR 601	39
II.	Costo del sistema de vigilancia propuesto	99

GLOSARIO

<i>Blanking</i>	Espacios en blanco encontrados en la codificación que sirven para colocar alguna información adicional o por si existe sobremodulación.
<i>Buffer</i>	Circuito con funciones de amplificación y/o separación de una señal, tanto de naturaleza lógica como digital.
CCTV	Circuito Cerrado de Televisión.
Codificación	Proceso mediante el cual ciertos datos se traducen a un protocolo de comunicación con el objetivo de interpretar, electrónicamente, la información transmitida a través de un medio electrónico.
Contraste	Razón entre los valores máximo y mínimo de luminancia en un elemento de visualización.
Convolución	Herramienta matemática que simplifica el cálculo de la respuesta de un sistema o función a la entrada de una variable arbitraria.

Corrección Gamma	Ajuste realizado a la información de color de la imagen, derivado de la incapacidad del ojo humano para captar los colores en forma lineal sino más bien en diferente porcentaje cada uno.
Dominio de Fourier	Representación matemática de una señal de video en el dominio de las frecuencias.
DTMF	Circuito integrado utilizado para marcaje telefónico con Frecuencia Múltiple y Tono Dual.
Filtro paso bajo	Circuito electrónico que se utiliza para eliminar, de una señal, las frecuencias superiores a una frecuencia de corte determinada.
Filtro paso alto	Circuito electrónico que se utiliza para eliminar, de una señal, las frecuencias inferiores a una frecuencia de corte determinada.
<i>FireWire</i> (IEEE 1394)	Puerto de computadora que se utiliza para una comunicación más veloz con elementos periféricos que manejan una gran cantidad de datos complejos, cámaras de video, discos duros, etc. También, llamado DV, digital video.
<i>Flicker</i>	Sensación visual molesta para el ojo humano, que consiste en parpadeo rápido y visible de alguna fuente de luz.

Histograma	Representación gráfica de la escala de colores de una imagen, desde la parte oscura a la parte clara.
Humedad Relativa	Humedad que contiene una masa de aire, en relación con la máxima humedad absoluta que podría admitir.
Iluminancia	Medida de la iluminación que incide sobre cada metro cuadrado de una superficie S y sobre la cual se distribuye, uniformemente, se calcula por medio de $E = \frac{\Phi}{S}$ donde Φ es potencia emitida como luz .
Lenguaje	Conjunto de normas lingüísticas que permiten escribir un programa para que sea entendido por la computadora y, además, pueda ser trasladado a otras computadoras para su funcionamiento en otros sistemas.
Medio termoelectrico	Elemento que convierte el calor en electricidad.
Modulación	Proceso mediante el cual la información se inserta en un soporte de transmisión.
Monocromático	Sinónimo de blanco y negro que se utiliza para referirse a imágenes y fotografías.
Muestreo	Proceso de toma de medidas instantáneas de una señal análoga cambiante en el tiempo.

NTSC	Norma de grabación de video creado en EEUU por el Comité Nacional de Sistemas de Televisión con velocidad de muestreo de 30 cuadros por segundo y 525 líneas por cuadro.
PAL	Norma de grabación de video derivado de NTSC y mejorado por medio de la inversión de la fase de la señal de color para eliminar errores, utilizado en Europa, con velocidad de muestreo de 25 cuadros por segundo y 625 líneas por cuadro.
PAM	Tren de impulsos con una frecuencia determinada por el teorema de Nyquist que toma muestras de una señal análoga.
Programa	Conjunto de instrucciones destinadas a ser utilizadas, directa o indirectamente, en un sistema informático, con el objetivo de realizar una función o una tarea o para obtener un resultado determinado.
SECAM	Primer norma de grabación de video a color utilizado en Europa y creado en Francia. Se elabora agregando información de color al formato de blanco y negro.

Series de Taylor

Herramienta matemática utilizada para la representación de funciones que son infinitamente diferenciables (que tienen todos los órdenes), por medio de la función $\sum_{n=0}^{+\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n$.

Threshold

Técnica para reducir el valor de la relación señal a ruido cuando se llega al umbral del brillo de la imagen.

YCrCb

Formato de señal de video a color en el cual la crominancia está formada por la resta de la luminancia a las señales de rojo y azul corregidas por un factor llamado Gamma.

YIQ

Formato de señal de video a color en el cual la crominancia está formada por la resta de la luminancia a las señales de rojo y azul corregidas en una fase de 33°.

YUV

Formato de señal de video a color en el cual la crominancia está formada por la resta de la luminancia a las señales de rojo y azul.

OBJETIVOS

General

Proponer un sistema de vigilancia utilizando la detección de movimiento en una secuencia de imágenes captadas por cámara de video digital.

Específicos

1. Mostrar los sistemas de vigilancia más comunes usados en el ámbito industrial y doméstico.
2. Indicar las principales características de video digital para comprender la forma en que se manipulan las imágenes, digitalmente.
3. Proporcionar los criterios técnicos para el procesamiento de secuencias de imágenes almacenadas y captadas por una cámara de video digital.
4. Puntualizar tres métodos de determinación del movimiento en secuencias de imágenes captadas por una cámara de video digital.
5. Diseñar un sistema de vigilancia con detección de movimiento en imágenes captadas por cámara de video digital, basado en las ventajas y desventajas de los métodos proporcionados.

RESUMEN

En Guatemala, existe mucha inseguridad, la cual ha sido provocada por una cultura de violencia que no todos compartimos y que poco a poco toca la puerta de cada guatemalteco. Derivado de lo cual, surge la necesidad de investigar y diseñar un sistema que vigile las propiedades o inmuebles que necesiten de protección para asegurar el patrimonio de cada persona.

Un sistema de vigilancia está constituido por tres tipos de medios: humanos, técnicos y organizativos; en los últimos tiempos, los sistemas de seguridad se han creado utilizando medios técnicos debido a la ineficiencia de los otros. Estos a su vez se dividen en Pasivos que están diseñados para retardar al intruso por medio de paredes, rejas, cristales blindados, puertas, etc. y en Activos o a lo que, también, se llama seguridad electrónica que son diseñados para alertar de un intento de violación a las medidas de seguridad. Un sistema electrónico está formado por: fuente de alimentación, cerebro, detectores, señalizadores y red.

El sistema de vigilancia propuesto, se basa en medios técnicos activos, utilizando el video digital para vigilar y detectar el movimiento de posibles intrusos en un área predefinida. Es usado este tipo de video por su fácil manejo, ya que, toma muestras de las imágenes y las divide en partes muy pequeñas que son convertirlas en código binario para computadoras, para, luego, procesar dichas imágenes. Este último procedimiento ayuda a eliminar el ruido, mejorar las imágenes y extraer información de interés de ellas. Dicha extracción es el verdadero fin del uso del video digital, puesto que, para el diseño propuesto, se necesita aislar objetos en movimiento.

Para determinar el movimiento en una secuencia de imágenes existen varios métodos, pero el llamado diferencial ejecuta la detección de una manera más eficiente, buscando la dirección del máximo cambio del brillo, gradiente, en cada punto de la imagen digital. Este método no utiliza el cálculo completo del gradiente, por lo difícil que resulta implementarlo en un lenguaje de programación de computadoras, por lo que en su lugar se utiliza un operador de Sobel para aproximarlos.

El diseño propuesto es muy confiable e impide pérdidas materiales y humanas; además, evita trabajos indeseables y tediosos, pero lo que es mejor, evade la realización de una tarea muy peligrosa: proveer de seguridad ante ataques de criminales.

INTRODUCCIÓN

La inseguridad en Guatemala es, actualmente, un flagelo insoportable, día a día se busca una forma nueva de proteger los bienes materiales y el hogar, ante los ataques del crimen organizado, pero, sobre todo, se busca algo que no tiene precio, la seguridad de nuestras propias familias. Debido a que se vive en una cultura de miedo y en una sociedad basada en la violencia, cada día se vuelve más necesario buscar protección. Lamentablemente, las autoridades no se dan a vasto para atender a cada individuo que necesite de su ayuda, por lo que cada uno debe velar por el cuidado de su familia y de sus bienes.

Proveer seguridad y protección es una tarea difícil y peligrosa, por lo que no cualquier individuo puede suministrarla, por lo que es necesario recurrir a la tecnología para poder lograrlo. En la actualidad, por medio de la tecnología, se han logrado producir varios tipos de sistemas de vigilancia pero su efectividad está limitada por su amplia dependencia hacia un operador o vigilante que las maneje o las inspeccione.

Ante dichas circunstancias, cobra relevancia estudiar el diseño de un sistema automático que, al vigilar por medio de una cámara digital, detecte el movimiento de algún sujeto que quebrante la seguridad de un establecimiento previamente determinado, para que, posteriormente, de la señal de dicho descubrimiento y pueda actuar en contra de la infracción.

Este estudio describe inicialmente, las partes principales de un sistema de vigilancia y los sistemas vigentes a la fecha con sus especificaciones y limitaciones.

Posteriormente, se presentan las características del video digital, comparándolas con las de video análogo, para así comprender la facilidad con que se pueden manejar las imágenes digitales; además, se distinguen las distintas formas para procesar imágenes, ya sea mejorándolas, eliminando el ruido o aislando algún objeto de interés en ellas; asimismo se hace énfasis en los métodos más eficientes y versátiles; a continuación, se muestran los métodos más beneficiosos para determinar el movimiento en secuencias de imágenes, especificando sus ventajas y sus debilidades.

Se concluye con la descripción de la propuesta del sistema de vigilancia, apoyándose en la información previa para detallar cada una de sus partes, mostrar sus especificaciones técnicas, presentar su costo total, explicar cómo funciona y en que ámbitos y para indicar su ventajas y flaquezas.

1. SISTEMAS DE VIGILANCIA

Un sistema de vigilancia es un conjunto de elementos interrelacionados entre sí, cuya empresa es resguardar un sitio determinado y alertar sobre violaciones a la seguridad del mismo. Dicha alerta es ejecutada con el fin de interceptar o interrumpir las acciones de un posible intruso de modo que se conserve la integridad física de dicho sitio.

Un sistema de seguridad o vigilancia está constituido por tres medios que deben aunarse como partes integrantes de un todo:

- a) Medios Humanos: constituidos por el personal de seguridad, tanto Pública, Institucional o Privada.
- b) Medios Técnicos: pasivos o físicos y activos o electrónicos.
- c) Medios Organizativos: planes, normas y estrategias.

De los anteriores cabe destacar que, los medios humanos y los medios organizativos no han cumplido con lo necesario para asegurar la integridad física de un hogar, negocio, industria o la integridad física y moral de personas individuales en los últimos años. Al mismo tiempo, se han alcanzado importantes progresos en el desarrollo de sistemas de computación capaces de procesar grandes cantidades de datos, en el perfeccionamiento de nuevos sensores que capten con mayor exactitud sus móviles de accionamiento y en lenguajes de programación que sean capaces de manejar con mayor eficiencia los datos y así manipularlos adecuadamente para obtener mejores resultados.

Tales progresos, junto con la creciente violación de los derechos humanos han revolucionado los sistemas de vigilancia basados en los Medios Técnicos. Actualmente existe una gran diversidad de sistemas que ayudan a proteger a las personas, los negocios, las empresas e inclusive los propios hogares. Además de la diversidad ya existente, se siguen investigando métodos e innovando tecnologías para crear nuevos sistemas que ayuden a simplificar el trabajo de las autoridades o que lo realicen en su totalidad.

1.1. Medios técnicos pasivos

Los medios técnicos pasivos están enfocados a disuadir, detener o al menos, retardar la progresión de una amenaza. El incremento del tiempo que estos elementos imponen a la acción agresora para alcanzar su objetivo resulta, en la mayoría de las ocasiones, imprescindible para que se produzca en tiempo adecuado a la reacción. El conjunto de medios pasivos también denominado seguridad física, puede estar constituido por:

- a) Protección perimetral: son elementos de carácter estático y permanente que conforman el cerrado de la instalación a proteger y el primer obstáculo que se presenta para la penetración de intrusos, como por ejemplo vallas, cercados, paredes, etc.
- b) Protección periférica: son elementos también estáticos y que impiden el acceso al propio edificio principal o núcleo de seguridad, y entre ellos se puede citar puertas, rejas, cristales, etc.
- c) Protección del bien: lo constituyen recintos o habitáculos cerrados (cajas fuertes, cámaras acorazadas, etc.).

1.1.1. Protección perimetral

Los elementos principales que la conforman están constituidos de:

- a) Mampostería: cercados fabricados con materiales de construcción, como por ejemplo muros y paredes.
- b) Metal: perímetros realizados por medio de cercas metálicas, verjas, vallas, alambradas acodadas en la parte superior, concertinas de alambre dentado, etc.

Un perímetro o cercado debe tener una altura mínima aconsejable de 3 metros; en objetivos de un nivel de riesgo elevado, se puede colocar un doble vallado perimetral paralelo, con una distancia prudencial entre ambos y en el pasillo interior instalar algunos elementos electrónicos de sistemas de seguridad activa; utilizando accesos a través puertas motorizadas con apertura a distancia. En estos cercados perimetrales, ya sean tanto sencillos como dobles, deben existir puertas peatonales para evitar aperturas continuadas de las puertas motorizadas, en el supuesto de paso de personal.

1.1.2. Protección periférica

Los principales elementos que conforman la protección periférica de las aberturas normales de la periferia de un edificio se pueden señalar como:

- a) Puertas: instaladas en los puntos principales de acceso al edificio o establecimiento. Según la seguridad que proporcionen, pueden distinguirse como de seguridad, blindadas y acorazadas.

- b) Sistemas de esclusas: instalados en dichos puntos de acceso, de forma que no pueda accederse directamente al interior. Esclusa es el conjunto de elementos fijos y móviles que forman un sistema de control de accesos para personas, vehículos u objetos bajo condiciones específicas de seguridad.
- c) Cristales blindados: colocados en ventanas de lugares sujetos a un riesgo especial, y del nivel que se considere conveniente.
- d) Rejas y contraventanas: instaladas en las ventanas, especialmente en aquellas de mayor accesibilidad, y en tragaluz.
- e) Rejillas y emparrillados: protectores de vacíos necesarios de ventilación.

Este conjunto de elementos no son los únicos posibles, ni necesariamente deben instalarse todos en todos los objetivos.

1.1.3. Protección del bien

En este apartado se deben incluir:

- a) Cajas fuertes: de la cuales existe una gran variedad en el mercado en cuanto a tamaños y sistemas de apertura.
- b) Cámaras acorazadas: construidas conforme a especificaciones reguladas reglamentariamente. Disponen de un acceso que puede tener dispositivo de bloqueo y estar temporizado.

1.2. Medios técnicos activos

La función de los medios técnicos activos es la de alertar local o remotamente de un intento de violación o sabotaje de las medidas de seguridad física establecidas. El conjunto de medios técnicos activos constituye lo que se denomina seguridad electrónica. Pueden utilizarse de forma oculta o pueden ser visibles. Sus funciones principales son:

- a) Detección de intrusos en el interior y en el exterior.
- b) Control de accesos y tráfico de personas, paquetes, correspondencia y vehículos.
- c) Vigilancia óptica por fotografía o circuito cerrado de televisión.
- d) Intercomunicación por megafonía.
- e) Protección de las comunicaciones.

Un sistema electrónico de seguridad está formado por un conjunto de elementos electromecánicos relacionados entre sí, que, a través de la información que nos proporcionan, contribuyen a la seguridad de un determinado entorno. Un sistema electrónico de seguridad consta de:

- a) Red.
- b) Fuente de alimentación: es el elemento de activación del sistema; además se debe disponer de una fuente que automatice el sistema ante posibles faltas de suministro casuales o intencionadas.

- c) Equipo de seguridad: es el cerebro de todo el sistema. Recibe los impulsos de los detectores y tras analizarlos, los transforma oportunamente en señales que envía a los señalizadores.
- d) Detectores: son dispositivos colocados tanto en el exterior como en el interior de objetivos con riesgo de intrusión, con la misión de informar a la central de las variaciones del estado ambiental de la zona que están protegiendo, indicando la intrusión en dichos objetivos.
- e) Señalizadores: representan una parte de vital importancia del sistema, puesto que si se consuma un intento de intrusión, se deberá conocer con precisión lo que está sucediendo y la ubicación de donde está sucediendo éste, para poder reaccionar con eficacia.

La fiabilidad de un sistema de protección es el grado de confianza que otorga el mismo en el cumplimiento de la misión para la que se ha establecido y viene determinada por los siguientes parámetros:

- a) Seguridad de reacción: por sí solos, los elementos que componen el sistema deben proporcionar tiempo y espacio para la reacción, especialmente los que constituyen el cierre perimetral de la instalación.
- b) Seguridad de falsas alarmas: los sistemas destinados a la vigilancia deben estar configurados de tal manera que descarten una violación cuando esta no amerite una alarma-reacción.
- c) Vulnerabilidad al sabotaje: disminuirá en razón al complemento de elementos activos que se hayan situado, así como por los puestos de vigilancia establecidos.

1.2.1. Detectores

Son los componentes básicos del sistema electrónico de seguridad. Son los iniciadores de la alarma y su función es vigilar un área determinada, para transmitir una señal al equipo de seguridad, cuando se detecta una situación de alarma. Su elección dependerá del área a controlar y del previsible agente causante de la intrusión.

Las causas desencadenantes que pueden activar un detector varían según su entorno, se pueden citar como ejemplo: movimiento del intruso, desplazamiento del detector, presión sobre el detector, rotura del objeto protegido, vibración, etc., por lo que es mejor dividirlos según el móvil de su accionamiento o según su ubicación. Por su ubicación pueden ser Interiores o Exteriores cuya diferencia no sólo radica por las inclemencias del tiempo que pueden soportar, sino más bien por la capacidad de distinguir las variaciones ambientales. Por su móvil de accionamiento pueden ser:

- a) Detector por contactos magnéticos: están compuestos de dos piezas enfrentadas; la ampolla (dos láminas flexibles dentro de una burbuja de cristal al vacío, formando el contacto), a cuyos extremos están soldados los hilos que forman el bucle de detección y un imán permanente cuyo campo magnético ejerce una fuerza magnética sobre los citados contactos cuando ambas piezas están enfrentadas.

Si se modifica la situación relativa de las mismas el campo magnético dejará de ejercer su acción sobre los contactos cerrándose o abriéndose. Su principal ventaja es su simplicidad de instalación, su bajo costo y bajo nivel de falsas alarmas.

- b) Detector por contactos mecánicos: se definen como aquellos que se fundamentan como contactos eléctricos con reposición. El ejemplo más claro es el interruptor colocado entre el cerco y la hoja de la puerta de un armario. Normalmente se utilizan en aquellos casos donde no existe espacio disponible para la instalación de un imán de contacto magnético.

- c) Barrera de rayos infrarrojos: detector que consta de un transmisor y un receptor entre los que se establece un haz (no visible) de infrarrojos. La interrupción de este haz provoca la alarma. El haz de infrarrojos no es totalmente lineal, sino que tiene cierta dispersión. La distancia máxima alcanzada es de 300 metros (mts), se utiliza en aquellos casos cuyo espacio protegido es largo y estrecho; su principal ventaja es el bajo precio y un bajo índice de falsas alarmas.

- d) Detectores inerciales: su funcionamiento se basa en la detección de las vibraciones de la superficie, mediante un sensor que en su interior dispone de elementos móviles que al producirse la agresión abren y cierran contactos eléctricos. Los más comunes son los contactos de péndulo, que disponen de una masa metálica soportada por una guía que, a través de un tornillo logra unirla a otra y en el caso que vibre cierra abre y cierra el circuito.

Otro modelo es el denominado de mercurio, en el cual existe una ampolla de vidrio conteniendo mercurio en su interior y en la que están inmersas las terminales del circuito detector actuando como un contacto normalmente cerrado. Al producirse las vibraciones, el mercurio cambia de posición en la ampolla y deja abierto el circuito.

- e) Detectores microfónicos: también llamados sísmicos, estos transforman las vibraciones mecánicas en una señal eléctrica a través de una cápsula piezoeléctrica, similar a las utilizadas en los micrófonos; que después de la ampliación y filtrado producen la señal de alarma. La sensibilidad de los detectores es regulable y en todo caso los detectores se fabrican de tal forma que las vibraciones ambientales no les influya, con el objeto de evitar las falsas alarmas.

- f) Alfombras de presión: están construidas por láminas o placas metálicas que entran en contacto al ser presionadas por el peso de la persona cerrando el circuito que forma. En su ventaja está que son baratas, pero con el inconveniente de su escasa duración y posible vulnerabilidad si se conoce su existencia.

- g) Redes conductoras: es un dispositivo de protección basado en la aplicación de una cinta o red conductora (adherida o embebida) a cualquier tipo de superficie, de tal forma que no pueda producirse el paso de una persona sin provocar la señal de alarma. La cinta se conecta al bucle de alarma manteniendo una continuidad eléctrica que cuando se pierda, por rotura o por puente eléctrico, entre ambos lados del bucle, se produce una situación de alarma.

- h) Detector por microondas: detectan movimiento, también son llamados "radares"; estos emiten energía electromagnética a cierta frecuencia, que tras rebotar y reflejarse en las paredes del recinto protegido, alcanzan la etapa receptora. Si en el recinto no hay ningún movimiento, las frecuencias de las señales emitidas y recibidas son las mismas; sin embargo si es el caso contrario, parte de la señal que llega al receptor posee diferente frecuencia que la original.

- i) Detector por ultrasonidos: basan su funcionamiento en el efecto Doppler, mediante la emisión y recepción de ondas ultrasónicas (entre 22 Kilohertz y 45 Kilohertz). Básicamente están formados por un transmisor de ultrasonidos, un receptor de ultrasonidos y un procesador de señales.
- j) Detector por infrarrojos pasivos: todos los cuerpos emiten radiaciones infrarrojas si están a una temperatura superior al cero absoluto. Este detector dispone de un sensor piroeléctrico, que genera en sus bornes una débil corriente cuando recibe una variación de radiación infrarroja. Se utiliza para detectar la presencia de un intruso percibiendo una variación en la cantidad de infrarrojos en el medio. Poseen el inconveniente de depender de la temperatura ambiental.
- k) Detectores de sonido: detectan sonidos que superan un cierto nivel de amplitud. Están prácticamente en desuso, debido a que solo deben instalarse en recintos dispuestos con un excelente aislamiento.
- l) Detectores capacitivos: captan la proximidad de un intruso a un objeto metálico, ya que varía la constante dieléctrica del ambiente y por tanto la capacidad eléctrica entre el intruso y la tierra de referencia. Necesita una instalación muy cuidadosa y pueden producir falsas alarmas por interferencias radioeléctricas.
- m) Detectores de vibración en vallas: protegen las vallas de los perímetros, detectan vibración, corte o movimiento de las mismas. Puede utilizar dos tipos de sensores, inerciales o continuos dependiendo del resultado deseado.

