

UNIVERSIDAD DEL VALLE DE GUATEMALA

Facultad de Ingeniería

Propuesta de un sistema de marcaje biométrico que emplea la manipulación de imágenes digitales captadas por cámara de video para el reconocimiento facial

Manuel Antonio Arredondo Escobar

Guatemala  
2007



Propuesta de un sistema de marcaje biométrico que emplea la manipulación de imágenes digitales captadas por cámara de video para el reconocimiento facial

UNIVERSIDAD DEL VALLE DE GUATEMALA

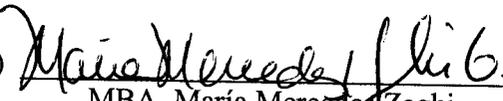
Facultad de Ingeniería

Propuesta de un sistema de marcaje biométrico que emplea la manipulación de imágenes digitales captadas por cámara de video para el reconocimiento facial

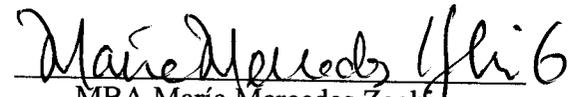
Trabajo de investigación presentado por  
Manuel Antonio Arredondo Escobar  
para optar al grado académico de  
Master en Tecnología y Ciencias de la Computación

Guatemala  
2007

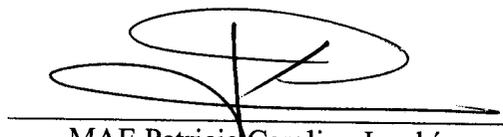
Vo. Bo.:

(f)   
MBA. María Mercedes Zaghi

Tribunal Examinador:

(f)   
MBA María Mercedes Zaghi

(f)   
MBA Noemí González Mérida

(f)   
MAE Patricia Carolina Joachín

Fecha de aprobación: Guatemala, 12 de diciembre de 2007

## ÍNDICE

ÍNDICE DE TABLAS.....	III
ÍNDICE DE ILUSTRACIONES.....	IV
RESUMEN .....	VI
1. INTRODUCCIÓN.....	1
2. SISTEMAS DE MARCAJE .....	3
2.1. Sistemas de marcaje tradicionales .....	4
2.2. Sistemas de marcaje biométricos .....	6
2.2.1. Arquitectura de un sistema biométrico.....	8
2.2.2. Fase operacional de un sistema de identificación personal .....	10
2.2.3. Exactitud en la identificación: medidas de desempeño .....	11
2.2.4. Sistemas biométricos actuales .....	13
3. CARACTERÍSTICAS DE IMAGEN DIGITAL Y SU PROCESAMIENTO .....	16
3.1. Video análogo.....	16
3.1.1. Proceso de exploración de la imagen .....	17
3.1.2. Las señales de color.....	19
3.2. Video digital.....	22
3.2.1. Muestreo .....	23
3.2.2. Cuantización.....	25
3.2.3. Codificación.....	27
3.2.3.1. Codificación de las señales compuestas .....	28
3.2.3.2. Codificación en componentes.....	29
3.3. Procesamiento de secuencias de imágenes.....	31
3.3.1. Fundamentos.....	31
3.3.2. Ruido en imágenes.....	34
3.3.3. Relaciones básicas entre píxeles.....	35
3.3.4. Pasos para el procesamiento de imágenes .....	37
3.3.5. Mejoramiento de Imágenes .....	38
3.3.5.1. Mejora mediante procesamiento puntual .....	38
3.3.5.2. Filtrado espacial.....	41
3.3.5.3. Mejora en el dominio de las frecuencias .....	42
3.3.6. Segmentación de imágenes .....	43

3.3.6.1. Regiones y bordes.....	44
3.3.6.2. Segmentación basada en umbralización .....	45
3.3.6.3. Segmentación basada en bordes .....	46
4. PROCESAMIENTO DE IMÁGENES PARA RECONOCIMIENTO FACIAL.....	53
4.1. Secuencias de histogramas.....	53
4.2. Modelo oculto de Markov.....	56
4.3. Vector Propio y Valor Propio.....	59
5. PROPUESTA DE DISEÑO.....	65
5.1. Elementos del sistema de marcaje .....	65
5.1.1. Fuente de alimentación .....	65
5.1.2. Equipo de comparación biométrica.....	66
5.1.3. Detectores .....	67
5.1.4. Visor de sucesos .....	67
5.1.5. Red.....	67
5.1.6. Diagrama del sistema propuesto .....	68
5.2. Descripción del diseño.....	69
5.2.1. Descripción del programa.....	70
5.2.2. Factores externos que causan distorsión en el sistema.....	71
5.2.3. Costo del sistema de vigilancia propuesto .....	72
5.2.4. Ventajas y desventajas del sistema de vigilancia propuesto .....	73
6. CONCLUSIONES Y RECOMENDACIONES .....	74
7. BIBLIOGRAFÍA.....	76

## ÍNDICE DE TABLAS

Tabla 1. Comparativa de las características de los sistemas biométricos.....	15
Tabla 2. Costo del sistema de marcaje propuesto .....	72

## ÍNDICE DE ILUSTRACIONES

Figura 1. Arquitectura de un sistema biométrico para identificación personal .....	9
Figura 2. Gráfica típica de la tasa de falso rechazo (FRR) y la de falsa aceptación (FAR) como funciones del umbral de aceptación $u$ para un sistema biométrico.....	12
Figura 3. Diversas características de reconocimiento en sistemas biométricos.....	13
Figura 4. División de las características biométricas para identificación personal.....	14
Figura 5. Variación infinita de un parámetro continuo en función del tiempo .....	16
Figura 6. Representación simplificada del barrido progresivo.....	17
Figura 7. Barrido entrelazado 2:1. Cuadro completo.....	18
Figura 8. Los campos de un entrelazado 2:1 .....	18
Figura 9. Obtención de la señal $Y$ .....	20
Figura 10. Valores de luminancia relativa .....	20
Figura 11. Representación de los colores en NTSC .....	22
Figura 12. Muestreo de una señal análoga $e(t)$ por un tren de impulsos $u(t)$ .....	23
Figura 13. Banda base de la señal de video .....	24
Figura 14. Espectro de una señal de video muestreada a la frecuencia $f_o$ .....	24
Figura 15. Cuando la frecuencia de muestreo es $f_o > 2f_s$ .....	25
Figura 16. Cuando la frecuencia de corte del filtro PB es superior a $f_o - f_s$ .....	25
Figura 17. Error de cuantización.....	26
Figura 18. Codificación de la señal compuesta.....	27
Figura 19. Codificación de componentes.....	27
Figura 20. Transición de análogo a digital de las señales compuestas.....	28
Figura 21. Cuantización de la señal de luminancia .....	30
Figura 22. Cuantización de la señal de crominancia.....	30
Figura 23. Ángulo de la irradiación y reflectancia de la imagen .....	32
Figura 24. Diagrama principal de una cámara digital .....	33
Figura 25. CCD .....	33
Figura 26. Célula Peltier .....	34
Figura 27. Pasos del procesamiento de imágenes .....	38
Figura 28. Ejemplos de histograma .....	40
Figura 29. Imagen con dos modas dominantes .....	45

Figura 30. Seguimiento de un contorno .....	47
Figura 31. Transformada de Hough. (a) Línea recta y (b) su transformada .....	48
Figura 32. Perfiles de bordes.....	49
Figura 33. Procesos de tratamiento de la imagen.....	54
Figura 34. Proceso de comparación de los histogramas .....	55
Figura 35. Transición de estados de un modelo oculto de Markov .....	57
Figura 36. Arquitectura de un Modelo oculto de Markov .....	58
Figura 37. Vector propio y valor propio .....	60
Figura 38. Estados de un electrón en un átomo de hidrógeno.....	64
Figura 39. Eigen caras .....	64
Figura 40. Elementos físicos del sistema de marcaje .....	68
Figura 41. Flujo de la información del sistema de marcaje .....	69

## RESUMEN

Responsable es aquel que conscientemente es la causa directa o indirecta de un hecho y que, por lo tanto, es imputable por las consecuencias de ese hecho. La responsabilidad es la posibilidad de ser considerado sujeto de una deuda u obligación.

Actualmente ese valor se ha perdido, debido a cambios culturales en la sociedad, ya no se puede confiar en que los individuos sean responsables y honrados y que cumplan las reglas en sus lugares de trabajo; es necesario vigilar el ingreso y egreso de cada persona de las instituciones de una forma más eficiente para que se respeten los horarios de trabajo, por lo que se convierte en una necesidad la utilización de herramientas de control que permitan un mejor control y más alta confiabilidad.

A continuación se presenta una propuesta para poder controlar los ingresos y egresos del personal de cualquier institución, a través de la manipulación de imágenes captadas por una cámara de video digital, almacenadas en una base de datos multimedia y que permiten reconocer a cada persona que ingresa a la institución o que sale de ella, pudiendo guardar escenas de las personas que no han sido reconocidas por alguna modificación importante en su apariencia física para una posterior retroalimentación al personal encargado del control de personal, con lo cual se reducen los errores y se aumenta la confiabilidad.

Dicho sistema puede ser implementado también como parte de sistemas de seguridad, en condominios o casas particulares para el acceso a los mismos a través del reconocimiento facial del núcleo familiar, pudiendo ser colocado también en garitas o ingreso de universidades.

## 1. INTRODUCCIÓN

Administrar el recurso humano en una empresa es una tarea difícil, debido a que es necesario tener muchos controles sobre el personal, pero al mismo tiempo se le debe hacer sentir que no está siendo controlado con el fin de que exista armonía y un buen ambiente de trabajo; actualmente se ha podido observar que cada persona trabaja de una forma más eficiente y con más motivación si su trabajo es interesante, si se les hace partícipes de la empresa, pero sobre todo, si su ambiente de trabajo es agradable y propicio para su desarrollo. Un ejemplo típico de estos controles, es el registro del ingreso y egreso del personal de la institución. Dicho registro puede ocasionar incomodidad y hasta problemas serios en el establecimiento si no se realiza adecuadamente.

Hoy por hoy existe un control de este tipo sobre los profesionales que laboran en muchas de las instituciones de Guatemala, cada persona debe registrar su ingreso y egreso de la institución por medio de un sistema de reconocimiento que registra ciertas medidas de las manos de cada trabajador o de las huellas dactilares, para posteriormente grabar la hora y la fecha en que se realiza dicho registro.

Este tipo de sistemas presentan varios inconvenientes en diversas situaciones, como cuando existe mucho calor o mucho frío y se observa un cambio en el volumen de las características corporales que serán medidas o si éstas cambian drásticamente, no permitiendo que algunas personas registren su entrada o salida de las instituciones, provocando problemas tanto para el personal como para los encargados de verificar ese registro, perdiéndose por completo el control. Estos inconvenientes evolucionan en problemas mayores, puesto que para el patrono el trabajador no se presentó a sus labores y para el trabajador el patrono no reconoce su tiempo trabajado.

Por lo tanto es necesario encontrar un método más confiable para el registro del ingreso y salida del personal de la institución, evitando así la pérdida del control y la incomodidad que surge con la pérdida del mismo.

Este trabajo consiste en una propuesta de diseño de un sistema de marcaje biométrico que evaluará los rostros del personal que registra el ingreso o egreso en una institución determinada para luego buscar en una pequeña base de datos la similitud de la imagen captada por una cámara de video con la imagen almacenada previamente en ella; luego de un reconocimiento afirmativo, el sistema guardará la hora exacta y el día en el que se está realizando ese ingreso o egreso.

Además el sistema tendrá la capacidad de grabar la imagen de las personas que hayan intentado registrarse sin ningún éxito, concatenando la hora y la fecha en que se dieron esos sucesos, con el fin de llevar un control más completo y evitar posteriores problemas por falta de evidencia en el registro del personal. Este tipo de control será menos incómodo y en cierta forma dará un aspecto más amigable.

## 2. SISTEMAS DE MARCAJE

Un sistema de marcaje es una herramienta administrativa que se utiliza para generar registro y control del personal, el cual consiste en la puesta en práctica de una serie de Instrumentos, con la finalidad de registrar y controlar al personal que labora en una determinada empresa o institución.

Para que las acciones o actividades empresariales se cumplan, es necesario que haya un adecuado registro y control del capital intelectual. Con el control y registro del personal, se trata de asegurar que las diversas unidades de la organización marchen de acuerdo con lo previsto. Los objetivos centrales de esta técnica son controlar las entradas y salida del personal, verificar el cumplimiento del horario de trabajo, controlar horas extras, registrar permisos, vacaciones, tardanzas, licencias, etc. Este proceso técnico se aplica desde el momento en que el colaborador ingresa a laborar a la institución, ya que su ingreso debe registrarse en una ficha personal preelaborada por el área de desarrollo de recursos humanos o en una base de datos destinada para tal tarea, complementándose inmediatamente con su elemento de identificación personal o con el ingreso inicial de información biométrica.

Los instrumentos técnicos de registro y control del desarrollo del recurso humano, son establecidos de acuerdo a las necesidades, naturaleza y exigencias de la empresa o institución, siendo dentro de los más utilizados los siguientes:

- Ficha personal.
- Ficha social.
- Tarjeta o registro de control diario de asistencia.
- Parte diario de asistencia.
- Papeletas de autorización de salidas.
- Tarjetas o registros de control de récord laboral.
- File personal.
- Rol vacacional.
- Cuadro de asignación de personal.
- Reglamento interno de trabajo

Los Instrumentos técnicos citados, sirven para controlar, registrar y manejar las ocurrencias laborales de los colaboradores, sean estos funcionarios, ejecutivos, administrativos, técnicos, auxiliares y obreros. El registro de personal es una función técnica importante para llevar a cabo la política de los recursos humanos, se lleva específicamente en el *file* o carpeta personal y, ficha familiar; este proceso técnico comienza en forma sucesiva, apenas el colaborador ingresa a laborar.

El registro y control de personal constituye una fuente importante de consulta sobre datos personales del colaborador, asistencia e inasistencia, puntualidad, vacaciones, licencias, permisos, ascensos y promociones entre otros, por tal motivo tiene que ser actualizada permanentemente. Su finalidad es servir de fuente de información para determinar la situación actual del colaborador, y para la toma de decisiones en cuanto a:

- Promoción o ascensos.
- Conceder becas.
- Otorgar derechos.
- Compensación por tiempo de servicios.
- Movimientos de rotación, o reubicación.

**2.1. Sistemas de marcaje tradicionales.** Los sistemas de marcaje tradicionales se han venido utilizando desde hace ya algunas décadas para controlar el horario de ingreso y egreso del personal laboral a un determinado sitio de trabajo. El primer sistema de este tipo que fue utilizado, es el de marcaje por tarjeta, el cual maneja una simple cartulina como herramienta de control, la cual cuenta con un formato impreso a cuadros que delimitan los diferentes días de cada mes, siendo estos disponibles para la impresión de la hora de entrada y salida del personal, la cual es proporcionada por un reloj de marcaje especial. Dicho sistema presenta muchos problemas, debido a la mala utilización del mismo y a las malas prácticas del personal, pues la tarjeta de cualquiera de los empleados podía ser impresa por algún otro compañero, alejándose así, el sistema de marcaje, de su objetivo primordial, controlar al personal.

