



**Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Ingeniería en Ciencias y Sistemas**

**MÉTODOS Y TÉCNICAS DE RECONOCIMIENTO DE ROSTROS
EN IMÁGENES DIGITALES BIDIMENSIONALES**

Jorge Rafael Valvert Gamboa

Asesorado por el Ing. Guillermo Sánchez Barrios

Guatemala, junio de 2006

**UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA**



NÓMINA DE JUNTA DIRECTIVA

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

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

| | |
|-------------------|---|
| DECANO | Ing. Murphy Olympo Paiz Recinos |
| EXAMINADOR | Ing. Edgar René Ornelis Hoíl |
| EXAMINADOR | Ing. Cesar Augusto Fernández Cáceres |
| EXAMINADOR | Inga. Vivian Damaris Campos Gonzáles |
| SECRETARIA | Inga. Marcia Ivonne Véliz Vargas |

HONORABLE TRIBUNAL EXAMINADOR

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

MÉTODOS Y TÉCNICAS DE RECONOCIMIENTO DE ROSTROS EN IMÁGENES DIGITALES BIDIMENSIONALES,

tema que me fuera asignado por la Coordinación de la Carrera de Ingeniería en Ciencias y Sistemas en febrero 2005.

Jorge Rafael Valvert Gamboa

ACTO QUE DEDICO A

| | |
|--|---|
| Mis padres | Por su apoyo incondicional, a ellos le debo este triunfo. |
| Mis hermanos | Por su apoyo, especialmente a Cynthia, que es un ejemplo de superación para mí. |
| Mi Madrina, Abuelita y mis primos | Por ser parte importante de mi vida y de motivación para completar mi carrera. |
| A mis amigos | Por apoyarme en momentos difíciles, y por brindarme una mano cuando más lo necesito. |
| La Universidad de San Carlos de Guatemala | Por albergarme en sus aulas y formarme como un profesional integro y con conciencia social. |

AGRADECIMIENTOS

A la Facultad de Ingeniería y la Escuela de Ingeniería en Ciencias y Sistemas, por el apoyo en la culminación de mi carrera y especialmente a mi asesor, por su tiempo dedicado a mi trabajo de graduación.

ÍNDICE GENERAL

| | |
|--|------|
| ÍNDICE DE ILUSTRACIONES | VII |
| GLOSARIO | IX |
| RESUMEN | XV |
| OBJETIVOS | XVII |
| INTRODUCCIÓN | XIX |
| | |
| 1. MARCO TEÓRICO | 1 |
| 1.1 Introducción..... | 1 |
| 1.2 Reconocimiento de Patrones..... | 2 |
| 1.2.1 Representación del conocimiento de formas..... | 2 |
| 1.2.2 Aprendizaje de Formas..... | 2 |
| 1.3 Reconocimiento de rostros..... | 3 |
| 1.3.1 Pre procesamiento de la imagen..... | 4 |
| 1.3.2 Extracción de Características..... | 4 |
| 1.3.3 Entrenamiento..... | 5 |
| 1.3.4 Reconocimiento..... | 5 |
| 1.4 Métodos de Aprendizaje de Rostros..... | 6 |
| 1.4.1 Redes Neuronales..... | 6 |
| 1.4.2 Algoritmos Genéticos..... | 8 |
| 1.5 Investigaciones actuales acerca de reconocimiento de rostros..... | 10 |

| | | |
|-----------|---|-----------|
| 1.5.1 | Sistema de reconocimiento, particularmente, para el reconocimiento de personas..... | 10 |
| 1.5.2 | Método de Heurística para Selección variable utilizando Redes neuronales..... | 10 |
| 2. | BIOMETRÍA..... | 13 |
| 2.1 | Introducción..... | 13 |
| 2.2 | Sistemas Biométricos..... | 13 |
| 2.3 | Indicadores Biométricos..... | 14 |
| 2.4 | Ventajas del reconocimiento de rostros en la Biometría..... | 18 |
| 3. | TÉCNICAS DE RECONOCIMIENTO DE ROSTROS. | 21 |
| 3.1 | Introducción..... | 21 |
| 3.2 | Descriptores invariantes..... | 24 |
| 3.2.1 | Descriptores de Contorno..... | 25 |
| 3.2.2 | Descriptores de región..... | 26 |
| 3.2.2.1 | Momentos geométricos de una región | 27 |
| 3.2.2.2 | Momentos geométricos invariantes de una región..... | 27 |
| 3.2.2.3 | Momentos de Yan..... | 31 |
| 3.2.2.4 | Vector descriptor del rostro..... | 34 |
| 3.3 | Reconocimiento de Rasgos..... | 35 |
| 3.3.1 | Uso de descriptores invariantes para el reconocimiento de rasgos faciales. | 36 |
| 3.4 | Preparación de imágenes para reconocimiento de rostros..... | 40 |

| | | |
|---------|---|----|
| 3.4.1 | <i>Portable Gray maph</i> -mapeo en escala de grises- | 40 |
| 3.4.1.1 | Descripción del formato..... | 41 |
| 3.5 | Eigen Caras..... | 44 |
| 3.5.1 | Construcción de un espacio Facial..... | 44 |
| 3.5.2 | Reconocimiento de rostros utilizando eigen caras..... | 47 |
| 3.5.2.1 | Etapas de Entrenamiento..... | 48 |
| 3.5.2.2 | Etapas de Reconocimiento..... | 49 |

| | | |
|---|--|----|
| 4. MÉTODOS DE CLASIFICACIÓN Y APRENDIZAJE PARA EL RECONOCIMIENTO DE ROSTROS..... | 51 | |
| 4.1 | Introducción..... | 51 |
| 4.2 | Clasificadores Lineales..... | 51 |
| 4.2.1 | Clasificador de distancia euclidiana..... | 52 |
| 4.2.2 | Clasificadores No Lineales..... | 55 |
| 4.2.2.1 | Redes neuronales..... | 56 |
| 4.2.2.1.1 | Fisiología de una neurona biológica..... | 58 |
| 4.2.2.1.2 | Redes neuronales artificiales..... | 60 |
| 4.2.2.1.2.1 | Funcionamiento de una RNA..... | 62 |
| 4.2.2.1.2.2 | Estructura básica de una RNA..... | 66 |
| 4.2.2.1.2.3 | Tipos de RNA..... | 67 |
| 4.2.2.1.2.4 | Aprendizaje de una RNA..... | 68 |
| 4.2.2.1.2.5 | Asociación de información en una RNA..... | 69 |
| 4.2.2.1.2.6 | Representación de la información de entrada y salida en una RNA..... | 70 |
| 4.2.2.1.3 | Regla delta generalizada..... | 71 |

| | | |
|-----------------|--|----|
| 4.2.2.1.3.1 | Funcionamiento de la regla delta generalizada..... | 73 |
| 4.2.2.1.3.2 | Momento en la regla delta generalizada..... | 74 |
| 4.2.2.1.3.3 | Algoritmo de propagación hacia atrás..... | 75 |
| 4.2.2.1.3.4 | Algoritmo de aprendizaje..... | 77 |
| 4.2.2.1.4 | Arquitectura de reconocimiento de rostros basada en redes neuronales electrónicas.. | 82 |
| 4.2.2.1.4.1 | Diseño de la arquitectura..... | 82 |
| 4.2.2.1.4.2 | Detección de las características visuales..... | 84 |
| 4.2.2.1.4.3 | Codificación de las características visuales..... | 86 |
| 4.2.2.1.4.4 | Fases de la arquitectura..... | 88 |
| 4.2.2.1.4.4.1 | Fase de Inicialización..... | 88 |
| 4.2.2.1.4.4.2 | Fase de Indexado..... | 92 |
| 4.2.2.1.4.4.2.1 | Selección de Candidatos..... | 92 |
| 4.2.2.1.4.4.2.2 | Reducción de Candidatos..... | 93 |
| 4.2.2.2 | Algoritmos genéticos..... | 95 |
| 4.2.2.2.1 | Introducción..... | 95 |
| 4.2.2.2.2 | Representación del cromosoma..... | 98 |
| 4.2.2.2.3 | Población inicial..... | 99 |
| 4.2.2.2.4 | Medida de evaluación..... | 99 |
| 4.2.2.2.5 | Técnicas de Selección..... | 99 |

| | | |
|-------------|---|------------|
| 4.2.2.2.6 | Técnicas de Cruzamiento..... | 103 |
| 4.2.2.2.7 | Técnicas de Mutación..... | 104 |
| 4.2.2.2.8 | Aplicación de un algoritmo genético en la detección de rostros..... | 107 |
| 4.2.2.2.8.1 | Características del algoritmo genético en la aplicación..... | 108 |
| 4.2.2.2.8.2 | Descripción del algoritmo genético | 110 |
| | | |
| 5. | APLICACIÓN DE UN SISTEMA DE RECONOCIMIENTO DE ROSTROS..... | 113 |
| 5.1 | Sistema de seguridad para el aeropuerto internacional La Aurora implementando un sistema de reconocimiento de rostros..... | 113 |
| 5.1.1 | Componentes del Sistema..... | 114 |
| 5.1.2 | Captura de Imágenes..... | 114 |
| 5.1.3 | Extracción y Almacenamiento de Imágenes Faciales.... | 115 |
| 5.1.4 | Normalización e Identificación de rostros de Personas.. | 115 |
| 5.1.5 | Sala de Control..... | 116 |
| 5.1.6 | Consideraciones Adicionales..... | 117 |
| 5.2 | Software para el reconocimiento de rostros..... | 118 |
| 5.2.1 | <i>FaceIT ARGUS</i> | 118 |
| 5.2.2 | <i>C-VIS Face Recognition</i> | 120 |
| | | |
| | CONCLUSIONES..... | 125 |
| | RECOMENDACIONES..... | 129 |
| | BIBLIOGRAFÍA..... | 131 |