Los sensores inerciales detectan las vibraciones de las vallas dispuestas en ellos y los sensores continuos se basan en el efecto llamado **Tribo Eléctrico**, es decir que detectan la producción de una corriente eléctrica al friccionar dos superficies conductoras con un aislante intercalado entre ellas.

- n) Detector por redes de fibra óptica: están basados en la transmisión de señales infrarrojas en una fibra óptica, que es una guía de luz cilíndrica compuesta de un alma de silicio con un índice de refracción muy elevado. La luz se transmite dentro del corazón de la fibra por reflexiones sucesivas sobre la superficie de separación entre éste y el revestimiento óptico. Cuando la fibra óptica es alterada (doblada o rota) la luz infrarroja se interrumpe provocando una señal de alarma.
- o) Detectores de presión diferencial: consiste en la instalación de dos tubos paralelos enterrados recorriendo el perímetro a proteger, rellenos de un fluido hidráulico en el que se mantienen presiones predeterminadas. Una diferencia en la presión es la que hace que se transmita la alarma. Debido a que la presión detectada no solo es función del peso del intruso sino de la intensidad de la misma, se puede detectar por pasos, saltos, o deslizamiento sobre el terreno.
- p) Detector por video sensor: son detectores que utilizan la señal de video procedente de una cámara de televisión; se activan al producirse una variación predeterminada del nivel de luminosidad en la zona vigilada. Tienen la ventaja de que utilizando las cámaras instaladas de CCTV pueden convertirse en sensores de protección de las zonas vigiladas. Se adaptan a cualquier tipo de terreno.

1.2.2. Señalizadores

Los señalizadores también se conocen con el nombre de avisadores, estos cumplen una doble función; un efecto psicológico para hacer huir al intruso y anunciar que se ha producido una intrusión en ese lugar. Los señalizadores pueden ser: sirenas electrónicas, sirenas mecánicas, iluminación súbita, luz lanza-destellos, flash, llamada telefónica, llamada por radio, máquina fotográfica, filmadora, cámaras digitales, cámaras web, etc.

Los requisitos fundamentales que debe cumplir un señalizador son:

- a) Imposibilidad de manipulación de los cables de conexión.
- b) Autoalimentación propia (batería).
- c) Actuación que impida la confusión con los señalizadores usados por las Fuerzas y Cuerpos de Seguridad, bomberos, ambulancia, etc.
- d) Bajo consumo.
- e) Potencia ajustada a la reglamentación vigente.
- f) Funcionamiento garantizado en condiciones atmosféricas adversas.
- g) Fiabilidad.

Según el riesgo del lugar a vigilar, las instalaciones deben ofrecer una seguridad diferente en contra de las puestas fuera de servicio mal intencionadas.

Las instalaciones con alto grado de seguridad disponen de circuitos especiales constantemente vigilados, con líneas de aviso de sabotaje, que vigilan todos los elementos de la instalación las 24 horas del día, en particular los dispositivos de alarma locales. El intento de poner fuera de servicio alguno de estos elementos, lleva consigo la activación de alarma interior en la central o de una alarma exterior.

1.3. Sistemas de vigilancia modernos

Actualmente existen muchas compañías que se dedican a la fabricación de sistemas de seguridad, especialmente aquellos que integran la seguridad industrial. Dichos sistemas se comercializan por muchos medios como la televisión, radiodifusoras y sobretodo por Internet, buscando un mismo objetivo: **Hacer más segura la vida de la humanidad**. Los avances en la tecnología han permitido tener sistemas más seguros y modernos, pero de todas las partes que integran un sistema de vigilancia, en la que se ha tenido más adelanto es en los detectores, dado que estos elementos pueden controlar una variedad más amplia de sucesos. Entre la diversidad de dispositivos actuales podemos encontrar los siguientes:

- a) Alarma de infrarrojos LAS3255: es una Alarma con mando a distancia como se observa en la Figura 1, posee una potente sirena, ideal para hogares y comercios; es de fácil instalación, detecta en un área de 7 metros y 135°, posee una función de back-up (guarda su configuración y sólo trabaja con lo esencial) cuando se conecta a la red y funciona en caso de fallo en el suministro eléctrico. Funciona a través de un alimentador 220 voltios de corriente alterna (v) o 6/15 voltios de corriente directa o con baterías.

Figura 1. Alarma de infrarrojos LAS3255



Fuente: www.seguridadplus.com

- b) Alarma de Seguridad LAS5201: es un avisador de presencia que proporciona una mayor seguridad contra los robos, ya que la intrusión de cualquier individuo será detectada inmediatamente. Posee un detector de movimiento por infrarrojos pasivos como se muestra en la Figura 2, que es capaz de percibir la presencia de cualquier persona situada en un radio de 10 mts. y una vez ha detectado alguna presencia, emitirá automáticamente un sonido de campana o de sirena que se prolongará durante 30 segundos.

Figura 2. Alarma de seguridad LAS5201



Fuente: www.seguridadplus.com

- c) Detector de Movimiento LAS1618: es un detector de movimiento con alarma y teléfono, su diseño incluye un marcador telefónico (como se observa en la Figura 3), que le permite programar hasta 3 llamadas con un número de hasta 16 dígitos por medio de su marcador DTMF (Dual Tone Multi Frequency) integrado, que marca cada dígito telefónico por medio de 2 tonos de diferente frecuencia, simultáneamente.

En caso de intrusión, el DTMF marcará el número de sus vecinos, familiares o su propio móvil. La central dispone de un sensor volumétrico con un área de detección de 180º y una distancia máxima de 6 mts.

Además incorpora una potente sirena piezoeléctrica Integrada en la carcasa, su capacidad es de 90 decibeles (dB). a un metro de distancia (aproximadamente como el ruido que ocasiona el tráfico en la ciudad) y se puede activar o desactivar la alarma tanto a través del teclado como desde el mando a distancia.

La pila se encarga de proporcionar autonomía en el sistema en caso de caída prevista o accidental de la energía eléctrica.

Figura 3. Detector de movimiento LAS1618



Fuente: www.seguridadplus.com

- d) Asegurador de puertas LAS5200: proporciona una alta seguridad contra los robos, ya que cualquier intruso será detectado inmediatamente. Posee una sirena de 115 dB que será activada con cualquier extraño que intente entrar por la puerta o ventana protegida; además el usuario puede programar el código de seguridad con hasta 10,000 combinaciones distintas en un teclado de fácil manejo como se observa en la Figura 4. Tiene 4 modos de funcionamiento: alarma, aviso, alarma temporizada y desactivación.

Figura 4. Asegurador de puertas LAS5200



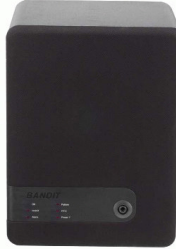
Fuente: www.seguridadplus.com

- e) Sistema de Niebla Activa LAS3209: es un sistema antirrobo que genera 28 metros cúbicos (mts^3) de niebla por segundo y provoca la invisibilidad del recinto, ganando el tiempo necesario para la llegada de las autoridades; además el intruso queda sin visión provocando su inmediata huida a causa de encontrarse indefenso.

El sistema de niebla se activa por el disparo de líquido a presión contenido en cartuchos que son recargables, dicha niebla no es tóxica y no deja residuos; cuando se dispara se disipa aproximadamente entre 15 y 20 minutos dependiendo de las dimensiones y la ventilación del recinto, que es aconsejable no supere los 140 m. cuadrados por 3m. de altura.

Este sistema permite asociar su funcionamiento a una alarma y tiene forma de altavoz como se observa en la figura 5 para pasar desapercibido en cualquier tipo de ambiente. En caso de falta de suministro eléctrico, se garantiza la operación del sistema electrónico por 24 horas (hrs.) y la generación de niebla por 2 hrs.

Figura 5. Sistema de niebla activa LAS3209



Fuente: www.seguridadplus.com

- f) Alarma de Tele-Asistencia LAS1495: es un sistema, que a través de la línea telefónica (posee un teléfono incorporado como se muestra en la Figura 6), permite controlar a distancia alguna intrusión, fuego o accidente de alguna persona. Avisa telefónicamente de todo lo ocurrido a un teléfono fijo o móvil, además se puede escuchar directamente qué es lo que está sucediendo y permite incluso hablar. El sistema de activación puede ser automático o utilizarse de forma manual, de esta manera la persona que porte el llavero lo puede utilizar de tele-asistencia.

El sistema realiza las llamadas pertinentes a ocho teléfonos predeterminados, hasta que conteste alguno de ellos, reproduciendo el mensaje sonoro que haya sido grabado, distinguiendo si la alarma es por fuego o intrusión. El sistema tiene sensores de movimiento y de fuego ajustables, que envían la señal vía radio para no complicarse con la instalación de cables. También, se puede disponer como opción de sensores para humo y gas. El sistema incluye baterías en caso de corte eléctrico. Los sensores de Humo, Gas y presencia así como la sirena son inalámbricos con un rango de sensibilidad confiable a 8 mts. con un ángulo de 90º horizontales y 23º verticales.

Figura 6. Alarmas de tele-asistencia LAS1495



Fuente: www.seguridadplus.com

Además de todos los sistemas anteriormente descritos, existen muchos sistemas con microcámaras, con los cuales se pueden grabar situaciones secretamente, sin necesidad preocuparse en instalaciones molestas y difíciles y sin ser obvios en cuanto a mostrar que se está protegiendo cierta área o cierto individuo. Por ejemplo, existe una cámara oculta en una cajetilla de tabaco que utiliza una batería de 9 v. y posee un transmisor vía radio de 1,2 gigahertz (GHz) incorporado.

Esta cámara tiene visión en un ángulo de 90°, con receptor y transmisor de 4 canales, tamaño de 90 x 55 x 22 milímetros (mm.) y un peso de 185 gramos (grs.). La autonomía estimada es de 90 minutos.

También existe una cámara diminuta a color con dimensiones de 14 mm. X 14 mm. X 13 mm., cuya fabricación es ideal para poderse ocultar en cualquier objeto o inclusive para ser llevadas encima con una batería, ya que además posee un consumo mínimo de tan solo 50 miliamperios (mA.) y su resolución es de 628 X 582 píxeles.

Figura 7. Microcámaras



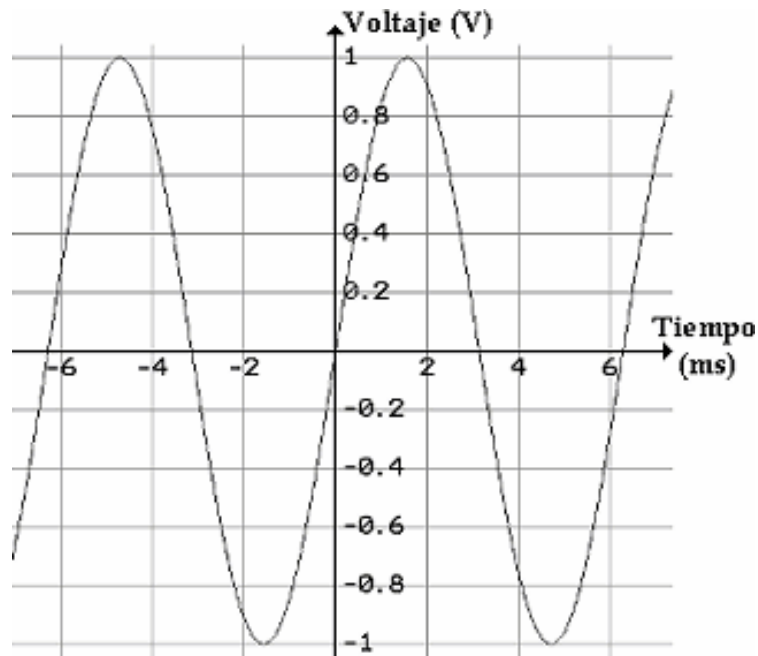
Fuente: www.seguridadplus.com

2. CARACTERÍSTICAS DE VIDEO DIGITAL

2.1 Video análogo

En un sistema análogo, la información se transmite mediante alguna variación infinita de un parámetro continuo como puede ser la tensión en un hilo o la intensidad de flujo de una cinta (véase la Figura 8). En un equipo de grabación, lo plasmado a lo largo del soporte físico es un elemento análogo; es decir que no importa en que punto se examine una grabación a lo largo de toda su extensión se encontrará un valor para la señal grabada. Dicho valor puede variar con una resolución infinita dentro de los límites físicos del sistema.

Figura 8. Variación infinita de un parámetro continuo en función del tiempo



Fuente: www.fuac.edu.com

Dichas características suponen la debilidad principal de las señales análogas. Dentro del ancho de banda permisible, cualquier forma de onda, incluso el ruido, es válida.

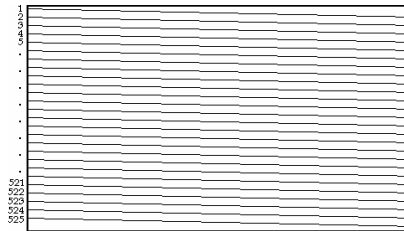
El ruido no puede detectarse en un sistema análogo; se puede tener la sospecha de que existe ruido, pero no se sabe qué proporción de la señal recibida corresponde al ruido y cual es la señal original. Si la función de transferencia de un sistema no es lineal, se producirán distorsiones, pero las formas de onda distorsionadas aún serán válidas; un sistema análogo es incapaz de detectar distorsiones.

Es característico de los sistemas análogos el hecho de que las degradaciones no puedan ser separadas de la señal original, al final en un sistema determinado, la señal estará formada por la suma de todas las degradaciones introducidas en cada etapa por las que haya pasado. Esto limita el número de etapas por las que una señal puede pasar sin que quede inutilizable.

2.1.1. Proceso de exploración de la imagen

Todas las normas vigentes de televisión en la actualidad, **NTSC** (*National Television Systems Comitee*, Comité Nacional de Sistemas de Televisión), **PAL** (*Phase Alternation Line*, Línea de Fase Alternada) y **SECAM** (*Systeme Electronique Color Avec Memoire*) se derivan, directa o indirectamente, de los estándares en blanco y negro definidos en los años 40 y 50. Estas primeras emisiones utilizaban un barrido progresivo (todas las líneas de la imagen se barren consecutivamente, como se puede ver en la Figura 9).

Figura 9. Representación simplificada del barrido progresivo

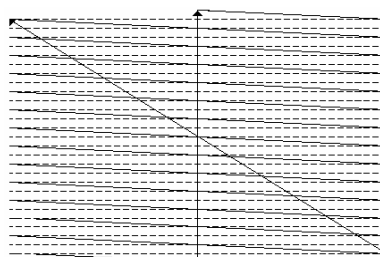


Fuente: www.fuac.edu.com

Por razones de orden práctico (radiaciones debidas a fugas magnéticas de los transformadores o filtrados imperfectos), fue indispensable utilizar una frecuencia de imagen que estuviera relacionada con la frecuencia de la red (60 Hz. en EE.UU., 50 Hz. en Europa) para minimizar el efecto visual de estas imperfecciones; la frecuencia de exploración fue, por tanto, de 30 imágenes/segundo en EE.UU. y de 25 imágenes/segundo en Europa. Estas primeras imágenes presentaban un parpadeo bastante molesto (también llamado *flicker* de campo). Tiempo después la captación de la imagen se hizo electrónica, haciendo que las definiciones alcanzaran un mayor número de líneas, esto gracias al barrido entrelazado.

El barrido entrelazado consiste en la transmisión de un primer campo compuesto por las líneas impares de la imagen y a continuación un segundo campo formado por las líneas pares, como se ve en la Figura 10.

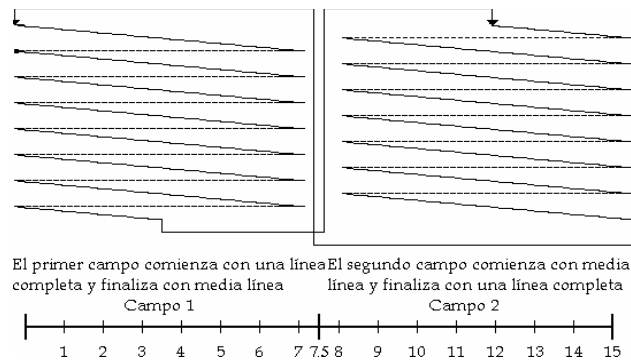
Figura 10. Barrido entrelazado 2:1. Cuadro completo



Fuente: www.fuac.edu.com

Esta forma de barrer la imagen, permite duplicar la frecuencia de refresco de la pantalla (50 ó 60 Hz., en lugar de los 25 ó 30 Hz.) sin aumentar el ancho de banda para un número de líneas dado. Como se ve en la Figura 11, el barrido entrelazado se obtiene utilizando un número impar de líneas, por ejemplo 525 ó 625 líneas que constituyen un cuadro, de manera que el primer campo comience en una línea completa, terminando en la mitad de otra línea, y el segundo campo comience en la mitad de una línea y finalice con una línea completa. En los países donde la frecuencia de la red es de 60 Hz., la velocidad de cuadro es de 30 Hz. por segundo y, por consiguiente, la frecuencia de campo es de 60Hz.

Figura 11. Los campos de un entrelazado 2:1



Fuente: www.fuac.edu.com

2.1.2. Frecuencias de exploración horizontal y vertical

La velocidad de campo de 60 Hz. es la frecuencia de exploración vertical. Este es el ritmo con que el haz electrónico completa su ciclo de movimiento vertical, desde la parte superior hasta la parte inferior de la pantalla para volver nuevamente a la parte superior. El número de líneas de exploración horizontal de un campo es la mitad del total de las 525 líneas de un cuadro completo (en el sistema NTSC), ya que un campo contiene la mitad de las líneas. Esto da por resultado 262.5 líneas horizontales para cada campo.

Como el tiempo que corresponde a un campo es 1/60 segundos y cada campo contiene 262.5 líneas, el número de líneas por segundo es: $262.5 \times 60 = 15750$ líneas/segundo. Esta frecuencia de 15750 Hz. es la velocidad con que el haz electrónico completa su ciclo de movimiento horizontal de izquierda a derecha por lo tanto, el tiempo durante el cual se realiza la exploración de una línea horizontal es de $63.5 \mu s$ ($1/15750$).

2.1.3. Las señales de color

El sistema para la televisión en color es el mismo que para la televisión monocromática excepto que también se utiliza la información de color. Esto se realiza considerando la información de imágenes en términos de rojo, verde y azul. Cuando es explorada la imagen en la cámara, se producen señales de video separadas para la información de rojo, verde y azul de la imagen. Filtros de color separan los colores para la cámara, sin embargo, para el canal estándar de 6 MHz. de televisión, las señales de video de rojo, verde y azul son combinadas de modo que se forman dos señales equivalentes, una correspondiente al brillo y otra para el color.

Dichas señales pueden tener tres distintos formatos, YIQ , YUV y $YCrCb$; dependiendo de cómo se representa la señal de Crominancia C , ya que la señal de Luminancia Y es igual para los tres. La señal de luminancia contiene solo variaciones de brillo de la información de la imagen, incluyendo los detalles finos, lo mismo que en una señal monocromática. La señal de luminancia o Y se utiliza para reproducir la imagen en blanco y negro, o monocroma y se forma combinando las señales roja (R), verde (G) y azul (B), en diferente porcentaje cada una, según la expresión:

$$Y = K_r * R + (1 - K_r - K_b)G + K_b * B \quad (2.1.3.1)$$

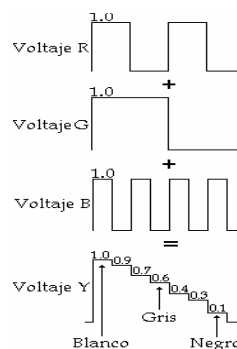
Donde K_r y K_b son coeficientes derivados de un ajuste que se hace a las señales R , G y B llamado **corrección Gamma**. Dicho ajuste es realizado en monitores, televisores o pantallas derivado de que el ojo humano no capta los colores en una forma lineal, sino que cada color en diferente porcentaje, por ejemplo, el ojo humano es más sensible al verde que al rojo o al azul. Según las normas estándar $K_r = 0.299$ y $K_b = 0.114$ por lo que la fórmula (2.1.3.1.) queda como:

$$Y = .0299 R + 0.587 G + 0.114 B \quad (2.1.3.2.)$$

Los porcentajes que se muestran en la ecuación corresponden a la brillantez relativa de los tres colores primarios. Por lo tanto, una escena reproducida en blanco y negro por la señal Y tiene exactamente la misma brillantez que la imagen original. La Figura 12 muestra como el voltaje de la señal Y se compone de varios valores de R , G y B . La señal Y tiene una máxima amplitud relativa de unidad, la cual es 100% blanca. Para los máximos valores de R , G y B (1V cada uno), el valor de brillantez es:

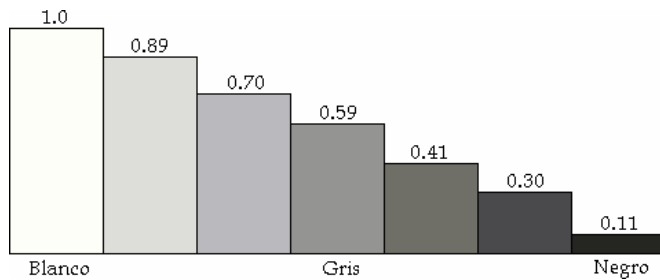
$$Y = .299(1) + 0.587(1) + 0.114(1) = 1\text{Lumen} \quad (2.1.3.3.)$$

Figura 12. Obtención de la señal Y



Fuente: www.fuac.edu.com

Figura 13. Valores de luminancia relativa



Fuente: www.fuac.edu.com

La señal de crominancia C , es la combinación de las señales de color y depende del formato de la señal utilizada. Según el formato YUV , C está formada por las señales $R-Y$ y $B-Y$ que se derivan de restar la Luminancia Y de las señales de color rojo R y azul B y se expresan como:

$$R - Y = 0.701R - 0.587G - 0.114B \quad (2.1.3.4.)$$

$$B - Y = -0.299R - 0.587G + 0.886B \quad (2.1.3.5.)$$

Según el formato YIQ , las señales I y Q son las que forman la crominancia C , que a su vez se derivan de las señales $R-Y$ y $B-Y$ y se utilizan en el formato de video NTSC. La señal I es la señal de color en fase y se expresa como:

$$I = \frac{(R - Y) * \cos 33^\circ}{1.14} + \frac{(B - Y) * \sin 33^\circ}{2.03} \quad (2.1.3.6.)$$

La señal Q es la señal de color en cuadratura y se genera de la misma manera, su expresión es:

$$Q = \frac{(R - Y) * \sin 33^\circ}{1.14} + \frac{(B - Y) * \cos 33^\circ}{2.03} \quad (2.1.3.7.)$$

Las señales I y Q se combinan para producir la señal C . Debido a que las señales I y Q están en cuadratura, la señal C o crominancia es la suma vectorial de estas, y su expresión es:

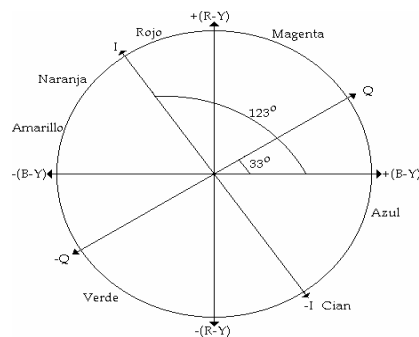
$$C = \sqrt{I^2 + Q^2} \quad (2.1.3.8.)$$

y su ángulo de fase está dado por

$$\theta = \tan^{-1} \frac{Q}{I} \quad (2.1.3.9.)$$

La Figura 14 muestra la rueda de colores para la radiodifusión de televisión. Las señales $R-Y$ y $B-Y$ se utilizan en la mayor parte de los receptores de televisión a color para demodular las señales de video R , G y B así como en el formato de video PAL. En el receptor, la señal C reproduce colores en proporción a las amplitudes de las señales I y Q . El matiz (o tono del color) se determina por la fase de la señal C y la profundidad o saturación es proporcional a la magnitud de la señal C . La parte exterior del círculo corresponde al valor relativo de 1.