A raíz de este tipo de fraudes y apoyándose en la tecnología, fueron diseñados nuevos sistemas que vendrían a eliminar los posibles engaños y a mejorar el control sobre el personal laboral, entre los cuales se pueden citar los siguientes:

- Tarjeta de banda magnética; es una tarjeta especial plástica con una banda magnética que guarda información sobre la persona que la porta; para realizar el marcaje, cada individuo debe deslizar la tarjeta sobre una ranura localizada en un reloj de marcaje, el cual al recibir los datos magnéticamente almacenados en la banda especial de la tarjeta, realiza un registro en una base de datos con los datos del individuo, la fecha y la hora en que se realizó dicha acción.
- Reloj de control con tarjeta de proximidad; utiliza la tecnología RFID en la cual se hace uso de un chip que almacena electrónicamente información que al pasar cerca del reloj de marcaje, este realiza una lectura del mismo por medio de radiofrecuencia, registrando la información que tiene grabada en una base de datos, junto con la fecha y la hora en que ocurre este evento.
- Tarjeta de código de barras; cada tarjeta guarda un número correlativo único impreso en código de barras que es necesario pasarlo en frente del lector (en este caso el reloj de marcaje) para que registre dicho código, así como la hora y la fecha de ese registro; el código es asignado a un único individuo para que pueda ser identificado y controlado. Reloj de marcaje de ingreso de datos digitales; este es un dispositivo en el cual se ingresan códigos específicos para cada trabajador mediante un teclado numérico y que, cuando se realiza el marcaje, guarda el código asociado a cada empleado junto con la fecha y la hora en que se digitaron los datos.

Estos sistemas utilizan la tecnología para cumplir su cometido y tienen la ventaja de poder controlar cerraduras eléctricas y así poder ingresar al lugar de trabajo creando una dependencia hacia los mismos para que cada individuo pueda cumplir con el horario asignado de trabajo, eliminando así el problema de los fraudes y engaños. A la vez surge una desventaja sobre la misma dependencia mencionada, resultando que por necesitar de estos dispositivos, el sistema puede perder de vista sus objetivos cuando el dispositivo de marcaje es olvidado, extraviado, averiado o inclusive robado, por lo que fue necesario crear sistemas que no dependan de ningún dispositivo extra que deba ser portado, sino que dependa de la medición de características físicas de cada individuo.

**2.2. Sistemas de marcaje biométricos.** Con la evolución de las tecnologías asociadas a la información las labores que tradicionalmente eran realizadas por seres humanos son, gracias a las mejoras tecnológicas, realizadas por sistemas automatizados. Dentro de la amplia gama de posibles actividades que pueden automatizarse, aquella relacionada con la capacidad para establecer la identidad de los individuos, ha cobrado una gran importancia, dicha actividad es denominada **biometría**. La biometría es la ciencia que se dedica a la identificación de individuos a partir de una característica anatómica o un rasgo de su comportamiento.

Una característica anatómica tiene la cualidad de ser relativamente estable en el tiempo, tal como una huella dactilar, la silueta de la mano, patrones de la retina o el iris o los rasgos faciales. Un rasgo del comportamiento es menos estable, pues depende de la disposición psicológica de la persona, por ejemplo la firma. No cualquier característica anatómica puede ser utilizada con éxito por un sistema biométrico. Para que esto así sea debe cumplir con las siguientes características:

- Universalidad: cualquier persona posee esa característica;
- Unicidad: la existencia de dos personas con una característica idéntica tiene una probabilidad muy pequeña;
- Permanencia: la característica no cambia en el tiempo; y
- Cuantificación: la característica puede ser medida en forma cuantitativa.

Para comprobar que dos rasgos medidos biométricamente corresponden o no a la misma persona se lleva a cabo un procedimiento que comienza con la clasificación de los rasgos y termina con el *matching* o comparación de ambos.

La clasificación corresponde a un análisis a escala "gruesa" de los patrones globales de los rasgos mencionados, que permiten asignarlos a un conjunto predeterminado o *clase*, lo que se traduce en una partición de la base de datos a ser revisada. Por otro lado, el *matching* es una comparación a escala "fina" de las características que están siendo medidas, con el fin de encontrar el grado de similitud entre ellas.

Un sistema de marcaje biométrico es un medio automatizado que realiza labores de biometría, es decir, un sistema que fundamenta sus decisiones de reconocimiento mediante una característica personal que puede ser reconocida o verificada de manera automatizada.

Cualquier proceso de identificación personal puede ser comprendido mediante un modelo simplificado que postula la existencia de tres indicadores de identidad que definen el proceso de identificación:

- Conocimiento: la persona tiene conocimiento (por ejemplo: un código),
- Posesión: la persona posee un objeto (por ejemplo: una tarjeta), y
- Característica: la persona tiene una característica que puede ser verificada (por ejemplo: una de sus huellas dactilares, rasgos de su cara, etc.).

Cada uno de los indicadores anteriores genera una estrategia básica para el proceso de identificación personal. Además pueden ser combinados con el objeto de alcanzar grados de seguridad más elevados y brindar, de esta forma, diferentes niveles de protección, es necesario tomar en cuenta que distintas situaciones requerirán diferentes soluciones para la labor de identificación personal, por ejemplo, con relación al grado de seguridad, se debe considerar el valor que está siendo protegido así como los diversos tipos de amenazas. También es importante considerar la reacción de los usuarios y el costo del proceso.

Las características básicas que un sistema biométrico para identificación personal debe cumplir pueden expresarse mediante las restricciones que deben ser satisfechas. Ellas apuntan, básicamente, a la obtención de un sistema biométrico con utilidad práctica, considerando lo siguiente:

- El desempeño, se refiere a la exactitud, la rapidez y la robustez alcanzada en la identificación, además de los recursos invertidos y el efecto de factores ambientales y/u operacionales. El objetivo de esta restricción es comprobar si el sistema posee una exactitud y rapidez aceptable con un requerimiento de recursos razonable.

- La aceptabilidad, que indica el grado en que la gente está dispuesta a aceptar un sistema biométrico en su vida diaria. Es claro que el sistema no debe representar peligro alguno para los usuarios y debe inspirar "confianza" a los mismos. Factores psicológicos pueden afectar esta última característica. Por ejemplo, el reconocimiento de una retina, que requiere un contacto cercano de la persona con el dispositivo de reconocimiento, puede desconcertar a ciertos individuos debido al hecho de tener su ojo sin protección frente a un "aparato". Sin embargo, las características anteriores están subordinadas a la aplicación específica. En efecto, para algunas aplicaciones el efecto psicológico de utilizar un sistema basado en el reconocimiento de características físicas será positivo, debido a que este método es eficaz implicando mayor seguridad.
- La fiabilidad, que refleja cuán difícil es burlar al sistema. El sistema biométrico debe reconocer características de una persona viva, pues es posible crear dedos de látex, grabaciones digitales de voz prótesis de ojos, etc. Algunos sistemas incorporan métodos para determinar si la característica bajo estudio corresponde o no a la de una persona viva. Por ejemplo, un sistema basado en el reconocimiento del iris revisa patrones característicos en las manchas de éste, un sistema infrarrojo para chequear las venas de la mano detecta flujos de sangre caliente y lectores de ultrasonido para huellas dactilares revisan estructuras subcutáneas de los dedos.

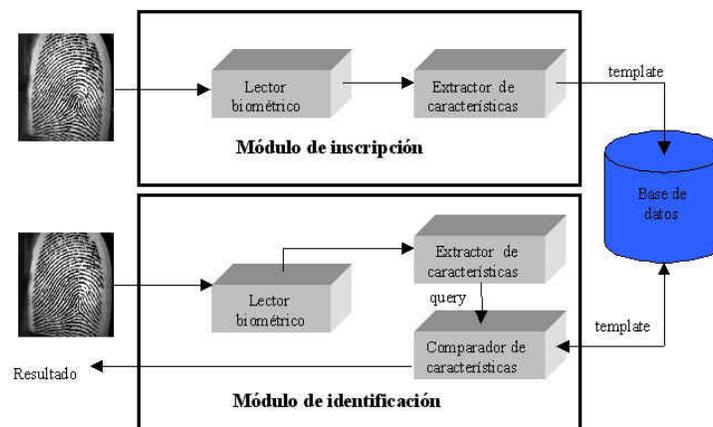
**2.2.1. Arquitectura de un sistema biométrico.** Los dispositivos biométricos poseen tres componentes básicos. El primero se encarga de la adquisición análoga o digital de algún indicador biométrico de una persona, como por ejemplo, la adquisición de la imagen de una huella dactilar. El segundo maneja la compresión, procesamiento, almacenamiento y comparación de los datos adquiridos con los datos almacenados. El tercer componente establece una interfaz con aplicaciones ubicadas en el mismo u otro sistema. La arquitectura típica de un sistema biométrico se presenta en la figura 1, que puede entenderse conceptualmente como dos módulos:

- Módulo de inscripción (*enrollment module*).
- Módulo de identificación (*identification module*).

El módulo de inscripción se encarga de adquirir y almacenar la información proveniente del indicador biométrico con el objeto de poder contrastar a ésta con la proporcionada en ingresos posteriores al sistema. Las labores ejecutadas por el módulo de inscripción son posibles gracias a la acción del lector biométrico y del extractor de características.

El primero se encarga de adquirir datos relativos al indicador biométrico elegido y entregar una representación en formato digital de éste, el segundo extrae, a partir de la salida del lector, características representativas del indicador. El conjunto de características anterior, que será almacenado en una base de datos central u otro medio como una tarjeta magnética, recibirá el nombre de *template*; en otras palabras un *template* es la información representativa del indicador biométrico que se encuentra almacenada y que será utilizada en las labores de identificación al ser comparada con la información proveniente del indicador biométrico en el punto de acceso.

**Figura 1. Arquitectura de un sistema biométrico para identificación personal.**



El módulo de identificación es el responsable del reconocimiento de individuos, por ejemplo en una aplicación de control de acceso; el proceso de identificación comienza cuando el lector biométrico captura la característica del individuo a ser identificado y la convierte a formato digital, para que a continuación el extractor de características produzca una representación compacta con el mismo formato de los *templates*; la representación resultante se denomina *query* y es enviada al comparador de *características* que confronta a éste con uno o varios *templates* para establecer la identidad.

El conjunto de procesos realizados por el módulo de inscripción recibe el nombre de *fase de inscripción*, mientras que los procesos realizados por el módulo de identificación reciben la denominación de *fase operacional*.

### **2.2.2. Fase operacional de un sistema de identificación personal.**

Un sistema biométrico en su fase operacional puede operar en dos modos:

- Modo de verificación
- Modo de identificación

Un sistema biométrico operando en el modo de verificación comprueba la identidad de algún individuo comparando la característica sólo con los templates del individuo, por ejemplo, si una persona ingresa su nombre de usuario entonces no será necesario revisar toda la base de datos buscando el template que más se asemeje al de él, sino que bastará con comparar la información de entrada sólo con el template que está asociado al usuario.

Esto conduce a una comparación uno-a-uno para determinar si la identidad reclamada por el individuo es verdadera o no. De manera más sencilla el modo de verificación responde a la pregunta: ¿eres tú quién dices ser?.

Un sistema biométrico operando en el modo de identificación descubre a un individuo mediante una búsqueda *exhaustiva* en la base de base de datos con los templates, lo que conduce a una comparación del tipo *uno-a-muchos* para establecer la identidad del individuo. En términos sencillos el sistema responde la pregunta: ¿quién eres tú?.

Generalmente es más difícil diseñar un sistema de identificación que uno de verificación. En ambos casos es importante la exactitud de la respuesta, sin embargo, para un sistema de identificación la rapidez también es un factor crítico. Un sistema de identificación necesita explorar toda la base de datos donde se almacenan los templates, a diferencia de un sistema verificador.

**2.2.3. Exactitud en la identificación: medidas de desempeño.** La información provista por los templates permite particionar la base de datos de acuerdo a la presencia o no de ciertos patrones particulares para cada indicador biométrico. Las "clases" así generadas permiten reducir el rango de búsqueda de algún template en la base de datos, sin embargo, los templates pertenecientes a una misma clase también presentarán diferencias conocidas como *variaciones intraclase*, estas variaciones implican que la identidad de una persona puede ser establecida sólo con un cierto nivel de confianza. Una decisión tomada por un sistema biométrico distingue "personal autorizado" o "impostor"; para cada tipo de decisión, existen dos posibles salidas, verdadero o falso, por lo tanto existe un total de cuatro posibles respuestas del sistema:

- Una persona autorizada es aceptada.
- Una persona autorizada es rechazada.
- Un impostor es rechazado.
- Un impostor es aceptado.

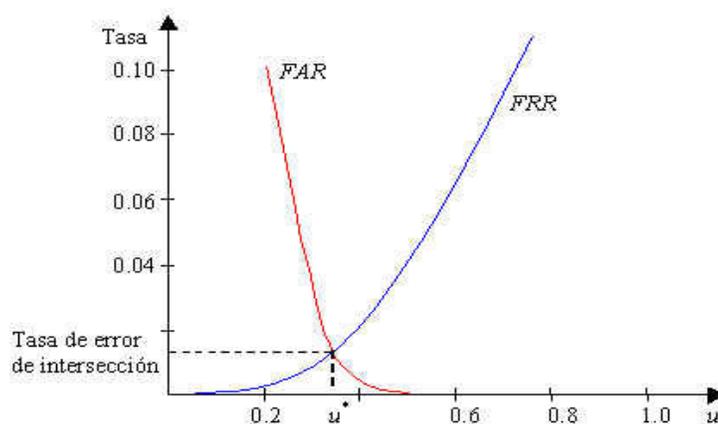
La primera y tercera salidas son correctas, mientras que la segunda y cuarta no lo son. El grado de confianza asociado a las diferentes decisiones puede ser caracterizado por la distribución estadística del número de personas autorizadas e impostores, en efecto, las estadísticas anteriores se utilizan para establecer dos tasas de errores:

- Tasa de falsa aceptación (FAR: *False Acceptance Rate*), que se define como la frecuencia relativa con que un impostor es aceptado como un individuo autorizado.
- Tasa de falso rechazo (FRR: *False Rejection Rate*), definida como la frecuencia relativa con que un individuo autorizado es rechazado como un impostor.

La FAR y la FRR son funciones del grado de seguridad deseado, usualmente el resultado del proceso de identificación será un número real normalizado en el intervalo  $[0, 1]$ , que indicará el "grado de parentesco" o correlación entre la característica biométrica proporcionada por el usuario y la almacenada en la base de datos.

Si, por ejemplo, para el ingreso a un recinto se exige un valor alto para el grado de parentesco (un valor cercano a 1), entonces pocos impostores serán aceptados como personal autorizado y muchas personas autorizadas serán rechazadas. Por otro lado, si el grado de parentesco requerido para permitir el acceso al recinto es pequeño, una fracción pequeña del personal autorizado será rechazada, mientras que un número mayor de impostores será aceptado. El ejemplo anterior muestra que la FAR y la FRR están íntimamente relacionadas, de hecho son duales una de la otra: una FRR pequeña usualmente entrega una FAR alta, y viceversa, como muestra la figura 2. El grado de seguridad deseado se define mediante el umbral de aceptación  $u$ , un número real perteneciente al intervalo  $[0,1]$  que indica el mínimo grado de parentesco permitido para autorizar el acceso del individuo.