ÍNDICE DE ILUSTRACIONES

FIGURAS

| | | |
|-----|--|----|
| 1. | Indicadores Biométricos..... | 15 |
| 2. | División de las características biométricas de identificación personal.... | 18 |
| 3. | Ejemplo del procesamiento de una imagen con fondo complejo..... | 20 |
| 4. | Ruido en el contorno de las imágenes. | 26 |
| 5. | Momentos centrales usados en el reconocimiento de rostros. | 29 |
| 6. | Momentos centrales normalizados. | 30 |
| 7. | Invariantes propuestos por M.K. Hu. | 32 |
| 8. | Momentos invariantes a cambios de contraste de Yan. | 34 |
| 9. | División de una imagen en celdas. | 39 |
| 10. | Ejemplo del formato de una imagen PGM. | 43 |
| 11. | Ejemplos de imágenes en formato PGM. | 43 |
| 12. | Ejemplo de imágenes procesadas con el método de eigen caras. | 45 |
| 13. | Algoritmo utilizado por el método de eigen caras. | 50 |
| 14. | Funcionamiento del clasificador euclidiano. | 55 |
| 15. | Neurona y sus partes constitutivas. | 59 |
| 16. | Estructura de una neurona artificial. | 62 |
| 17. | Funcionamiento de una neurona artificial aislada. | 64 |
| 18. | Funciones típicas de activación de una neurona artificial. | 65 |
| 19. | Red neuronal compuesta de 3 capas. | 67 |

| | | |
|-----|---|-----|
| 20. | Arquitectura de la red neuronal..... | 83 |
| 21. | Fases de la arquitectura de la red neuronal. | 84 |
| 22. | Puntos de interés de un rostro. | 85 |
| 23. | Configuraciones de puntos para obtener vectores de atributos. | 87 |
| 24. | Algoritmo de inicialización de la arquitectura de la red neuronal. | 89 |
| 25. | Algoritmo utilizado para la ponderación de mapas de memoria. | 91 |
| 26. | Fase de selección de candidatos. | 94 |
| 27. | Algoritmo para la detección de rostros. | 94 |
| 28. | Objeto de interés dividido en celdas. | 108 |
| 29. | Cámara IP con sensor de movimiento. | 115 |
| 30. | Una pantalla del sistema Facelt. | 119 |
| 31. | Configuración de una aplicación de reconocimiento de rostros. | 119 |
| 32. | Clasificación de imágenes de C-Vis. | 121 |
| 33. | Un ejemplo del funcionamiento de Face Snap Recorder. | 123 |

TABLAS

| | | |
|----|---|----|
| I. | Comparativa entre las técnicas de biometría más utilizadas..... | 17 |
|----|---|----|

GLOSARIO

| | |
|---------------------------|---|
| ALGORITMO GENÉTICO | Algoritmos introducidos por John Holland en 1970, cuyo propósito es resolver problemas complejos, están inspirados en el proceso de evolución natural de los seres vivos. |
| ART | <i>Adaptative Resonance Theory</i> . Modelo inicialmente propuesto por G. A. Carpenter and S. Grossberg en 1986 para extender de los mecanismos de aprendizaje competitivos de contra propagación y mapas auto organizados. |
| ART-2 | Ver ART. El 2 de ART 2 indica que la red puede procesar entradas analógicas, a diferencia de la ART 1 que solo es capaz de procesar entradas binarias. |
| BIOMETRÍA | Ciencia que se dedica a el estudio del cuerpo humano como medio de identificación y autenticación de algún sistema. |
| CAPA DE ENTRADA | Capa de una red neuronal artificial multicapa, que recibe la información de entrada a ser procesada por la red. |

| | |
|------------------------------------|---|
| CAPA OCULTA | En una red neuronal artificial multicapa, capa conformada por un conjunto de neuronas que no están en contacto directo con las entradas y salidas. |
| CAPA DE SALIDA | Capa de una red neuronal artificial multicapa, que transfiere hacia el exterior el resultado de la información procesada. |
| CARACTERÍSTICA VISUAL | Característica importante del rostro la cual posee un conjunto puntos de interés, de los cuales se forman vectores de característica. |
| CELDA | Región lógica de una imagen que ha sido procesada y es útil para el reconocimiento de patrones en una imagen. |
| CROMOSOMA | Lugar donde se encuentra codificada toda la información del ser vivo. |
| CÚMULO | Conjunto agrupado de patrones. |
| DESCRIPTOR INVARIANTE DE HU | Es una técnica muy utilizada para el reconocimiento de patrones desarrollada por M.K Hu. Se le denomina invariante, ya que, presenta invarianza a cambios en tamaño y traslación de las imágenes. |

| | |
|-------------------------------------|---|
| DESCRIPTOR INVARIANTE DE YAN | Es una técnica desarrollada por W. Yan que es utilizada en el reconocimiento de patrones en imágenes, presentando invarianza ante los cambios de contraste. |
| EIGEN CARA | Imagen que contiene las características básicas de un rostro, obtenidas de la imagen de un rostro. |
| EIGEN VECTOR | Vectores utilizados en álgebra lineal para cálculos con significancia física. Se usan en varios campos de la física y en general para cálculos relacionados con la naturaleza. Son parte básica del cálculo de las eigen caras sobre un espacio facial. |
| ESPACIO FACIAL | Cuando se calcula una eigen cara, se obtiene este espacio para la contraposición con otras imágenes de rostros conocidos. En este espacio se calculan los eigen vectores. |
| FEED BACK | Alimentación hacia atrás. En una red neuronal se refiere a que información de salida será utilizada en las entradas. |

| | |
|--------------------------|---|
| FEED FORWARD | Alimentación hacia delante. En una red neuronal artificial se refiere a que el flujo de la información será de entrada a salida. |
| FENOTIPO | Características individuales producidas por los genes, que distinguen a un individuo de otro. |
| FITNESS | Función de evaluación de cromosomas en un algoritmo genético. |
| FORMATO PGM | Formato de mapeo de escala de grises. Utilizado para simplificar cualquier imagen en un equivalente solo con escala de gris. |
| MAPA DE MOMENTOS | Es un conjunto de celdas creadas en una imagen para encontrar momentos estándar las cuales son parte de los descriptores invariantes -ver descriptor invariante de Hu y Yan-. |
| MOMENTOS ESTÁNDAR | Son utilizados para el cálculo de los descriptores invariantes de HU. Se basan en una región de la imagen -ver celda- dividida en partes lógicas. |
| MUTACIÓN | Variación de la información contenida en el código genético, habitualmente el cambio de un gen a otro, producido por un factor externo a un algoritmo genético. |

| | |
|------------------------------------|---|
| PATRÓN | Conjunto de características únicas, que son utilizadas para reconocer un objeto. |
| PESO | En redes neuronales, se le llama peso a la fuerza de conexión de entrada expresada por un número real. |
| PERCEPTRÓN | Modelo básico de una red neuronal. |
| PGM | <i>Portable Gray Map</i> -Mapeo en escala de grises- Ver Formato PGM. |
| RED NEURONAL ARTIFICIAL | Redes de neuronas que imitan la forma de actuar de un sistema nervioso biológico. Sus partes principales son: los elementos de procesamiento o neuronas, estado de activación, funciones de salida y conexión con otras neuronas. |
| RNA | Red Neuronal artificial. |
| RNA MONOCAPA | En este tipo de red existe una sola capa de neuronas las cuales tienen entrada y salida. |
| RNA MULTICAPA | En este tipo de red la distribución de neuronas se lleva a cabo formando niveles, existen 3 capas, de entrada, oculta y de salida. |

| | |
|----------------------------------|--|
| SINAPSIS | Conexión entre 2 neuronas. |
| STRING BINARIO | Representación de un cromosoma en forma de un listado de unos y ceros. |
| UMBRAL | Nivel mínimo necesario para la activación o toma de decisión de alguna función. |
| VECTOR DE CARACTERÍSTICAS | Es un vector que contiene momentos invariantes (De HU o de YAN) que es utilizado para describir un objeto en particular, en nuestro caso un rasgo facial. |
| SISTEMA BIOMÉTRICO | Sistema que se basa en una característica del cuerpo humano, la cual debe cumplir con los siguientes elementos: Universalidad, Unicidad, Permanencia y Cuantificación. |

RESUMEN

El reconocimiento de rostros humanos en imágenes digitales se compone de un conjunto de técnicas para la preparación de imágenes y reconocimiento de patrones, así como, también, métodos de aprendizaje de rostros.

Para reconocer un rostro de una imagen digital, primero se debe preparar, simplificándola, esto se logra convirtiéndola a un formato en escala de grises o aplicando el método de eigen caras.

Después de simplificar la imagen, se deberá localizar el rostro dentro de la imagen, esta parte requiere el uso de algoritmos que reconocen patrones, como por ejemplo los descriptores de contorno.

La fase de aprendizaje, es la parte más crítica del proceso de reconocimiento de rostros, ya que, en ella, se debe almacenar información sobre el rostro, que después deberá ser usada para el reconocimiento. En esta etapa se procesan los rostros, haciéndolos invariantes al tamaño y a los cambios de contraste, utilizando los algoritmos de los momentos invariantes de Hu y de Yan. Para el aprendizaje y reconocimiento de rostros, se utilizan clasificadores no lineales de patrones, los cuales pueden ser calculados por métodos de simulación o algún método no matemático, en este caso, las redes neuronales y los algoritmos genéticos.