Figura 14. Representación de los colores en NTSC



Fuente: www.fuac.edu.com

Según el formato $YCrCb$ la crominancia C está formada por Cr y por Cb que a su vez también se derivan de las señales $R-Y$ y $B-Y$ y están dadas por:

$$Cr = \frac{0.5 * (R - Y)}{(1 - Kr)} \quad (2.1.3.10.)$$

$$Cb = \frac{0.5 * (B - Y)}{(1 - Kb)} \quad (2.1.3.11.)$$

donde Kr y Kb son los coeficientes derivados de la corrección **Gamma**.

Así se consigue que los sistemas de color y monocromáticos sean completamente compatibles.

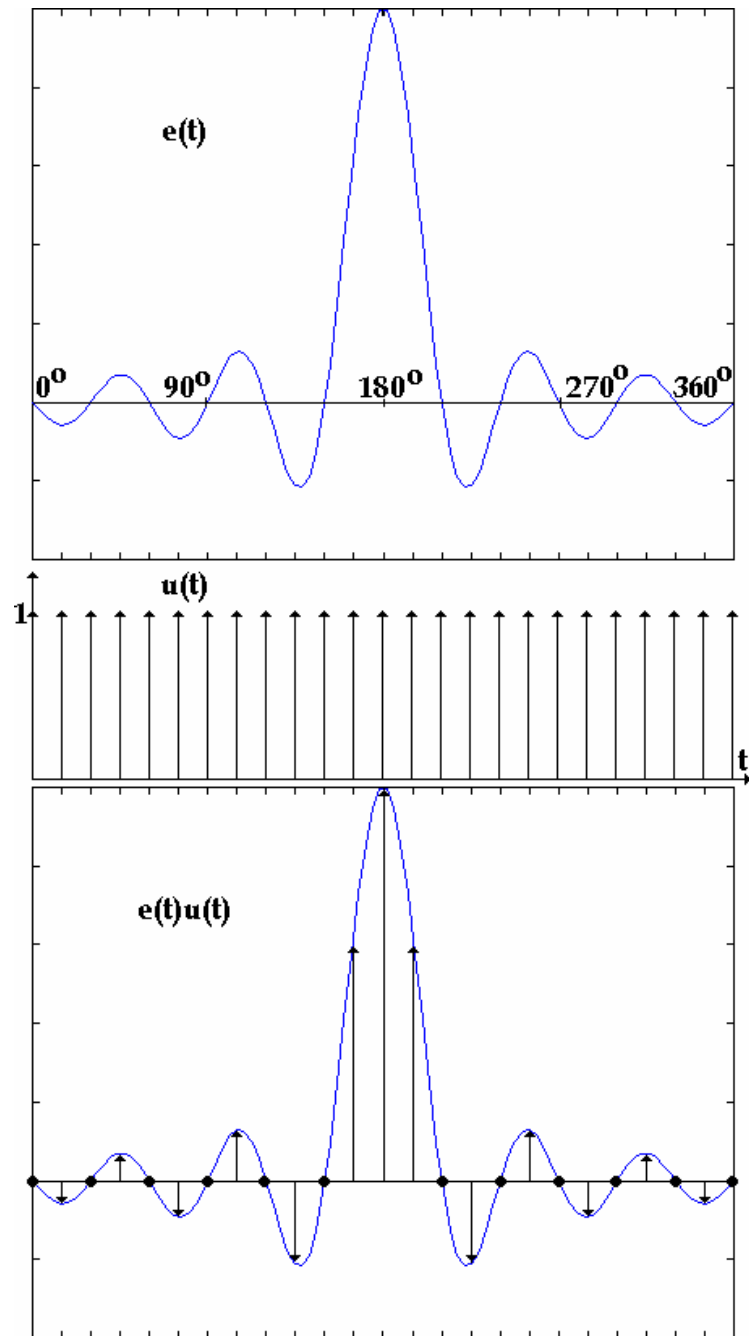
2.2. Video digital

La digitalización de una señal de video tiene lugar en tres pasos: Muestreo, Cuantización y Codificación.

2.2.1. Muestreo

Sea una señal análoga $e(t)$ como la representada en el Figura 15. Se toman muestras breves de $e(t)$ cada 15° a partir de $t = 0$. En 360° se habrán explorado 24 muestras. El resultado será una serie de impulsos cortos cuyas amplitudes siguen a la señal análoga. A este tren de impulsos modulados en amplitud por la señal análoga se le denomina señal PAM (*Pulse Amplitude Modulation* o Modulación por Amplitud de Pulsos). Este muestreo puede representarse por la multiplicación de la señal análoga $e(t)$ por un tren de impulsos $u(t)$, dando por resultado la señal de la parte inferior de la Figura 15.

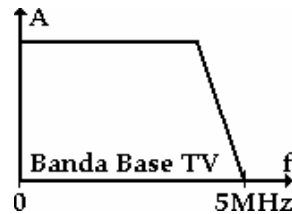
Figura 15. Muestreo de una señal analógica $e(t)$ por un tren de impulsos $u(t)$



Fuente: www.fuac.edu.com

Ahora bien, una señal de video está compuesta por un gran número de frecuencias formando un espectro continuo que va desde 0 a unos 5 MHz. como se representa en la Figura 16.

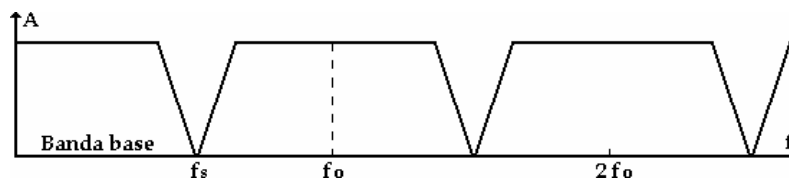
Figura 16. Banda base de la señal de video



Fuente: www.fuac.edu.com

Al muestrear esta señal, cada frecuencia de video aparecerá en las bandas laterales superiores e inferiores de cada armónico de la frecuencia de muestreo, incluyendo naturalmente la banda base, esto es, el armónico cero. El espectro de la señal muestreada se presentará por tanto, como se ve en la Figura 17. De esta misma figura se deduce una condición elemental que debe cumplirse: que $f_o > 2f_s$ para que la banda lateral inferior de la frecuencia de muestreo y la banda base no se superpongan.

Figura 17. Espectro de una señal de video muestreada a la frecuencia f_o



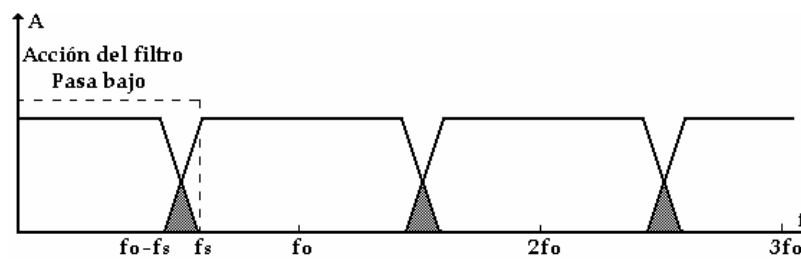
Fuente: www.fuac.edu.com

Este razonamiento fue deducido por Nyquist-Shannon, al establecer que para conseguir un muestreo-recuperación sin distorsión, se requiere que la frecuencia de muestreo f_o sea al menos dos veces más elevada que la frecuencia máxima presente en la señal analoga muestreada.

La recuperación de la banda base se realizaría con un filtro pasa bajo que corte todas las frecuencias superiores a $f_o/2$.

De no cumplirse el teorema del muestreo de Nyquist, el filtro dejaría pasar frecuencias pertenecientes a la banda lateral inferior contaminantes de la banda base, que producirían solapamientos con las frecuencias más altas de la misma. Este efecto se denomina *aliasing* (ver la Figura 18).

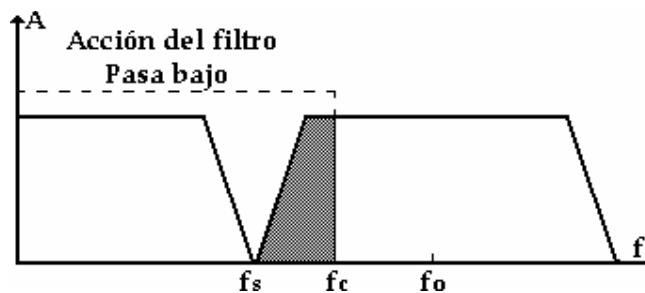
Figura 18. Cuando la frecuencia de muestreo es $f_o > 2f_s$



Fuente: www.fuac.edu.com

Otro motivo de *aliasing* se produce cuando el filtro no está bien calculado y permite el paso de frecuencias de la banda lateral inferior, aunque no estén solapadas con la banda base (ver la Figura 19).

Figura 19. Cuando la frecuencia de corte del filtro PB es superior a $f_o - f_s$

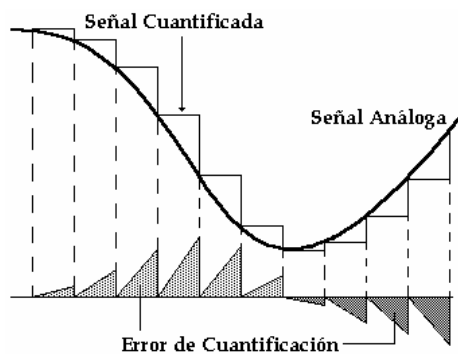


Fuente: www.fuac.edu.com

2.2.2. Cuantización

Así se denomina al proceso mediante el cual se atribuye a cada muestra un valor de amplitud dentro de un rango de niveles predeterminado. Este valor se representa por un número que será convertido a un código de ceros y unos en el proceso de codificación. Por razones de facilidad en los cálculos, el número de niveles se hace coincidir con una potencia de dos y los impulsos de la señal PAM se redondean al valor superior o inferior según sobrepasen o no la mitad del ancho del nivel en que se encuentran. El error que se produce con estas aproximaciones equivale a sumar una señal errónea a los valores exactos de las muestras, como se ve en la Figura 20.

Figura 20. Error de cuantización



Fuente: www.fuac.edu.com

Esta señal errónea aparecerá en el proceso de recuperación después de la decodificación digital-análoga, en forma de ruido visible. Se habla así de "ruido de cuantización" que dependerá obviamente del número N de niveles empleados en el proceso. Cuanto más niveles existan menor será el ruido generado. La relación señal/ruido de cuantización es:

$$\frac{S}{C} = (20\text{Log}N + 10.8)\text{dB} \quad (2.2.2.1.)$$

de cuyo resultado se sacan las siguientes conclusiones:

- a) La relación señal/ruido de cuantización depende únicamente del número de niveles N en que se divide el trazo completo de la señal.
- b) Existe un sumando constante 10.8 dB. que tiene su origen en la misma definición de señal/ruido en televisión, donde se toma para la señal el valor pico a pico y para el ruido su valor eficaz.

Es evidente que usando codificación binaria resulta $N = 2^m$, donde m = número de bits, por tanto:

$$\frac{S}{C} = (6(m) + 10.8) = 58.8dB \quad (2.2.2.2.)$$

La anterior ecuación es valida para la digitalización de una señal monocroma o para cada componente de color. Se adoptaron 8bits para la digitalización de la señal de video, por lo que la relación señal/ruido de cuantización queda como:

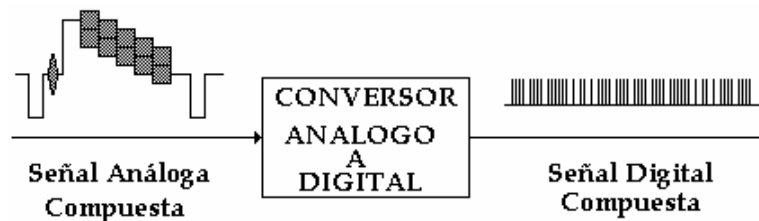
$$\frac{S}{C} = (6(8) + 10.8) = 58.8dB \quad (2.2.2.3.)$$

2.2.3. Codificación

La codificación final de la señal de salida de un equipo depende de su aplicación; puede usarse por ejemplo un código binario puro o un código de complemento a dos para aplicaciones locales, pero cuando se trata de aplicaciones específicas, la codificación se puede dar en forma especial.

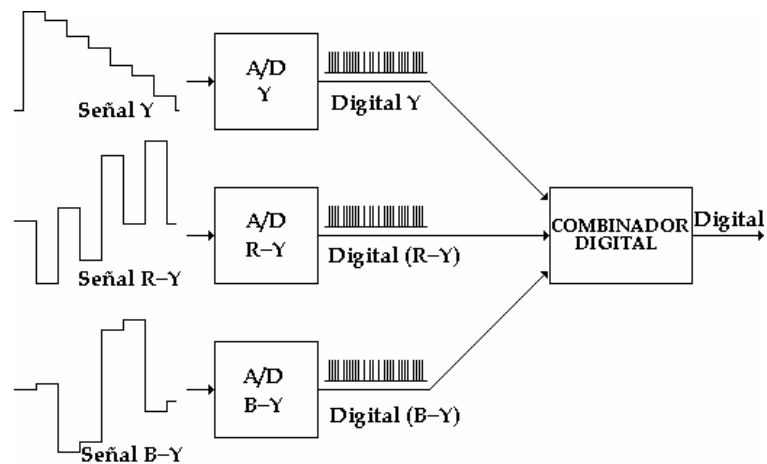
Dos planteamientos aparentemente contradictorios se mantienen aún hoy día acerca de la digitalización de la señal de televisión en color, la codificación de señales compuestas (ver la Figura 21) y la codificación de componentes (ver la Figura 22).

Figura 21. Codificación de la señal compuesta



Fuente: www.fuac.edu.com

Figura 22. Codificación de componentes



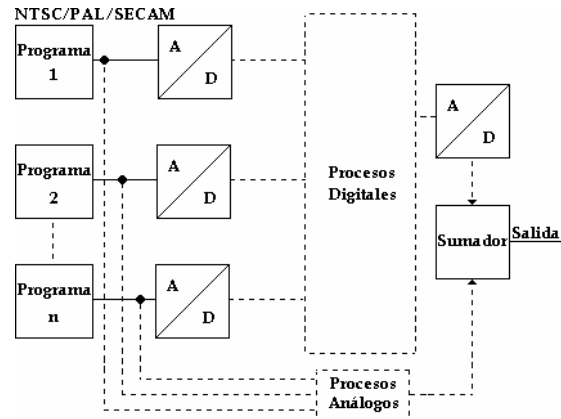
Fuente: www.fuac.edu.com

2.2.3.1. Codificación de las señales compuestas

Esta propuesta consiste en digitalizar directamente las señales compuestas existentes (NTSC, PAL, SECAM). Con ello persiste el problema de la incompatibilidad de las distintas normas internacionales, aun manteniendo la misma frecuencia de muestreo y codificación.

La decodificación devolvería las señales NTSC, PAL o SECAM, respectivamente. La ventaja fundamental de digitalizar la señal compuesta radica en que el equipo puede incluirse como una unidad mas en los estudios análogos actualmente en servicio, sin necesidad de codificar o decodificar el NTSC, PAL o SECAM. La Figura 23 muestra como opera el tratamiento de imágenes análogas durante la transición de la televisión análoga a digital, para el caso de codificación de señales compuestas.

Figura 23. Transición de análogo a digital de las señales compuestas



Fuente: www.fuac.edu.com

Pasada la transición, se tiene otra ventaja, que es el tratamiento de una señal única de video como ocurre actualmente en los estudios análogos. Para los casos NTSC y PAL que modulan en amplitud a la subportadora de color, el muestreo, la mezcla y el encadenado corresponderán a una sencilla multiplicación de todas las muestras por un factor situado entre 0 y 1. Pero en el caso del SECAM, es necesario descomponer primero la señal de video en sus componentes Y , $R-Y$ y $B-Y$ antes de la mezcla. Este problema elimina esta ventaja para el SECAM y en todo caso, cada fuente de video digital tendría que disponer de codificación y decodificación NTSC, PAL y SECAM, lo que representa una degradación de las imágenes por causa de los sucesivos procesos de codificación-decodificación.

2.2.3.2. Codificación en componentes

Por este método se digitalizan las tres señales Y , $K_1(R-Y)$ y $K_2(B-Y)$ donde K_1 y K_2 son factores de ponderación que impone el sistema digital. Estos factores no tienen los mismos valores que los coeficientes ponderados de NTSC, PAL o SECAM. La primera y gran ventaja que se deriva de esta codificación es que siendo estas tres señales comunes a todos los sistemas, la compatibilidad puede alcanzarse por regulación internacional de los parámetros de muestreo, cuantización y codificación. En tal sentido el CCIR (*Comité Consultatif International des Radiocommunications* o Comité Consultivo Internacional de Radio Comunicaciones) emitió en 1982 la norma 4:2:2 CCIR 601 de televisión digital en componentes.

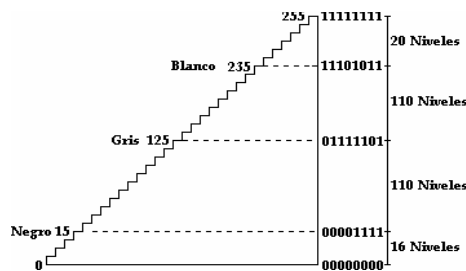
Esta norma define los parámetros básicos del sistema de televisión digital que aseguran la mayor compatibilidad mundial. Se basa en una señal Y , Cr y Cb en el formato llamado 4:2:2 (4 muestreos Y por 2 muestreos Cr y 2 muestreos Cb), con una digitalización sobre 8 bits, con posibilidad de ampliarla a 10 bits para aplicaciones más exigentes.

Cualquiera que sea el estándar de barrido, la frecuencia de muestreo es de 13.5 MHz para la luminancia Y . Para las señales de crominancia Cr y Cb , dado su ancho de banda más limitado se muestrean a la mitad de la frecuencia de la luminancia, es decir, 6.75 MHz. Lo que corresponde a una definición de 720 muestreos por línea en luminancia y de 360 muestreos por línea de crominancia.

Para esta frecuencia de muestreo, el número de muestras por línea es de 864 y 858 para la luminancia y de 432 y 429 para las diferencias de color (sistema de 625 y 525 líneas respectivamente).

La estructura de muestreo es ortogonal, a consecuencia de que la frecuencia de muestreo es un múltiplo entero de la frecuencia de líneas. Las muestras de las señales diferencias de color se hacen coincidir con las muestras impares de la luminancia, o sea 1ª, 3ª, 5ª, etc. Como el número de bits/muestra es de 8, tanto para la luminancia como para las señales diferencias de color, lo que corresponde a $2^8 = 256$ niveles de cuantización. La luminancia utiliza 220 niveles a partir del 16 que corresponde al nivel de negro, hasta el 235 correspondiente al nivel de blanco. Se acepta una pequeña reserva del 10% para la eventualidad de que ocurran sobremodulaciones (ver la Figura 24).

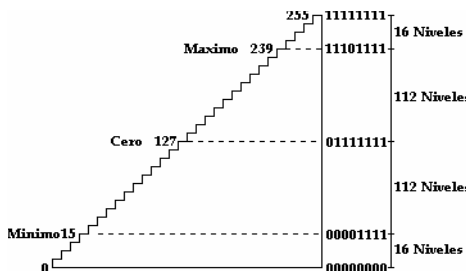
Figura 24. Cuantización de la señal de luminancia



Fuente: www.fuac.edu.com

Para las señales diferencias de color se utilizan 224 niveles, que se reparten a ambos lados del cero análogo, que se hace corresponder con el número digital 128. La señal variará entre los valores extremos $128+112=240$ y $128-112=16$, con una reserva de 16 niveles a ambos lados (ver la Figura 25).

Figura 25. Cuantización de la señal de crominancia



Fuente: www.fuac.edu.com

Dado que las señales Cr y Cb están disponibles simultáneamente en cada línea, la definición vertical es idéntica tanto para luminancia como para crominancia, y se corresponde con el número de líneas útiles del estándar de exploración de partida (480 para los de 525 líneas, 576 para los de 625 líneas). El flujo bruto resultante es $(13.5 * 8) + (2 * 6.75 * 8) = 216 \text{ Mbit/s}$. Además, la digitalización de la parte útil de la señal de video solo requiere 166 Megabit/segundo (Mb/s), si se tiene en cuenta la inutilidad de digitalizar los intervalos de supresión del haz (también llamados "*blanking*") de línea y campo. Por tanto, estos tiempos libres pueden aprovecharse para transportar los canales de sonido digital, así como datos de servicio u otros.

Tabla I. Parámetros de la norma 4:2:2 CCIR 601

Parámetros	Sistemas	
	NTSC 525 líneas 60 campos	PAL/SECAM 625 líneas 50 campos
Señales codificadas	Y, Cr, Cb	
Número de muestras por línea completa <ul style="list-style-type: none"> • Luminancia • Crominancia 	858 429	864 432
Estructura de muestreo	Ortogonal, estructura idéntica de todos los campos y cuadros. Las señales Cr y Cb se muestrean simultáneamente con las muestras impares de la luminancia.	
Frecuencia de muestreo <ul style="list-style-type: none"> • Luminancia • Crominancia 	13.5 MHz 6.75 MHz	
Codificación	Cuantización lineal. Codificación con 8 bits por muestra para la luminancia y cada señal de crominancia.	
Número de muestras activas por líneas digital: <ul style="list-style-type: none"> • Luminancia • Crominancia 	720 360	
Correspondencia entre los niveles de video y de cuantización: <ul style="list-style-type: none"> • Luminancia • Crominancia 	220 niveles de cuantización. El nivel negro corresponde al número digital 16; el nivel nominal de blanco al número 235. 224 niveles de cuantización en la parte central del margen de cuantización. El nivel de video cero corresponde al número 128.	