**Figura 2. Gráfica típica de la tasa de falso rechazo (FRR) y la de falsa aceptación (FAR) como funciones del umbral de aceptación  $u$  para un sistema biométrico.**

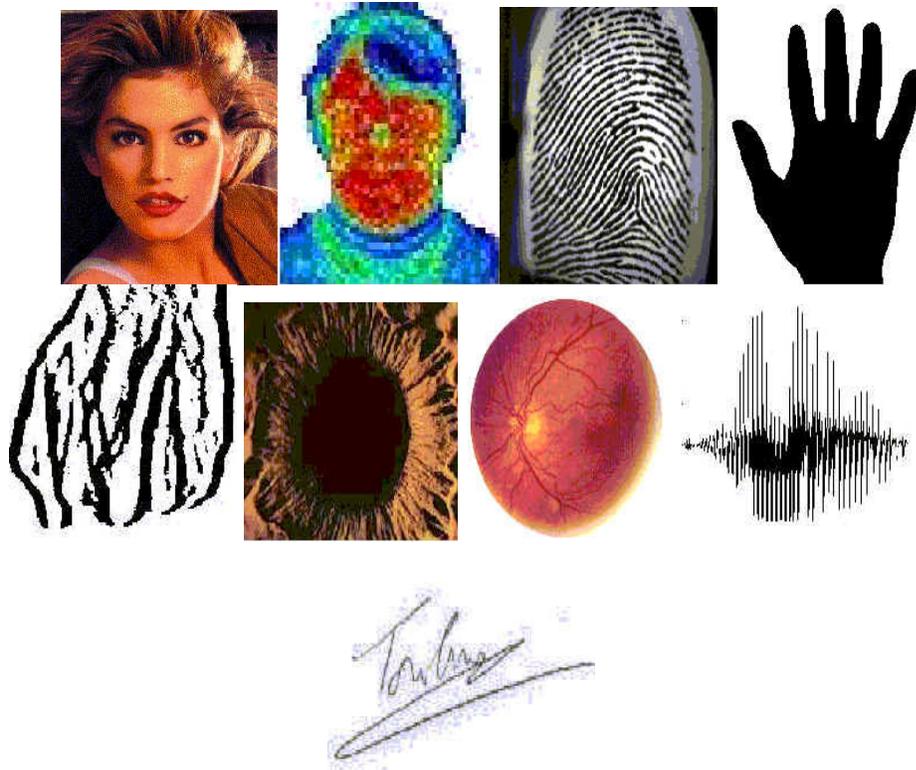


La FRR es una función estrictamente creciente y la FAR una estrictamente decreciente en  $u$ , al ser ambas modeladas como función del umbral de aceptación tienen por dominio al intervalo real  $[0,1]$ , que es además su recorrido, puesto que representan frecuencias relativas. En la figura 2 puede apreciarse un umbral de aceptación particular, denotado por  $u^*$ , donde la FRR y la FAR toman el mismo valor, el cual recibe el nombre de *tasa de error de intersección* (*cross-over error rate*) y puede ser utilizado como medida única para caracterizar el *grado de seguridad* de un sistema biométrico.

**2.2.4. Sistemas biométricos actuales.** En la actualidad existen sistemas biométricos que basan su acción en el reconocimiento de diversas características, como puede apreciarse en la figura 3. Las técnicas biométricas más conocidas son nueve y están basadas en los siguientes indicadores biométricos:

- Rostro.
- Termograma del rostro.
- Huellas dactilares.
- Geometría de la mano.
- Venas de las manos.
- Iris.
- Patrones de la retina.
- Voz.
- Firma.

**Figura 3. Diversas características de reconocimiento en sistemas biométricos.**

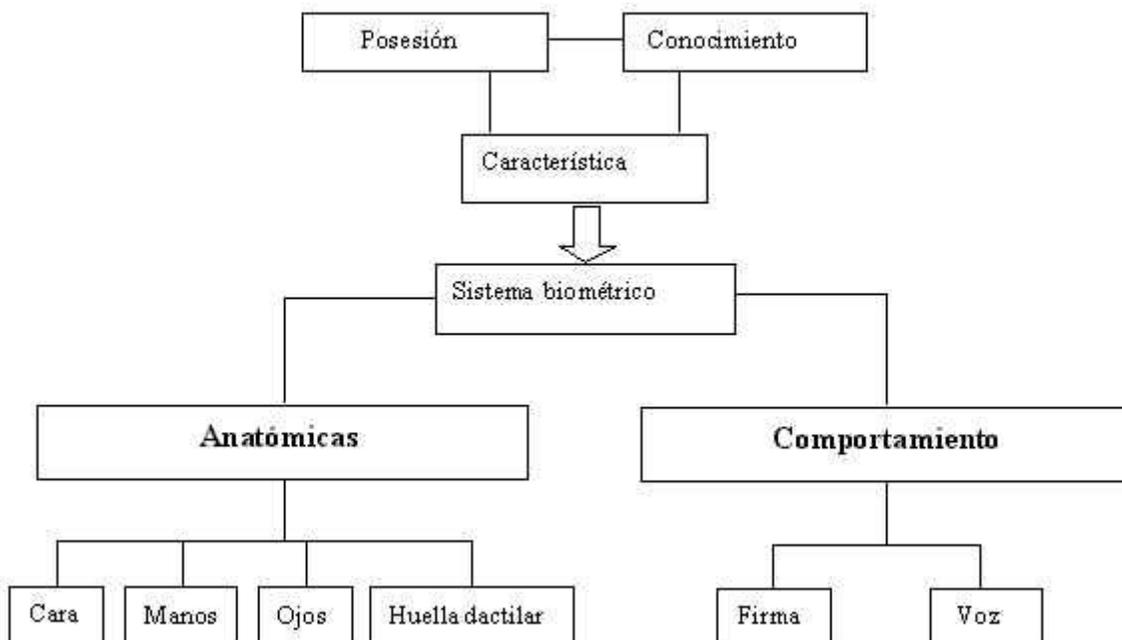


Cada una de las técnicas anteriores posee ventajas y desventajas comparativas, las cuales deben tenerse en consideración al momento de decidir que técnica utilizar para una aplicación específica. En particular deben considerarse las diferencias entre los métodos anatómicos y los de comportamiento.

Una huella dactilar, salvo daño físico, es la misma día a día, a diferencia de una firma que puede ser influenciada tanto por factores controlables como por psicológicos no intencionales; también las máquinas que miden características físicas tienden a ser más grandes y costosas que las que detectan comportamientos.

Debido a diferencias como las señaladas, no existe un único sistema biométrico que sea capaz de satisfacer todas las necesidades, una compañía puede incluso decidir el uso de distintas técnicas en distintos ámbitos. Los resultados alcanzados por el sistema conjunto son mejores que los obtenidos por sus partes por separado, las limitaciones de las alternativas por separado son soslayadas, logrando además respuestas exactas con un tiempo de proceso adecuado. En la figura 4 se presenta un esquema de división de las características biométricas.

**Figura 4. División de las características biométricas para identificación personal**



A continuación se puede ver una tabla comparativa de los distintos factores que se deben analizar a la hora de decidirse por un tipo de biometría en concreto:

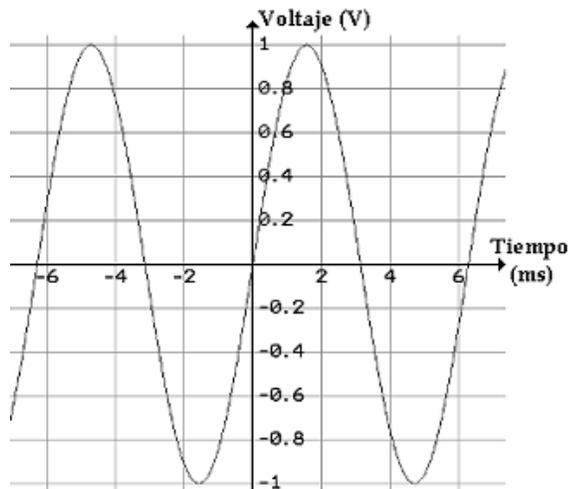
**Tabla 1. Comparativa de las características de los sistemas biométricos.**

<i>Característica</i>	<i>Huellas digitales</i>	<i>Geometría de la mano</i>	<i>Retina</i>	<i>Iris</i>	<i>Cara</i>	<i>Firma</i>	<i>Voz</i>
<i>Facilidad de uso</i>	Alto	Alto	Bajo	Medio	Medio	Alto	Alto
<i>Causas de error</i>	Sequedad, suciedad, edad	Heridas, edad	Gafas	Baja iluminación	Iluminación, edad, gafas, cabello	Variabilidad de la firma	Ruido, resfriados, estado de ánimo
<i>Precisión</i>	Alta	Alta	Muy alta	Muy alta	Alta	Alta	Alta
<i>Coste</i>	Medio	Medio	Alto	Alto	Medio	Medio	Bajo
<i>Aceptación del usuario</i>	Medio	Medio	Medio	Medio	Medio	Medio	Alto
<i>Nivel de seguridad requerido</i>	Alto	Medio	Alto	Muy alto	Medio	Medio	Medio
<i>Estabilidad con el tiempo</i>	Alto	Medio	Alto	Alto	Medio	Medio	Medio

### 3. CARACTERÍSTICAS DE IMAGEN DIGITAL Y SU PROCESAMIENTO

**3.1. 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 5). 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 5. Variación infinita de un parámetro continuo en función del tiempo



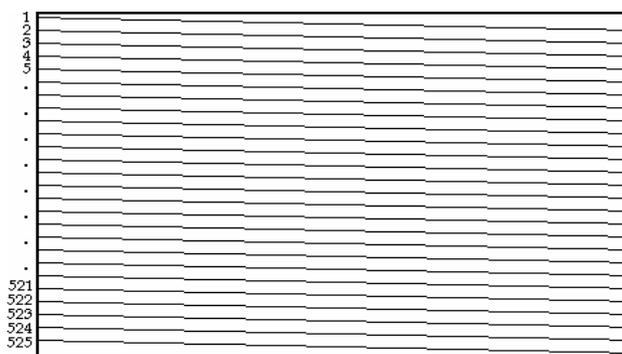
Dichas características suponen la debilidad principal de las señales análogas. Dentro del ancho de banda permisible, cualquier forma de onda, incluso el ruido, es válida. El ruido no puede detectarse en un sistema análogo; se puede tener la sospecha de que existe ruido, pero no se sabe qué proporción 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.

**3.1.1. 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 6).

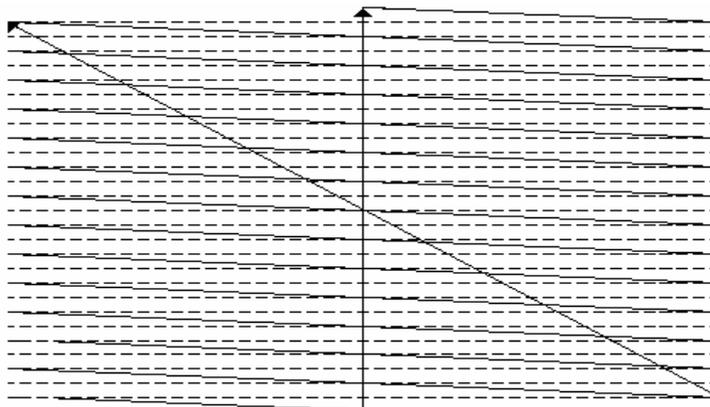
**Figura 6. Representación simplificada del barrido progresivo**



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.

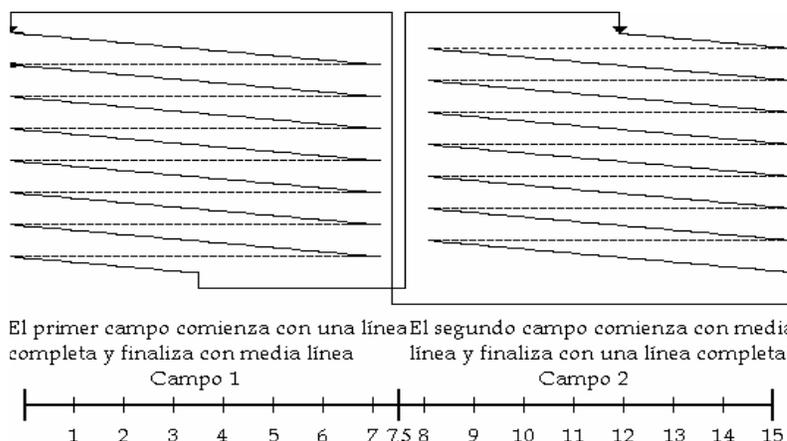
El barrido entrelazado consiste en la transmisión de un primer campo compuesto por las líneas impares de la imagen y a continuación un segundo campo formado por las líneas pares, como se ve en la Figura 7.

**Figura 7. Barrido entrelazado 2:1. Cuadro completo**



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 8, 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 8. Los campos de un entrelazado 2:1**



**3.1.2. 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:

$$Y = Kr * R + (1 - Kr - Kb)G + Kb * B \quad (3.1.2.1)$$

Donde  $Kr$  y  $Kb$  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  $Kr = 0.299$  y  $Kb = 0.114$  por lo que la fórmula (3.1.2.1.) queda como:

$$Y = .0299 R + 0.587 G + 0.114 B \quad (3.1.2.2.)$$

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 9 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:

$$Y = .299(1) + 0.587(1) + 0.114(1) = 1 \text{Lumen} \quad (3.1.2.3.)$$

Figura 9. Obtención de la señal  $Y$

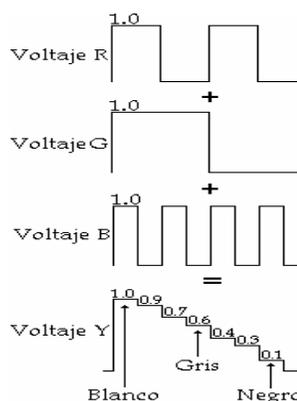
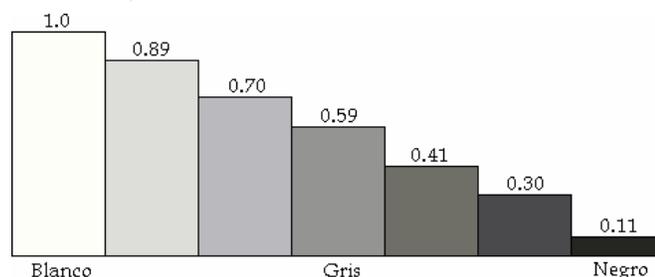


Figura 10. Valores de luminancia relativa



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:

$$R - Y = 0.701R - 0.587G - 0.114B \quad (3.1.2.4.)$$

$$B - Y = -0.299R - 0.587G + 0.886B \quad (3.1.2.5.)$$

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:

$$I = \frac{(R-Y) \cdot \cos 33^\circ}{1.14} + \frac{(B-Y) \cdot \sin 33^\circ}{2.03} \quad (3.1.2.6.)$$

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

$$Q = \frac{(R-Y) \cdot \sin 33^\circ}{1.14} + \frac{(B-Y) \cdot \cos 33^\circ}{2.03} \quad (3.1.2.7.)$$

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:

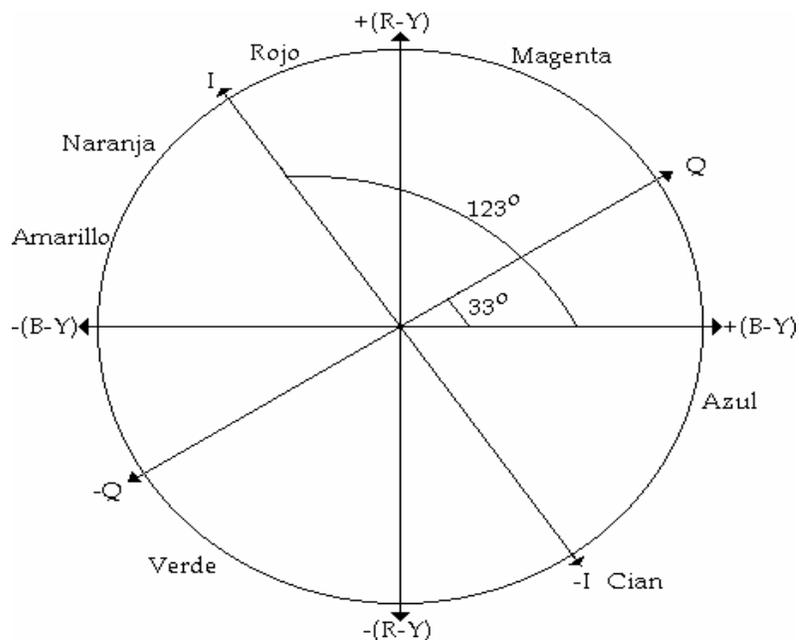
$$C = \sqrt{I^2 + Q^2} \quad (3.1.2.8.)$$

y su ángulo de fase está dado por

$$\theta = \tan^{-1} \frac{Q}{I} \quad (3.1.2.9.)$$

La Figura 11 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 11. Representación de los colores en NTSC



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:

$$Cr = \frac{0.5 * (R - Y)}{(1 - Kr)} \quad (3.1.2.10.)$$

$$Cb = \frac{0.5 * (B - Y)}{(1 - Kb)} \quad (3.1.2.11.)$$

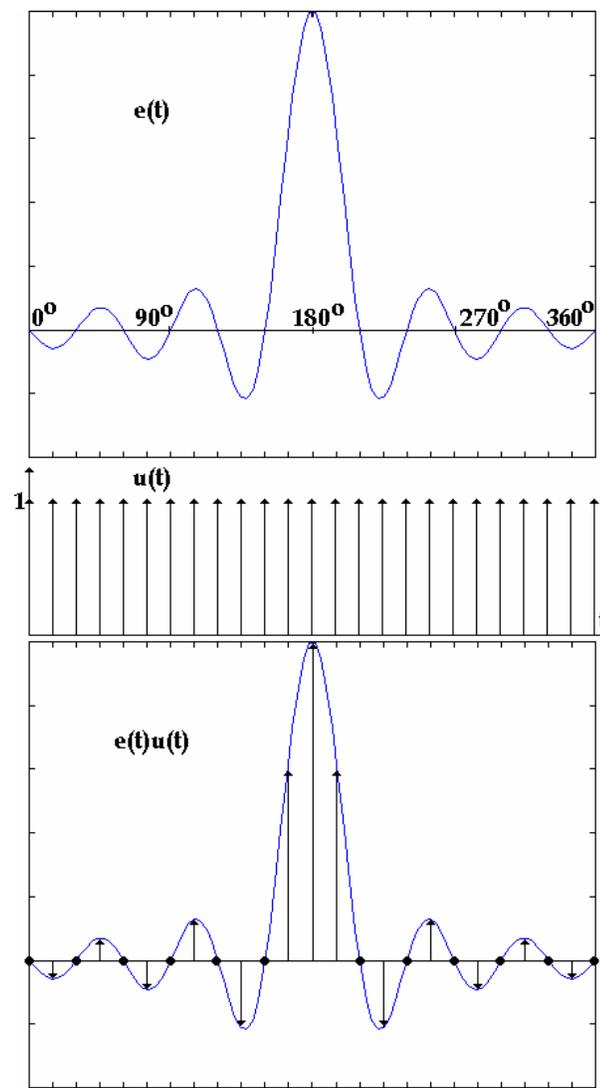
donde  $Kr$  y  $Kb$  son los coeficientes derivados de la corrección Gamma.

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

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

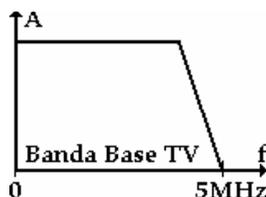
**3.2.1. Muestreo.** Sea una señal análoga  $e(t)$  como la representada en el Figura 12. Se toman muestras breves de  $e(t)$  cada  $15^\circ$  a partir de  $t = 0$ . En  $360^\circ$  se habrán explorado 24 muestras. El resultado será una serie de impulsos cortos cuyas amplitudes siguen a la señal análoga. 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 12.

Figura 12. Muestreo de una señal análoga  $e(t)$  por un tren de impulsos  $u(t)$



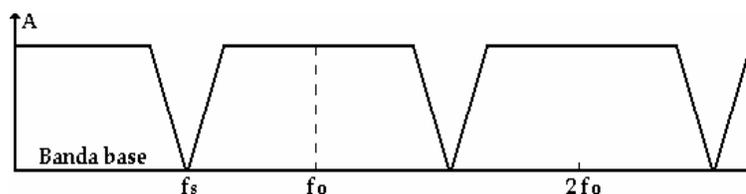
Ahora bien, una señal de video está compuesta por un gran número de frecuencias formando un espectro continuo que va desde 0 a unos 5 MHz. como se representa en la Figura 13.

Figura 13. Banda base de la señal de video



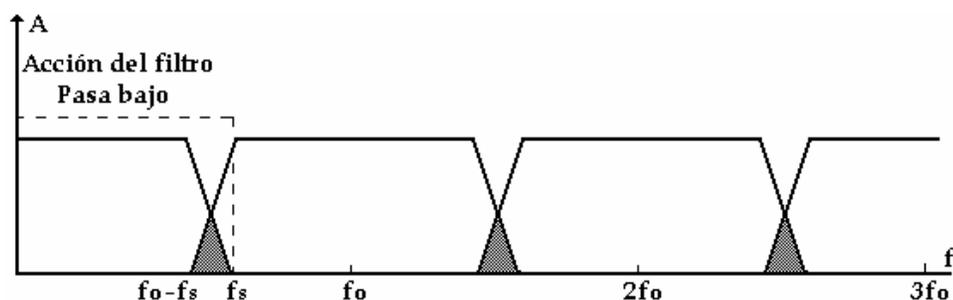
Al muestrear esta señal, cada frecuencia de video aparecerá en las bandas laterales superiores e inferiores de cada armónico de la frecuencia de muestreo, incluyendo naturalmente la banda base, esto es, el armónico cero. El espectro de la señal muestreada se presentará por tanto, como se ve en la Figura 14. 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 14. Espectro de una señal de video muestreada a la frecuencia  $f_o$



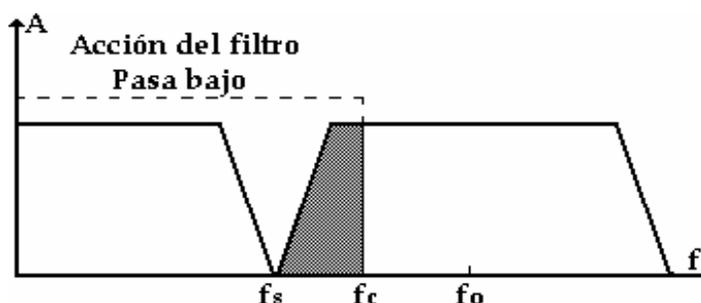
Este razonamiento fue deducido por Nyquist-Shannon, al establecer que para conseguir un muestreo-recuperación sin distorsión, se requiere que la frecuencia de muestreo  $f_o$  sea al menos dos veces más elevada que la frecuencia máxima presente en la señal análoga muestreada. La recuperación de la banda 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 15).

Figura 15. Cuando la frecuencia de muestreo es  $f_o > 2f_s$



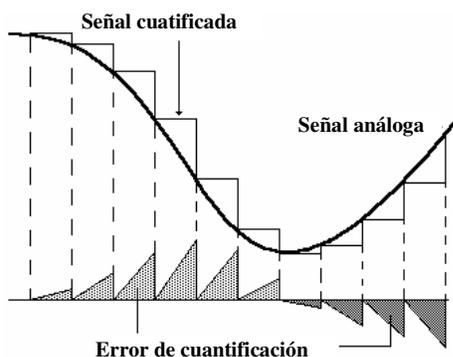
Otro motivo de *aliasing* se produce cuando el filtro no está bien calculado y permite el paso de frecuencias de la banda lateral inferior, aunque no estén solapadas con la banda base (ver la Figura 16).

Figura 16. Cuando la frecuencia de corte del filtro PB es superior a  $f_o - f_s$



**3.2.2. 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 17.

Figura 17. Error de cuantización



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:

$$\frac{S}{C} = (20 \text{Log} N + 10.8) \text{dB} \quad (3.2.2.1.)$$

de cuyo resultado se sacan las siguientes conclusiones:

- La relación señal/ruido de cuantización depende únicamente del número de niveles  $N$  en que se divide el trazo completo de la señal.
- 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:

$$\frac{S}{C} = (6(m) + 10.8) = 58.8 \text{dB} \quad (3.2.2.2.)$$

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:

$$\frac{S}{C} = (6(8)+10.8) = 58.8dB \quad (3.2.2.3.)$$

**3.2.3. 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 18) y la codificación de componentes (ver la Figura 19).

Figura 18. Codificación de la señal compuesta

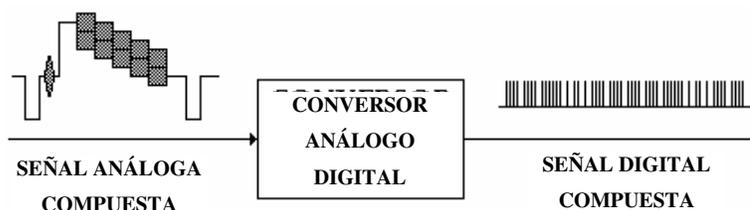
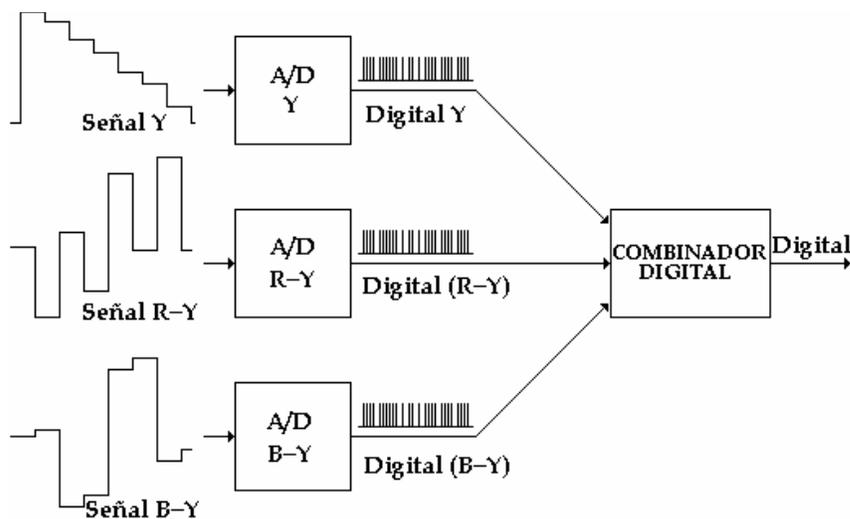


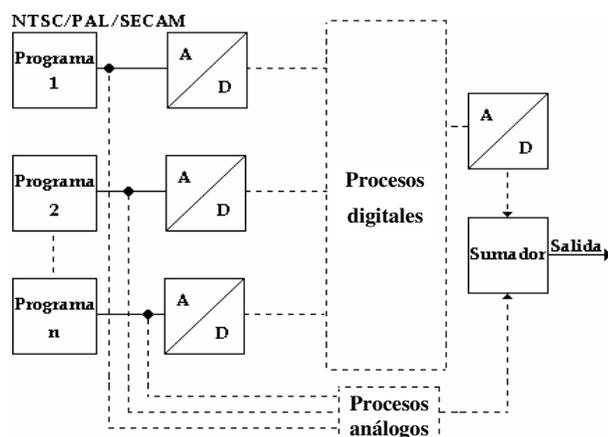
Figura 19. Codificación de componentes



**3.2.3.1. 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 20 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 20. Transición de análogo a digital de las señales compuestas**



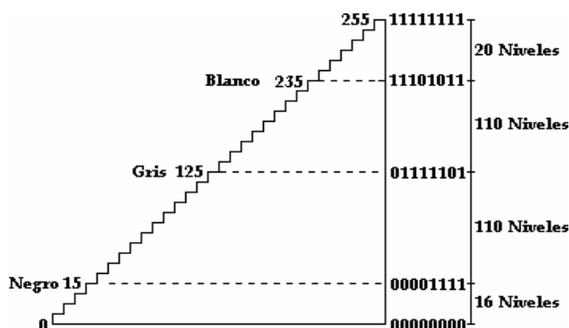
Pasada la transición, se tiene otra ventaja, que es el tratamiento de una señal única de video como ocurre actualmente en los estudios análogos. Para los casos NTSC y PAL que modulan en amplitud a la subportadora de color, el muestreo, la mezcla y el encadenado corresponderán a una sencilla multiplicación 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.

**3.2.3.2. 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.

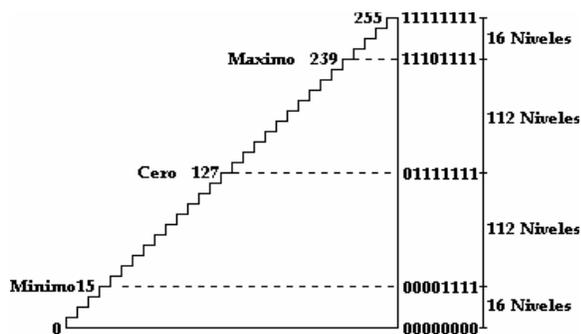
Para esta frecuencia de muestreo, el número de muestras por línea es de 864 y 858 para la luminancia y de 432 y 429 para las diferencias de color (sistema de 625 y 525 líneas respectivamente). La estructura de muestreo es ortogonal, a consecuencia de que la frecuencia de muestreo es un múltiplo entero 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 21).

Figura 21. Cuantización de la señal de luminancia



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 22).

Figura 22. Cuantización de la señal de crominancia



Dado que las señales  $C_r$  y  $C_b$  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.

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.

**3.3. 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.

**3.3.1. 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  $0 < r(x, y) < 1$ .

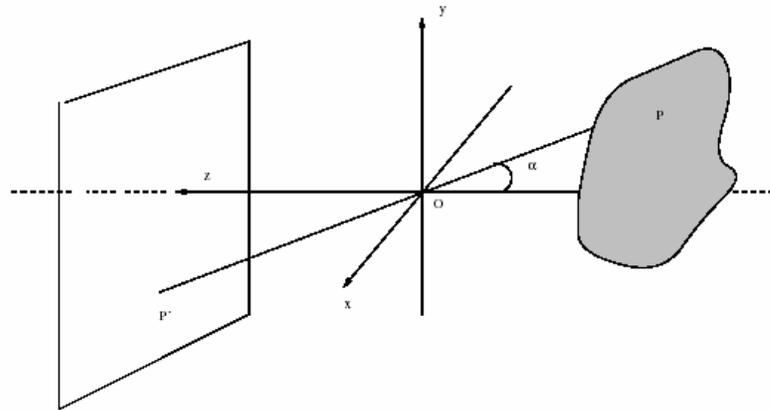
Las ecuaciones anteriores indican que la reflectancia está acotada por 0 (absorción total de luz) y 1 (reflexión total de luz). La naturaleza de  $e(x, y)$  la determina la fuente de luz y la de  $r(x, y)$  es determinada por las características del objeto.