Básicamente, una red neuronal se compone de varias neuronas individuales interconectadas entre si. Cada neurona recibe información de entrada o de otras neuronas, en el caso de la primera capa sería la imagen procesada y, en base a pesos en sus entradas y un valor umbral, se transmite la información a la salida de la neurona, siendo las posibles salidas uno o cero, las cuales alimentan las entradas de otras neuronas. Este proceso devolverá la posición del rostro almacenado en la base de datos, en base a una función de salida.

Para que una red aprenda se deben modificar los pesos de las conexiones en las neuronas, en dependencia de la información de entrada. El criterio para modificar estas entradas se realiza a través de métodos de aprendizaje supervisado, como por ejemplo: el método de aprendizaje por corrección de error, aprendizaje por refuerzo y aprendizaje estocástico.

La aplicación del reconocimiento de rostros humanos en imágenes digitales, es muy útil en entornos en los que es imposible el monitoreo humano de personas, facilitando la detección de personas que se encuentren en alguna base de datos de sospechosos. La identificación de rostros es, también, un buen método biométrico, el cual se puede utilizar en controles de acceso y áreas restringidas.

OBJETIVOS

General

Dar a conocer los métodos y técnicas para el aprendizaje y reconocimiento de rostros humanos en una imagen digital bidimensional, basándose en rasgos característicos del rostro humano.

Específicos

1. Dar a conocer y exponer las técnicas más utilizadas para el procesamiento de una imagen digital que contiene un rostro humano, así como las técnicas de identificación de rasgos básicos de un rostro humano en una imagen digital.
2. Dar a conocer métodos para el aprendizaje de rostros utilizando herramientas de inteligencia artificial como redes neuronales y algoritmos genéticos, basándose en los rasgos básicos de un rostro identificados en una imagen digital.
3. Dar a conocer como el reconocimiento de rostros humanos, puede ser útil para la biometría.
4. Proponer la implementación de un sistema de reconocimiento de rostros en un espacio público.

5. Dar a conocer aplicaciones actuales biométricas que se basen en el reconocimiento de rostros.

INTRODUCCIÓN

El reconocimiento de rostros y la visión por medio de una computadora es un tema que ha tomado mucho auge en los últimos años, debido a la importancia que tiene este en la construcción de robots con más capacidades, métodos de identificación aun más confiables o la identificación de personas en lugares públicos como aeropuertos, cárceles etc.

En este trabajo el reconocimiento de rostros es el tema central y se describen las fases que se deben cubrir para el reconocimiento de un rostro en una imagen digital bidimensional.

A lo largo de esta investigación, se podrá conocer los métodos de inteligencia artificial para el reconocimiento y aprendizaje de rostros humanos, además de las técnicas de identificación de patrones, de reconocimiento de rostros humanos en una imagen a través de rasgos característicos aprendidos.

En el capítulo uno se describe el contenido teórico de la investigación como base para los siguientes capítulos. Se hace una breve referencia también a algunos trabajos realizados por investigadores, acerca del reconocimiento de rostros.

En el capítulo dos se dedica a dar una introducción a la biometría y a las diferentes técnicas existentes y su comparación con el reconocimiento de rostros.

En el capítulo tres se describen las técnicas para obtención de las características básicas de un rostro humano bidimensional en una imagen digital para su reconocimiento, así como la preparación de esta a un formato que simplifique la tarea de los algoritmos, pero que, a la vez, no se pierda ningún rasgo característico del rostro. También, se expone el método de eigen caras, que es una técnica especial que preprocesa y reconoce un rostro humano en una imagen.

En capítulo cuatro se exponen los métodos para el aprendizaje y reconocimiento de rostros, principalmente, enfocado en redes neuronales y algoritmos genéticos.

Finalmente, en el capítulo cinco se incluye una aplicación para el aeropuerto nacional y, también, se describen dos aplicaciones comerciales existentes y sus características básicas.

1. MARCO TEÓRICO

1.1 Introducción

El reconocimiento de rostros es un tema fascinante en lo que respecta a las ciencias de la computación y particularmente a la inteligencia artificial.

Para el cerebro humano (un sistema intrincado que aun no se conoce en su totalidad) es muy sencillo aprender rostros de personas y reconocerlos, indiferentemente del ángulo de visión este es capaz de identificarlos.

Pero ¿como el cerebro almacena la información de un rostro? , y ¿como puede reconocerlo, independientemente si ve el rostro completo, si hay poca luz o desde muchos ángulos diferentes que por primera vez ve?, estas son preguntas que los investigadores se han hecho por décadas.

Como resultado de estas investigaciones se han desarrollado técnicas y métodos en el área de la inteligencia artificial para emular la forma en que se cree que el cerebro funciona. Partiendo de esto se han desarrollado las redes neuronales, los algoritmos genéticos y además de técnicas especializadas para el tratamiento de imágenes digitales.

1.2 Reconocimiento de Patrones

En este sub tema se trata de forma general el reconocimiento de patrones. Este puede considerarse como parte de la inteligencia artificial. De

esta afirmación se pueden mencionar algunas características del reconocimiento de patrones.

1.2.1 Representación del conocimiento de formas

El reconocimiento de formas se interpreta como la adquisición de patrones de clase conocida y su almacenamiento para establecer un prototipo del patrón de cada clase. En otras palabras, las plantillas que tienen características básicas, servirán para el reconocimiento de un objeto, por ejemplo sabemos que la plantilla para una letra A mayúscula, es que son dos líneas inclinadas unidas en un vértice y atravesadas por una línea horizontal por la mitad.

De la descripción anterior se pueden decir que cualquier letra A podrá ser reconocida, indiferentemente de su tamaño, color o grosor.

1.2.2 Aprendizaje de Formas

El aprendizaje implica cambios en el sistema. En un sistema de reconocimiento de Patrones, dependiendo del método de aprendizaje, se trata de calcular el patrón o prototipo o el conjunto de patrones prototipo que caracterizan cada una de las clases a discriminar o filtrar.

Cuando se establece la manera de representar el conocimiento, se podría formular un modelo de aprendizaje inductivo, a partir de un conjunto de

entrenamiento del patrón (conjunto de patrones) prototipo utilizando un algoritmo de aprendizaje.

1.3 Reconocimiento de Rostros

El reconocimiento de rostros es un área del reconocimiento de patrones que ha sido investigada desde hace varios años. Durante todo este tiempo se han propuesto técnicas diferentes para la resolución de esta tarea, entre las que se pueden mencionar:

- Igualamiento de plantillas
- Cálculo de eigen caras Características Geométricas
- Métodos que utilizan redes neuronales

Independientemente de la técnica que sea implementada, se utilizan siempre dos conjuntos de datos:

- El primero es utilizado siempre para la etapa de aprendizaje, el cual es llamado conjunto de entrenamiento. Se debe tratar de que los patrones que integran este conjunto sean lo más diferente posible entre si, y que además, representen al problema, para poder tener un buen porcentaje de generalización.
- El segundo conjunto de patrones, es llamado conjunto de prueba, y es utilizado en la etapa de reconocimiento.

Por ultimo, para reconocer un rostro se deben realizar los siguientes pasos generales:

- Preprocesamiento de la imagen
- Extracción de Características
- Entrenamiento
- Reconocimiento

1.3.1 Pre procesamiento de la imagen

Como se puede intuir las imágenes de una misma persona son captadas en momentos diferentes, lo que conlleva a que cada imagen sea diferente, debido a que puede variar la iluminación, el ángulo de enfoque y el tamaño del rostro (profundidad). Por esto es necesario preprocesar la imagen.

Entre las tareas más comunes de preprocesamiento se puede mencionar:

- Extracción de información de la imagen de un rostro, que sea irrelevante.
- Normalización del tamaño de la imagen (todos los rostros de un tamaño similar).
- Aplicación de un método de filtrado (en esta investigación la conversión a formato PGM (*Portable Gray Maph*))

1.3.2 Extracción de Características

La extracción de características es una de las etapas de la cual depende el desempeño del sistema de reconocimiento de rostros. El objetivo principal de esta etapa es extraer la información más discriminante de un rostro, eliminando la información innecesaria. Las técnicas más comunes son:

- Extracción de componentes básicos del rostro
- Relaciones de distancia entre ojos, nariz, boca y cejas.
- Descriptores Invariantes y vectores de característica

En esta investigación se describirán los descriptores invariantes y el trazado de vectores de característica en el capítulo tres.

1.3.3 Entrenamiento

El entrenamiento básicamente consiste en utilizar alguna forma de aprendizaje que le permita al sistema aprender rostros que formarán el conjunto de entrenamiento. El tipo de entrenamiento a utilizar dependerá de la metodología utilizada para el reconocimiento.

Para la metodología de reconocimiento de redes neuronales, por ejemplo, el entrenamiento consiste en obtener valores correspondientes a cada una de las conexiones (pesos) las cuales forman la red neuronal. El algoritmo de retro propagación es uno de los más comunes, y es expuesto en el capítulo cuatro.

1.3.4 Reconocimiento

La última etapa, en un sistema de reconocimiento de rostros es la etapa de reconocimiento. Consiste en alimentar al sistema con imágenes de rostros diferentes a las utilizadas durante el entrenamiento, esperando obtener como resultado, alguna forma de codificación que nos permita saber de que rostro se trata.

1.4 Métodos de Aprendizaje de Rostros

1.4.1 Aprendizaje utilizando Redes Neuronales

Las redes neuronales están inspiradas en el comportamiento del cerebro humano. Esto significa que cuando se desarrolla un modelo de una red neuronal se está tratando de imitar el comportamiento de el cerebro humano.

Una red neuronal artificial está compuesta por un conjunto de elementos que imitan los componentes neuronales biológicos, se puede decir que están compuestas por neuronas artificiales.