La segunda ventaja de esta codificación es que una vez alcanzada la digitalización, sólo se requiere un paso final de conversión D/A y una codificación NTSC, PAL o SECAM según el sistema adoptado de transmisión. La tercera ventaja es que el tratamiento digital en componentes elimina los efectos perturbadores mutuos de luminancia y crominancia a la vez que en edición electrónica desaparecen los problemas derivados de la estructura de 4 y 8 campos NTSC y PAL respectivamente. Sólo habría de tenerse en cuenta la estructura de dos campos entrelazados como en televisión en blanco y negro.

2.2.5. Ventajas del video digital

La calidad de reproducción de un sistema digital de video bien diseñado es independiente del medio y depende únicamente de la calidad de los procesos de conversión. Cuando se copia una grabación digital, aparecen los mismos números en la copia, no se trata de un duplicado, sino de una clonación. Si no es posible distinguir la copia del original, no se habrá producido ninguna pérdida en la generación. Las grabaciones digitales pueden copiarse indefinidamente sin que haya pérdida en la calidad.

Una de las mayores ventajas que presenta la tecnología digital es su bajo costo. Si la realización de copias no ocasiona pérdidas de calidad, los equipos de grabación no tienen por qué ser mejor de lo necesario. No hay necesidad del consumo de cinta tan grande y excesiva que tienen los equipos de grabación análogos. Cuando la información que se ha de grabar adopta la forma de números discretos, estos pueden empaquetarse densamente en un soporte sin pérdida de la calidad. De darse el caso que algunos bits estén defectuosos por causa del ruido o de pérdidas de señal, el sistema de corrección de errores puede restituir el valor original.

Las redes de comunicaciones desarrolladas para manejar datos pueden llevar perfectamente video digital acompañado también de audio a distancias indefinidas sin pérdidas de calidad. La difusión de televisión digital emplea estas técnicas para eliminar las interferencias, así como los problemas de atenuación de señales y de recepción de camino múltiple propio de las emisiones análogas. Al mismo tiempo, se hace un uso más eficaz del ancho de banda disponible. Los equipos digitales pueden llevar incorporados equipos de autodiagnóstico. El costo de mantenimiento se reduce.

3. PROCESAMIENTO DE SECUENCIAS DE IMÁGENES

El procesamiento de imágenes se logra a partir de la implementación del video digital, es decir que gracias a la utilización de los instrumentos digitales se pueden realizar cambios en las secuencias de imágenes, con el fin de mejorarlas o para obtener y aislar datos de interés de ellas. Dichos cambios son por lo tanto son muy importantes en el campo de la robótica y la visión artificial.

3.1. Fundamentos

Para explicar mejor la captación de una imagen la cual es dependiente de la luz, existe el Modelo de Imagen Simple. Éste se basa en tres definiciones importantes; estas son la **radiancia**, la **reflectancia** y la **irradiación**. El papel más importante, lo juega la **radiancia** de la escena que se denota como $I(x, y)$ y está dada por el producto de dos componentes, la cantidad de luz que incide en la escena que se está observando y la cantidad de luz que reflejan los objetos en la escena. Estos conceptos reciben el nombre de **irradiación** y **reflectancia** y se denotan $e(x, y)$ y $r(x, y)$ respectivamente, es decir $I(x, y) = e(x, y)r(x, y)$ donde $0 < e(x, y) < \infty$ y $0 < r(x, y) < 1$.

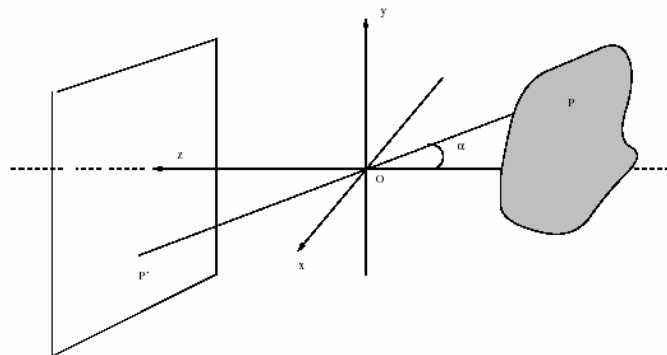
Las ecuaciones anteriores indican que la reflectancia está acotada por 0 (absorción total de luz) y 1 (reflexión total de luz). La naturaleza de $e(x, y)$ la determina la fuente de luz y la de $r(x, y)$ es determinada por las características del objeto.

La radiancia también está afectada por la distancia del punto de captación de la imagen (f) y por el diámetro del lente con que se capta la misma (d), y está dada por:

$$e = l \frac{\pi}{4} \left(\frac{d}{f}\right)^2 \cos^4 \alpha$$

donde α es el ángulo del rayo desde el objeto hasta el centro de la lente como se muestra en la Figura 26.

Figura 26. Angulo de la irradiación y reflectancia de la imagen



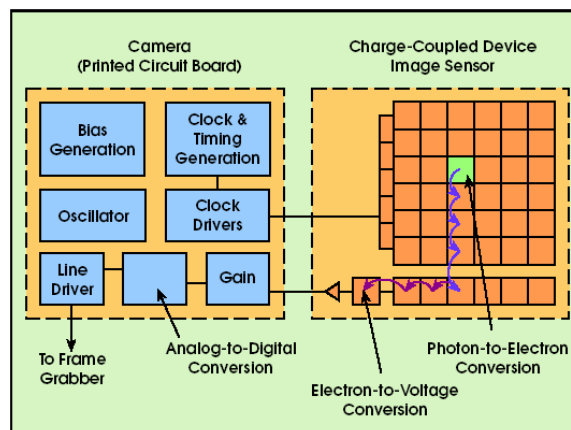
Fuente: Introducción al Procesamiento y Análisis de Imágenes Digitales

La captura de las imágenes se realiza en dos dimensiones, pero con cierto grado de distorsión, debido a que los sistemas de grabación no son perfectos e introducen provocadas por los efectos de muestreo y la cuantificación. El muestreo de una imagen tiene el efecto de reducir la resolución espacial de la misma, y el efecto de cuantificación viene dado por la imposibilidad de tener un rango infinito de valores de medida para la intensidad de brillo de los píxeles. La tecnología actual permite en algunos casos llegar hasta 10 bits, aunque lo más común es tener 8 bits que equivale a 256 niveles de gris para codificar este valor lumínico.

La imagen es captada por una cámara de video digital en una forma muy sencilla, que depende en su mayoría de un CCD que es como el corazón de la cámara. Un CCD es un circuito integrado cuyas siglas significan Dispositivo de Carga Acoplada (*Charge-Coupled Device*), que transforma la información óptica en eléctrica. La fotoelectricidad es la propiedad de la cual se valen los CCD para realizar la función transductora, es decir la transformación de un tipo de energía en otro, de la luz en electricidad.

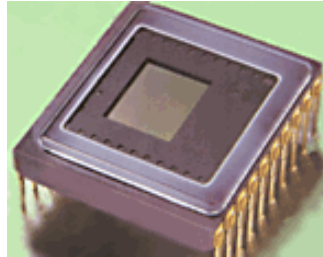
Un CCD capta la imagen por medio del *Target* que es una placa compuesta por sensores diminutos que son los encargados de convertir las imágenes en impulsos eléctricos. Cuando la luz reflejada por los objetos o por seres vivos pasa a través del sistema de lentes de la cámara y se enfoca sobre la placa, ésta origina una reacción eléctrica en función de la intensidad de la luz en cada punto (sensor), obteniendo corriente que a su vez es convertida a voltaje, luego ese voltaje es pasado por un amplificador y convertido de análogo a digital para luego ser enviado a un medio óptico o magnético para su almacenamiento. El diagrama del circuito principal de una cámara de video se muestra en la Figura 27 y físicamente un CCD se muestra en la figura 28.

Figura 27. Diagrama principal de una cámara digital



Fuente: CCD vrs. CMOS, DALSA Technology

Figura 28. CCD



Fuente: CCD vrs. CMOS, DALSA Technology

Cada uno de esos puntos de información digital, derivados de imágenes captadas por la cámara, es llamado Píxel. Los Píxeles constituyen la información fundamental para el procesamiento de imágenes o de secuencia de ellas.

3.1.1. Ruido en imágenes

La captura de imágenes por medio de cámaras de video ejecutada de tal forma que su diferencia con la realidad sea mínima, es una tarea difícil o casi imposible de realizar; siendo el responsable de este problema, el ruido.

El ruido es causado por los dispositivos que ejecutan la captura, procesamiento y almacenamiento de las imágenes y por la interacción de estos con el medio del cual se extraen; entre ellos distintos se pueden citar: el ruido térmico, el ruido de lectura, el ruido por transferencia de cuadros, el ruido por sobre exposición, el ruido por cuantización, etc. Por otra parte el movimiento brusco de los objetos o de la misma cámara, poca luminosidad, luz excesiva, humedad del ambiente y reflejos contribuyen, junto con los distintos tipos de ruido mencionados, a las degradaciones que hacen que la imagen tan sólo sea una aproximación de la escena real.

El ruido térmico es generado a causa del flujo de electrones de los semiconductores, en este caso del CCD. Para lograr eliminar este ruido es necesario enfriar el CCD.

El enfriamiento de este dispositivo puede llevarse a cabo por medios líquidos o por medios termoeléctricos, de los cuales los medios líquidos son más eficientes, aunque a un costo más alto, ya que se utilizan líquidos especiales a bajas temperaturas que necesitan un equipo de almacenamiento de alto costo; sin embargo, el proceso más utilizado es el termoeléctrico que utiliza el método de células de Peltier, junto con disipadores de calor y ventiladores. Estas células trabajan como una bomba de extracción de calor que al pasar una tensión continua a través de las superficies de contacto de dos sustancias semiconductoras distintas (una tipo P una tipo N) en un lado se genera calor y por el otro se absorbe tal y como se observa en la Figura 29.

Figura 29. Célula Peltier



Fuente: www.metamonds.net

El ruido de lectura es generado por los circuitos y amplificadores que son la etapa siguiente del CCD y el ruido de transferencia de cuadros se forma por la transferencia del área de almacenamiento de cada cuadro; estas dos formas de ruido pueden eliminarse por filtros eléctricos aunque no en su totalidad, por lo que es necesario hacer uso del procesamiento digital de imágenes.

El ruido por sobre exposición se da cuando la capacidad de almacenamiento de los píxeles se satura en un CCD por lo que los píxeles vecinos son invadidos generando una mancha blanca, generalmente en sentido vertical; a este efecto se le denomina **blooming**.

El ruido de Cuantización ya ha sido explicado en el capítulo 2 pero no está de más aclarar que también puede ser eliminado por filtros eléctricos pero de una forma más eficiente por medio de procesamiento digital de imágenes.

3.1.2. Relaciones básicas entre píxeles

En una imagen digitalizada, cada píxel de coordenadas (x, y) tiene cuatro vecinos horizontales y verticales vecinos cuyas coordenadas son

$$(x+1, y), (x-1, y), (x, y+1), (x, y-1) \quad (3.1.2.1.)$$

este conjunto de píxeles recibe el nombre de 4-vecinos de p y se notan $N_4(p)$. Cada píxel está a distancia unitaria de (x, y) . Los píxeles diagonales vecinos tienen coordenadas

$$(x+1, y+1), (x-1, y-1), (x-1, y+1), (x+1, y-1) \quad (3.1.2.2.)$$

y se notan $N_D(p)$. Estos puntos, junto con los 4-vecinos se llaman 8-vecinos de p y se notan $N_8(p)$.

La conectividad entre píxeles es un concepto importante usado para establecer las fronteras de objetos y las regiones componentes de una imagen.

Para establecer si dos píxeles están conectados se debe establecer si son adyacentes en algún sentido, por ejemplo si son 4-vecinos y sus niveles de gris cumplen algún criterio de similaridad.

Las operaciones Aritmético / Lógicas entre píxeles se usan muy frecuentemente en el procesamiento de imágenes. Las operaciones aritméticas son muy utilizadas, tanto en el cálculo de funciones estadísticas, como en el cálculo de rutinas con matrices. Las operaciones fundamentales entre dos píxeles p y q , se definen así:

- a) Suma $p + q$: esta operación da como resultado un aumento de la brillantez del píxel p en un factor q .
- b) Diferencia $p - q$: este tipo de función, disminuye la brillantez del píxel p en un factor q .
- c) Multiplicación $p * q$ o pq o $p \times q$: comprende el escalonamiento del píxel p en si mismo, q veces.
- d) División $p \div q$: es la operación contraria a la multiplicación e implica disminuir el píxel p en q veces.

Las principales operaciones lógicas son:

- a) AND $p \text{ AND } q$: de esta operación se obtiene como resultado un píxel que contiene los datos p y que al mismo tiempo están contenidos en el píxel q y es importante en la búsqueda de bordes en imágenes.

- b) OR $pORq$: el resultado de esta operación es un píxel con los datos que están contenidos en p y no en q y viceversa y se utiliza en el mejoramiento de imágenes.
- c) COMPLEMENTO $NOTp$ o $NOTq$: esta operación es importante en el proceso de morfología, utilizando la parte que complementa a cada píxel.

Es importante notar que las operaciones aritméticas se aplican a imágenes multivaluadas, es decir a imágenes cuyos píxeles pueden tomar diferentes valores, mientras que las lógicas se refieren a imágenes bivaluadas tomando los valores 0 ó 1.

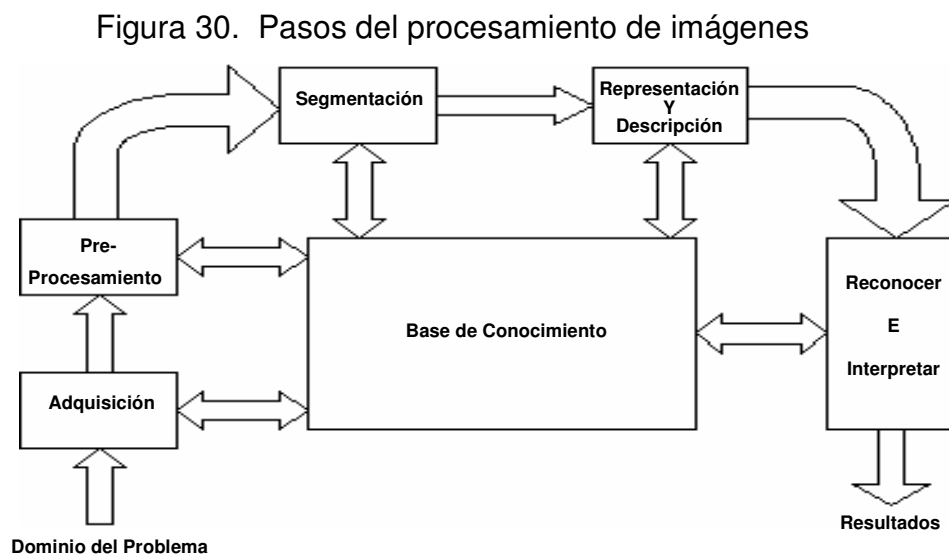
3.1.3. Pasos para el procesamiento de imágenes

El primer paso en el proceso es adquirir la imagen digital. Para ello se necesitan sensores y la capacidad para digitalizar la señal producida por el sensor, el cual puede ser una cámara color o monocromática que produce una imagen completa cada 1/30 segundos. Si la salida de la cámara no está en formato digital, es necesario usar un conversor analógico digital en la salida de la misma.

Una vez que la imagen digitalizada ha sido obtenida, el segundo paso consiste en el preprocesamiento de dicha imagen. El objetivo de este paso es mejorar la imagen en forma que nuestro objetivo final tenga mayores posibilidades de éxito, es necesario mejorar el contraste, suprimir el ruido y aislar regiones importantes.

El tercer paso es la segmentación, cuyo objetivo es dividir la imagen en las partes que la constituyen o en los objetos que la forman. Este es uno de los problemas más difíciles en el procesamiento de la imagen, pero por otro lado una buena segmentación facilitará mucho cualquier procesamiento que se desee aplicar. La salida del proceso de segmentación es una imagen de datos que contiene la frontera de la región o los puntos de ella misma. Luego se necesita especificar un método que extraiga los datos de interés o rasgos que producen alguna información cuantitativa y que pueden ser básicos para diferenciar una clase de objetos de otra.

El último paso es el reconocimiento y la interpretación. El reconocimiento es el proceso que asigna una etiqueta a un objeto basada en la información que proporcionan los descriptores. La interpretación lleva a asignar significado al conjunto de objetos reconocidos. La necesidad de conocimiento a priori también es parte de este último paso, que no es más que la interacción entre el reconocimiento y una base de datos. El esquema del proceso completo se muestra en la Figura 30.



Fuente: Introducción al Procesamiento de Imágenes

3.2. Mejoramiento de Imágenes

El principal objetivo de las técnicas de mejora de imágenes es procesar una imagen dada de forma que la imagen resultante sea más apropiada que la imagen original para unas aplicaciones específicas.

3.2.1. Mejora mediante procesamiento puntual

Podría decirse que este tipo de transformaciones son operaciones de memoria cero, es decir que no necesitan más que transformarse de un valor a otro; entre ellas están: ampliación de contraste, corte y umbralización, negativo digital, extracción de bits, compresión del rango, potenciación, diferencia y promedio de imágenes y modificación del histograma. Las imágenes con poco contraste aparecen a menudo debido a iluminación pobre o no uniforme o bien a la no linealidad o pequeño rango dinámico de los sensores de las imágenes, por lo que es necesario aplicarles una **ampliación de contraste**, que se logra ampliando los intervalos de nivel de gris en las regiones necesitadas.

La aplicación de **corte**, equivale a llevar la imagen completa de una escala de grises a otro y la **umbralización** corresponde a llevar la imagen completa a los umbrales del nivel de gris, lo que sería convertirla a Bitmap. Las imágenes médicas, o las placas de una imprenta, necesitan negativos de éstas para llevar a cabo su cometido, por lo que el **negativo digital** se debe aplicar. Éste consiste en transformar la imagen restando cada uno de sus niveles de gris del máximo valor. Para determinar el número de bits que son visualmente significativos es necesario utilizar la **extracción de bits**, que se logra poniendo una restricción para que se extraiga únicamente el enésimo bit comparando la restricción con cada uno de los bits de la imagen.

La **compresión del rango** se utiliza cuando los niveles de gris de una imagen suele ser muy grande, por ejemplo las imágenes astronómicas, por lo que es necesario hacer una transformación mediante la ecuación

$$v = c \log_{10}(1 + u) \quad (3.2.1.1)$$

donde u es la señal original, v es la señal resultante y c es una constante para la escala. Este tipo de transformaciones realiza las pequeñas magnitudes de los píxeles comparada con las grandes magnitudes.

Otro tipo de transformación que puede ser útil y que en algunos casos corresponde la respuesta de determinados medios de captación de imágenes es la transformación llamada **potenciación** que es de la forma

$$v = c^1 (u/c)^\lambda \quad (3.2.1.2)$$

donde c^1 es una constante de escalado mientras que c lleva el rango inicial u al intervalo $[0,1]$ y $\lambda \in [0, \infty]$. Si $\lambda < 1$ amplía los niveles de gris bajos, mientras que si $\lambda > 1$ el efecto es el contrario.

La **diferencia de imágenes** y el **promedio de imágenes** sólo puede suceder entre dos o más imágenes, la primera tiene aplicaciones en segmentación y realce; la segunda se utiliza para obtener una versión de la imagen original con menos ruido, utilizando distintas tomas de la misma escena, basándose en la disminución de la desviación típica del ruido.

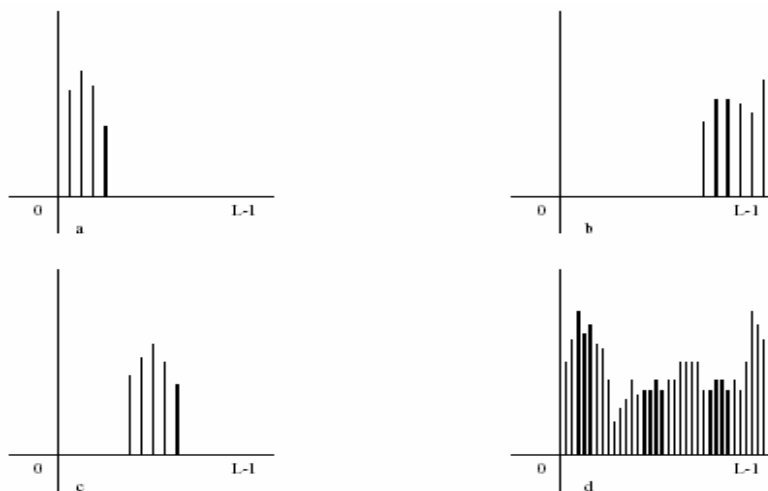
La **modificación del histograma** puede ser de tres formas: uniformándolo, especificándolo o haciendo mejoras locales.

El histograma de una imagen digital con niveles de gris en el rango $[0, L - 1]$ es una función discreta dada por

$$p(r_k) = \frac{n_k}{n} \quad (3.2.1.3)$$

donde r_k es el k -ésimo nivel de gris, n_k es el número de píxeles en la imagen con ese nivel de gris, n es el número total de píxeles en la imagen y $k = 0, 1, 2, \dots, L - 1$. Entonces podría decirse que la función $p(r_k)$ proporciona una estimación de la probabilidad de ocurrencia del nivel de gris r_k . En la Figura 30 se pueden observar ejemplos de histograma.

Figura 31. Ejemplos de histograma



Fuente: Introducción al Procesamiento de Imágenes

El histograma mostrado en la Figura 30(a) muestra que los niveles de gris están concentrados en la zona oscura de la escala, ocurriendo lo contrario en la Figura 30(b); en la Figura 30(c) se muestra un histograma con poco rango dinámico y por tanto de bajo contraste, siendo la Figura 30(d) lo contrario, teniendo un alto contraste.

3.2.2. Filtrado espacial

El uso de plantillas (máscaras) espaciales para el procesamiento de imágenes recibe usualmente el nombre de **filtrado espacial** y las máscaras reciben el nombre de filtros espaciales. Existen dos tipos de filtros espaciales, los lineales y los no lineales para mejorar una imagen; los filtros lineales están basados en transformaciones de la forma

$$g(i, j) = \sum_{m,n \in O} h(i-m, j-n) f(m, n) \quad (3.2.2.1.)$$

donde $f(i, j)$ es la imagen de entrada, $g(i, j)$ es la imagen de salida y $h(i, j)$ es el llamado núcleo de convolución. Una matriz O rectangular es lo usual, con número impar de filas y columnas, asignándole valores también al píxel central, su tamaño puede variar sustancialmente en función del tamaño de los objetos de la imagen.

Los llamados filtros de paso bajo atenúan o eliminan componentes de alta frecuencia en el dominio de Fourier mientras que dejan las frecuencias bajas sin alterarlas. Las frecuencias altas caracterizan las fronteras y detalles con alta variación en la imagen. El efecto contrario se obtiene con los filtros de paso alto. Los filtros no lineales también operan sobre los vecinos, pero en general sus operaciones están basadas directamente en los valores de los píxeles que se considera y no usan coeficientes.