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:

$$e = l \frac{\pi}{4} \left(\frac{d}{f}\right)^2 \cos^4 \alpha \quad (3.3.1.1)$$

donde  $\alpha$  es el ángulo del rayo desde el objeto hasta el centro de la lente como se muestra en la Figura 23.

**Figura 23. Ángulo de la irradiación y reflectancia de la imagen**

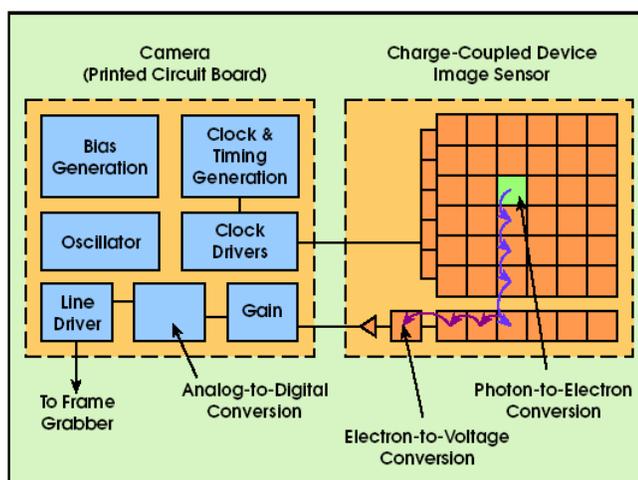


La captura de las imágenes se realiza en dos dimensiones, pero con cierto grado de distorsión, debido a que los sistemas de grabación no son perfectos e introducen distorsiones provocadas por los efectos de muestreo y la cuantificación. El muestreo de una imagen tiene el efecto de reducir la resolución 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.

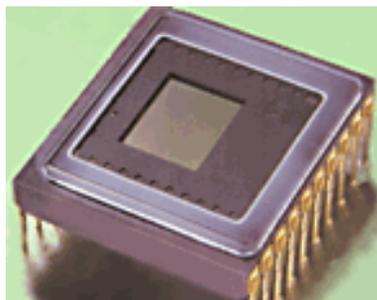
La imagen es captada por una cámara de video digital en una forma muy sencilla, que depende en su mayoría de un CCD que es como el corazón de la cámara. Un CCD es un circuito integrado cuyas siglas significan Dispositivo de Carga Acoplada (*Charge-Coupled Device*), que transforma la información óptica 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 24 y físicamente un CCD se muestra en la figura 25.

**Figura 24. Diagrama principal de una cámara digital**



**Figura 25. CCD**

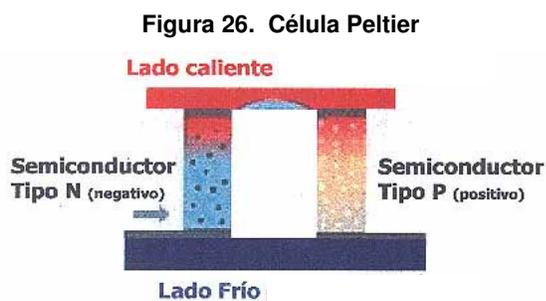


Cada uno de esos puntos de información digital, derivados de imágenes captadas por la cámara, es llamado Píxel. Los Píxeles constituyen la información fundamental para el procesamiento de imágenes o de secuencia de ellas.

**3.3.2. 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 26.



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 anteriormente, 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.

**3.3.3. Relaciones básicas entre píxeles.** En una imagen digitalizada, cada píxel de coordenadas  $(x, y)$  tiene cuatro vecinos horizontales y verticales vecinos cuyas coordenadas son

$$(x+1, y), (x-1, y), (x, y+1), (x, y-1) \quad (3.3.3.1.)$$

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

$$(x+1, y+1), (x-1, y-1), (x-1, y+1), (x+1, y-1) \quad (3.3.3.2.)$$

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  $pANDq$ : 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.

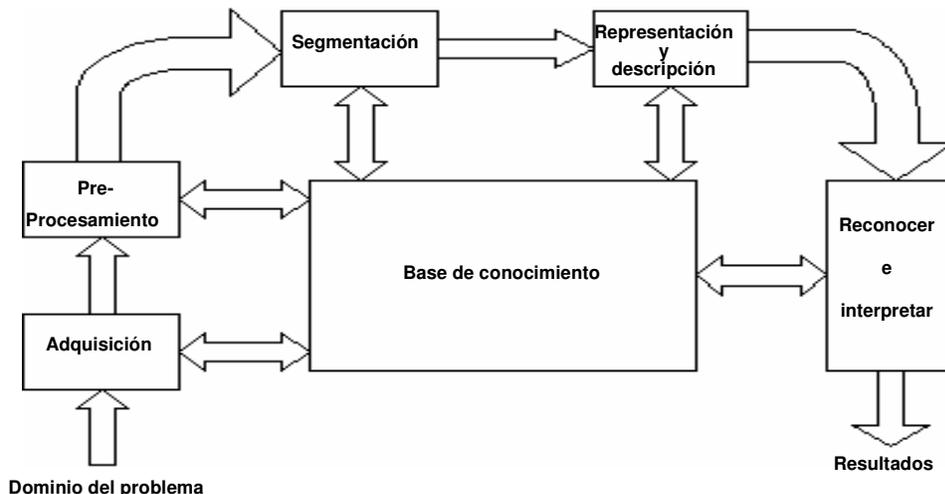
**3.3.4. Pasos para el procesamiento de imágenes.** El primer paso en el proceso es adquirir la imagen digital. Para ello se necesitan sensores y la capacidad para digitalizar la señal producida por el sensor, el cual puede ser una cámara color o monocromática que produce una imagen completa cada 1/30 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. El esquema del proceso completo se muestra en la Figura 27.

Figura 27. Pasos del procesamiento de imágenes



**3.3.5. 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.

**3.3.5.1. 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

$$v = c \log_{10}(1 + u) \quad (3.3.5.1.1)$$

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

$$v = c^1 (u/c)^\lambda \quad (3.3.5.1.2)$$

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.

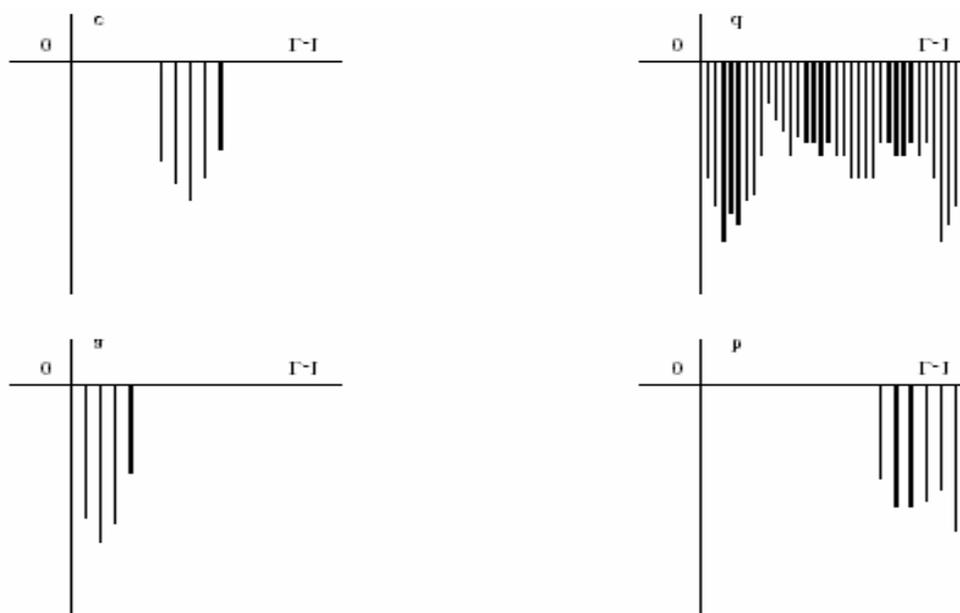
La **diferencia de imágenes** y el **promedio de imágenes** sólo puede suceder entre dos o más imágenes, la primera tiene aplicaciones en segmentación y realce; la segunda se utiliza para obtener una versión de la imagen original con menos ruido, utilizando distintas tomas de la misma escena, basándose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

$$p(r_k) = \frac{n_k}{n} \quad (3.3.5.1.3.)$$

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 28 se pueden observar ejemplos de histograma.

Figura 28. Ejemplos de histograma



El histograma mostrado en la Figura 28(a) muestra que los niveles de gris están concentrados en la zona oscura de la escala, ocurriendo lo contrario en la Figura 28(b); en la Figura 28(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.

**3.3.5.2. 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

$$g(i, j) = \sum_{m,n \in O} h(i-m, j-n) f(m, n) \quad (3.3.5.2.1.)$$

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

$\begin{matrix} 1 & 1 & 1 \\ 1/9 * 1 & 1 & 1 \\ 1 & 1 & 1 \end{matrix}$	$\begin{matrix} 1 & 1 & 1 \\ 1/10 * 1 & 2 & 1 \\ 1 & 1 & 1 \end{matrix}$	$\begin{matrix} 1 & 2 & 1 \\ 1/16 * 2 & 4 & 2 \\ 1 & 2 & 1 \end{matrix}$
---	--	--

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.

**3.3.5.3. 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:

$$g(x, y) = (h * f)(x, y) \quad (3.3.5.3.1.)$$

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

$$G(u, v) = H(u, v)F(u, v) \quad (3.3.5.3.2.)$$

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.

**3.3.6. 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.

**3.3.6.1. 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.

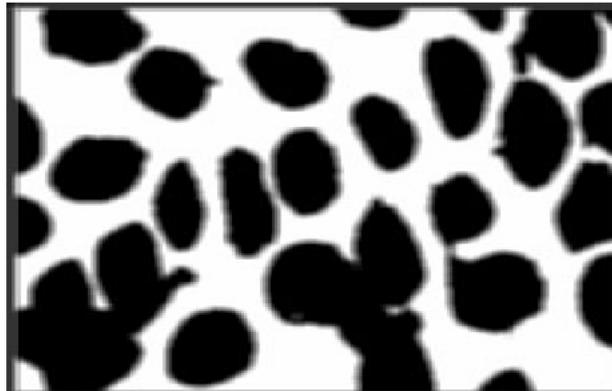
La segmentación también puede realizarse mediante el proceso denominado **disparidad**, el cual se basa en encontrar los píxeles que se encuentran en las fronteras de las regiones; estos píxeles llamados bordes, pueden localizarse examinando los píxeles vecinos, midiendo las diferencias de la intensidad 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.

**3.3.6.2. Segmentación basada en umbralización.** Si se tiene una imagen como en la Figura 29, 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.

$$g(i, j) = \begin{cases} 1 & \text{si } f(i, j) \geq T \\ 0 & \text{si } f(i, j) < T \end{cases} \quad (3.3.6.2.1.)$$

Figura 29. Imagen con dos modas dominantes



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 aislen 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.

**3.3.6.3. 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.

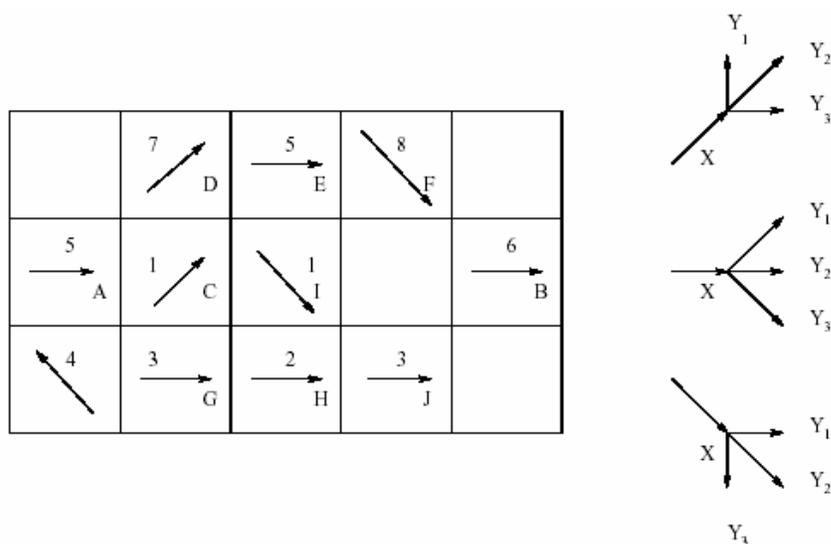
Existen diferentes métodos para la segmentación basada en bordes que difieren en la estrategia para la construcción de la frontera, así como en la cantidad de información previa que se incorpora en el método. Entre las más sencillas están la **Extracción de Frontera, Transformada de Hough y Métodos de 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 30 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 30. Seguimiento de un contorno



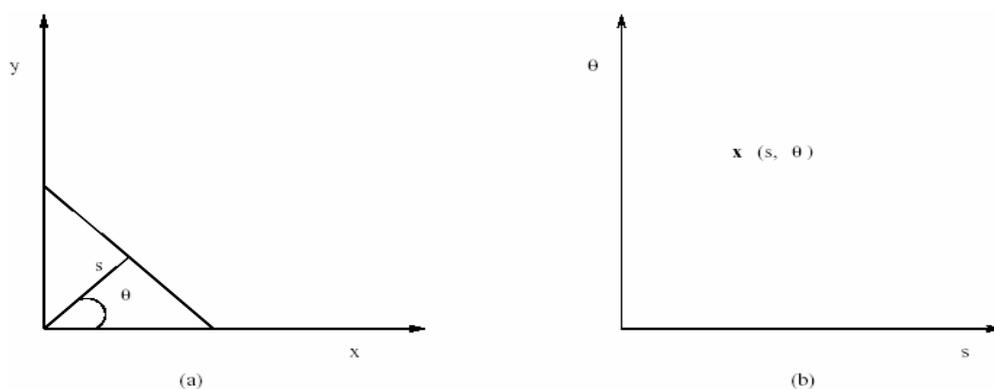
Una frontera puede ser también vista como un camino que se forma a través de un grafo construido uniendo los elementos bordes o es una discontinuidad de algún tipo en la función de intensidad de la imagen.

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  $\theta$  puede representarse mediante

$$s = x \cos \theta + y \sin \theta \quad (3.3.6.3.1.)$$

La transformada de Hough de esta línea es un punto en el plano  $(s, \theta)$ , esto es, todos los puntos de esta línea se aplican en un único punto (ver Figura 30). Este hecho puede usarse para detectar líneas rectas dado un conjunto de puntos frontera.