De manera general las redes neuronales tienen tres características principales:

- Aprendizaje
- Generalización
- Adaptación

La característica de aprendizaje se refiere a que una red neuronal tiene la capacidad de almacenar conocimiento por medio de un proceso de entrenamiento. Este conocimiento es almacenado por los pesos de las conexiones entre las neuronas que conforman la red.

La Generalización significa que se pueden obtener salidas razonables cuando se usan entradas diferentes a las utilizadas durante el proceso de entrenamiento.

La característica de adaptabilidad se refiere a que una red neuronal puede ser re entrenada para funcionar adecuadamente ante cambios en el medio ambiente.

Existen diferentes arquitecturas de redes neuronales entre las cuales se pueden mencionar:

- Redes alimentadas hacia delante de capa simple
- Redes alimentadas hacia delante de multicapas
- Redes Recurrentes.

Una de las arquitecturas más utilizadas para el reconocimiento de rostros es la red alimentada hacia delante multicapas o red de perceptrones multicapas (MLP), ya que una red multicapa es capaz de resolver problemas más complejos que una red monocapa.

El tipo de redes multicapa constan de tres componentes principales:

- Una capa de entrada
- Una o más capas ocultas
- Una capa de salida

Estas redes regularmente son entrenadas por medio de algoritmos de retro propagación del error que consta de dos fases:

- La fase hacia delante
- La fase hacia atrás

Estas fases son explicadas más a detalle en el capítulo cuatro.

1.4.2 Aprendizaje utilizando Algoritmos Genéticos

Un algoritmo genético es un método de optimización basado en los mecanismos de la evolución natural para resolver problemas de búsqueda y aprendizaje.

La información de las características de un ser humano, están codificadas en los cromosomas. Los algoritmos genéticos se basan en esta idea para crear técnicas aplicables a una computadora para resolver los problemas mencionados anteriormente y además con un nivel de complejidad alto.

Está fundamentado en el la ley natural de la “supervivencia del más apto”.

Para aplicar un Algoritmo Genético a algún problema, es necesario definir dos elementos importantes: la codificación de los cromosomas de los individuos considerados como soluciones y, la función de aptitud o supervivencia de éstos. Además, se debe trabajar con los diferentes operadores genéticos.

Los algoritmos genéticos establecen una analogía entre el conjunto de soluciones de un problema y el conjunto de individuos de una población natural, codificando la información de cada solución en un string (vector binario) a modo de cromosoma.

John Holland, creador de los algoritmos, dice: “se pueden encontrar soluciones aproximadas a problemas de gran complejidad computacional mediante un proceso de evolución simulada”.

A tal efecto se introduce una función de evaluación de los cromosomas, llamada calidad ("*fitness*") y que está basada en la función objetivo del problema. Igualmente se introduce un mecanismo de selección de manera que los cromosomas con mejor evaluación sean escogidos para "reproducirse" más a menudo que los que la tienen peor.

Los algoritmos desarrollados por Holland inicialmente eran sencillos pero dieron buenos resultados en problemas considerados difíciles. Los algoritmos Genéticos están basados en integrar e implementar eficientemente dos ideas fundamentales: Las representaciones simples como strings binarios de las soluciones del problema y la realización de transformaciones simples para modificar y mejorar estas representaciones.

Para llevar a la práctica el esquema anterior y concretarlo en un algoritmo, hay que especificar los siguientes elementos:

- Representación cromosómica
- Población inicial
- Medida de evaluación
- Criterio de selección / eliminación de cromosomas
- Técnicas de recombinación
- Técnicas de mutación

En el capítulo cuatro se extienden estos conceptos y se expone el funcionamiento de los algoritmos genéticos para el aprendizaje de rostros.

1.5 Investigaciones actuales sobre reconocimiento de rostros

En la actualidad algunos científicos han propuesto métodos para el reconocimiento de rostros humanos, según una recopilación en esta investigación, se menciona de manera general algunos de estos trabajos.

1.5.1 Sistema de reconocimiento, particularmente para el reconocimiento de personas.

Roberto Brunelli compara dos estrategias sobre una base de datos común (imágenes frontales de 47 personas, cuatro imágenes por persona). Este científico desarrolló dos nuevos algoritmos: el primero consiste en el cálculo de un conjunto de características geométricas, tales como ancho y altura de la nariz, posición de la boca y forma de la barbilla.

El segundo método está basado en el igualamiento de plantillas en escala de grises. Según Brunelli, se obtuvo un 90% de reconocimiento satisfactorio con el primer método, y con el segundo se logró un reconocimiento perfecto -0 errores de reconocimiento-.

1.5.2 Método de reconocimiento basado en heurística para la selección variable, utilizando redes neuronales

Meziane Yacoub, propone un método para la selección de características para detectar e identificar zonas discriminatorias en la capa de

entrada de un modelo conexionista. El objetivo es seleccionar características simples que tengan un buen poder discriminatorio. El método que proponen es el HVS extendido que está basado en el HVS(Heurística para la selección variable) que ha mostrado resultados exitosos sobre redes de Perceptrones Multicapa para la selección variable.

HVS permite identificar y seleccionar las variables más importantes en el espacio de características. Para el reconocimiento utilizaron un modelo de Redes neuronales llamado TDNN (*Time Delay Neural Networks*). En este modelo se utiliza una base de datos que consiste en 92 imágenes (4 personas diferentes, 23 imágenes por persona). La arquitectura utilizada posee dos módulos:

El primero es un extractor de características con dos capas, con pesos compartidos produciendo en la última capa oculta un mapa de características de la imagen presentada en la entrada (para más información sobre esta terminología consulte el capítulo cuatro).

El segundo modulo es un clasificador que recibe como entrada el mapa de características obtenido anteriormente. La salida del clasificador es la identidad de la imagen presentada. La capa de salida tiene tantas unidades como personas que se desee identificar.

Estas investigaciones reflejan los últimos avances en reconocimiento de rostros.

Con este marco teórico, se pretende que el lector se familiarice con los métodos y técnicas utilizadas en el reconocimiento de rostros, y tenga un panorama completo del contenido de los demás capítulos de esta investigación.

2. BIOMETRÍA

2.1 Introducción

La biometría o biométrica, ha sido definida como la captura y uso de características fisiológicas o de comportamiento para la identificación personal y /o propósitos de verificación individual. La biometría es extensamente usada por los forenses, para controles de acceso y comunicaciones seguras. El reconocimiento de rostros es un método biométrico por el cual un humano identifica a otro. Los seres humanos son capaces de detectar e identificar rostros sin ningún esfuerzo. Construir sistemas automatizados para llevar a cabo esta tarea, sin embargo, es una tarea difícil, debido a la significativa variabilidad encontrada durante el proceso de la formación de la imagen.

2.2 Sistemas biométricos

Con la evolución de las tecnologías asociadas a la información, la sociedad está cada día más conectada electrónicamente.

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 importancia y como consecuencia directa, la *biometría* ha emergido.

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, como se menciono anteriormente. Una característica anatómica tiene la cualidad de ser relativamente estable en el tiempo, tal como un rostro humano, huella dactilar, la silueta de la mano, patrones de la retina o el iris. 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, Unicidad, Permanencia y Cuantificación* .

2.3 Indicadores biométricos

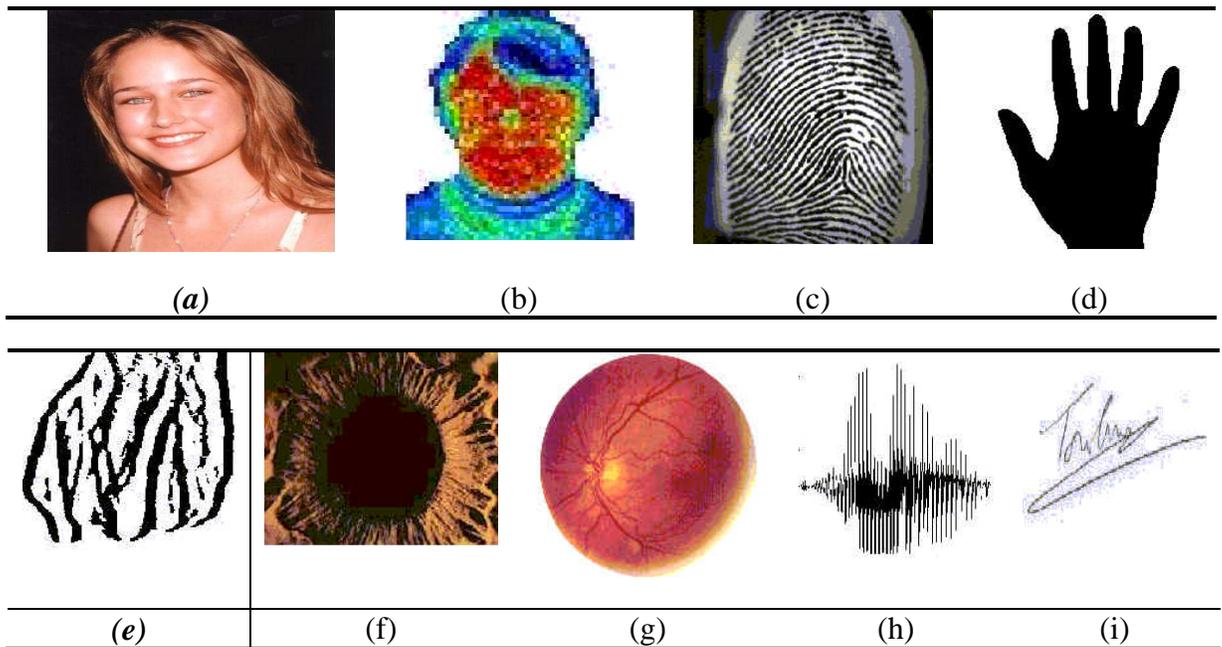
En la actualidad existen sistemas biométricos que basan su acción en el reconocimiento de diversas características, como puede apreciarse en la figura uno.