Los filtros de alisamiento se usan con el objetivo reducir el ruido, suprimiendo detalles pequeños antes de la extracción de objetos usualmente grandes, también puede usarse para unir segmentos.

Es usual utilizar filtros de tamaño 3x3 o bien máscaras con forma de estrella siendo muy importante que la suma de los coeficientes sea la unidad. A veces se aumenta la fuerza del píxel central y la de aquellos que están más próximos, como por ejemplo

$$\begin{array}{ccc}
 \begin{array}{ccc} 1 & 1 & 1 \\ 1/9 * 1 & 1 & 1 \\ 1 & 1 & 1 \end{array} &
 \begin{array}{ccc} 1 & 1 & 1 \\ 1/10 * 1 & 2 & 1 \\ 1 & 1 & 1 \end{array} &
 \begin{array}{ccc} 1 & 2 & 1 \\ 1/16 * 2 & 4 & 2 \\ 1 & 2 & 1 \end{array}
 \end{array}$$

El alisamiento también se puede hacer por medianas, donde la mediana es el valor central de un conjunto de valores ordenados. El filtrado por mediana reduce el emborronamiento de las fronteras, es decir que reduce las indefiniciones en los bordes de las imágenes causadas por movimiento de la cámara o del motivo. La idea es reemplazar el valor de la imagen en un punto por la mediana de los niveles de brillo de sus vecinos. Uno de sus mayores problemas es que si se utiliza una matriz rectangular se dañan las líneas finas y los bordes de los objetos.

Para destacar los detalles finos de una imagen o mejorar detalles que han sido emborronados por errores o por problemas con la adquisición de la imagen se utilizan los **filtros de realce**. Este tipo de filtro se puede obtener si se colocan coeficientes positivos en el centro y negativos en la periferia de la matriz, siendo importante que la suma de estos sea cero.

3.2.3. Mejora en el dominio de las frecuencias

El fundamento de las técnicas basadas en el dominio de las frecuencias es el teorema de convolución que dice que si $g(x, y)$ es la resultante de la convolución dada por:

$$g(x, y) = (h * f)(x, y) \quad (3.2.3.1.)$$

donde $f(x, y)$ es la imagen y $h(x, y)$ es un operador lineal, entonces se verifica la siguiente relación en el dominio de las frecuencias

$$G(u, v) = H(u, v)F(u, v) \quad (3.2.3.2.)$$

donde G , H y F son las transformadas de Fourier de g , h y f respectivamente. Quiere decir entonces que los problemas de mejora de imágenes pueden ser formulados en el sentido de buscar un operador que al ser multiplicado con la imagen original en el dominio de la frecuencia, el resultado sea la mejora deseada; por ejemplo, los bordes de la imagen $f(x, y)$ pueden ser acentuados usando una función que enfatice las frecuencias altas de la transformada $F(u, v)$.

En la práctica, las máscaras espaciales pequeñas se usan más que la transformada de Fourier por su simplicidad de implementación y velocidad en las operaciones; sin embargo, la comprensión de los conceptos en el dominio de las frecuencias es esencial para la solución de problemas no fácilmente comprensibles en el dominio espacial.

El ruido en los niveles de gris de una imagen contribuyen significativamente al contenido de las frecuencias altas en el dominio de Fourier; por lo tanto, el emborronamiento se alcanza en el dominio de Fourier atenuando un rango especificado de componentes de altas frecuencias en la transformada de una imagen dada, por lo tanto es necesario utilizar un **filtro de paso bajo**.

Un filtro de paso bajo posee un punto de transición o un límite en el cual no deja pasar cierta frecuencia, la cual es llamada frecuencia de corte; y esta es el parámetro que se debe definir para conseguir el emborronamiento deseado en la imagen. Si por el contrario se tiene una imagen emborronada que atenúa las componentes de frecuencia alta en su transformada de Fourier, entonces el realce de imágenes puede alcanzarse mediante un **filtro paso alto** que atenúe las frecuencias bajas sin distorsionar la información en las frecuencias altas en el dominio de Fourier.

3.3. Segmentación de imágenes

El primer paso para hacer análisis de imágenes es segmentar la imagen en regiones; una región es un grupo de píxeles conectados que tienen propiedades similares. Es obvio que las regiones son importantes para la interpretación de las imágenes, pues pueden corresponder a objetos en la escena. Una imagen puede contener varios objetos y además cada objeto puede contener varias regiones que corresponden a partes del mismo.

Para que una imagen pueda ser interpretada correctamente, tiene que ser dividida en regiones que correspondan a objetos o partes de ellos. Sin embargo, debido a los errores de segmentación, la correspondencia entre regiones y objetos no será perfecta y se debe utilizar conocimiento específico en etapas posteriores de la interpretación de las imágenes.

3.3.1. Regiones y bordes

El primer paso en el análisis para entender una imagen es particionar la imagen de forma que las regiones que representan diferentes objetos son marcadas explícitamente.

Estas particiones pueden ser obtenidas a partir de características de los niveles de gris en la imagen, o utilizando alguna otra. Existen dos aproximaciones para particionar una imagen en regiones: segmentación basada en regiones y estimación de la frontera utilizando detectores de bordes.

El sistema de aproximación basada en regiones, coloca todos los píxeles que corresponden a un objeto, los agrupa y los marca para indicar que pertenecen a una región. Este proceso también recibe el nombre de segmentación.

Dos principios muy importantes en la segmentación son el valor de similitud y la proximidad espacial. Dos píxeles pueden ser asignados a la misma región si tienen características similares de intensidad y si están próximos. La varianza de los niveles de gris en una región y la compactación de una región pueden ser también usadas como valores de similitud y proximidad respectivamente. Los principios de similitud y proximidad aparecen por la hipótesis de que los puntos de un mismo objeto producirán en la imagen píxeles con niveles de gris similares y próximos espacialmente, aunque esto no es siempre cierto, cumple en gran parte con las expectativas.

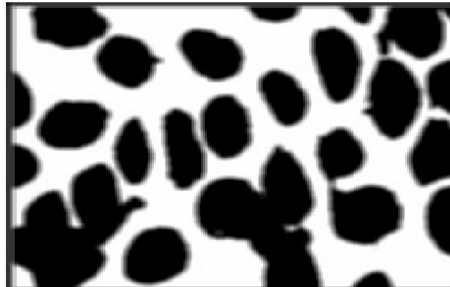
La segmentación también puede realizarse mediante el proceso denominado **disparidad**, el cual se basa en encontrar los píxeles que se encuentran en las fronteras de las regiones; estos píxeles llamados bordes, pueden localizarse examinando los píxeles vecinos, midiendo las diferencias de la intensidad para su detección así como la textura y el movimiento. En una imagen ideal, las fronteras (cerradas) podrían obtenerse entonces a partir de las regiones utilizando algoritmos de seguimiento de fronteras y análogamente, las regiones pueden obtenerse a partir de bordes usando un algoritmo de relleno. Por desgracia, en las imágenes reales es raro obtener esta reciprocidad.

3.3.2. Segmentación basada en umbralización

Si se tiene una imagen como en la Figura 31, que se compone de objetos oscuros sobre fondo luminoso de tal modo que los niveles de gris se ubican en dos modas dominantes; existe una forma obvia de extraer los objetos del fondo, seleccionar un umbral T que separe dichas modas.

$$g(i, j) = \begin{cases} 1 & \text{si } f(i, j) \geq T \\ 0 & \text{si } f(i, j) < T \end{cases} \quad (3.3.2.1.)$$

Figura 32. Imagen con dos modas dominantes



Fuente: Introducción al Procesamiento de Imágenes

También pueden ser utilizados umbrales múltiples pero este tipo de umbralización es generalmente menos fiable que la umbralización en dos regiones, porque es difícil establecer múltiples umbrales que aíslen regiones correctamente, especialmente cuando hay muchas modas.

Si la iluminación de la escena es desigual, el método simple de modas no es apropiado. La iluminación desigual puede ser debida a sombras o la propia dirección de la iluminación. En estos casos, el mismo umbral no puede ser usado en toda la imagen, por lo que se utilizan los métodos **adaptativos**, en los que se examinan regiones pequeñas de la imagen para obtener un umbral de sólo esa región; partiendo la imagen en una matriz de $m \times m$ subimágenes.

Uno de los aspectos más importantes de la selección del umbral es su capacidad para identificar de una forma consistente los picos del histograma. Esta característica es particularmente importante para la umbralización en las situaciones en las que las características de la imagen pueden cambiar sobre un amplio rango de intensidades.

Es claro que las posibilidades de seleccionar un buen umbral aumentan si los picos del histograma son altos, estrechos, simétricos y separados por valles profundos.

3.3.3. Segmentación basada en bordes

Estas técnicas de segmentación se basan en los bordes encontrados en la imagen por detectores especializados. El objetivo final es alcanzar por lo menos una segmentación parcial; es decir, agrupar cadenas de bordes en una imagen, que tengan correspondencia con objetos de la misma.

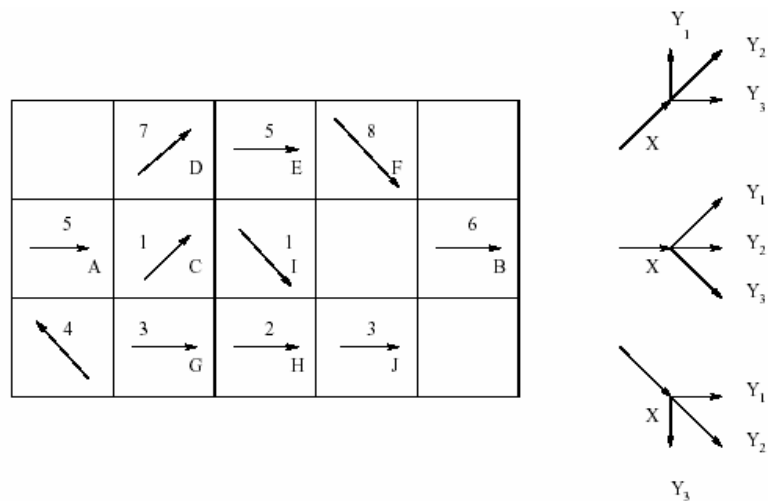
Existen diferentes métodos para la segmentación basada en bordes que difieren en la estrategia para la construcción de la frontera, así como en la cantidad de información previa que se incorpora en el método. Entre las más sencillas están la **Extracción de Frontera**, **Transformada de Hough** y **Métodos de Gradiente**.

Las fronteras son bordes unidos que caracterizan la forma de un objeto. Son, por tanto, útiles para calcular rasgos geométricos como tamaño y orientación. Conceptualmente, las fronteras pueden encontrarse trazando los bordes conectados, sin embargo, hay dificultades asociadas con estas definiciones de conectividad. Un algoritmo simple para trazar fronteras cerradas en imágenes binarias es el siguiente:

- a) Empezar dentro de una región A, (por ejemplo, el primer píxel que se encuentre de la región cuando se hace un rastreo por filas).
- b) Girar a la izquierda y pasar al píxel siguiente, si está dentro de la región A, en caso contrario girar a la derecha y pasar al píxel siguiente.
- c) Continuar hasta que se llegue al punto del que se partió.

La Figura 32 muestra un ejemplo de cómo se sigue un contorno. El contorno puede definirse como los puntos por donde se entra o sale de la región.

Figura 33. Seguimiento de un contorno



Fuente: Introducción al Procesamiento de Imágenes

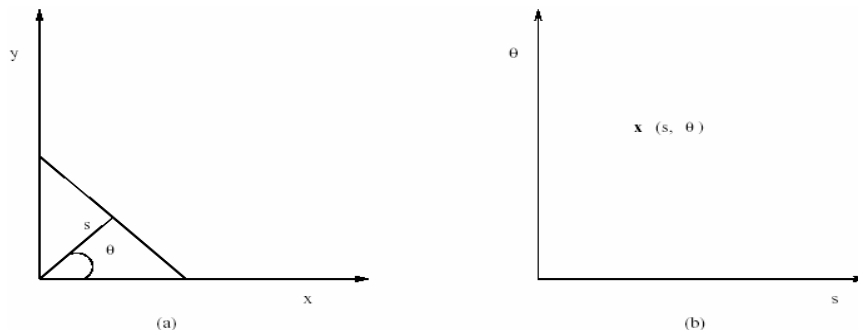
Una frontera puede ser también vista como un camino que se forma a través de un grafo construido uniendo los elementos bordes o es una discontinuidad de algún tipo en la función de intensidad de la imagen.

En la transformada de Hough se considera la unión de dos puntos examinando si están o no en una curva de una forma específica. Este método es obviamente una forma de considerar relaciones globales entre píxeles; ya que se consideran los n puntos de una imagen en la que se quieren encontrar subconjuntos de ellos que estén sobre una línea recta a distancia s y con orientación θ puede representarse mediante

$$s = x \cos \theta + y \sin \theta \quad (3.3.3.1)$$

La transformada de Hough de esta línea es un punto en el plano (s, θ) , esto es, todos los puntos de esta línea se aplican en un único punto (ver Figura 33). Este hecho puede usarse para detectar líneas rectas dado un conjunto de puntos frontera.

Figura 34. Transformada de Hough. (a) Línea recta y (b) su transformada



Fuente: Introducción al Procesamiento de Imágenes

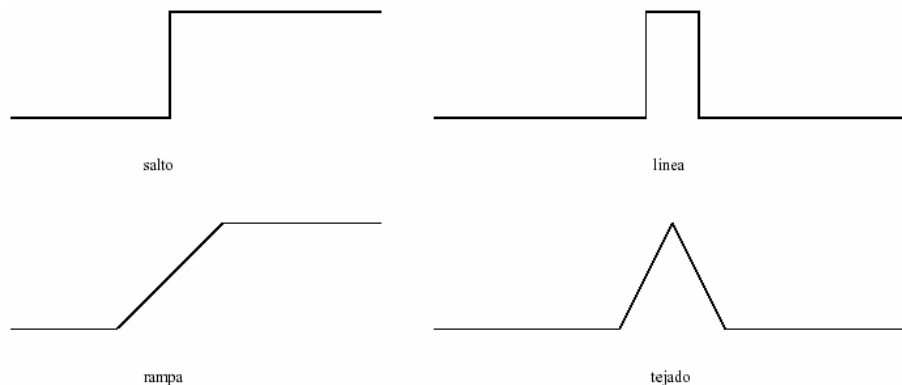
Si se supone que los puntos frontera $(x_i, y_i), i = 1, 2, \dots, n$ se aplican a algunos valores cuantificados de los parámetros s y θ en el espacio (s, θ) y se cuenta $C(s, \theta)$, el número de puntos frontera que se aplican en una determinada localización (s, θ) , es decir

$$C(s_k, \theta_l) = C(s_k, \theta_l) + 1 \text{ si } s_k = x_i \cos \theta_l + y_i \sin \theta_l \quad (3.3.3.2.)$$

Entonces los máximos locales de $C(s, \theta)$ proporcionan los diferentes segmentos de líneas rectas que pasan a través de un borde dado. Existen muchos tipos de bordes diferentes, algunos son debidos a las sombras que se producen sobre los objetos, otros a la variación de la reflectancia de los objetos o incluso por la textura de los objetos, por lo que en su mayoría resulta difícil encontrarlos, teniendo que utilizar métodos más complejos para lograr ese objetivo.

Un método más complejo pero a su vez más efectivo es el de Gradientes, en el cual se definen los bordes como un cambio local “significativo” en la intensidad de la imagen; dichos cambios los agrupa en dos tipos: de **salto**, donde la intensidad cambia abruptamente de una lado a otro valor diferente en el lado opuesto y de **línea**, donde las intensidades cambian abruptamente volviendo en poco espacio a su valor inicial. Los saltos y pasos son raros en las imágenes, debido en gran parte al proceso de digitalización, y lo que se observa son rampas y tejados como se observa en la Figura 34.

Figura 35. Perfiles de bordes.



Fuente: Introducción al Procesamiento de Imágenes

En una dimensión, los cambios de salto, están asociados con un pico local de la primera derivada. Un operador isotrópico es un detector de bordes que tiene como objetivo determinar la fuerza de un borde punto por punto con independencia de su dirección en la imagen. Si se considera una imagen de intensidades $f(x, y)$ con coordenadas x e y y se rota en θ produciendo (x^1, y^1) según la fórmula

$$\begin{aligned} x^1 &= x \cos \theta + y \sin \theta \\ y^1 &= -x \sin \theta + y \cos \theta \end{aligned} \quad (3.3.3.)$$

las derivadas de los dos sistemas se relacionan mediante

$$\begin{aligned} \frac{\partial f}{\partial x^1} &= \frac{\partial f}{\partial x} \cos \theta + \frac{\partial f}{\partial y} \sin \theta \\ \frac{\partial f}{\partial y^1} &= -\frac{\partial f}{\partial x} \sin \theta + \frac{\partial f}{\partial y} \cos \theta \end{aligned} \quad (3.3.3.4.)$$

Está comprobado que un operador isotrópico es la suma de los cuadrados de las derivadas parciales en las direcciones x e y , ya que

$$\left(\frac{\partial f}{\partial x^1}\right)^2 + \left(\frac{\partial f}{\partial y^1}\right)^2 = \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 \quad (3.3.3.5.)$$

Para obtener la dirección en la que se produce el cambio se debe obtener el θ para el que la derivada parcial se maximiza como sigue

$$d\left(\frac{\partial f}{\partial x^1}\right) / d\theta = 0 \quad (3.3.3.6.)$$

de donde

$$\theta = \arctan\left(\frac{\frac{\partial f}{\partial y}}{\frac{\partial f}{\partial x}}\right) \quad (3.3.3.7.)$$

sustituyendo este θ en la fórmula de la derivada parcial se tiene que el máximo cambio vale

$$M = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \quad (3.3.3.8.)$$

Las expresiones anteriores tienen operaciones de raíces cuadradas que podrían ser computacionalmente lentas, por lo que se utilizan aproximaciones que hacen la tarea más rápida. Lo anterior se logra convolucionando la imagen con los operadores de **Roberts**, **Sobel** y el **Laplaciano**.

El operador de Roberts es uno de los más antiguos y es muy fácil de calcular, ya que usa sólo una ventana de tamaño 2x2. Sus máscaras de convolución son

$$G_x = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad G_y = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (3.3.3.9.)$$

la magnitud del gradiente se calcula entonces mediante la siguiente aproximación:

$$G(f(i, j)) = |f(i, j) - f(i+1, j+1)| + |f(i, j+1) - f(i+1, j)| \quad (3.3.3.10.)$$

Una gran desventaja del operador de Roberts es su sensibilidad al ruido, ya que se usan muy pocos píxeles para aproximar el gradiente.

El operador de Sobel usa máscaras de tamaño 3x3, evitando que el gradiente se calcule en un punto intermedio como ocurre con el operador de Roberts. El operador de Sobel es la magnitud del gradiente calculado mediante:

$$M = \sqrt{s_x^2 + s_y^2} \quad (3.3.3.11.)$$

donde las derivadas parciales se calculan mediante

$$\begin{aligned} s_x &= (a_2 + ca_3 + a_4) - (a_0 + ca_7 + a_6) \\ s_y &= (a_0 + ca_1 + a_2) - (a_6 + ca_5 + a_4) \end{aligned} \quad (3.3.3.12.)$$

Es importante notar que este operador le da más peso a los píxeles más cercanos al centro de la máscara. La máscara que utiliza está dada por

$$\begin{bmatrix} a_0 & a_1 & a_2 \\ a_7 & (i, j) & a_3 \\ a_6 & a_5 & a_4 \end{bmatrix} \quad (3.3.3.13.)$$

En los operadores anteriores calculan el gradiente y si su valor está por encima de un umbral, se supone la existencia de un punto de borde, pero como resultado se obtienen muchos puntos. Una mejor aproximación sería encontrar sólo los puntos que tienen máximos locales en el gradiente y considerarlos como los puntos de borde. Esto implica utilizar los cruces por cero de la segunda derivada en las intensidades de la imagen. El operador Laplaciano es el equivalente bidimensional de la segunda derivada y está dado por

$$\Delta^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (3.3.3.14.)$$

$$\frac{\partial^2 f}{\partial x^2} = f(i+2, j) - 2f(i+1, j) + f(i, j) \quad (3.3.3.15.)$$

$$\frac{\partial^2 f}{\partial y^2} = f(i, j+2) - 2f(i, j+1) + f(i, j) \quad (3.3.3.16.)$$

de donde resulta la siguiente máscara

$$\Delta^2 \approx \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (3.3.3.17.)$$

El operador laplaciano marca la presencia de un borde cuando la salida del operador realiza una transición por cero.

3.4. Morfología matemática

La palabra morfología denota la rama de la Biología que se dedica al estudio de la forma y estructura de los animales y plantas. La morfología matemática se basa en la geometría y la forma para simplificar imágenes y conservar las principales características de formas de los objetos.

Un sistema de operadores como los de la morfología matemática es útil porque pueden formarse composiciones de sus operadores, que cuando actúan sobre formas complejas, son capaces de descomponerlas en partes que tienen sentido y separarlas de las partes que le son extrañas. También permite que las formas subyacentes sean identificadas y reconstruidas de forma óptima a partir de sus formas distorsionadas y ruidosas. La morfología matemática se puede usar, entre otros, con los siguientes objetivos:

- a) Preprocesamiento de imágenes (supresión de ruido, simplificación de formas).
- b) Destacar la estructura de los objetos (extraer el esqueleto, marcado de objetos, envolvente convexa, ampliación, reducción).
- c) Descripción cualitativa de objetos (área, perímetro, etc).

3.4.1. Operaciones básicas sobre conjuntos

El lenguaje de la morfología matemática binaria es el de la teoría de conjuntos. Sean A y B conjuntos en un n -espacio E^n con elementos $a = (a_1, \dots, a_n)$ y $b = (b_1, \dots, b_n)$ respectivamente.

- a) La traslación de A por $x \in E^n$ que se nota A_x se define como $A_x = \{c/c = a + x, \text{ para } a \in A\}$.
- b) La reflexión de B notada \hat{B} se define como $\hat{B} = \{x/x = -b, \text{ para } b \in B\}$.
- c) La diferencia de dos conjuntos A y B , notada $A - B$, se define mediante $A - B = \{x/x \in A, x \notin B\}$.

3.4.2. Operaciones morfológicas

Entre las operaciones morfológicas se pueden encontrar: **Dilatación**, **Erosión**, **Apertura** y **Clausura**.