Figura 31. Transformada de Hough. (a) Línea recta y (b) su transformada



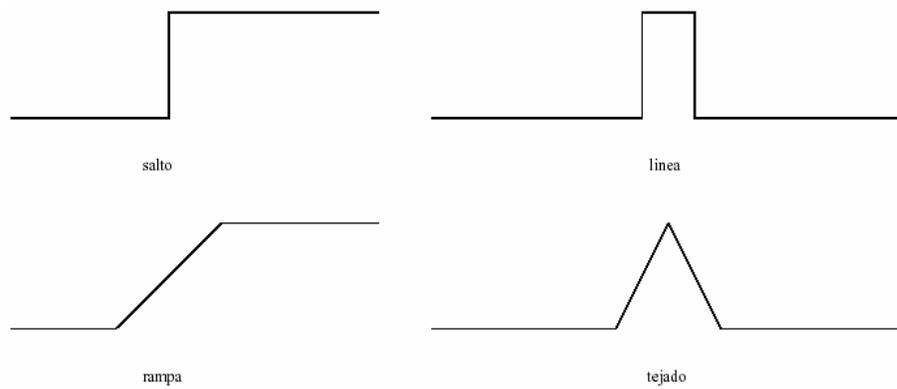
Si se supone que los puntos frontera  $(x_i, y_i), i = 1, 2, \dots, n$  se aplican a algunos valores cuantificados de los parámetros  $s$  y  $\theta$  en el espacio  $(s, \theta)$  y se cuenta  $C(s, \theta)$ , el número de puntos frontera que se aplican en una determinada localización  $(s, \theta)$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

$$C(s_k, \theta_l) = C(s_k, \theta_l) + 1 \text{ si } s_k = x_i \cos \theta_l + y_i \sin \theta_l \quad (3.3.6.3.2.)$$

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 un 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, lo que se observa son rampas y tejados como se ve en la Figura 31.

Figura 32. Perfiles de bordes



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  $\theta$  produciendo  $(x^1, y^1)$  según la fórmula

$$\begin{aligned} x^1 &= x \cos \theta + y \sin \theta \\ y^1 &= -x \sin \theta + y \cos \theta \end{aligned} \quad (3.3.6.3.3.)$$

las derivadas de los dos sistemas se relacionan mediante

$$\begin{aligned} \frac{\partial f}{\partial x^1} &= \frac{\partial f}{\partial x} \cos \theta + \frac{\partial f}{\partial y} \sin \theta \\ \frac{\partial f}{\partial y^1} &= -\frac{\partial f}{\partial x} \sin \theta + \frac{\partial f}{\partial y} \cos \theta \end{aligned} \quad (3.3.6.3.4.)$$

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

$$\left( \frac{\partial f}{\partial x^1} \right)^2 + \left( \frac{\partial f}{\partial y^1} \right)^2 = \left( \frac{\partial f}{\partial x} \right)^2 + \left( \frac{\partial f}{\partial y} \right)^2 \quad (3.3.6.3.5.)$$

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

$$d\left(\frac{\partial f}{\partial x^1}\right) / d\theta = 0 \quad (3.3.6.3.6.)$$

de donde

$$\theta = \arctan\left(\frac{\partial f}{\partial y} / \frac{\partial f}{\partial x}\right) \quad (3.3.6.3.7.)$$

sustituyendo este  $\theta$  en la fórmula de la derivada parcial se tiene que el máximo cambio vale

$$M = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \quad (3.3.6.3.8.)$$

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

$$G_x = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad G_y = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (3.3.6.3.9.)$$

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

$$G(f(i, j)) = |f(i, j) - f(i+1, j+1)| + |f(i, j+1) - f(i+1, j)| \quad (3.3.6.3.10.)$$

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:

$$M = \sqrt{s_x^2 + s_y^2} \quad (3.3.6.3.11.)$$

donde las derivadas parciales se calculan mediante

$$\begin{aligned} s_x &= (a_2 + ca_3 + a_4) - (a_0 + ca_7 + a_6) \\ s_y &= (a_0 + ca_1 + a_2) - (a_6 + ca_5 + a_4) \end{aligned} \quad (3.3.6.3.12.)$$

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

$$\begin{bmatrix} a_0 & a_1 & a_2 \\ a_7 & (i, j) & a_3 \\ a_6 & a_5 & a_4 \end{bmatrix} \quad (3.3.6.3.13.)$$

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

$$\Delta^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (3.3.6.3.14.)$$

$$\frac{\partial^2 f}{\partial x^2} = f(i+2, j) - 2f(i+1, j) + f(i, j) \quad (3.3.6.3.15.)$$

$$\frac{\partial^2 f}{\partial y^2} = f(i, j+2) - 2f(i, j+1) + f(i, j) \quad (3.3.6.3.16.)$$

de donde resulta la siguiente máscara

$$\Delta^2 \approx \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (3.3.6.3.17.)$$

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

## 4. PROCESAMIENTO DE IMÁGENES PARA RECONOCIMIENTO FACIAL

El reconocimiento de rostros a través de medios informáticos es un problema que puede ser abordado utilizando diferentes tipos de características faciales. El principal problema que aparece es que las características varían con pequeñas modificaciones de la posición o el gesto del rostro.

El reconocimiento facial de humanos es difícil de lograr debido a que las representaciones abstractas de las imágenes reales como por ejemplo los negativos fotográficos o las imágenes digitales son afectadas por el movimiento o el cambio en la dirección de iluminación; dicho problema se observa cuando el reconocimiento de rostros se asocia al reconocimiento de imágenes, como se ha hecho tradicionalmente, con lo cual se apunta hacia una simple percepción y se dejan de lado los aspectos del conocimiento.

Reconocer rostros es una actividad que debe realizarse bajo retroalimentación de conocimiento previo, que se aparta mucho del reconocimiento de objetos. En realidad, la especie humana muestra la sociabilidad como una característica especial en su evolución, la identificación de rostros (y expresiones) es fundamental en la integración de un grupo social.

Los niños aprenden a distinguir rostros mucho antes que a distinguir objetos, además el 'talento' de distinguir rostros se desarrolla en forma no supervisada ya que nuestra especie lo desarrolló mucho antes de desarrollar el lenguaje.

Este tipo de consideraciones llevan a plantear métodos de reconocimiento de rostros en los cuales se elimine o por lo menos se disminuyan los efectos dañinos asociados con cambio de iluminación o de posición.

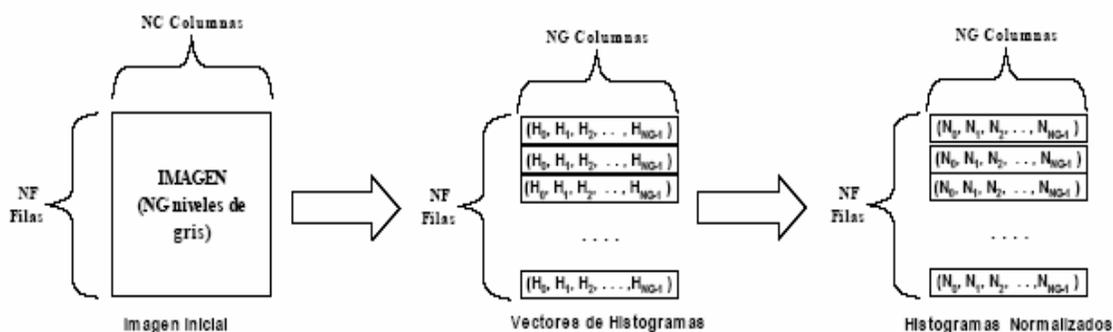
**4.1. Secuencias de histogramas.** En estadística, un histograma es una representación gráfica de una variable en forma de barras, donde la superficie de cada barra es proporcional a la frecuencia de los valores representados.

Se utiliza cuando se estudia una variable continua, como franjas de edades o altura, y por comodidad, sus valores se agrupan en clases, es decir, valores contiguos. El histograma de una imagen representa la frecuencia relativa de los niveles de gris de la imagen. Las técnicas de modificación del histograma de una imagen son útiles para aumentar el contraste de imágenes con histogramas muy concentrados.

Una solución al reconocimiento de rostros es convertir la imagen bidimensional en una secuencia espacio-temporal de vectores. Se trata de partir desde una imagen sobre la cual posteriormente se calcula un histograma de cada una de las filas de píxeles que la componen. Una imagen de  $NF$  filas y  $NC$  columnas con  $NG$  niveles de gris, dará lugar a  $NF$  histogramas, cada uno correspondiendo a una fila, de  $NG$  componentes cada uno. La secuencia ordenada de estos  $NF$  histogramas puede utilizarse para identificar una imagen. Aunque en teoría es posible que dos imágenes diferentes tengan la misma secuencia ordenada de histogramas, la posibilidad de que esta situación se produzca en un proceso de identificación es casi nula. En una secuencia de histogramas normalizados, cada histograma corresponde a una fila de la imagen; en ella se ofrece una información bastante robusta acerca de la luminosidad de la misma, por lo que se convierte en un análisis del rostro como una secuencia que muestra niveles de luminosidad desde la parte superior a la parte inferior de la imagen.

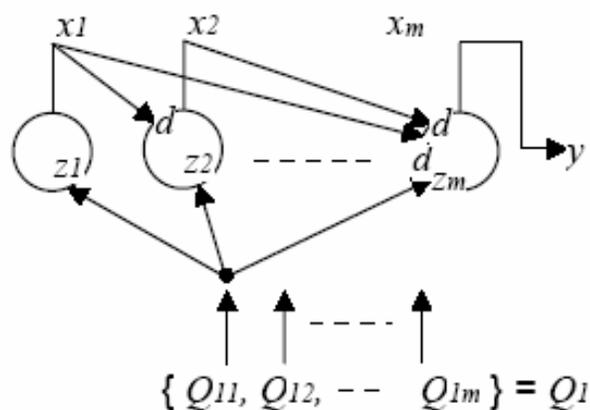
La imagen es la materia prima sobre la cual se calculan los histogramas correspondientes a cada una de las filas de la imagen, y procede a continuación a normalizar dichos histogramas. Finalmente, un último proceso tiene como función elegir el rostro indicado de entre todos los rostros aprendidos.

**Figura 33. Procesos de tratamiento de la imagen**



Cada imagen de entrada se hace corresponder con una secuencia de histogramas (vectores) normalizados; la secuencia de histogramas en el tiempo identificará a la imagen. Sea  $\{Q_{11}, Q_{12}, \dots, Q_{1n}\}$  la secuencia de histogramas normalizados que entra para el reconocimiento de un rostro, entonces el proceso viene dado como se muestra en la figura 34.

Figura 34. Proceso de comparación de los histogramas



Cada uno de los  $Q_{1i}$  es un vector normalizado y se aplica a las entradas de todas y cada una de las unidades; se les permite permanecer allí durante un tiempo  $t$ ; después se aplica el vector siguiente  $Q_{1i+1}$ , y así sucesivamente.

Una capa de entrada con un único elemento de proceso transforma la imagen de entrada, una capa interna, que responde a una red espacio-temporal, construye una secuencia de vectores normalizados y la capa de salida asocia a la imagen con un rostro de entre los aprendidos. La capa de entrada realiza la misión de transformar una imagen de entrada en una secuencia de vectores de histogramas normalizados. La imagen consta de  $NF$  filas y  $NC$  columnas, con  $NG$  niveles de gris. Esta capa calcula el histograma de cada una de las  $NF$  filas, dando como resultado una secuencia de  $NF$  histogramas de  $NG$  componentes cada uno.

Por último, se requieren que su entrada esté normalizada, se normaliza cada uno de estos **NF** histogramas según la ecuación:

$$X = \left( \frac{x_1}{\sqrt{x_1^2 + \dots + x_{NG}^2}}, \frac{x_2}{\sqrt{x_1^2 + \dots + x_{NG}^2}}, \dots, \frac{x_{NG}}{\sqrt{x_1^2 + \dots + x_{NG}^2}} \right) \quad (4.1.1.)$$

La capa intermedia corresponde a una red espacio-temporal, compuesta de **NI\*NF** unidades, donde **NI** es el número de imágenes (rostros) a aprender, y **NF** es el número de elementos (histogramas) de que consta una secuencia.

La capa interna está dividida en grupos de **NF** unidades, cada una de las cuales se encarga de aprender y reconocer los **NF** histogramas correspondientes a una imagen. La salida de la última unidad de cada grupo será el valor de salida del grupo. La unidad de salida de cada grupo que tenga mayor respuesta será la única que tenga salida no nula.

Finalmente, la capa de salida consta de **NF** elementos de proceso que alcanzan rápidamente un valor de equilibrio, es decir que se comparan los elementos procesados con los elementos que se encuentran almacenados, localizando el elemento que se asemeje más a la serie de histogramas o simplemente colocando un valor nulo como consecuencia de no encontrar semejanzas de peso.

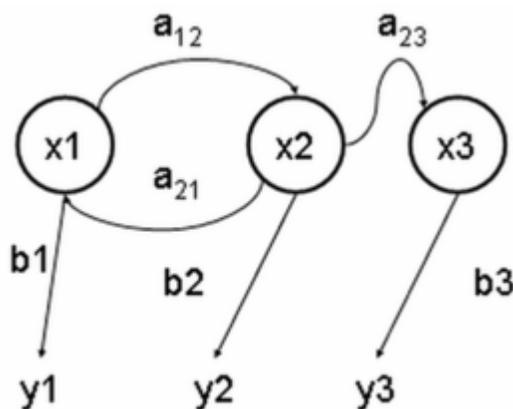
**4.2. Modelo oculto de Markov.** Un modelo oculto de Markov o HMM (por sus siglas del inglés, *Hidden Markov Model*) es un modelo estadístico en el que se asume que el sistema a modelar es un proceso de Markov de parámetros desconocidos.

El objetivo es determinar los parámetros desconocidos (u ocultos, de ahí el nombre) de dicha cadena a partir de los parámetros observables. Los parámetros extraídos se pueden emplear para llevar a cabo sucesivos análisis, por ejemplo en aplicaciones de reconocimiento de patrones.

En un modelo de Markov normal, el estado es visible directamente para el observador, por lo que las probabilidades de transición entre los diversos estados son los únicos parámetros.

En un modelo oculto de Markov, el estado no es visible directamente, sino que sólo lo son las variables que a su vez son influidas por cada estado. Cada estado tiene una distribución de probabilidad sobre los posibles símbolos de salida, como consecuencia, la secuencia de símbolos generada por un HMM proporciona cierta información acerca de la secuencia de los estados. Los modelos ocultos de Markov son especialmente aplicados a reconocimiento de formas temporales, como reconocimiento del habla, de escritura manual, de gestos o bioinformática.

**Figura 35. Transición de estados de un modelo oculto de Markov**

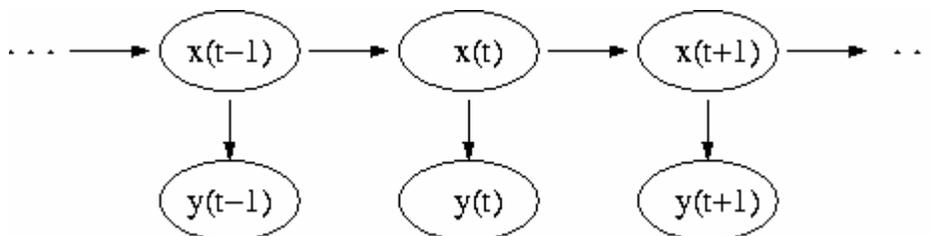


- x — estados ocultos
- y — salidas observables
- a — probabilidades de transición
- b — probabilidades de salida

El diagrama que se encuentra en la figura 36 muestra la arquitectura general de un HMM. Cada óvalo representa una variable aleatoria que puede tomar determinados valores. La variable aleatoria  $x(t)$  es el valor de la variable oculta en el instante de tiempo  $t$ . La variable aleatoria  $y(t)$  es el valor de la variable observada en el mismo instante de tiempo  $t$ . Las flechas indican dependencias condicionales.

El valor de la variable oculta  $x(t)$  (en el instante  $t$ ) solo depende del valor de la variable oculta  $x(t-1)$  (en el instante  $t-1$ ). A esto se le llama la propiedad de Markov. De forma similar, el valor de la variable observada  $y(t)$  solo depende del valor de la variable oculta  $x(t)$  (ambas en el instante  $t$ ).