Las técnicas biométricas más conocidas son nueve y están basadas en los siguientes indicadores biométricos:

- a) Rostro.
- b) Termograma del rostro.
- c) Huellas dactilares.
- d) Geometría de la mano.
- e) Venas de las manos.
- f) Iris.
- g) Patrones de la retina.
- h) Voz.

- i) Firma.
- j) Secuencias de DNA.

Figura 1. Indicadores Biométricos



(a) Rostro (b) Termograma Facial, (c) Huella dactilar (d) Geometría de la mano, (e) Venas de la mano, (f) Iris, (g) Patrones de la retina, (h) Voz (i) Firma.

Cada una de las técnicas anteriores tienen ventajas y desventajas comparativas, las cuales deben tenerse en consideración al momento de decidir que técnica utilizar para una aplicación específica la tabla 1 muestra una comparativa de entre las técnicas más utilizadas.

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, como por ejemplo que la persona este nerviosa y le tiemble el pulso al firmar.

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. Más aún, existen esquemas que utilizan de manera integrada más de una característica para la identificación.

Algunas veces se integran el reconocimiento de rostros y huellas dactilares. La razón es que el reconocimiento de rostros es rápido pero no extremadamente confiable, mientras que la identificación mediante huellas dactilares es confiable pero no eficiente en consultas a bases de datos. Lo anterior sugiere el utilizar el reconocimiento de rostros para particionar la base de datos. Luego de esto comienza la identificación de la huella.

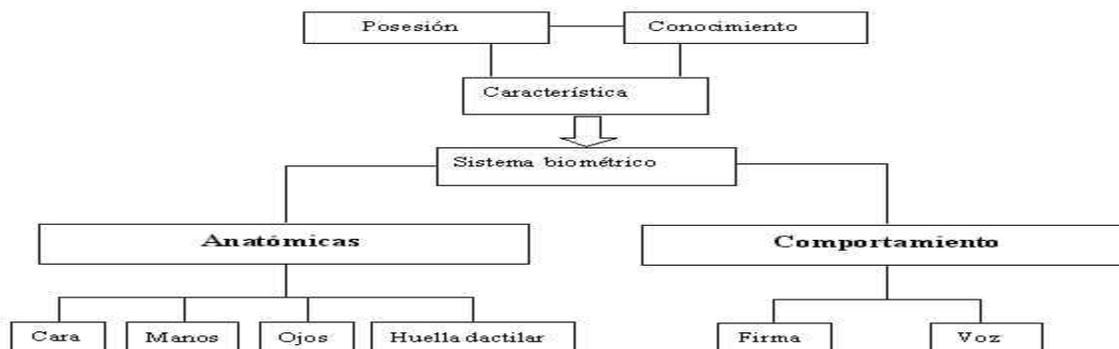
TABLA I. Comparativa entre las técnicas de biometría más utilizadas

| Técnica | Requiere solicitar la muestra | Contacto del aparato de medición con el cuerpo | Coste | Requiere Mantenimiento | Fortaleza | Debilidad |
|------------------------------|-------------------------------|--|-------|------------------------|--|---|
| Reconocimiento Facial | NO | NO | MEDIO | NO | Toma de muestra sin contacto físico y acuerdo con la persona | Lentitud en el procesamiento de la imagen. |
| Huellas digitales | SI | SI | BAJO | SI | Rapidez de la búsqueda entre miles de huellas | Lentitud en la toma de la muestra |
| Geometría de la mano | SI | SI | BAJO | SI | Difícil de falsificar y rapidez de búsqueda | Lentitud en la toma de la muestra |
| Iris | SI | SI | ALTO | SI | Muy exacto | Intrusión en el ojo |
| Retina | SI | MUY CERCA | ALTO | SI | Muy exacto | Intrusión en el ojo |
| Voz | NO | NO | BAJO | POCO | Bajo costo | Falla en entornos ruidosos |
| Firma | SI | NO | BAJO | SI | Bajo costo. | Estado psicológico de la persona, puede variar la muestra |

Los resultados alcanzados por el sistema conjunto son mejores que los obtenidos por sus partes por separado.

La figura 2 muestra la división de características biométricas.

Figura 2. División de las características biométricas para identificación personal



2.4 Ventajas del reconocimiento de rostros en la biometría

Mientras que algunos indicadores requieren alguna acción voluntaria, el reconocimiento de rostros puede ser usado pasivamente. Esto es una ventaja para hacer más fácil su utilización y ser convertido, por ejemplo en un sistema de vigilancia policial.

En lo que concierne a la adquisición de datos, es más fácil de obtener buenas imágenes faciales que buenas huellas digitales. Por ejemplo un 5% de personas no pueden proveer una suficientemente buena imagen para un lector que es utilizado para autenticación. Las razones incluyen piel cortada, dedos vendados, callosidades, piel seca o húmeda etc. Otra desventaja que se encuentra en la detección de huellas es que los scanner de huella digital o de palma de la mano pueden transmitir gérmenes a la mano.

En contraste con el sistema de reconocimiento de rostros es totalmente higiénico y no requiere mantenimiento debido a que es utilizado a distancia. Los scanner de iris pueden proveer una alta efectividad para la identificación de personas.

Sin embargo, debido a que el iris es muy pequeño, necesita una cámara especial de alta resolución para que este sea capturado. Además la cámara deberá estar cerca para capturar toda la resolución del iris, todo este proceso es altamente sensitivo a los movimientos del cuerpo y trae como consecuencia que cualquier persona debe estar muy quieta para no ser rechazada.

Los lectores de retina se basan en los patrones de las venas en la parte de atrás de un ojo. Estos lectores son muy exactos, pero debido a que su funcionamiento requiere de intrusión en el ojo, la mayoría de las personas se resisten a ello.

Además tanto los lectores de iris y de retina fallan cuando se usan lentes de contacto, o cuando se parpadea cuando se toma la imagen. Los lentes normales también son un problema para estos scanner.

En contraste, un sistema automatizado de reconocimiento de rostros, solamente requiere de una o dos cámaras comunes, y sin necesidad de algún sistema especial ya que el campo de visión es lo suficientemente grande para cubrir el rango de las diferentes estaturas de las personas, independientemente si están paradas o sentadas.

Un buen algoritmo de reconocimiento de rostros trabaja inclusive con reflejos provenientes de lentes o con los ojos cerrados.

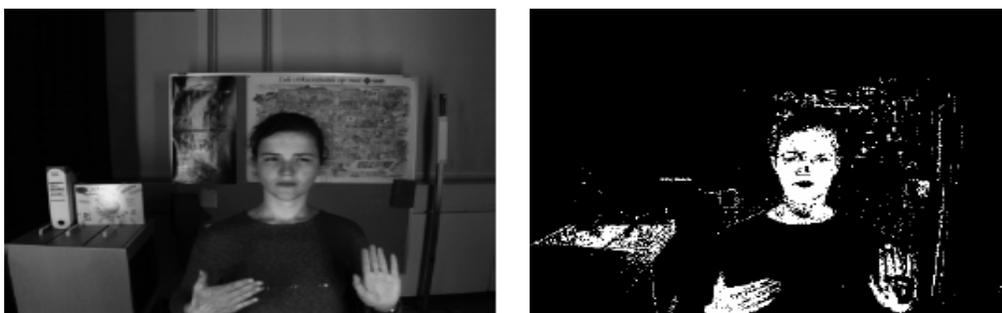
El reconocimiento de voz para propósitos de vigilancia, no es muy confiable en entornos ruidosos como lugares públicos o a través de líneas telefónicas con acústica variable.

Los sistemas de reconocimiento de voz también son sensitivos a fallar en condiciones en la que la persona está enferma con gripe. Una cinta grabada con la voz de una persona puede falsear a los sistemas de reconocimiento de voz, que no tienen un método de integridad de la voz.

La firma es utilizada para legalizar documentos, pero usualmente las personas varían significativamente su firma cada vez, debido a el bolígrafo utilizado, entonces cada vez que se lee una firma es difícil saber si la persona que la coloca es la dueña de ella, esto reduce la confiabilidad de los sistemas de identificación de firmas.

El reconocimiento de rostros, es muy fácil de operar en interiores y exteriores detectando rostros en fondos complejos (ver figura tres).

Figura 3. Detección de Rostros en una imagen con fondo complejo



3. TÉCNICAS DE RECONOCIMIENTO DE ROSTROS

3.1 Introducción

El *rostro* es una de las partes más importantes en el ser humano puesto que es el factor principal de influencia en las relaciones humanas. A éste le son asociadas al menos seis emociones que se reflejan en las expresiones faciales, además de constituir el elemento fundamental para la identidad de las personas. El problema del reconocimiento de rostros puede ser dividido en dos problemas generales:

- *detección y localización* del o de los rostros en la imagen
- *determinación de la identidad* de dichos rostros.

Un sistema completo para la detección, localización y reconocimiento o identificación de rostros debe integrar las siguientes tareas.

Dada una imagen bidimensional:

- Determinar si la imagen contiene rostros. Si es así, determinar la posición y el tamaño.
- Determinar la identidad de cada rostro presente en la imagen.

- Proporcionar una descripción acerca del rostro (alegría, enojo, tristeza, entre otros).
- Entregar información extra, por ejemplo, hacia donde está mirando la persona, la probable edad, el sexo, si porta lentes o sombrero.