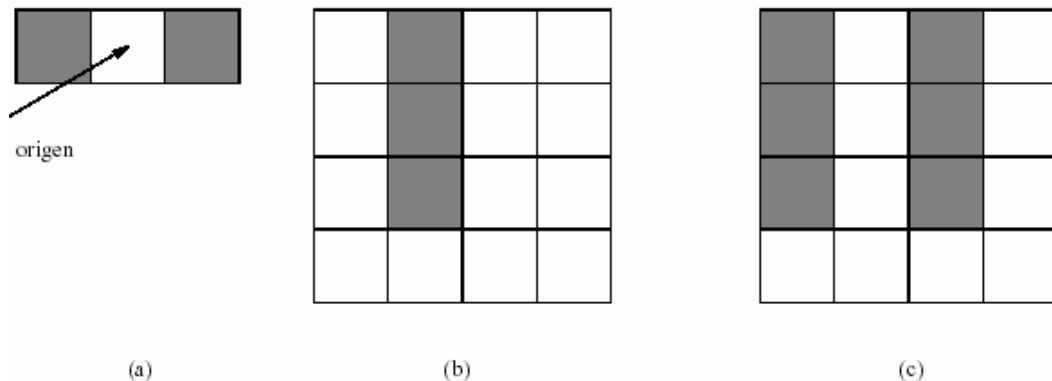
La Dilatación es la transformación morfológica que combina dos vectores utilizando la suma, por lo tanto es el conjunto de todos los posibles vectores que son suma de pares de elementos, uno de A y otro de B . Formalmente la dilatación de A y B , se nota $A \oplus B$ y se define mediante

$$A \oplus B = \{c \in E^n / c = a + b \text{ para } a \in A \text{ y } b \in B\} \quad (3.4.2.1.)$$

El primer elemento de la dilatación A , está asociado con la imagen que se está procesando y el segundo B recibe el nombre de elemento estructural, por la forma que actúa sobre A .

Cuando se realiza una dilatación con un elemento estructural que contiene cero, lo que resulta es una expansión de la imagen.

Figura 36. Dilatación. (a) elemento estructural, (b) imagen, (c) resultado



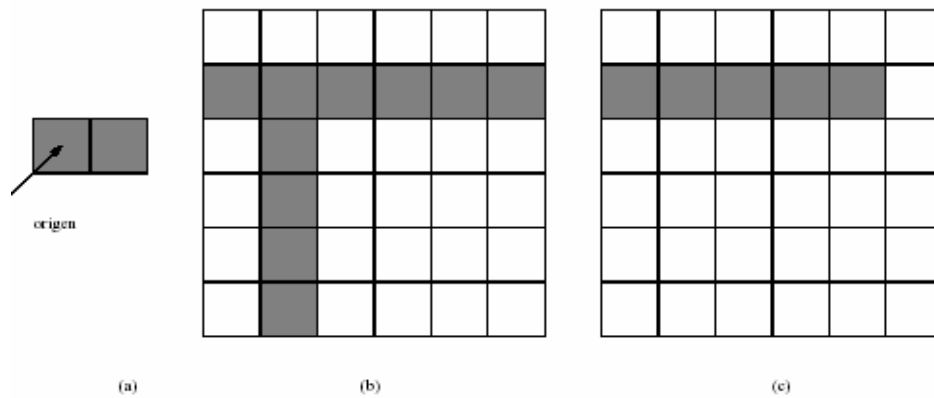
Fuente: Introducción al Procesamiento de Imágenes

La erosión es la operación morfológica que combina dos conjuntos usando el concepto de inclusión. Si A y B son conjuntos en el espacio, entonces la erosión de A y B es el conjunto de todos los elementos x para los que $x + b \in A$ para todo $b \in B$. La erosión de A por B se nota por:

$$A \ominus B = \{x \in E^n / x + b \in A \text{ para } b \in B\} \quad (3.4.2.2.)$$

La utilidad de la erosión puede apreciarse mejor cuando ésta se expresa de forma diferente. La erosión de una imagen A por un elemento estructural B , es el conjunto de todos los elementos $x \in E^n$ para los cuales B trasladado por x está contenido en A .

Figura 37. Erosión. (a) elemento estructural, (b) imagen, (c) resultado



Fuente: Introducción al Procesamiento de Imágenes

La erosión se concibe usualmente como una reducción de la imagen original, prácticamente es la operación dual de la dilatación, porque lo que la dilatación hace al objeto, la erosión lo hace al fondo.

La Apertura y la Clausura tienen su origen en la dilatación y la erosión, puesto que son combinaciones de ellas y se definen respectivamente como

$$A \circ B = (A \ominus B) \oplus B \quad (3.4.2.3.)$$

$$A \bullet B = (A \oplus B) \ominus B \quad (3.4.2.4.)$$

La apertura se utiliza para seleccionar los puntos de **A** que se acoplen a **B**, se pueden suavizar contornos, romper uniones estrechas y eliminar salientes estrechos. La clausura tiende a suavizar las secciones de contornos pero en sentido inverso, uniendo separaciones estrechas, eliminando huecos y pequeñas discontinuidades en las fronteras.

4. DETERMINACIÓN DE MOVIMIENTO

La percepción visual del movimiento es una función esencial del sistema visual humano. Su importancia es central en la estimación del movimiento relativo entre objetos, en la estimación del movimiento propio, en los procesos de orientación y en la discriminación entre objetos espacialmente separados.

En lo que concierne a la interpretación pasiva de escenas, el análisis visual del movimiento es aplicado en el contexto de la reconstrucción de escenas; puede ser utilizado en tareas de segmentación o diferenciación entre los objetos frontales y los de fondo o en el análisis de la forma de los mismos. De igual manera el tipo de movimiento de un objeto puede indicar su carácter rígido o flexible.

En lo que concierne a la interpretación activa de escenas experimentada por un ojo en movimiento, la percepción visual del movimiento puede proporcionar información sobre la distancia y velocidad relativa de los objetos presentes en el campo visual, permitiendo la predecir y esquivar los posibles obstáculos.

Existen numerosos estudios psicológicos y psicofísicos donde se reconoce la importancia de la percepción visual en las tareas mencionadas pero a pesar de ellos no existe una constancia explícita de exactamente cual es el método de estimación visual del movimiento utilizado por los humanos y otros animales. No obstante, estos estudios están siendo de gran relevancia en la creación de modelos visuales cada vez más completos y que al menos revelan y gozan propiedades y comportamientos próximos a los de un sistema visual real.

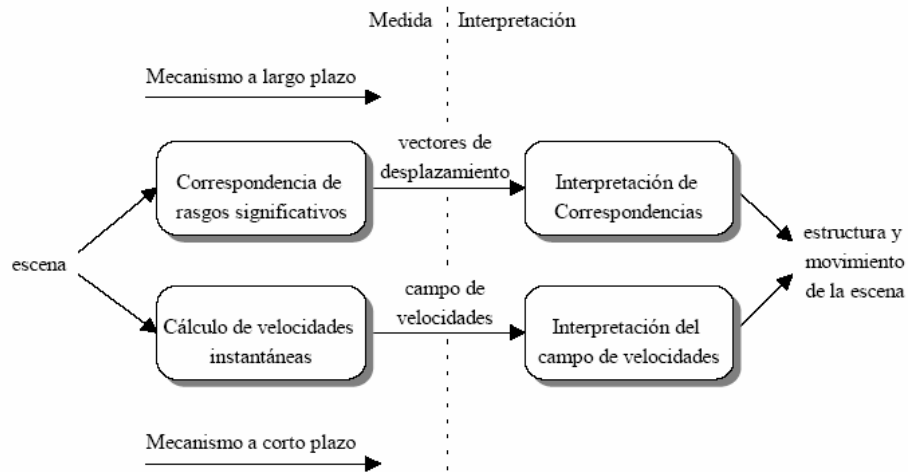
Una vez establecido un modelo más o menos acertado del sistema de percepción visual del movimiento usado por los humanos, se planea la posibilidad de crear un sistema artificial semejante, sino en estructura, al menos en comportamiento y funcionalidad. Añadida a la complejidad inherente a la percepción del movimiento, se unen restricciones tecnológicas que limitan la cantidad de procesamiento que se puede efectuar sobre las imágenes, teniendo esto, consecuencias directas sobre la resolución y las medidas de movimiento que se involucran en los cálculos a efectuar.

En el proceso de detección existe la posibilidad de que los animales avanzados puedan utilizar dos mecanismos de medida del movimiento bidimensional:

1. Un mecanismo a largo plazo donde se relacionan objetos o elementos significativos como esquinas o contornos, comúnmente llamados rasgos, detectados en intervalos de tiempo relativamente grande y asociando un movimiento bidimensional a los grupos de rasgos en instantes diferentes.
2. Un mecanismo a corto plazo donde se efectúa una estimación de la velocidad instantánea de un conjunto de puntos pertenecientes al plano de la imagen para formar un campo de movimiento sobre el mismo.

Se podría pensar que ambos procesos son concurrentes y compatibles, pero en realidad existen posturas que defienden uno frente al otro, refiriéndose al primero como el método de **vectores de desplazamiento** y al segundo como el método de **flujo óptico**.

Figura 38. Mecanismos de medida del movimiento



Fuente: Visión artificial dinámica

El método de Vectores de Desplazamiento utiliza la segmentación de imágenes como una herramienta para la detección del movimiento, buscando puntos de correspondencia entre los segmentos encontrados entre dos cuadros tomados en diferente intervalo de tiempo sobre la misma imagen; lo cual da como resultado una difícil tarea para a nivel de programación, ya que los tiempos manejados dentro de un programa que realice esta labor, resultarían relativamente grandes con respecto al tiempo de captura de cada cuadro de la secuencia de imágenes, pues se debe incluir un conocimiento del entorno y se debe manejar cierta información global, volviendo la detección de movimiento un trabajo inexacto y poco confiable.

Por otro lado el método de Flujo óptico utiliza la variabilidad del brillo de la imagen a nivel de cada píxel para detectar si existe movimiento siendo más flexible en cuanto al tiempo que se debe manejar en la comparación entre imágenes, utilizando información más específica y siendo innecesario el conocimiento previo del entorno sobre el cual se encuentra el objetivo.

4.1. Flujo óptico

El flujo óptico es la velocidad aparente de las estructuras de niveles de gris, que puede considerarse como una aproximación del campo de movimiento y puede ser obtenido de una secuencia de imágenes; por lo tanto el cálculo del flujo óptico consiste en encontrar el desplazamiento aparente de los objetos en una secuencia de imágenes. Recuperar este desplazamiento es uno de los problemas claves en visión por ordenador y se han realizado muchos trabajos de investigación sobre este campo durante los últimos años.

Si se representa una imagen como una aplicación $I(x, y, t)$, donde (x, y) representa la coordenada espacial de la imagen y t el tiempo, se puede observar a una secuencia de imágenes como la variación de la intensidad en las coordenadas de la imagen a través del tiempo así:

$$I(x, y, t) = I(x + \delta_x, y + \delta_y, t + \delta_t) \quad (4.1.1.)$$

donde $\vec{v} = (\delta_x, \delta_y)$ es el desplazamiento de la imagen en el punto (x, y) después de dt unidades de tiempo. El desarrollo en series de Taylor da como resultado:

$$I(x, y, t) = I(x, y, t) + \delta_x \frac{\partial I}{\partial x} + \delta_y \frac{\partial I}{\partial y} + \delta_t \frac{\partial I}{\partial t} \quad (4.1.2.)$$

si se sustrae $I(x, y, t)$ en ambos lados de la ecuación y se divide por δ_t se tiene:

$$\frac{\delta_x}{\delta_t} \frac{\partial I}{\partial x} + \frac{\delta_y}{\delta_t} \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} = 0 \quad (4.1.3.)$$

Haciendo que el límite $\delta_t \rightarrow 0$, se obtiene la ecuación **restringida del flujo óptico** de Horn y Schunk (1981)

$$\frac{dx}{dt} \frac{\partial I}{\partial x} + \frac{dy}{dt} \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} = 0 \quad (4.1.4.)$$

Esta ecuación relaciona la velocidad de los puntos en una imagen con la variación del patrón de grises de la misma. A partir de esta ecuación no es posible determinar $\vec{v} = (\delta_x, \delta_y)$, tan solo se puede estimar $\vec{v} \perp$, la componente en la dirección normal a la curva en la dirección del gradiente, es decir el componente de flujo óptico en la dirección del gradiente del brillo de la imagen. La información disponible en cada punto de una secuencia de imágenes es sólo la componente del campo de movimiento en la dirección del vector gradiente del patrón de brillo.

$$\vec{v} \perp = \frac{-I_t \nabla I(x, y, t)}{\|\nabla I(x, y, t)\|^2} \quad (4.1.5.)$$

Este fenómeno se conoce como el **problema de la apertura** y es debido a que no es posible obtener una única solución a causa de que esta ecuación contiene dos incógnitas. Para tratar este problema se suele imponer alguna restricción sobre el flujo óptico, como por ejemplo restringir las intensidades de los píxeles correspondientes, para que no varíen de una imagen a otra.

4.2. Factores que influyen en la detección del movimiento

El cálculo del flujo óptico es un problema mal condicionado en visión artificial y requiere de suposiciones adicionales de regularización.

Pero además de este problema existe una serie de factores que lo convierten en un problema mayor. Estos factores están condicionados por el hecho de que las imágenes que se utilizan son una representación de escenas de espacio tridimensional y se pueden clasificar en dos tipos principales: **Factores fotométricos** que son los que se deben a las fuentes de iluminación y **Factores geométricos** que son derivados de la geometría de la imagen.

4.2.1. Factores fotométricos

Entre ellos se pueden considerar las principales fuentes de iluminación, las propiedades de reflectancia de las distintas superficies y las sombras. Toda superficie posee unas propiedades de reflectancia características, que de forma simplificada se pueden dividir en superficies **lambertianas** que son aquellas cuyo brillo es igual, independientemente desde la dirección en que se mire y en superficies **especulares** que son aquellas que reflejan toda la luz en una misma dirección, generalmente en forma perpendicular a la superficie. En la mayoría de casos las superficies se acercan más al tipo lambertiana, por lo que en la mayoría de los métodos para el cálculo del flujo óptico se suele suponer que las superficies tienen un ese tipo.

Pero en la realidad esta suposición no es correcta. El efecto de los brillos hace que los niveles de intensidades de las superficies cambien de una imagen a otra; esto se debe a que los objetos al cambiar de posición no reflejan la luz de igual manera, o si la fuente de luz se desplaza, los brillos cambian en todos los objetos de la imagen, y se apreciaría un flujo que no correspondería a un movimiento real. Aparte de los problemas antes mencionados, también existen los que vienen dados por las condiciones de luminosidad de la escena, por ejemplo las sombras que hacen difícil el trabajo cuando un objeto pasa de estar en la sombra a ser iluminado o viceversa, pues los niveles de gris cambia.

Otro ejemplo son los reflejos y las transparencias, pues los reflejos se deben a la luz reflejada de un objeto a otro, haciendo que cambien su intensidad de grises y las transparencias tienen un comportamiento especular y provocan que objetos situados detrás de ellos no sean detectados.

Si no existe una iluminación uniforme en los objetos de la escena, es difícil o impreciso calcular el flujo óptico y por lo tanto detectar el movimiento.

4.2.2. Factores geométricos

Son aquellos que dependen de la forma y posición de los objetos en la escena, de las relaciones que existen entre ellos y de las transformaciones que sufren al pasar de una imagen a otra. Entre los factores que se puede enumerar están:

1. Oclusiones: Cuando un objeto tapa a otro u oculta un trozo de otro al desplazarse.
2. Objetos que aparecen y desaparecen: El tamaño de las imágenes es limitado, luego, puede ocurrir que un objeto que están en una parte de la secuencia, termine desapareciendo o apareciendo a partir de un cuadro determinado.
3. Objetos no rígidos: Son objetos que pueden cambiar de forma. Es más complicado encontrar modelos que ayuden a identificar el desplazamiento del objeto, así como resulta difícil determinar de que manera se deforma.

4. Proyección perspectiva: La transformación en perspectiva es bastante compleja, pues los objetos de las imágenes pueden cambiar bastante su forma dependiendo de la trayectoria que tengan en la escena y de su posición con respecto a la cámara.
5. Trayectorias arbitrarias: El movimiento más adecuado para el cálculo de flujo óptico es el paralelo al plano de la imagen. Cuando se tiene movimiento de rotación, escalado o cualquier otro que sea complejo, el desplazamiento de los píxeles del objeto no es igual.

Además existen otros problemas que están relacionados con la cámara, como por ejemplo el ruido debido a los sensores de obtención de imágenes, distorsiones que dependen de los parámetros intrínsecos de las misma y el efecto de la lente.

4.3. Métodos de determinación del movimiento

Durante el paso de los años se ha dedicado mucho esfuerzo al estudio del problema del flujo óptico y se han realizado algunas clasificaciones. La siguiente clasificación se basa en los trabajos realizados por Mitiche y Bouthemy (1996), por Beauchemin y Barron (1995), y por Fleet (1992).

4.3.1. Método diferencial

En este método lo que se persigue es calcular el flujo óptico a través de las derivadas espaciales y temporales de las intensidades de la imagen; para este método se parte de la ecuación de restricción del Flujo Óptico (4.1.4.) y fue propuesto por Horn y Schunk (1981).

En este método se supone que la intensidad de la imagen en un punto o una función de la misma, se conserva en el tiempo. De la ecuación de flujo óptico se tiene:

$$I_x v_x + I_y v_y + I_t = 0 \quad (4.3.1.1.)$$

pero como ya se mencionó, esta ecuación contiene dos incógnitas, por lo que no es posible obtener una única solución. Escribiéndola de otra manera:

$$(I_x, I_y) \cdot (v_x, v_y) = \nabla I \cdot \vec{V} = -I_t \quad (4.3.1.2.)$$

y como antes ya se mencionó se puede determinar la componente del flujo óptico en la dirección del gradiente del brillo de la imagen y se puede expresar como:

$$\|\nabla I\| V_{\perp} = -I_t \quad (4.3.1.3.)$$

$$V_{\perp} = \frac{-I_t}{\|\nabla I\|} = -\frac{I_t}{\sqrt{I_x^2 + I_y^2}} \quad (4.3.1.4.)$$

donde $\|\nabla I\|$ es la norma del vector gradiente del brillo y V_{\perp} es la norma del vector velocidad en la dirección del gradiente del brillo. Expresado en forma vectorial se tiene:

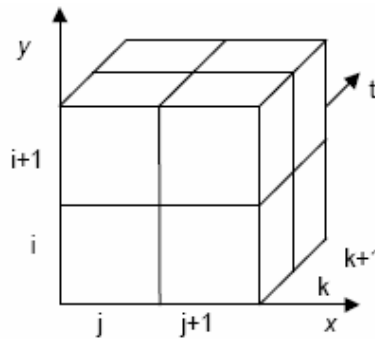
$$\vec{V}_{\perp} = -\frac{I_t}{\sqrt{I_x^2 + I_y^2}} \hat{e} \quad (4.3.1.5.)$$

donde $\hat{e} = \frac{\nabla I}{\|\nabla I\|}$ es el vector unitario en la dirección del gradiente de brillo.

La determinación del flujo óptico pasa, por tanto, por el paso previo de determinar los gradientes temporales y espaciales del patrón de grises de la imagen. Horn y Schunk proponen hacer esto con una convolución tridimensional con un cubo de 2x2x2:

$$\begin{aligned}
 I_x &= \frac{1}{4} (I_{i,j+1,k} - I_{i,j,k} + I_{i+1,j+1,k} - I_{i+1,j,k} + I_{i,j+1,k+1} - I_{i,j,k+1} + I_{i+1,j+1,k+1} - I_{i+1,j,k+1}) \\
 I_y &= \frac{1}{4} (I_{i+1,j,k} - I_{i,j,k} + I_{i+1,j+1,k} - I_{i,j+1,k} + I_{i+1,j,k+1} - I_{i,j,k+1} + I_{i+1,j+1,k+1} - I_{i,j+1,k+1}) \\
 I_t &= \frac{1}{4} (I_{i,j,k+1} - I_{i,j,k} + I_{i+1,j,k+1} - I_{i+1,j,k} + I_{i,j+1,k+1} - I_{i,j+1,k} + I_{i+1,j+1,k+1} - I_{i+1,j+1,k})
 \end{aligned}
 \tag{4.3.1.6.}$$

Figura 39. Cubo de convolución tridimensional



Fuente: Visión artificial dinámica

Para encontrar las dos componentes del flujo óptico es preciso calcular, a partir de los gradientes obtenidos, modificando la ecuación (4.3.1.5.):

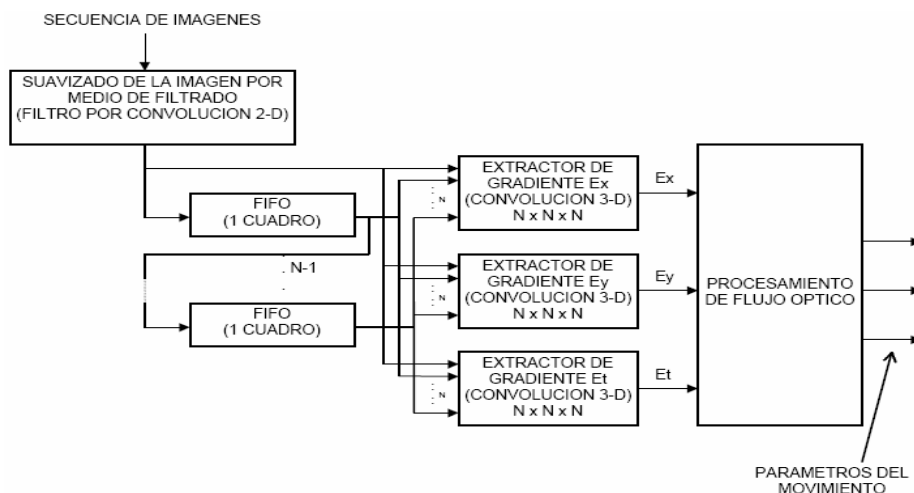
$$v_{\perp x} = - \frac{I_t I_x}{\sqrt{I_x^2 + I_y^2}} \tag{4.3.1.7.}$$

$$v_{\perp y} = - \frac{I_t I_y}{\sqrt{I_x^2 + I_y^2}} \tag{4.3.1.8.}$$

Para determinar el valor del flujo óptico en las áreas donde falla la ecuación de restricción de flujo óptico, Horn y Schunk agregan una condición adicional: los cambios en el flujo óptico deben ser suaves. En otras palabras, no deben existir cambios bruscos en el movimiento entre puntos cercanos de la imagen, por lo que se deben utilizar filtros por convolución como el gaussiano para que suavice la imagen antes de proceder a calcular el gradiente.

La convolución es una de las operaciones más útiles y utilizadas en el procesamiento de imágenes. En estos casos se relacionan dos matrices, una de ellas con la imagen original y la otra es la llamada **matriz de convolución**. Esta matriz se superpone sobre la matriz de imagen, haciendo coincidir su elemento de referencia (usualmente el elemento central), con un elemento en la matriz de imagen, luego se calcula la suma de los productos de los elementos superpuestos entre ambas matrices, para obtener el valor de la imagen resultante. La determinación del movimiento por métodos diferenciales tiene el inconveniente de ser muy cargado en cuanto al desarrollo de la programación que necesita.