Figura 36. Arquitectura de un Modelo oculto de Markov



La probabilidad de observar la secuencia  $Y = y(0), y(1), \dots, y(L-1)$  de longitud  $L$  está dada por  $P(Y) = \sum_x P(Y|X)P(X)$  donde la sumatoria se realiza sobre los nodos ocultos  $X = x(0), x(1), \dots, x(L-1)$ .

Los modelos ocultos de Markov se utilizan entonces para encontrar la secuencia más probable de ciertos estados ocultos que puedan haber generado una secuencia de salida dada, a partir de los parámetros del modelo. Por ejemplo, si a un individuo le interesa realizar tres actividades en un determinado día: caminar por la plaza, salir de compras y limpiar su departamento. Lo que él hace depende exclusivamente del estado del tiempo en ese día; no se tiene la información clara acerca del estado del tiempo de donde el individuo vive, pero se conocen las tendencias generales, basándose en lo que hizo en el día, se intenta adivinar el estado del tiempo.

Si el estado del tiempo se comporta como una cadena de Markov discreta, existen dos estados, "Lluvioso" y "Soleado", pero no se les puede observar directamente, es decir, están ocultos. Por lo tanto, existe también una cierta posibilidad de que este individuo haga una de sus actividades cada día, dependiendo del estado del tiempo: "caminar", "comprar" o "limpiar"; dado que el individuo cuenta cuales fueron sus actividades de ese día, esas son las observaciones y el sistema completo es un modelo oculto de Markov.

De esta manera se pueden utilizar los modelos ocultos de Markov para el reconocimiento de rostros, tomando la imagen inicial de entrada que ha sido previamente capturada, como si fueran una secuencia de píxeles de salida que son generados a partir de una imagen determinada, la cual se encuentran dentro de una base de datos pero no se sabe con exactitud cual es y que se desea encontrar y luego evaluando la mejor probabilidad de observación de la secuencia. Este procedimiento es constituido por el algoritmo de Viterbi.

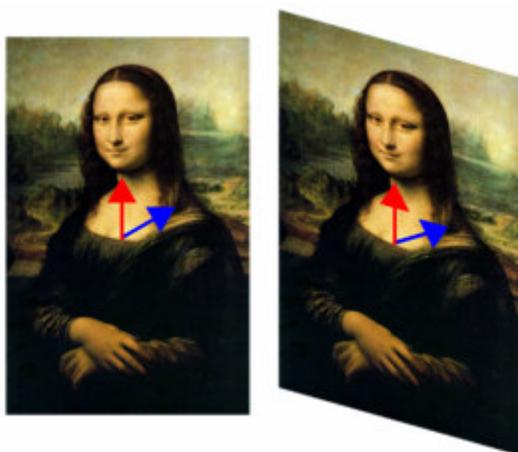
**4.3. Vector propio y valor propio.** Los vectores propios, autovectores o eigenvectores de un operador lineal son los vectores no nulos que, cuando son transformados por un operador, dan lugar a un múltiplo escalar de sí mismos, con lo que no cambian su dirección; este escalar  $\lambda$  recibe el nombre valor propio, autovalor, valor característico o eigenvalor. Un espacio propio, autoespacio o eigenespacio es el conjunto de vectores propios con un valor propio común.

La palabra alemana eigen, que se traduce en español como propio se usó por primera vez en este contexto por David Hilbert en 1904. Eigen se ha traducido también como inherente, característico o el prefijo auto-, donde se logra apreciar la importancia de los valores propios para definir la naturaleza de un evento determinado.

Por ejemplo en la figura 37 se logra observar una transformación en la Mona Lisa, la imagen se ha deformado de tal forma que su eje vertical no ha cambiado, simplemente se han recortado las esquinas en la imagen de la derecha.

El vector azul, representado por la flecha azul que va desde el pecho hasta el hombro, ha cambiado de dirección, mientras que el rojo, representado por la flecha roja, no ha cambiado; este vector rojo es entonces un vector propio de la transformación, mientras que el azul no lo es, ya que el vector rojo no ha cambiado de longitud, su valor propio es 1 y todos los vectores de esta misma dirección son vectores propios, con el mismo valor propio y forman el espacio propio de este valor propio.

**Figura 37. Vector propio y valor propio**



Las transformaciones lineales del espacio como la rotación, la reflexión, el ensanchamiento, o cualquier combinación de las anteriores; pueden interpretarse mediante el efecto que producen en los vectores. Los vectores pueden visualizarse como flechas de una cierta longitud apuntando en una dirección y sentido determinados.

Los vectores propios de las transformaciones lineales son vectores que, o no se ven afectados por la transformación o se ven multiplicados por un escalar que no varía su dirección.

El valor propio de un vector propio es el factor de escala por el que ha sido multiplicado y un espacio propio es un espacio formado por todos los vectores propios del mismo valor propio, además del vector nulo, que no es un vector propio. La multiplicidad de un valor propio es la dimensión del espacio propio asociado; mientras que el espectro de una transformación en espacios vectoriales finitos es el conjunto de todos sus valores propios.

Por ejemplo, un vector propio de una rotación en tres dimensiones es un vector situado en el eje de rotación sobre el cual se realiza la rotación, el valor propio correspondiente es 1 y el espacio propio contiene a todos los vectores paralelos a ese eje, y como es un espacio de una dimensión, su multiplicidad geométrica es uno.

Formalmente, se definen los vectores propios y valores propios de la siguiente manera: si  $A$  es un operador lineal en un cierto espacio vectorial  $V$ ,  $v$  es un vector diferente de cero en  $V$  y  $c$  es un escalar (posiblemente cero) tales que

$$Av = cv \quad (4.3.1.)$$

Entonces se dice que  $v$  es un vector propio del operador  $A$ , y su valor propio asociado es  $c$ . Debe observarse que si  $v$  es un vector propio con el valor propio  $c$  entonces cualquier múltiplo diferente de cero de  $v$  es también un vector propio con el valor propio  $c$ .

A medida que la tierra rota, los vectores en el eje de rotación permanecen invariantes, si se considera la transformación lineal que sufre la tierra tras una hora de rotación, una flecha que partiera del centro de la tierra al Polo Sur sería un vector propio de esta transformación, pero una flecha que partiera del centro a un punto del ecuador no sería un vector propio; debido a que la flecha que apunta al polo no cambia de longitud por la rotación, su valor propio es 1.

Otro ejemplo sería una lámina de metal que se expandiera uniformemente a partir de un punto de tal manera que las distancias desde cualquier punto al punto fijo se duplicasen, esta expansión es una transformación con valor propio 2. Cada vector desde el punto fijo a cualquier otro es un vector propio, y el espacio propio es el conjunto de todos esos vectores.

Para las transformaciones lineales del espacio de dos dimensiones  $\mathbb{R}^2$ , los vectores propios son:

- rotación: ningún vector propio de valores reales.
- reflexión: los vectores propios son perpendiculares y paralelos al eje de simetría, los valores propios son -1 y 1, respectivamente.
- escalado uniforme: todos los vectores son vectores propios, y el valor propio es el factor de escala.

- proyección sobre una recta: los vectores propios con el valor propio 1 son paralelos a la línea, vectores propios con el valor propio 0 son paralelos a la dirección de la proyección

Si se desean calcular los valores propios de una matriz dada y ésta es pequeña, se puede calcular simbólicamente usando el polinomio característico, sin embargo, la mayoría de veces resulta imposible para matrices extensas, por lo que se debe usar un método numérico.

Para encontrar valores propios de matrices cuadradas se utiliza el polinomio característico: decir que  $\lambda$  es un valor propio de  $A$  es equivalente a decir que el sistema de ecuaciones lineales  $(A - \lambda I)v = 0$  (donde  $I$  es la matriz identidad) tiene una solución no nula  $v$  (un vector propio), y de esta forma es equivalente al determinante:

$$\det(A - \lambda I) = 0 \quad (4.3.2.)$$

La función  $p(\lambda) = \det(A - \lambda I) = 0$  es un polinomio de  $\lambda$  pues el determinante se define como suma de productos y los valores propios de una matriz son los ceros de su polinomio característico. Todos los valores propios de una matriz  $A$  pueden calcularse resolviendo la ecuación  $p_A(\lambda) = 0$ .

Si  $A$  es una matriz  $n \times n$ , entonces  $p_A$  tiene grado  $n$  y  $A$  tiene al menos  $n$  valores propios.

El teorema fundamental del álgebra dice que esta ecuación tiene exactamente  $n$  raíces (ceros), teniendo en cuenta su multiplicidad. Todos los polinomios reales de grado impar tienen un número real como raíz, así que para  $n$  impar toda matriz real tiene al menos valor propio real.

Una vez que se conocen los valores propios  $\lambda$ , los vectores propios se pueden hallar resolviendo  $(A - \lambda I)v = 0$ .

En la práctica, los valores propios de las matrices extensas no se calculan usando el polinomio característico, ya que calcular el polinomio resulta muy costoso, y extraer las raíces exactas de un polinomio de grado alto puede ser difícil de calcular y expresar. El teorema de Abel-Ruffini implica que las raíces de los polinomios de grado alto (5 o superior) no pueden expresarse usándose simplemente raíces enésimas, por lo que existen algoritmos eficientes para aproximar raíces de polinomios, pero pequeños errores en la estimación de los valores propios pueden dar lugar a errores grandes en los vectores propios.

Como consecuencia, los algoritmos generales para encontrar vectores propios y valores propios son iterativos. La manera más fácil es usar el método de las potencias, en el que se escoge un vector aleatorio  $v$  y se calcula una secuencia de vectores unitarios es

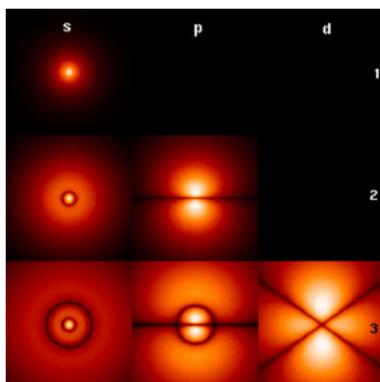
$$\frac{Av}{\|Av\|}, \frac{A^2v}{\|A^2v\|}, \frac{A^3v}{\|A^3v\|}, \dots \quad (4.3.3.)$$

Esta secuencia casi siempre convergerá a un vector propio correspondiente al mayor valor propio.

Un ejemplo interesante se muestra en la función de onda asociada a los estados ligados de un electrón en un átomo de hidrógeno puede verse como los vectores propios del átomo de hidrógeno hamiltoniano así como al operador momento angular.

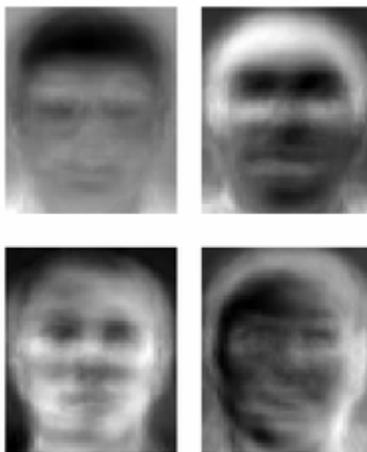
La función está asociada a los valores propios interpretados como sus energías (incrementándose según  $n=1,2,3,\dots$ ) y al momento angular (incrementándose según  $s,p,d,\dots$ ), en la figura 38 se muestra el cuadrado del valor absoluto de las funciones de onda. Las áreas más iluminadas corresponden a densidades de probabilidad más altas para una posición. El centro de cada figura es el núcleo atómico, un protón.

**Figura 38. Estados de un electrón en un átomo de hidrógeno**



En cuanto al procesamiento de imágenes, las imágenes de caras pueden verse como vectores cuyas componentes son la luminancia de cada píxel. La dimensión de este espacio vectorial es el número de píxeles. Los vectores propios de la matriz de covarianza asociada a un conjunto amplio de imágenes normalizadas de rostros se llaman caras propias. Son muy útiles para expresar una imagen de un rostro como la combinación lineal de otras. Las caras propias proporcionan un medio de aplicar compresión de datos a los rostros, para propósitos de biometría.

**Figura 39. Eigen caras**



## 5. PROPUESTA DE DISEÑO

Los temas anteriormente desarrollados son la base para proponer un sistema de marcaje efectivo y seguro apoyándose en los avances tecnológicos y garantizando el registro exacto de los eventos que se desean controlar. Este sistema de marcaje utiliza imágenes faciales como factor para establecer la identidad de los individuos, es decir para implementar el sistema biométrico, de tal forma que se pueda identificar a los sujetos de una forma más cómoda, efectiva y segura, manejando una mínima cantidad de errores, que a la vez podrán ser disipados a través de la decisión humana con la ayuda de evidencia guardada cuando cada error sea generado; su móvil de accionamiento es automático, activado por la aproximación de cada individuo a la cámara del sistema, grabando la imagen facial del mismo y propiciando una serie de procesamientos a ésta para luego compararla con las imágenes encontradas en la base de datos, generando un registro de fecha y hora para esa identidad si es que la comparación entrega un resultado positivo; en caso la comparación obtuviera resultado negativo, la imagen es grabada y un registro de error es generado. Estos registros de error podrán ser visualizados por los auditores del sistema.

**5.1. Elementos del sistema de marcaje.** Los elementos que conforman el sistema de marcaje propuesto están clasificados en la siguiente forma:

- Fuente de alimentación.
- Equipo de comparación biométrica.
- Detectores.
- Visor de sucesos.
- Red.

**5.1.1. Fuente de alimentación.** La fuente de alimentación es 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 posibilidad de registrar los marcajes del personal de la institución.

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 del sistema.

La duración de la energía continua fuera de línea puede ser aumentada al utilizar un banco de baterías en lugar de una sola, dependiendo de la variación que se desee aumentar al tiempo fuera de línea.

**5.1.2. Equipo de comparación biométrica.** El equipo de comparación biométrica es el cerebro del sistema de marcaje y es la parte más importante de éste. De él depende un reconocimiento facial exitoso de los sujetos de los cuales se desea registrar la entrada y salida de la institución, ya que puede reportar errores y fracasos en el intento del reconocimiento si es que no se encuentra correctamente configurado.

El cerebro utilizado en el sistema propuesto es un CPU (*Central Processing Unit*, Unidad Central de Procesamiento) con un microprocesador de 2.4 Gigahertz (Ghz), 512 Megabytes (Mb) de memoria RAM (*Random Acces Memory*, Memoria de Acceso Aleatorio), un disco duro de 160 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 Visual C++, acompañado de las librerías IPL 2.5v (*Image Processing Library*, librería de procesamiento de imágenes) y OpenCV (Open Source Computer Vision, Librería para la implementación de Visión Artificial) de Intel; las cuales se utilizan para el tratamiento y procesamiento de imágenes así como para capturar las mismas por medio de cámaras digitales. Además es necesario como mínimo el sistema operativo Windows XP.

**5.1.3. Detectores.** Son los elementos que tienen la misión de entregar la materia prima al equipo de comparación biométrica para que éste efectúe su cometido; es decir, el accionamiento del sistema a través de proximidad y la digitalización las imágenes para su posterior tratamiento o para el ingreso de un nuevo individuo a la base de datos a través de una cámara de video digital.

Los elementos utilizados en la propuesta son una cámara digital A4 TECH PC con puerto USB, captura de imágenes a 30 cuadros por segundo, de 640x480 píxeles de resolución, con un ángulo de visibilidad de 54 grados, con autofocus y de 42 centímetros cúbicos ( $cm^3$ ) de volumen y un detector de proximidad de Infrarrojos que detecta el movimiento únicamente de personas con una longitud de onda de  $880\text{ nm}$ , utilizan pulsos de  $600\ \mu s$  en una portadora de  $40\text{ Khz}$ .