Obviamente, esta tarea es extremadamente compleja, puesto que en el caso general no se conoce ni el número de rostros, ni sus correspondientes posiciones, tamaños y orientaciones en la imagen. Además, es muy difícil resolver el problema cuando se cuenta con una base de modelos con muchos rostros. Otros factores que dificultan la solución de este problema son la eventual presencia de adornos, por ejemplo: lentes, sombrero, barba o las normales variantes según el sexo, la edad y la expresión facial.

Desde sus inicios a la fecha, se han propuesto diversas técnicas para dar solución al problema del reconocimiento de rostros, ya sea en tiempo real o no. Sin embargo, este problema es todavía un reto para investigadores y científicos debido a que aún se desconoce en gran medida cómo el ser humano procesa la información durante el reconocimiento de un objeto, en particular de un rostro.

Los trabajos desarrollados para el reconocimiento automático de rostros en gran parte se basan en la información proporcionada por una imagen bidimensional. El proceso se lleva a cabo a través de la extracción de un conjunto de características globales aplicando transformaciones de imagen sobre la región enmarcada por el rostro, o locales, por medio del análisis

particular de diferentes componentes del rostro, tales como ojos, nariz y boca, para la conformación de un modelo representativo del rostro en su totalidad. Los resultados producto de este modelo son usados posteriormente como índices para la búsqueda del rostro en una base de datos.

El reconocimiento en forma automática de rostros representa uno de los desafíos más grandes en el área del reconocimiento de patrones; sin embargo, hoy en día es una de las tareas que más se abordan en la investigación.

Un sistema para el reconocimiento de rostros por computadora puede ser auxiliado por otros sistemas, como por ejemplo, alguno para el reconocimiento del habla, o algún otro para el control de un determinado proceso, con el fin de generar eficiencia y rapidez en el reconocimiento.

Prácticamente todos los trabajos relacionados con el reconocimiento de rostros han sido desarrollados tomando en cuenta la información contenida en una fotografía. Por lo general se ha trabajado con pequeñas bases de datos. Entre los más relevantes, que utilizan información bidimensional, se encuentran los siguientes:

- Aproximaciones, donde la forma del objeto está definida en términos de mini-patrones; éstos son descripciones abstractas de rasgos geométricos como arcos y esquinas.
- Plantillas.

- Aprendizaje por análisis multi resolución de imágenes digitales, empleando como clasificador un Perceptrón multicapa -MLP-
- Reconocimiento automático de objetos utilizando descriptores invariantes y algoritmos genéticos.
- Redes Neuronales.

Esta investigación está basada en los métodos de descriptores invariantes, redes neuronales y algoritmos genéticos. Debido a que son las más utilizadas en aplicaciones de reconocimiento de rostros, ya que por la naturaleza del reconocimiento de un rostro, es buena idea copiar el funcionamiento de la biología en un algoritmo de computadora.

3.2 Descriptores Invariantes

Las fases involucradas en el procesamiento de imagen para su reconocimiento son: la adquisición o captura, el pre procesamiento, la segmentación, la representación y descripción, y el reconocimiento e interpretación de la imagen.

Después de la segmentación de una imagen en regiones por cualquier método que se elija, el conjunto de píxeles dentro de la región segmentada debe describirse de forma tal que pueda ser procesado posteriormente por una computadora.

La descripción de la región de interés puede llevarse a cabo en términos de características externas, esto es, la información contenida en el contorno o frontera; o en términos de características internas, obtenidas a partir de los píxeles dentro del contorno de la región de interés.

En los siguientes subtemas, se explicaran dos tipos de descriptores invariantes, el de Hu, que es utilizado para lograr invarianza en rotaciones, cambios de tamaño y traslaciones (obviamente útiles cuando se trata de una imagen digital) y el de Yan que nos permite manejar información del rostro invariante a los cambios de contraste moderados en la imagen.

3.2.1 Descriptores de contorno

Para describir el contorno de una imagen se pueden utilizar los descriptores simples, números de forma, descriptores de Fourier o momentos.

Uno de los problemas que presentan las descripciones de contorno es la sensibilidad al ruido. En diversas aplicaciones es mejor emplear descriptores basados en la información de los píxeles comprendidos en una región (utilizando descriptores de región).

Como se puede observar en la figura cuatro, el ruido en el contorno de una imagen, hace difícil su identificación, por lo que es preferible descriptores de región en vez de contorno.

Figura 4. Ruido en el contorno de una imagen



3.2.2 Descriptores de región

Para describir un objeto se pueden utilizar los descriptores de región; existen descriptores simples, topológicos, de textura y momentos geométricos.

Un conjunto de descriptores que han sido usados por más de treinta años en el modelado y el reconocimiento de objetos en visión por computadora son los momentos geométricos de Hu, en los cuales nos basaremos en esta investigación.

En el siguiente subtema se da una explicación general de cómo se calcula el contorno de un rostro utilizando descriptores de región. Se debe resaltar que el lector deberá estar familiarizado con funciones de cálculo matemático básicas para entender de la mejor forma posible la exposición.

3.2.2.1 Momentos geométricos de una región

Los rasgos que describen un objeto pueden ser derivados a partir del teorema de unicidad de Popouulis. Este teorema afirma que si $f(x,y)$ es una función continua con valores no cero ($f(x,y) \neq 0$), solamente en una región finita del plano xy , un conjunto infinito de momentos de esa región m_{pq} , con $p, q = 0, 1, 2, \dots$ existe, el cual está determinado de manera única por $f(x,y)$. Inversamente, la secuencia de momentos m_{pq} determina de manera única la función $f(x,y)$.

Si $f(x,y)$ es una función continua, positiva y acotada, definida sobre una región R del plano xy , entonces el momento de orden $(p+q)$ de $f(x,y)$ se define como:

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy$$

Ecuación 3.1

3.2.2.2 Momentos geométricos invariantes de una región -Momentos de Hu-

Los momentos invariantes pueden ser obtenidos de los momentos definidos en la figura cinco. Los momentos presentan invarianza a

transformaciones lineales tales como translaciones, rotaciones, cambios de escala.

En los últimos años se han usado los momentos invariantes puesto que permiten caracterizar la geometría de un objeto para permitir así su posterior reconocimiento.

Los momentos centrales de orden $(p+q)$ de una región en el caso continuo, son invariantes a traslaciones y pueden definirse como:

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy$$

Ecuación 3.2

Donde

$$\bar{x} = \frac{m_{10}}{m_{00}} \quad \text{Y} \quad \bar{y} = \frac{m_{01}}{m_{00}}$$

Para el reconocimiento digital se prefiere convertir las integrales en la ecuación de la figura tres en sumatorias. De esta forma, la ecuación de la figura tres se expresa como:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy$$

Ecuación 3.3.

En la práctica, únicamente los diez primeros momentos son usados en tareas de reconocimiento de rostros. A continuación se muestran algunos de ellos:

Figura 5. Momentos centrales utilizados para el reconocimiento de rostros

$$\begin{aligned} \mu_{10} &= m_{10} - \frac{m_{10}}{m_{00}} (m_{00}) = 0 \\ \mu_{11} &= m_{11} - \frac{m_{10}m_{01}}{m_{00}} = m_{11} - \bar{y}m_{10} \\ \mu_{20} &= m_{20} - \frac{2m_{10}^2}{m_{00}} + \frac{m_{10}^2}{m_{00}} = m_{20} - \bar{x}m_{10} \\ \mu_{02} &= m_{02} - \frac{m_{01}^2}{m_{00}} = m_{02} - \bar{y}m_{01} \\ \mu_{30} &= m_{30} - 3\bar{x}m_{20} + 2\bar{x}^2m_{10} \\ \mu_{12} &= m_{12} - 2\bar{y}m_{11} - \bar{x}m_{02} + 2\bar{y}^2m_{10} \\ \mu_{21} &= m_{21} - 2\bar{x}m_{11} - \bar{y}m_{20} + 2\bar{x}^2m_{01} \\ \mu_{03} &= m_{03} - 3\bar{y}m_{02} + 2\bar{y}^2m_{01} \end{aligned}$$

Los momentos centrales pueden ser normalizados para que sus valores también sean invariantes a cambios de escala. Esto se puede lograr al dividir los momentos centrales por un factor proporcionado por el tamaño del objeto de interés.

Los momentos centrales normalizados se definen como:

Figura 6. Momentos centrales normalizados

| | |
|-------|--|
| | $\eta_{pq} = \frac{\mu_{pq}}{\mu^{\gamma}_{00}}$ |
| Donde | |
| | $\gamma = \frac{p+q}{2} + 1$ |
| Para | $p + q = 2, 3, \dots$ |

Según la teoría de momentos, para obtener los momentos invariantes a traslación, cambios de escala y rotación se deben obtener funciones no lineales cuando se combinan los momentos centrales normalizados, η_{pq} .

En 1962, Hu propuso siete invariantes a partir de los momentos normalizados η_{pq} . Estos, como se sabe, tienen la propiedad de ser también invariantes a rotaciones del objeto de interés. Los siete momentos invariantes se definen a partir de los momentos de orden dos y tres como se expone en la figura siete.

Los momentos de Hu son útiles para realizar una descripción del contorno del rostro, para luego crear un vector de característica, cuyas componentes serán los momentos de Hu.

Como se planteo anteriormente los momentos de Hu poseen la característica de invarianza ante las rotaciones, traslaciones y cambios de escala en una imagen.

También es necesario utilizar un descriptor el cual presente invarianza al contraste en la imagen, para esto son utilizados los momentos de Yan.