Figura 40. Arquitectura para detección de movimiento por gradientes



Fuente: Visión artificial dinámica

4.3.2. Método de correlación

Este método se basa en la búsqueda de correspondencias utilizando ventanas o patrones alrededor de cada píxel. La particularidad especial de este método es que se utiliza información de los vecinos que en muchos casos hacen que la búsqueda de la correspondencia sea más efectiva.

Este método busca asociar píxeles a través de regiones similares en las imágenes que se obtienen por maximización de alguna medida de similaridad; es decir que la imagen se separa en partes y luego se determina el movimiento de cada una de esas partes, con algoritmos generalmente denominados como “segmentación de imágenes”.

Existen tres algoritmos generales para segmentar una imagen los cuales son: El **Algoritmos de segmentación de líneas**, **Algoritmos de segmentación en bloques** y **Algoritmos de segmentación de grupos** (*clusters*).

Los Algoritmos de segmentación en líneas extraen en primer lugar el contorno de los objetos contenidos en la imagen, utilizando operadores por convolución u operadores morfológicos detectores de bordes.

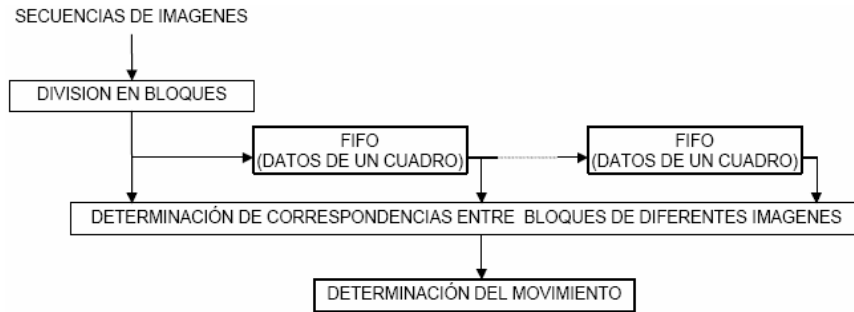
El siguiente paso es determinar las coordenadas de cada uno de los segmentos de líneas que conforman los objetos en las imágenes, determinan la correspondencia entre los segmentos de las imágenes consecutivas y por último, determinan el movimiento en base al desplazamiento de los segmentos entre las sucesivas imágenes.

Los algoritmos de segmentación en bloques, dividen las imágenes en una serie de rectángulos u otros polígonos de tamaños variables que se adapten a las áreas de nivel de gris o color más o menos homogéneo llamados **bloques semilla**, y luego vuelven a dividir estos bloques en bloques más pequeños si es que no se cumplen ciertas restricciones en cuanto a la dispersión de niveles de gris o de color dentro de los cuadros. El proceso de división continúa hasta cumplir con dichas restricciones. El siguiente paso es buscar las correspondencias entre los bloques de dos imágenes consecutivas y por último, se determina el movimiento de aquellos bloques que han cambiado de posición.

Los algoritmos de segmentación en grupos o *clusters*, utilizan procesos de simplificación de las imágenes en áreas de color o textura uniforme, tales como los operadores morfológicos. Las imágenes quedan divididas en áreas uniformes, luego se procede a determinar los centros de cada área, hallar las correspondencias entre las áreas de las imágenes consecutivas y por último se determina el movimiento entre los centros de áreas correspondientes de dichas imágenes.

Todos los algoritmos de segmentación tienen la característica común de generar un campo vectorial de flujo óptico poco denso: un vector por segmento, cuadro o área. Es posible asignar el mismo vector a todos los píxeles dentro de un cuadro o área, pero es una generalización que puede no ser útil en todos los casos. Los algoritmos de segmentación adolecen el problema de ser computacionalmente muy exigentes. Los procesos de buscar las correspondencias entre segmentos, bloques o áreas de una imagen a otra no son sencillos. Además cuando una imagen está compuesta por múltiples objetos, las segmentaciones suelen ser numerosas y por tanto, el cálculo extremadamente laborioso.

Figura 41. Arquitectura para detección de movimiento por gradientes



Fuente: Visión artificial dinámica

4.3.3. Método de error

Este método se basa en la obtención de una estimación de velocidad a partir una medida de confianza, es decir que se estima el vector de velocidad confrontando los píxeles con una medida comparativa. Esa medida de comparación suele ser extraída de la función de error de la medida del flujo óptico, el cual está dado por:

$$E_c(u, v) = (I_x v_x + I_y v_y + I_t)^2 \quad (4.3.3.1.)$$

Debido a que el objetivo es encontrar una medida que permita descartar vectores de flujo óptico erróneos; se necesita entonces que ésta valide el cálculo realizado. Esa medida es la varianza que se extrae del error del cálculo del flujo óptico y está dado por:

$$\sigma_b^2 = E(e^T e) \quad (4.3.3.2.)$$

donde e está dado por

$$e = b - A\hat{v} \quad (4.3.3.3.)$$

y donde b es un vector de dimensión $M \times 1$ que contiene las derivadas de la imagen en el tiempo:

$$b = -(I_t(x_1) \dots I_t(x_M))^T \quad (4.3.3.4.)$$

A es una matriz de dimensión $M \times 2$ que contiene las derivadas parciales de la imagen en el tiempo

$$A = \begin{bmatrix} I_x(x_1) \dots & I_x(x_1) \\ I_x(x_1) \dots & I_x(x_1) \end{bmatrix}^T \quad (4.3.3.5.)$$

y \hat{v} es el cálculo del flujo óptico dado por $\hat{v} = (A^T A)^{-1} A^T b$.

Luego de obtener este parámetro y de obtenerse el flujo óptico, lo que sigue es verificar si el mismo flujo calculado píxel por píxel, está dentro del rango permitido según la desviación. La desventaja de este método es también la carga de programación que implica, pero ésta es compensada con una mejor y más exacta detección debido a ese parámetro extra de comparación.

5. PROPUESTA DE DISEÑO

Los temas anteriormente desarrollados son la base para proponer un sistema de vigilancia efectivo y seguro apoyado en los avances tecnológicos garantizando, así, la integridad física de las personas. Este sistema de vigilancia utiliza **medios técnicos activos** tanto para alertar de la violación de la seguridad, como para accionar en contra de ella; su móvil de accionamiento es el movimiento, el cual es detectado en secuencias de imágenes captadas por cámaras de video.

5.1. Elementos del sistema de vigilancia

Los elementos que conforman el sistema de vigilancia propuesto están clasificados en la siguiente forma:

- a) Fuente de alimentación.
- b) Equipo de seguridad.
- c) Detectores.
- d) Señalizadores.
- e) Red.

tal y como se clasifican los elementos de los **Medios Técnicos Activos** del modelo de sistemas de vigilancia mostrado en el capítulo 1. Cabe mencionar que los elementos utilizados como fuente de alimentación, señalizadores y red suelen ser muy comunes en los diferentes sistemas ya existentes, a diferencia del equipo de seguridad y de los detectores que son elementos muy variables.

5.1.1. Fuente de alimentación

La Fuente de alimentación es, como ya se mencionó, el elemento de activación del sistema; es decir que, constituye la fuente de energía con la cual éste trabaja, por lo tanto, es necesario utilizar una fuente auxiliar que la sustituya en los momentos en que sea interrumpida accidental o deliberadamente, para que el sistema sea confiable en cualquier momento y se disponga siempre con la debida seguridad.

Para el efecto del sistema propuesto se utiliza un UPS (*Uninterruptible Power Supply*, Fuente de energía sin interrupción) con voltaje de salida de $115\pm 5\%$ Voltios, capacidad de 410 Watts y una duración continua de energía de 28 minutos fuera de línea; estos parámetros deben tomarse como los mínimos para un mejor desempeño del sistema. Se puede sustituir el UPS por un inversor de 24 Vdc, 60Hz., 230 Vca. y 900 Watts dependiendo de la carga que sea agregada al sistema, que se relacionada con la cantidad de señalizadores y sensores que se utilicen. Si se cuenta solamente con una cámara y un señalizador, como en el sistema planteado, bastará utilizar un UPS.

5.1.2. Equipo de seguridad

El equipo de seguridad es el cerebro del sistema de vigilancia y es la parte más importante de éste. De él depende una detección exitosa de cualquier intruso, ya que puede reportar falsas alarmas u omitirlas, si no se encuentra correctamente configurado. Además el tipo de móvil de accionamiento que el sistema sea capaz de detectar también depende de la potencial del cerebro, es decir que, un equipo de seguridad cuyo móvil de accionamiento sea la apertura de una puerta, seguramente no será capaz de detectar movimiento y mostrar el lugar en el cual se ha detectado.

El cerebro utilizado en el sistema propuesto es una CPU (*Central Processing Unit*, Unidad Central de Procesamiento) con un microprocesador de 2.4 Gigahertz (Ghz), 128 Megabytes (Mb) de memoria RAM (*Random Access Memory*, Memoria de Acceso Aleatorio), un disco duro de 80 Gigabytes (Gb), un puerto USB 2.0 (*Universal Serial Bus*, Bus Serial Universal) y/o un puerto FIREWIRE (IEEE 1394, para mayor rapidez en la transferencia de datos); siendo estos los requerimientos mínimos del sistema debido a la gran cantidad de archivos de imagen que deben utilizarse y almacenarse.

Además es necesario utilizar software especial para la captura, el tratamiento y el procesamiento de las imágenes, por lo que se emplea el lenguaje de programación Delphi 6, que es el lenguaje Pascal orientado a objetos, acompañado de las librerías IPL 2.5v (*Image Processing Library*, librería de procesamiento de imágenes) de Intel, Imaging Toolkit y CapturePro; de las cuales las dos primeras se utilizan para el tratamiento y procesamiento de imágenes y la última se emplea para capturar las mismas por medio de cámaras digitales. Además es necesario el sistema operativo Windows XP.

5.1.3. Detectores

Son los elementos que tienen la misión de informar a la central de las variaciones del estado de la zona que se está protegiendo, indicando cambios en el ambiente de la misma. El elemento utilizado en el la propuesta es una cámara digital Panasonic modelo PV-GS12 con formato de grabación digital PCM, salida de datos a puerto USB 2.0 y FIREWIRE (IEEE 1394), captura de imágenes a 60 cuadros por segundo, consumiendo 7.8 Vdc. y 1.4 A., y de 673 centímetros³ (cm³) de volumen. Los requerimientos mínimos para la cámara son: salida de datos digital a puerto USB 2.0, grabación de imágenes digitales, captura de imágenes a 30 cuadros por segundo y volumen máximo de 680 cm³.

5.1.4. Señalizadores

Son los elementos encargados de avisar de la ocurrencia de alguna violación a la seguridad del lugar que está bajo protección. Normalmente el señalizador de un sistema de vigilancia lo constituye una sirena o un altoparlante; para el caso de esta propuesta, además de la sirena, se utiliza una pantalla por medio de la cual se puede observar el objeto o la persona en movimiento; siendo señalizado y resaltado para una mejor claridad.

La pantalla utilizada es un monitor LCD (*Liquid Crystal Display*, Pantalla de cristal líquido) de 17" a color; la cual es colocada, junto con el Equipo de Seguridad o cerebro, en un lugar designado como estación de vigilancia y se emplea para monitorear y observar las condiciones en las que se presenta la violación a la seguridad detectada. Como requerimiento mínimo para la pantalla se sugiere un monitor de 15" a color. La sirena que se utiliza está diseñada para interiores, con capacidad para 109 dB, 8 ohm de impedancia y consumo de 12 Vdc. y 500 mA.

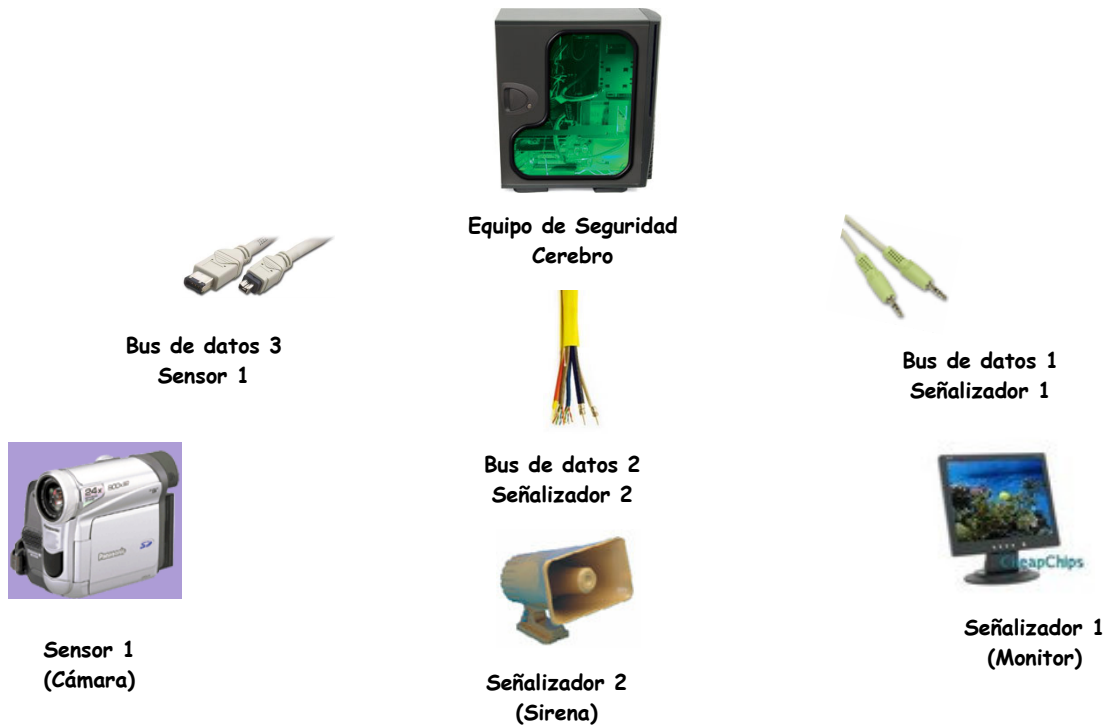
5.1.5. Red

La red esta definida como los canales por medio de los cuales, los datos recabados por los sensores y los datos enviados a los señalizadores, son transportados por el cerebro del sistema de vigilancia. Para el sistema propuesto, se utiliza como bus de datos del sensor (cámara de video) hacia el cerebro, un Cable de interfase DV (IEEE 1394) de 4 pines a 4 pines, de un largo de 10 metros como mínimo. Como Bus de datos del cerebro para la sirena, se utiliza un cable calibre 16 de 5 metros de largo, junto con un sistema de optoacopladores para su activación y desactivación en caso de violación de la seguridad.

5.1.6. Diagrama del sistema propuesto

El sistema de vigilancia consta físicamente de 7 partes como se puede observar en la figura 42.

Figura 42. Diagrama del sistema de vigilancia propuesto



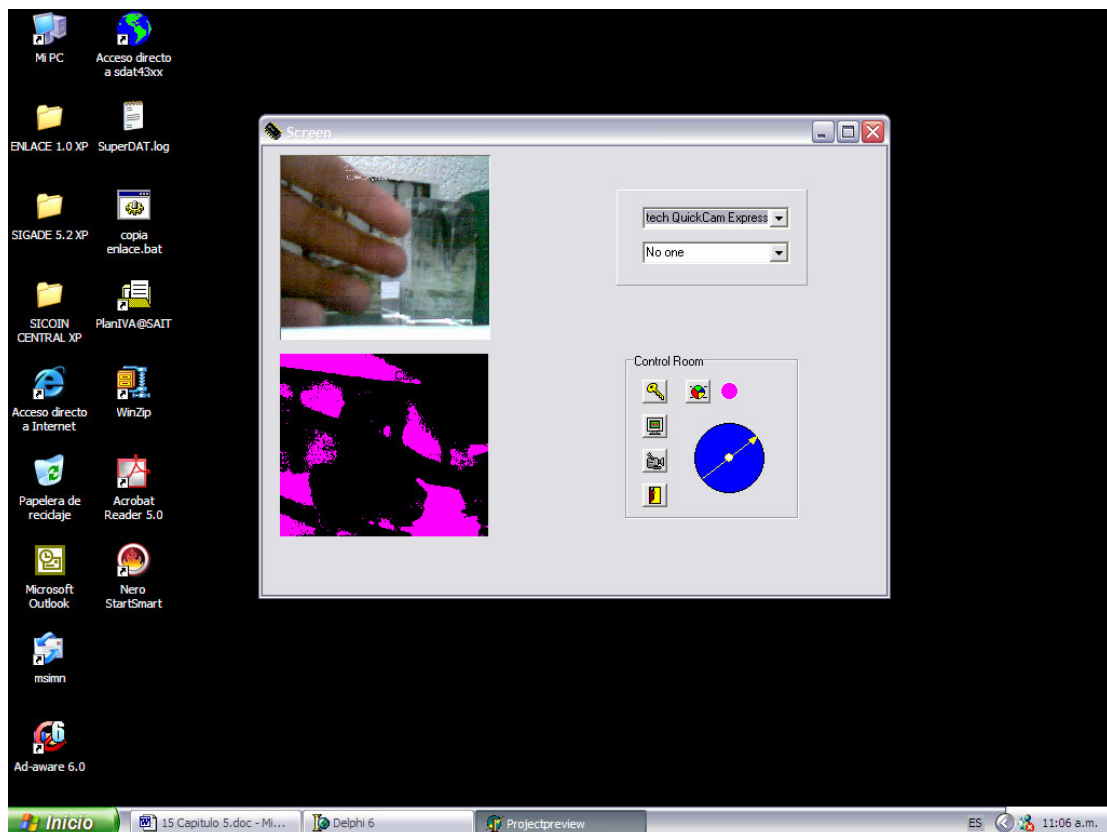
Este diagrama contiene los requerimientos mínimos para el sistema propuesto, pudiendo aumentarse el número de sensores, así como el número de señalizadores, dependiendo del área que se desee ser protegida y de las personas que monitoreen dicha área. La instalación de los sensores (cámaras) debe realizarse en un lugar donde no sean visibles a los intrusos, para que no sean eliminadas de la red, al igual que el bus de datos que va hacia el cerebro; si fuera posible debiera utilizarse una cámara oculta para no despertar sospechas en cuanto a su posición.

5.2. Descripción del diseño

La propuesta del sistema de vigilancia utiliza la cámara de video como un sensor, haciendo que ésta capte las imágenes a través de sus lentes cuadro por cuadro 60 veces por segundo, es decir que capta las imágenes como si tomara una fotografía cada 16.66 milisegundos; luego estas imágenes son convertidas en información digital y enviadas al cerebro del sistema.

El cerebro utiliza el programa de ejecución “Screen” para ejecutar la captura y el procesamiento de las imágenes así como la detección del movimiento, el cual es elaborado por medio del lenguaje de programación Delphi, cuya pantalla de interacción es mostrada en la Figura 43.

Figura 43. Pantalla de interacción del sistema de vigilancia



El programa hace posible la visualización previa, por medio del monitor, en tiempo real del entorno en el cual se aplica el sistema de vigilancia; además en él se consigue elegir los efectos que se pueden aplicar a la imagen. Con un clic se puede iniciar el proceso de vigilancia para la detección de movimiento, el cual al ser descubierto, es mostrado por medio del monitor localizado en la estación de vigilancia y señalado por la sirena para que las autoridades respectivas tomen la decisión de actuar o ignorar la acción detectada.

Además el sistema puede convertirse en un medio completamente automático, al ser conectado a un sistema de protección periférica para que cierre por completo el lugar protegido en el momento de detectar objetos o personas con determinadas características previamente programadas y al utilizar la opción de llamada de alerta, en la cual después de efectuar el cierre periférico, se vincula un programa adicional que realiza una llamada telefónica, avisando sobre la trasgresión del espacio protegido.

5.2.1. Descripción del programa

El programa en lenguaje Delphi, hace uso de dos librerías especiales, IPL o Imaging Toolkit y CapturePro. Por medio de la librería CapturePro se efectúa la detección automática de la cámara utilizada como sensor de movimiento y la captura (traslado) de imágenes en tiempo real de la cámara hacia el CPU, por medio de instrucciones específicas de la misma, las cuales son grabadas cuadro por cuadro en una memoria tipo buffer (se graba la imagen en un archivo y es eliminada hasta que se grabe otra) y posteriormente convertidas a escala de grises. Seguidamente se utiliza la librería IPL, que puede ser sustituida por la librería Imaging Toolkit, para realizar todas las demás tareas, siendo la primera de ellas convertir la imagen a un formato accesible para la misma y así poder manipular las imágenes.

La segunda tarea es detectar el movimiento en la secuencia de imágenes ya almacenada. Para llevar a cabo dicha tarea es necesario primero hacer la extracción de fondo, que corresponde a detectar los píxeles de la imagen que no son parte de algún objeto en movimiento frente a la cámara y crear una MHI (*Motion History Image*, Historia de movimiento de la imagen); esto se logra mediante la ejecución de una comparación, cuadro por cuadro, del cambio de brillo de la imagen en el tiempo, por medio de la siguiente fórmula:

$$MHI = pixelB + |pixelB - pixelA| \quad (5.2.1.1)$$

en la cual se calcula el MHI por medio del valor absoluto de la diferencia de brillo entre el píxel de un cuadro B de la imagen y el píxel del cuadro A anterior a este, en seguida este valor absoluto es sumado al mismo píxel del cuadro B pero con menor umbral (se disminuye la característica de *Threshold*), a continuación se calcula la diferencia entre el píxel de otro cuadro posterior y el resultado de la ecuación, haciéndolo durante todo el proceso.

El anterior método puede ser sustituido por la técnica de desviación estándar; la cual calcula la desviación estándar del brillo de los píxeles de un conjunto de cuadros de la imagen y luego compara ese dato con los píxeles de cada uno de esos cuadros, para definir cuales tienen un cambio significativo en su brillo y así decidir si está en movimiento o no lo está. Este método se puede lograr utilizando la siguiente fórmula:

$$desv.std. = \sqrt{\frac{\sum(x, y)^2}{N} - \left(\frac{\sum(x, y)}{N}\right)^2} \quad (5.2.1.2.)$$

siendo N el número de cuadros de la imagen que se utilizan para hacer las sumatorias de los píxeles.

Seguidamente el programa se da a la tarea de obtener los bordes de los elementos en movimiento mediante la utilización del gradiente. Si se utilizaran las fórmulas de cálculo para realizar esta tarea, el programa gastaría demasiados recursos y demasiado tiempo, por lo que fue necesario utilizar otros medios para poder lograrlo.

La convolución de la imagen con el operador de Sobel da como resultado una aproximación al cálculo del gradiente de una imagen, por lo tanto y derivado a su relativa simplicidad, se utilizó este método para dichos efectos.