Los requerimientos mínimos para la cámara digital son: puerto de salida USB y adquisición de imágenes a 30 cuadros por segundo.

**5.1.4. Visor de sucesos.** Es el elemento por medio del cual se logran visualizar las grabaciones de los individuos que no se han logrado identificar. Para tal tarea se utiliza una pantalla por medio de la cual se puede observar dicho evento.

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 comparación o cerebro, en un lugar designado y se emplea para monitorear al personal no reconocido por el sistema. Como requerimiento mínimo para la pantalla se sugiere un monitor de 15" a color.

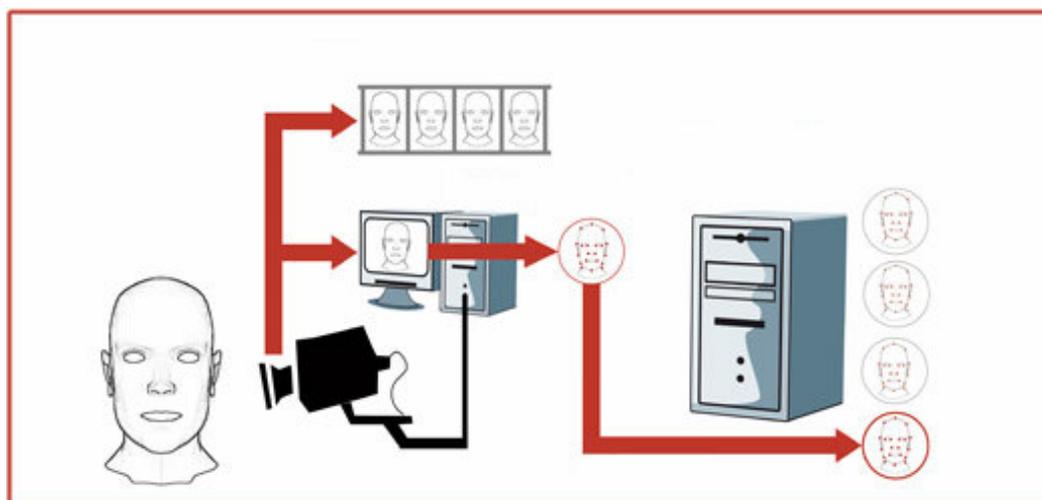
**5.1.5. Red.** La red está definida como los canales por medio de los cuales, los datos recabados por los detectores son transportados por el cerebro del sistema de marcaje. Para el sistema propuesto, se utiliza como bus de datos de la cámara de video hacia el cerebro, un Cable de interfase USB 2.0, de un largo de 2 metros como mínimo y para el detector de proximidad se utiliza un cable paralelo de 2 metros.

**5.1.6. Diagrama del sistema propuesto.** El sistema de marcaje se muestra físicamente en la figura 40 y en forma de flujo en la figura 41.

**Figura 40. Elementos físicos del sistema de marcaje**



Figura 41. Flujo de la información del sistema de marcaje



**5.2. Descripción del diseño.** La propuesta del sistema de marcaje utiliza el detector de infrarrojos como un sensor, haciendo que la cámara de video capte las imágenes a través de sus lentes a razón de 30 cuadros por segundo, como si tomara una fotografía cada 33.33 milisegundos, después de detectar la proximidad de una persona; ya que el detector de infrarrojos se gradúa a una longitud de onda de  $830\text{ nm}$  que hace que los infrarrojos reboten y regresen al sensor solamente en humanos; luego estas imágenes son convertidas en información digital y enviadas al cerebro del sistema, en donde son procesadas y comparadas con la base de datos de rostros, para luego reflejar un resultado de detección positiva o negativa, registrando ambos efectos, la hora en que se dieron y algunos otros datos necesarios.

El cerebro utiliza el programa de ejecución "Face Recognition" para ejecutar la captura y el procesamiento de las imágenes así como el reconocimiento facial, el cual es elaborado por medio del lenguaje de programación Visual C++ utilizando el método de valor propio y vector propio.

Además se utiliza una base de datos SQL para grabar los eventos ya mencionados, de tal forma que se generan registros de marcaje y registros de error, que además apuntan hacia un archivo de video que muestra a la persona que no ha sido reconocida.

El programa hace posible la visualización de los sucesos del reconocimiento en tiempo real dejando mostrar el rostro que se está intentando reconocer, la imagen de las eigenfaces y la imagen del rostro que concuerda con el rostro que ingresa a través de la cámara; en caso que no se tenga éxito en el reconocimiento, aparecerá una pantalla indicando el fallo y la hora de este.

**5.2.1. Descripción del programa.** El programa en lenguaje Visual C++, hace uso de dos librerías especiales, IPL y OpenCV. Cuando el sistema es iniciado, carga los drivers de la cámara comprobando que la misma funcione y adecuando el dispositivo para que las imágenes puedan ser cargadas, luego el sistema pregunta si se desea iniciar la secuencia de escaneo o si se desea ingresar registros de un nuevo individuo. Al elegir el ingreso de una persona a la base de datos el sistema abre una ventana en la cual existe un botón de captura que al ser presionado con un clic del Mouse, muestra un mensaje de conteo hasta diez, que al finalizar hace una captura de la imagen del individuo, pasándola primero por un filtro de realce, luego convirtiéndola a niveles de gris y pasándola por un filtro pasa bajos para posteriormente guardarla en la base de datos de imágenes mostrando un mensaje de éxito al guardarla.

Al finalizar el proceso de aprendizaje, el sistema vuelve a realizar la pregunta de inicio de escaneo o ingreso de registros; a elegir el inicio de escaneo el equipo de comparación se sitúa en un modo de espera, es decir aguardando una señal de interrupción en el puerto paralelo que es generada por el sensor de infrarrojos al detectar a una persona a 50 cm, al suceder esta interrupción la cámara es activada para que inicie la secuencia de grabación de imágenes, capturando exactamente 5 cuadros con imágenes de la persona que se encuentra delante de ella.

Posteriormente las imágenes son pasadas por filtros de realce para destacar los detalles finos de las mismas y mejorar los detalles que han sido emborronados convolucionando cada cuadro de la secuencia con una matriz que contiene coeficientes positivos en el centro y negativos en la periferia de la misma, haciendo siempre que la suma de estos sea cero; luego de pasar por estos filtros, las imágenes son transformadas a niveles de grises para poder utilizar el método de reconocimiento facial, pero pasando antes cada cuadro por un filtro pasa bajos, debido a que la mayor parte del ruido en imágenes a niveles de gris se encuentra en las frecuencias altas.

Al contar ya con una secuencia de imágenes mejoradas, el sistema procede a utilizar el método de vector propio y valor propio para el reconocimiento facial, ejecutando el cálculo de las eigenfaces o eigenvalores de las mismas a través de una función especial de la librería OpenCV; el sistema procede de tal forma que los cinco cuadros que han sido capturados ingresen en esa función para que se calculen su eigenvalores y luego cada valor se compare con las imágenes de la base de datos.

Si el sistema realizó una comparación exitosa, se muestra un mensaje de éxito y se inserta un registro en la base de datos con el nombre de la persona que hizo dicho registro, la hora y la fecha en que lo realizó; si por el contrario no se tuvo éxito, un pequeño video del individuo que realizó el intento de marcaje es grabado para que la persona encargada del registro y control del marcaje pueda hacer una corrección en la base de datos, generando manualmente un registro a la persona que el sistema no pudo reconocer. El sistema cuenta con un reporte mensual por día de los registros de marcaje que se generaron durante un mes determinado.

#### **5.2.2. Factores externos que causan distorsión en el sistema.**

Existen tres factores principales que pueden causar distorsión en el sistema; estos son: la cantidad de luz del entorno, las condiciones climáticas del medio ambiente y la irregularidad en el flujo de energía eléctrica.

En cuanto a la cantidad de luz, es obvia la distorsión, debido a que las imágenes en sí dependen de la cantidad de luz que se refleje en los objetos que serán captados por las cámaras; por lo tanto si existe una mala fuente de luminosidad, las personas 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; aunque dicho problema puede ser solucionado al colocar una fuente de luz cerca de la cámara como parte del sistema.

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 irregularidad en el flujo de energía eléctrica como picos o caídas de voltaje pueden afectar la captura de imágenes pues no propicia el buen funcionamiento de los aparatos electrónicos; además pueden existir sobre corrientes que pueden dañar el sistema si es que no se cuenta con la protección adecuada.

El problema es fácilmente solucionable, ya que se puede utilizar una fuente de energía que incorpore un protector contra picos de voltaje, un regulador de voltaje y fusibles para las sobre corrientes.

**5.2.3. Costo del sistema propuesto.** El costo total del sistema de marcaje con los requerimientos mínimos, a un tipo de cambio de Q.7.80 por USD.1.00, es de Q.7,580.00 distribuido de la siguiente forma:

**Tabla 2. Costo del sistema de marcaje propuesto**

Lugar de cotización	Elementos del sistema propuesto	Costo en Quetzales
Bmart	1 UPS	Q.1,200.00
Bmart	1 CPU	Q.3,000.00
Bmart	1 Monitor	Q.1,800.00
Electrónica Panamericana	1 Sensor infrarojo	Q.150.00
Electrónica Panamericana	1 Cable paralelo para el sensor	Q.50.00
Electrónica Panamericana	1 Cable USB para la cámara	Q.100.00
Bmart	1 Cámara de video	Q.280.00
Bmart	1 Licencia Windows XP	Q.400.00
www.microsoft.com	1 Licencia Visual C ++	Q.600.00
www.intel.com	1 Licencia OpenCV para C++	Gratis
www.intel.com	1 Librería IPL para C++	Gratis
		<b>Q.7,580.00</b>

**5.2.4. Ventajas y desventajas del sistema propuesto.** El sistema de marcaje 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 en imágenes dependiendo un poco de la luz y las condiciones climáticas del medio ambiente y los cambios drásticos en el rostro de las personas, como por ejemplo la utilización de bigote o barba, utilización de anteojos, operaciones faciales de cambio de imagen, etc., pudiendo evitarse en una pequeña porción estos problemas con la actualización diaria de las imágenes de la base de datos.

Entre las ventajas que se tienen con respecto a otros sistemas están: la posibilidad de grabar los sucesos ocurridos; la minimización de errores y aumento de la exactitud, la posibilidad de identificar posteriormente a los individuos que no fueron reconocidos por el sistema para corregir los errores en el registro de ingresos y egresos y para actualizar la base de datos de rostros; una mayor comodidad para los individuos que deben ser controlados; un mayor control sobre el personal de las instituciones; una mucho mayor dificultad para permitir engaños o trasgresión de las reglas y un 99% de confiabilidad.

## 6. CONCLUSIONES Y RECOMENDACIONES

El sistema de marcaje propuesto que utiliza el reconocimiento facial como método biométrico puede producir como resultado del intento de la búsqueda, cierta cantidad de fracasos, aunque en menor medida que los demás sistemas de marcaje ya existentes que utilizan otros métodos biométricos, la diferencia sustancial o mayor valor agregado que brinda dicho sistema, se encuentra en la capacidad de mostrar imágenes grabadas acerca del momento en que se generó el error o fracaso en el sistema, consignando fecha y hora del acontecimiento, con lo que se le da una herramienta poderosa al ente rector de los registros de marcaje de la institución, para poder eliminar cualquier problema en el registro diario de control del personal.

El manejo de imágenes en la base de datos suele producir procesos más lentos y puede llegar a consumir muchos recursos del sistema, dependiendo de la cantidad de archivos de imágenes que es estén manejando y almacenan en la base de datos; la propuesta fue diseñada para el manejo de pocas imágenes, pensando en una institución que maneja un listado de personal entre 60 y 80 personas; por lo que se recomienda utilizar bases de datos multimedia avanzadas con algoritmos de búsqueda más potentes y rápidos en caso que el listado de personal sea mayor a 100 personas, para que el proceso de registro no sea lento y que no consuma todos los recursos disponibles.

Los cambios de iluminación o movimiento en los objetos que se desea capturar son los que más ruido producen en las imágenes por lo que método de procesamiento que se utiliza en el sistema debe presentar la mayor inmunidad posible ante este tipo de cambios bruscos; el método de reconocimiento facial que se utiliza en el sistema, vector propio y valor propio, es el más confiable e inmune a esos cambios, ya que todos los píxeles de cada imagen se toman como vectores y no se ven afectados por la transformación o el movimiento en las imágenes reales, así mismo cada píxel puede ser multiplicado por un escalar ya que no variaría su dirección, dando lugar así a un margen de error más pequeño.

Las imágenes son producto de la luz que incide sobre los objetos reales que se están capturando en la cámara de video, por lo que es imposible que la imagen se

digitalice sin que exista luz en el medio, debiendo existir siempre una fuente de luz para que el sistema pueda funcionar; lo que suele suceder es que existiendo una fuente de luz continua las imágenes pueden tener mucho ruido o pueden tener un exceso de luz (sobrecarga) y uno de los primeros pasos para poder realizar un sistema de este tipo, es convivir con esta dificultad al utilizar filtros y tratamiento especial en las imágenes que serán utilizadas para el sistema.

## 7. BIBLIOGRAFÍA

1. Dalsa Technology With Vision. *CCD vs. CMOS: facts and fiction*. Canada; enero, 2001.
2. De la Escalera Hueso, Arturo. *Visión por computador: Fundamentos y métodos*. España: Editorial Prentice Hall, 2001. 304 pp.
3. Intel Corporation. *Open source computer vision library, reference manual*. USA; diciembre, 2001.
4. Jain, Anil K. *Fundamentals of digital image processing*. USA: Editorial Prentice Hall, 1988. 592 pp.
5. Lawrence Livermore National Laboratory. *Color spaces in digital video*. USA; agosto, 1997.
6. 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.
7. Pajares Martínsanz, Gonzalo y Jesús M. de la Cruz Gracá. *Visión por computador: Imágenes digitales y aplicaciones*. España: Editorial Rama, 2001. 800 pp.
8. Umbaugh, Scott E. *Computer imaging: Digital image analysis and processing*. USA: Editorial Prentice Hall, 2004. 659 pp.
9. Walker, Bruce H. *Optical design for visual systems*. USA: Editorial SPIE Press, 2000. 325 pp.

## Referencias de internet

1. *Basic introduction to OpenCV Image Processing.* <http://www.discover.uottawa.ca/~qchen>, Octubre, 2007.
2. *Face Detection.* <http://opencvlibrary.sourceforge.net/FaceDetection>, Octubre, 2007.
3. *Face Recognition.* <http://www.face-rec.org/>, Agosto, 2007.
4. *Ingeniería de Sistemas Biométricos.* [http://www2.ing.puc.cl/~iing/ed429/sistemas\\_biometricos.htm](http://www2.ing.puc.cl/~iing/ed429/sistemas_biometricos.htm), Septiembre, 2007.
5. *Sistemas Biométricos.* [http://www.biometricgroup.com/in\\_the\\_news](http://www.biometricgroup.com/in_the_news), Agosto, 2007.
6. *Sistemas Biométricos Avanzados.* <http://www.ppgia.pucpr.br/~alekoe/Papers>, Agosto, 2007.
7. *Visión Artificial.* <http://www.escet.urjc.es/~visiona>, Septiembre, 2007.