3.2.2.3 Momentos de Yan

Estos descriptores pueden ser usados para describir la información contenida en una región R de manera indirecta por medio de un histograma de niveles de gris de esa región. El histograma de una imagen digital (región) con

niveles de gris de los píxeles en el rango $[0, L-1]$ es una función discreta $h(x) = n(x)/n$.

Figura 7. Invariantes propuestos por M.K. Hu

$$\begin{aligned}\Phi_1 &= \eta_{20} + \eta_{02} \\ \Phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ \Phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ \Phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ \Phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \\ &\quad [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \Phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \Phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03}) \\ &\quad [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]\end{aligned}$$

Donde x es el x -ésimo nivel de gris, $n(x)$ el número de píxeles de la región de interés con ese nivel de gris, y n el número total de píxeles de la imagen para $x=0,1,\dots,L-1$.

El histograma de una imagen o región proporciona una descripción global de dicha imagen o región. Cada $h(x)$ representa la probabilidad en la cual el nivel de gris k se encuentra presente en la imagen.

El histograma de una imagen (región) es invariante a traslaciones, rotaciones y cambios de escala. El momento uno dimensional de orden p del histograma de una imagen (región) está definido como:

$$m_p = \sum_{x=0}^{L-1} x^p h(x)$$

De igual manera, los momentos centrales de $h(x)$ pueden definirse como:

$$\mu_p = \sum_{x=0}^{L-1} (x - \bar{x})^p h(x)$$

Donde,

$$\bar{x} = m_1 / m_0$$

De forma similar, los momentos centrales normalizados se definen como:

$$\eta_k = \mu^\gamma / \mu_0, \quad \gamma = k + 1$$

De los momentos centrales normalizados definidos por la ecuación anterior se pueden derivar los siguientes tres momentos invariantes a cambios de contraste.

Figura 8. Momentos invariantes a cambios de contraste de Yan

$$\beta_1 = \eta_4 / \eta_2^2$$

$$\beta_2 = \eta_5 / \eta_2 \eta_3$$

$$\beta_3 = \eta_6 / \eta_2 \eta_4$$

El vector que describe un patrón rostro o uno no rostro está formado por siete invariantes. Este vector presenta invarianza ante traslaciones, rotaciones y cambios de tamaño del objeto en cuestión. Adicionalmente, los tres invariantes de Yan proporcionan al vector que ha de describir al rostro robustez ante cambios de contraste. Esto es útil cuando las condiciones de iluminación cambian en forma moderada.

3.2.2.4 Vector descriptor del rostro

El vector que describe un patrón rostro o uno no rostro está formado por siete invariantes, como se indico en los sub temas anteriores. Para explicar el vector descriptor del rostro usaremos los cuatro primeros momentos invariantes de Hu y los tres de Yan.

De lo expuesto en el sub tema 3.2.2.3, se puede decir que este vector presenta invarianza ante traslaciones, rotaciones y cambios de tamaño del objeto en cuestión.

Adicionalmente, los tres invariantes de Yan proporcionan al vector que ha de describir al rostro robustez ante cambios de contraste. Esto es útil cuando las condiciones de iluminación cambian en forma moderada. El vector descriptor viene dado como sigue:

$$V = [\Phi_1, \Phi_2, \Phi_3, \Phi_4, \beta_1, \beta_2, \beta_3]$$

Ecuación 3.4.

Los invariantes que forman este vector se usan como índice para la comparación entre el vector descriptor del rostro y el vector descriptor de la imagen a explorar, y de esta forma realizar la clasificación de estos patrones por medio de la distancia euclidiana y una red neuronal.

3.3 Reconocimiento de rasgos.

Luego de reconocer que existe un rostro en la imagen, es fundamental reconocer características clave del rostro. No se puede basar en todas las características que posee un rostro humano, si no que en las que podrían haber patrones, como es el caso de ojos, boca y nariz.

El rostro de una persona puede cambiar debido a varios factores como la edad, operaciones, o simples cambios cosméticos como dejarse crecer la barba o quitársela, o reducirse o aumentarse las cejas.

Existen 2 características faciales, o rasgos faciales, que muy difícilmente pueden cambiar debido a los factores mencionados anteriormente, los ojos y la boca.

En este punto, explicaremos como se lleva a cabo el reconocimiento de estos dos rasgos faciales.

3.3.1 Uso de descriptores invariantes para el reconocimiento de rasgos faciales

Como se explico en el sub tema 3.2, los descriptores invariantes dan una, como su nombre lo indica, descripción matemática de un patrón en particular, en este caso un rostro humano en una imagen. Esta técnica puede ser utilizada también para que, luego de detectada la imagen, sobre ella se pueda hacer una detección de un rasgo facial utilizando estos descriptores.

Los descriptores que utilizaremos serán los propuestos por HU. Como se menciono en el sub-tema 3.2, estos descriptores han sido utilizados en el reconocimiento de rostros, por que se ha comprobado que son eficaces en el reconocimiento de patrones. Para esta exposición usaremos los primeros 4 descriptores de Hu.

Entonces de acuerdo con la metodología que proponemos, cada característica facial (ojo izquierdo, ojo derecho y boca) puede ser descrita como sigue:

$$F_i = [\phi_1 \phi_2 \phi_3 \phi_4], i = 1,2,3$$

Ecuación 3.5

En la ecuación anterior se define el vector descriptor para los rasgos, ojos y boca, cada uno con una iteración diferente, estos invariantes se pueden observar en la figura siete.

Las características (ϕ_n), serán descritas en términos de momentos ordinarios y no en términos de momentos centrales.

La razón es que los momentos ordinarios pueden ser sumados directamente cuando se refieren a un mismo punto (el origen del marco de coordenadas de la imagen procesada), mientras que los momentos centrales se refieren a puntos diferentes (al centroide de cada celda).

Si usáramos momentos invariantes para cada celda directamente para obtener momentos invariantes de regiones más grandes, seria necesario recalcular los centroides de cada celda y también los momentos invariantes para estos nuevos centroides, lo que afectaría el rendimiento del reconocimiento en general.

Si una región R es dividida en sub regiones $R_i = 1, \dots, r$, los momentos estándar de R pueden ser obtenidos como la sumatoria de los momentos estándar de cada sub región como:

$$m_{pq}(R) = \sum (m_{pq}(R_i)) \quad i = 1 \dots r$$

Ecuación 3.6

Para cada sub región, los momentos estándar son calculados por la siguiente expresión:

$$m_{pq}(R_i) = \sum \sum x^p y^q f(x,y)$$

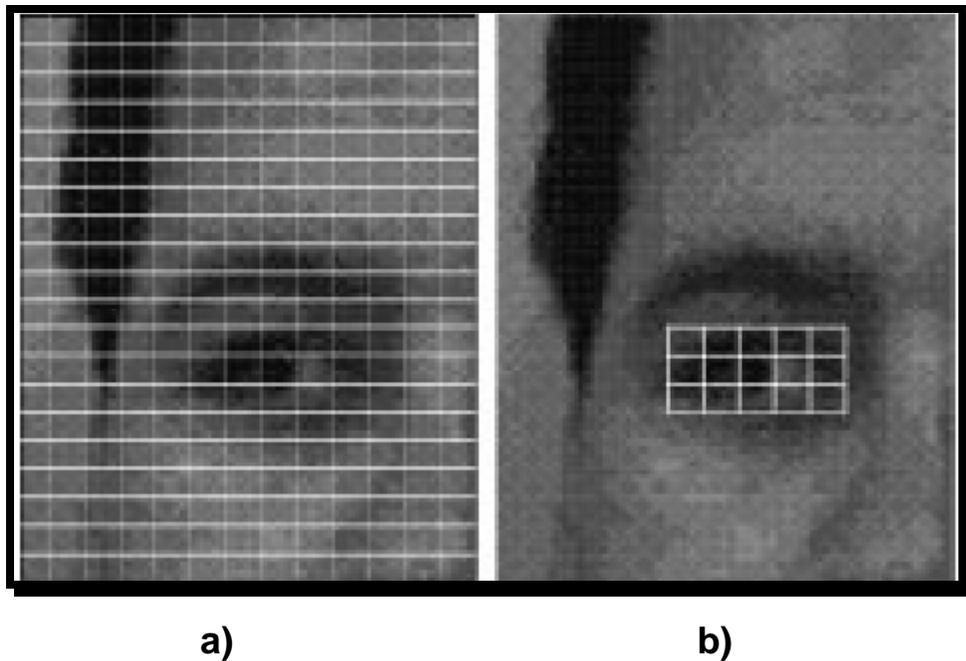
Ecuación 3.7.

Con las ecuaciones anteriores se pueden encontrar patrones, mapeando de una imagen dividida en regiones (con una cuadrícula fija aplicada a la imagen como se muestra en la figura nueve).

En la figura nueve (a) se muestra una imagen dividida en regiones. Cada cuadro es llamado celda y los momentos estándar son calculados usando un mapa de momentos. Un objeto en la imagen, puede ser subdividido en regiones cuando se aplica una cuadrícula (ver figura nueve (b)). De esta forma el mismo objeto a tamaños diferentes puede incluir más o menos celdas.

En el capítulo cuatro se aplicaran estas técnicas utilizando algoritmos genéticos y redes neuronales. En el siguiente sub tema se explicaran dos métodos para el procesamiento de imágenes para que se puedan realizar los cálculos aquí expuestos.

Figura 9. División de una imagen en celdas



La figura nueve nos muestra como se divide una imagen por medio de celdas. La figura A muestra una a imagen dividida en celdas del mismo tamaño utilizando una cuadrícula fija. La figura nueve B Una característica facial en particular sub dividida en celdas.