El cálculo del gradiente aproximado se logra de la siguiente forma:

$$\frac{df}{dx} = (I_{1,3} + I_{2,3} + I_{3,3}) - (I_{1,1} + I_{2,1} + I_{3,1}) \quad (5.2.1.3.)$$

$$\frac{df}{dy} = (I_{3,1} + I_{3,2} + I_{3,3}) - (I_{1,1} + I_{1,2} + I_{1,3}) \quad (5.2.1.4.)$$

donde el operador de Sobel está dado por

$$S = \begin{bmatrix} I_{1,1} & I_{1,2} & I_{1,3} \\ I_{2,1} & I_{2,2} & I_{2,3} \\ I_{3,1} & I_{3,2} & I_{3,3} \end{bmatrix} \quad \text{donde} \quad S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad S_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (5.2.1.5.)$$

Después de obtener los bordes del objeto o persona en movimiento, se incorpora la tercera y última tarea, señalar el descubrimiento.

La señalización del descubrimiento consiste en la activación de la sirena, mandando una señal a los optoacopladores que se encargan de pasar corriente a la misma, por medio del puerto paralelo; que realmente consiste en una señal de digital (1 ó 0) con dos estados, encendido y apagado. Posteriormente aparece la silueta del objeto o persona cuyo movimiento ha sido detectado, en el monitor, para que las autoridades respectivas analicen si es necesario interceptarlo o ignorarlo.

5.2.2. Factores que causan distorsión en el sistema

Existen 3 factores principales que no contribuyen a poder lograr una detección exitosa; estos son: **la cantidad de luz del entorno, las condiciones climáticas del medio ambiente y el índice de la matriz del operador Sobel.**

En cuanto a la cantidad de luz, es obvia la distorsión, debido a que la esencia del flujo óptico son las variaciones de luz en la escala de grises y la dirección en que estas ocurren; por lo tanto si existe una mala fuente de luminosidad, los cuerpos no reflejarán una buena cantidad de luz y por consiguiente la cámara de video no podrá captar las imágenes con claridad debido al ruido insertado por la oscuridad del ambiente. La iluminancia necesaria para que el sistema propuesto trabaje adecuadamente se encuentra entre los niveles de 500-1000 Lux.

En lo que concierne a las condiciones climáticas del medio ambiente también se producen distorsiones, adhiriendo ruido a las imágenes, puesto que la humedad excesiva, el calor abundante, la aridez o el extremo frío, producen situaciones que producen variación en las imágenes (como la vaporización o la bruma). El sistema trabaja eficientemente en un ambiente cuya temperatura se encuentra en un rango de 5°C-40°C con humedad relativa entre 30%-60%.

La utilización de una matriz de 3x3 como operador de Sobel está diseñada de manera que elimine cierta cantidad del ruido de las imágenes. Es decir que elimina el ruido pero no lo hace en su totalidad, por lo que pueden presentarse problemas de distorsión durante el proceso de búsqueda y detección en el sistema de vigilancia. El resultado de convolucionar la imagen con un operador Sobel de matriz 3X3 devuelve el borde de la imagen con aproximadamente 50% de ruido, es decir que el 50% del resultado no son bordes en la imagen. Para lograr eliminar mayor cantidad de ruido es necesario umbralizar la imagen o utilizar un operador de Sobel mayor (9X9 por ejemplo).

5.3. Costo del sistema de vigilancia propuesto

El costo total del sistema de vigilancia con los requerimientos mínimos, a un tipo de cambio de Q.7.50 por USD.1.00, es de Q.13,495.00 distribuido de la siguiente forma:

Tabla II. Costo del sistema de vigilancia propuesto

Lugar de Cotización	Elementos del Sistema Propuesto	Costo en Quetzales
Compu City	1 UPS	Q.1,800.00
Compu City	1 CPU	Q.3,500.00
Compu City	1 Monitor	Q.2,000.00
Electrónica CEF	1 Sirena	Q.150.00
Electrónica CEF	1 Sistema de Optoacopladores (2 integrados)	Q.100.00
Electrónica CEF	1 Bus de datos para los Señalizadores	Q.100.00
www. Panasonic.com	1 Cámara de Video	Q.4,500.00
Electrónica CEF	1 Bus de datos para el Sensor	Q.250.00
Compu City	1 Licencia Windows XP	Q.373.00
www.borland.com	1 Licencia Delphi	Q.722.00
www.intel.com	1 Librería IPL para Delphi	Gratis
www.mcm-design.com	1 Librería Imaging Toolkit	Gratis
www.pegasusimaging.com	1 Librería CapturePro para Delphi	Gratis
		Q.13,495.00

5.4. Ventajas y desventajas del sistema de vigilancia propuesto

El sistema de vigilancia propuesto tiene muchas ventajas y pocas desventajas, por lo que promete ser un sistema confiable y exacto. Las desventajas son el costo inicial alto, la sensibilidad al ruido dependiendo de la luz y las condiciones climáticas del medio ambiente.

Entre las ventajas que se tienen con respecto a otros sistemas están: la posibilidad de grabar los sucesos ocurridos; la posibilidad de identificar a los intrusos y las condiciones en que se presenta la violación; casi no existen falsas alarmas, se puede conectar con otros sistemas; pueden aumentarse los sensores sin aumentar demasiado el costo global; puede ser utilizado para obtener, por parte de las autoridades, un panorama de la situación para planear una estrategia de captura; el sistema puede ser monitoreado desde varias partes al mismo tiempo; se puede volver completamente automático conectándolo a un sistema de cierre periférico y notificar por medio de llamadas telefónicas; los sensores pueden ser casi invisibles a los intrusos y por sobre todo, garantiza la seguridad en un 99%.

CONCLUSIONES

1. Los sistemas de vigilancia más usados, en el ámbito industrial y doméstico en nuestros días, son construidos con medios técnicos que se basan en sensores cuyo movimiento registra, únicamente, dos estados, activo o inactivo y depende de éstos el determinar o no una violación a la seguridad del sitio protegido, sin proveer de mayor información referente a dicha violación.
2. Las características del video digital permiten que las imágenes sean tratadas y mejoradas con facilidad, debido a que, además de utilizar la división de éstas en varios cuadros como se efectúa en video análogo, también, se divide cada cuadro en píxeles que pueden ser transformados en datos digitales y, posteriormente, ser manipulados por medio del empleo de lenguajes de programación de computadoras.
3. El criterio técnico más eficaz utilizado para procesar y arreglar, adecuadamente, una secuencia de imágenes es hacer uso de máscaras aplicadas a la imagen por medio de la convolución, debido a que es una herramienta versátil, pues se pueden efectuar una gran cantidad de procesos y arreglos, dependiendo del tipo de máscara que se utilice, además es sencilla y rápida de aplicar porque al ser utilizada se elimina la necesidad de una enorme cantidad de memoria y se reduce el tiempo de cálculo de los programas de ejecución.

4. Los métodos más utilizados para determinar el movimiento en una secuencia de imágenes digitales son: el método diferencial, el método de correlación y el método de error; los cuales se basan en el cálculo del flujo óptico, cuyo cometido es captar la velocidad aparente de los niveles de gris de las estructuras de una imagen.

5. El diseño efectivo de un sistema de vigilancia que detecte el movimiento en una secuencia de imágenes captadas por cámara de video digital, se logra basándose en las grandes ventajas que proporciona el método diferencial para la determinación del movimiento, como los son: su pequeño margen de error y la poca sensibilidad a los cambios de luz; las cuales se logran con la utilización del gradiente para obtener el flujo óptico, puesto que la efectividad se aumenta en tanto se aumente el orden de la derivada que se aplique.

RECOMENDACIONES

1. Los sistemas de vigilancia más completos son los que están compuestos por los tres tipos medios de un sistema de vigilancia básico; por lo que en este sentido, es necesario que la propuesta de diseño además de contar con medios técnicos, también sea diseñada para que cuente con medios organizativos y medios humanos.
2. Los cuadros de una imagen pueden ser simplificados, de tal manera que el sistema no sea afectado por la lentitud de procesar píxel por píxel; dicho trabajo se logra agrupando píxeles en áreas de un tamaño determinado y así manipular con facilidad y en menor tiempo las imágenes capturadas.
3. El tratamiento de imágenes por medio de convolución con máscaras es una herramienta muy extensa y merece mayor atención para poder utilizarla más ampliamente, ya que, con las máscaras se pueden filtrar las imágenes, mejorarlas y extraer ciertas características u objetos de una imagen necesarios para la tarea que se este implementando.
4. Además de estudiar métodos que se basan en el flujo óptico, se sugiere examinar los que emplean vectores de desplazamiento, con el fin de diseñar un sistema que mezcle ambos tipos, para lograr mejores resultados y así obtener un sistema con mayor inmunidad al ruido.
5. El método utilizado en el diseño debe cambiarse a conveniencia según las necesidades presentadas, debido a que éste funciona bien para las condiciones manejadas pero puede no hacerlo si estas son modificadas.

BIBLIOGRAFÍA

1. De la Escalera Hueso, Arturo. **Visión por computador: Fundamentos y métodos.** España: Editorial Prentice Hall, 2001. 304 pp.
2. Jain, Anil K. **Fundamentals of digital image processing.** USA: Editorial Prentice Hall, 1988. 592 pp.
3. Pajares Martínsanz, Gonzalo y Jesús M. de la Cruz Gracá. **Visión por computador: Imágenes digitales y aplicaciones.** España: Editorial Ra-ma, 2001. 800 pp.
4. Umbaugh, Scott E. **Computer imaging: Digital image analysis and processing.** USA: Editorial Prentice Hall, 2004. 659 pp.
5. Walker, Bruce H. **Optical design for visual systems.** USA: Editorial SPIE Press, 2000. 325 pp.
6. Izaguirre, Aitzol Zuloaga. **Visión artificial dinámica, determinación de movimiento a partir de secuencias de imágenes.** Tesis Doctorado en Tecnologías de la Información, universidad Del País Vasco, España, 1998. 31 pp.
7. Molina, Roberto. **Introducción al procesamiento y análisis de imágenes digitales.** Tesis Licenciatura en Ciencias de la Computación e I.A., universidad de Granada, España, 1998. 355 pp.

8. Sánchez Pérez, Javier. **Estimación del flujo óptico en secuencias de imágenes y de la carta de disparidad en pares estéreo: aplicación a la reconstrucción tridimensional.** Tesis Doctorado en Informática y Sistemas, universidad de Las Palmas de Gran Canaria, España, 2001. 229 pp.
9. Dalsa Technology With Vision. **CCD vrs. CMOS: facts and fiction.** Canada; enero, 2001.
10. Department of Computer Science, National Tsing Hua University. **Computation of optical flow under non-uniform brightness variations.** Taiwan; febrero, 2004.
11. Instituto Nacional de Automática, Universidad de San Juan. **Estimación del flujo óptico utilizando fusión de datos.** Argentina; enero, 2004.
12. Intel Corporation. **Intel image processing library, reference manual.** USA; agosto, 2000.
13. Intel Corporation. **Open source computer vision library, reference manual.** USA; diciembre, 2001.
14. Lawrence Livermore National Laboratory. **Color spaces in digital video.** USA; agosto, 1997.
15. School of Electrical and Computer Engineering, Purdue University. **Robust motion estimation under varying illumination.** USA; mayo, 2004.
16. Swiss Center for Electronics and Microtechnology. **Motion estimation based on the direction of intensity gradient.** Suiza; enero, 2002.

Bibliografía electrónica

1. Fundamentos del video análogo y digital. <http://www.fuac.edu.co>. Mayo, 2005.
2. Introducción a la refrigeración, Células Peltier. <http://www.metamonds.net>. Julio, 2005.
3. Medios técnicos en seguridad y protección. <http://www.monografías.com>. Mayo, 2005.
4. SeguridadPlus. <http://www.seguridadplus.com>. Mayo, 2005.

Apéndice 1:

**Código Fuente del programa
de detección de movimiento.**

APÉNDICE 1

Código Fuente del programa de detección de movimiento

```
unit Preview;

interface

uses
  Windows, Messages, SysUtils, Variants, Classes, Graphics, Controls, Forms,
  Dialogs, StdCtrls, Buttons, CapPro30v, CAPTUREPRO3Lib_TLB, ExtCtrls,
  mcmlImageTypeDef, mcmlImage, mcmlImageColor, mcmlImageKernel, mcmlImageMath,
  mcmShape, mcmlImageFilter;

type
  TScreen = class(TForm)
    Cap1: TCapPRO;
    Combo1: TComboBox;
    Combo2: TComboBox;
    Panel1: TPanel;
    GroupBox1: TGroupBox;
    SpeedDetect: TSpeedButton;
    SpeedClose: TSpeedButton;
    SpeedConect: TSpeedButton;
    SpeedInicioC: TSpeedButton;
    Timer1: TTimer;
    ICView: TmcmlImageCtrl;
    ImageMath: TmcmlImageMath;
    ImageColor: TmcmlImageColor;
    Timer2: TTimer;
    mcmS: TmcmShape;
    mcmS1: TmcmShape;
    SpeedColor: TSpeedButton;
    Shape1: TShape;
    ImageFilter: TmcmlImageFilter;
    procedure FormCreate(Sender: TObject);
    procedure Cap1FilterListChanged(Sender: TObject;
      FilterCategory: String; ChangeType: Cardinal; IFilterIndex: Integer;
      FilterName: String);
    procedure Combo2Change(Sender: TObject);
    procedure SpeedCloseClick(Sender: TObject);
    procedure SpeedConectClick(Sender: TObject);
    procedure SpeedDetectClick(Sender: TObject);
    procedure Timer1Timer(Sender: TObject);
    procedure SpeedColorClick(Sender: TObject);
    procedure Timer2Timer(Sender: TObject);
    procedure SpeedInicioCClick(Sender: TObject);
  private
    Procedure ReloadVideoDevices;
    Procedure ReloadVideoEffects;
```

```

    { Private declarations }
  public
  { Public declarations }
end;

var
  Screen: TScreen;
  dude,detect: boolean;
  Bitmap1,Bitmap2: TBitMap;
  Imagelook1,Imagelook2,Imageresult: TmcmlImage;
  Filter1,Filter2: TmcmlImage;
  Colores: PPaletteEntry;
  Colorblanco,angulo: integer;
  rojo,verde,azul,threecolor: byte;
  puntero,puntero1,punterosuma,punterototal: real;

implementation

{$R *.dfm}

procedure TScreen.ReloadVideoDevices;{Reconoce los dispositivos conectados a la computadora}
var
  I: integer;
begin
  with Combo1 do
    begin
      Clear;
      Assert(Cap1.NumDevices = Cap1.NumVideoDevices);
      for I := 0 to Cap1.NumVideoDevices - 1 do
        begin
          Items.Add(Cap1.ObtainVideoDeviceName(I));
          Assert(Items[I] = Cap1.ObtainDeviceName(I));
        end;
      if Items.Count > 0 then
        begin
          Enabled := not Cap1.IsConnected;
        end
      else
        begin
          Enabled := False;
          Items.Add('Source');
        end;
      SpeedConect.Enabled:=Enabled;
      SpeedInicioC.Enabled:=Enabled;
      if Cap1.IsConnected then
        ItemIndex := Cap1.VideoDeviceIndex
      else
        ItemIndex := 0;
    end;
  end;
end;

Procedure TScreen.ReloadVideoEffects;{Carga los efectos de video disponibles para cada dispositivo}
var
  I, VideoEffectCount: integer;
Begin
  With Combo2 do
    begin
      clear;

```



```

if Cap1.NumDevices > 0 then
begin
Enabled := True;
Items.Add('No one');
VideoEffectCount := Screen.Cap1.ObtainNumFilters('Video Effect 1');
if VideoEffectCount > 0 then
begin
for I := 0 to VideoEffectCount - 1 do
Items.Add(Screen.Cap1.ObtainFilterName('Video Effect 1', I));
ItemIndex := Screen.Cap1.ObtainVideoProcFilterIndex(0) + 1;
end
else
ItemIndex := 0;
end
else
begin
Enabled := False;
Items.Add('Effects');
ItemIndex := 0;
end;
end;
end;

```

procedure TScreen.FormCreate(Sender: TObject);{inicia las variables y verifica si existen dispositivos conectados}

```

var
oldcursor: TCursor;
begin
ThreeColor:=0;
Colorblanco:=1;
rojo:=0;
verde:=255;
azul:=0;
dude:=true;
detect:=True;
Shape1.Brush.Color:=RGB(rojo,verde,azul);
ReloadVideoDevices;
OldCursor := Screen.Cursor;
Screen.Cursor := crHourGlass;
Try
Cap1.OnFilterListChanged:=Cap1FilterListChanged;
Finally
Screen.Cursor:=OldCursor;
end;
end;

```

```

procedure TScreen.Cap1FilterListChanged(Sender: TObject;
FilterCategory: String; ChangeType: Cardinal; IFilterIndex: Integer;
FilterName: String);{verifica si existió un cambio en la conexión de dispositivos a la computadora}
begin
If FilterCategory='Video Capture Sources' then
if(ChangeType = 0) or (changeType = 1) then
ReloadVideoDevices;
end;
end;

```

```

procedure TScreen.Combo2Change(Sender: TObject);{aplica el efecto de video seleccionado}
begin
Screen.Cap1.SetVideoProcFilter(0, 'Video Effect 1', Combo2.ItemIndex-1);
end;

```

```

Screen.Cap1.PreviewHeight:=169;
Screen.Cap1.PreviewWidth:=193;
end;

```

```

procedure TScreen.SpeedCloseClick(Sender: TObject);{cierra la aplicación}
begin
Close;
end;

```

```

procedure TScreen.SpeedConectClick(Sender: TObject);{transfiere las imágenes del dispositivo de video a
la ventana de visualización del programa}

```

```

begin
  If dude then
    Begin
      Screen.Cap1.Connect(Combo1.ItemIndex);
      Screen.Cap1.Preview:=true;
      Screen.Cap1.PreviewHeight:=169;
      Screen.Cap1.PreviewWidth:=193;
      dude:=False;
      ReloadVideoEffects;
      SpeedDetect.Enabled:= True;
      SpeedColor.Enabled:=True;
    end
  else
    begin
      Screen.Cap1.Disconnect;
      Screen.Cap1.Preview:=False;
      dude:=true;
      Combo2.Enabled:=False;
      Combo2.Text:='Effects';
      SpeedDetect.Enabled:= False;
      SpeedColor.Enabled:=False;
    end;
end;

```

```

procedure TScreen.SpeedDetectClick(Sender: TObject);{habilita los cronómetros para realizar las tareas
de detección}

```

```

begin
  If detect then
    begin
      Timer1.Enabled:=True;
      Timer2.Enabled:=True;
      detect:=False;
    end
  else
    begin
      Timer1.Enabled:=False;
      Timer2.Enabled:=True;
      detect:=True;
    end;
end;

```

```

procedure TScreen.Timer1Timer(Sender: TObject);{cronómetro que resta la diferencia en las imágenes y
la muestra con colores diferentes}

```

```

begin
  Cap1.FrameFile:='c:\Frames\framesum.bmp';
  Cap1.CaptureFrame;
  Bitmap1:=TBitmap.Create;

```

```

    Imagemask1:=TmcImage.Create;
    Bitmap1.LoadFromFile('c:\Frames\frameini.bmp');
    Imagemask1.DibHandle:=Bitmap1.ReleaseHandle;
    Bitmap2:=TBitMap.Create;
    Imagemask2:=TmcImage.Create;
    Bitmap2.LoadFromFile('c:\Frames\framesum.bmp');
    Imagemask2.DibHandle:=Bitmap2.ReleaseHandle;
    ImageMath.SourceImage[0]:=Imagemask1;
    ImageMath.SourceImage[1]:=Imagemask2;
    ImageMath.Execute(IM_SUB);
    ImageColor.SourceImage[0]:=ImageMath.ResultImage;
    ImageColor.ConvertTo(IF_GREY8,false);
    ImageColor.SourceImage[0]:=ImageColor.ResultImage;
    ImageColor.BrightContrast(100,0);
    ImageMath.SourceImage[0]:=ImageColor.ResultImage;
    ImageMath.Scale[0]:=1.4;
    ImageMath.Execute(IM_DIV);
    ImageMath.SourceImage[0]:=Imageresult;
    ImageMath.SourceImage[1]:=ImageMath.ResultImage;
    ImageMath.Execute(IM_ADD);
    Imageresult:=ImageMath.ResultImage;
    ImageColor.SourceImage[0]:=ImageMath.ResultImage;
    ImageColor.ConvertTo(IF_BW,false);
    ICView.Image.DibHandle:=ImageColor.ResultImage.ReleaseHandle;
    Getmem(Colores,256*SizeOf(TPaletteEntry));
    Colores.peRed:=rojo;
    Colores.peGreen:=verde;
    Colores.peBlue:=azul;
    ICView.Image.SetPaletteEntry(Colorblanco,Colores);
end;

```

```

procedure TScreen.SpeedColorClick(Sender: TObject);{Cambia de color las diferencias entre imágenes}
begin

```

```

    ThreeColor:=ThreeColor+1;
    if ThreeColor=1 then
        begin
            rojo:=255;
            verde:=0;
            azul:=0;
            Shape1.Brush.Color:=RGB(rojo,verde,azul);
        end;
    if ThreeColor=2 then
        begin
            rojo:=255;
            verde:=0;
            azul:=255;
            Shape1.Brush.Color:=RGB(rojo,verde,azul);
        end;
    if ThreeColor=3 then
        begin
            rojo:=0;
            verde:=0;
            azul:=255;
            Shape1.Brush.Color:=RGB(rojo,verde,azul);
        end;
    if ThreeColor=4 then
        begin
            rojo:=0;

```

```

    verde:=255;
    azul:=0;
    Shape1.Brush.Color:=RGB(rojo,verde,azul);
    ThreeColor:=0;
end;
end;

```

procedure TScreen.Timer2Timer(Sender: TObject);{cronómetro que expone el ángulo promedio de movimiento de la imagen}

```

Var
X,Y: integer;
begin
    ImageFilter.SourceImage[0]:=Imageresult;
    ImageFilter.Filter(FLT_SOBELNS);
    Filter1:=ImageFilter.ResultImage;
    ImageFilter.Filter(FLT_SOBELEW);
    Filter2:=ImageFilter.ResultImage;
    Punterototal:=0;
    punterosuma:=0;
    For Y:=1 to Filter1.Height-1 do
        begin
            For X:=1 to Filter1.Width-1 do
                begin
                    if GetGValue(Filter2.Pixel[X,Y])=0 then
                        puntero:=pi/2
                    else
                        begin
                            puntero:=arctan(GetGValue(Filter1.Pixel[X,Y])/GetGValue(Filter2.Pixel[X,Y]));
                            punterosuma:=punterosuma+puntero;
                        end;
                    end;
                end;
            punterototal:=punterosuma/(Filter1.Height*Filter1.Width);
            angulo:=trunc(punterototal*180/pi);
            mcmS1.Angle:=angulo;
            mcmS1.Refresh;
        end;
    end;
end;

```

procedure TScreen.SpeedInicioCClick(Sender: TObject);{captura el entorno inicial}

```

begin
    Cap1.FrameFile:='c:\Frames\frameini.bmp';
    Cap1.CaptureFrame;
end;

```

end.