3.4 Preparación de imágenes para el reconocimiento de rostros.

Antes de realizar cualquier tarea de reconocimiento de un rostro en una imagen, es necesario convertir esta a un formato el cual sea útil para generar patrones utilizando las técnicas expuestas anteriormente.

En este sub tema se explicara el formato PGM, que es utilizado para simplificar gráficos complejos en una escala de gris.

Una técnica más avanzada que utiliza otro pre procesamiento de imagen más complejo es el cálculo de eigen caras, la cual simplifica la imagen y aísla las características de interés del rostro. Esta técnica es utilizada para obtener rostros en una imagen con un fondo complejo y se analizara con detalle en el sub tema 3.5.

3.4.1 Mapeo en escala de grises (Portable Gray maph)

Para llevar a cabo una serie de técnicas complejas es necesario empezar desde lo simple. Cuando se habla de reconocer rostros de una imagen, es común pensar que se debe leer una imagen y compararla con otra.

Esta forma de pensar al parecer simple, no es muy efectiva cuando se piensa que un rostro puede cambiar mucho de imagen a imagen, lo cual haría que cualquier cambio en el ángulo de la imagen, de luz, gesto etc, resultaría en una identificación negativa.

Es necesario entonces pre procesar la imagen para el reconocimiento, para que las técnicas que conlleven a este sean más eficaces.

La técnica del mapeo en escala de grises es una forma fácil de convertir un mapa de bits en una imagen de escala de grises.

3.4.1.1 Descripción del formato PGM

El formato de mapeo en escala de grises, como su nombre lo indica es un mapeo de una imagen en su equivalente en nivel de gris. Existen varios convertidores de formato de imagen a pgm, entre los que se pueden mencionar Lead Tools y Rea Converter Pro.

Un archivo pgm, es un archivo plano en código ASCII, el cual tiene un formato específico para el mapeo.

El formato de un archivo PGM es el siguiente (ver figura diez):

- Un identificador del archivo pgm, que es llamado “número mágico”, que consta de dos caracteres.
- Luego un espacio en blanco (espacio, Tabulador, CR y LF)
- En ASCII decimal el ancho.
- Espacio en blanco
- En ASCII decimal la altura
- Espacio en Blanco
- El máximo valor de gris en ASCII decimal
- Espacio en Blanco
- Luego se encuentra la información de la imagen, de la siguiente forma:

- Cada decimal ASCII separado por espacios en blanco representa un píxel y su valor de gris.
 - La cantidad de decimales ASCII encontrados serán los especificados en el ancho y alto de la imagen.
 - El valor del decimal ASCII, estará entre 0 y el valor máximo especificado en la cabecera.
 - Cada fila tendrá la cantidad de decimales ASCII según el valor del ancho especificado en la cabecera.
 - Cada columna tendrá la cantidad de decimales ASCII según el valor del alto especificado en la cabecera.
 - El orden de los decimales ASCII, será un equivalente a la posición de los píxeles en la imagen.
 - Un decimal ASCII 0 , significa NEGRO.
 - Un decimal ASCII con el máximo valor especificado, significa BLANCO.
- Caracteres después de un símbolo # son ignorados (comentarios).
 - Ninguna línea debe ser mayor a 70 caracteres.

Existe una variante en el formato adicional para pgm, llamado de *RAWBITS*, el cual tiene las siguientes diferencias:

- El número mágico es P5 en lugar de P2.
- Los valores de gris son guardados como bytes en lugar de decimales ASCII.

- No se permiten espacios en blanco en la sección de definición de valores de gris, solo un carácter de espacio es permitido luego del valor máximo de gris en la cabecera.
- Los archivos son más pequeños y más rápidos de leer y escribir debido a que son bytes los que se manejan.
- Este formato solo permite que el valor máximo de gris sea 255 (8 bits).

Figura 10. Ejemplo del formato de una imagen PGM

```
P2
# grafica.pgm
24 7
15
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 3 3 3 3 0 0 7 7 7 7 0 0 11 11 11 11 0 0 15 15 15 15 0
0 3 0 0 0 0 0 7 0 0 0 0 0 11 0 0 0 0 0 15 0 0 15 0
0 3 3 3 0 0 0 7 7 7 0 0 0 11 11 11 0 0 0 15 15 15 15 0
0 3 0 0 0 0 0 7 0 0 0 0 0 11 0 0 0 0 0 15 0 0 0 0
0 3 0 0 0 0 0 7 7 7 7 0 0 11 11 11 11 0 0 15 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

Figura 11. Ejemplos de Imágenes en formato PGM



3.5 Eigen Caras

Otro método utilizado para procesar una imagen es el cálculo de eigen caras. El término eigen proviene del término alemán que significa característica inherente. Este método está basado en los eigen vectores, que son usados en el álgebra lineal para diversos cálculos relacionados con la naturaleza.

La idea principal es descomponer una imagen facial en un conjunto pequeño de imágenes con características faciales, llamadas eigen caras (*eigenfaces*), las cuales podría decirse que son los principales componentes característicos de la imagen original.

Estas eigen caras funcionan como vectores básicos ortogonales (*eigenvector*) de un sub espacio lineal llamado “Espacio Facial”. El reconocimiento del rostro se realiza proyectando una nueva imagen de un rostro en un espacio facial y comparando su posición en el espacio facial con alguna cara conocida por el sistema.

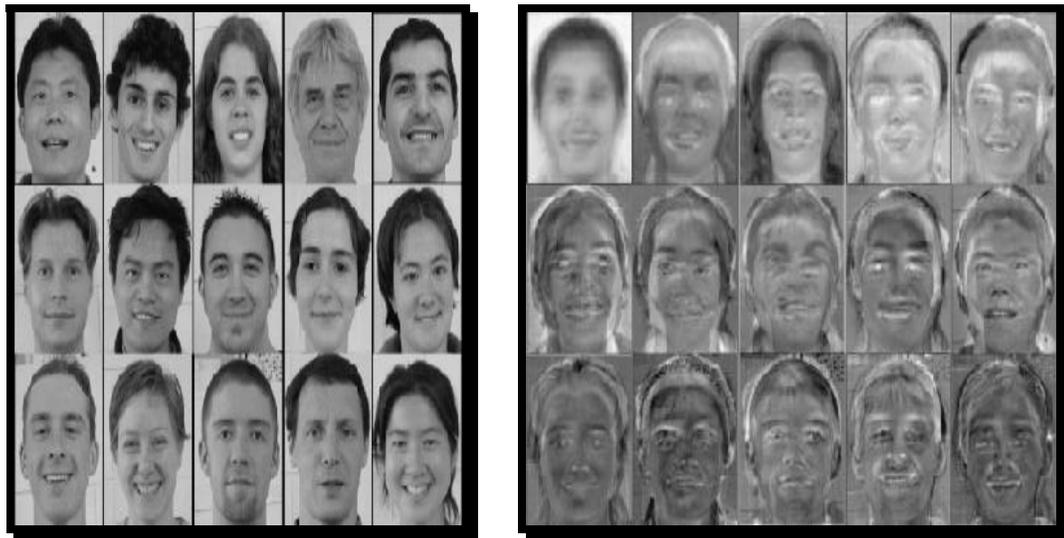
En la figura doce se observa una comparativa de quince imágenes a las que se les ha aplicado el algoritmo de eigen cara y los espacios faciales resultantes.

3.5.1 Construcción de un espacio facial

El fundamento de una eigen cara es la construcción de una imagen facial. A continuación se da una definición matemática de la construcción del espacio facial. Los vectores que se utilizan para el cálculo de estos espacios

son llamados eigenvectores y su valor escalar, valor eigen (*eigenvector* y *eigenvalue*).

Figura 12. Ejemplo de imágenes procesadas con el método de eigen caras



(a) Imágenes de Entrada

(b) Eigen Caras

Supongamos que tenemos una imagen facial de N píxeles. Esta será representada por el vector \mathbf{I} de dimensión N . $S\{\mathbf{I}_i | i = 1, \dots, M\}$ es el conjunto de imágenes a entrenar. La cara promedio de estas M imágenes está dada por:

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i.$$

Ecuación 3.8

Entonces cada cara Γ_i difiere de la cara promedio Ψ por Φ_i

$$\Phi_i = \Gamma_i - \Psi; i = 1, \dots, M.$$

Ecuación 3.9

La matriz de covarianza para las imágenes con las que se entrenara al sistema se construye de la siguiente forma:

$$\mathbf{C} = \mathbf{A}\mathbf{A}^T.$$

Ecuación 3.10

Donde $\mathbf{A} = [\Phi_1, \dots, \Phi_M]$, son los vectores base en el espacio facial, y las eigen caras son entonces eigen vectores ortogonales de la matriz de covarianza \mathbf{C} .

Buscar los eigen vectores de la matriz \mathbf{C} de N por N es una tarea muy difícil en una imagen típica, se debe simplificar la imagen. Debido a que el número de imágenes de entreno es usualmente menor al número de píxeles en una imagen, habrá solo $M-1$, en vez de N , eigen vectores con significado. De está cuenta las eigen caras son calculadas primero buscando los eigen vectores

$$\mathbf{v}_l (l = 1, \dots, M) \quad M (L):$$

$$\mathbf{L} = \mathbf{A}^T \mathbf{A}.$$

Ecuación 3.11

Los eigen vectores $\mathbf{v}_l (l = 1, \dots, M)$ de la matriz \mathbf{C} son expresados por una combinación lineal de las diferencias entre las caras de las imágenes con peso $\Phi_i (i = 1, \dots, M)$ $\mathbf{v}_l (l = 1, \dots, M)$

$$\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_M] = [\Phi_1, \dots, \Phi_M] [\mathbf{v}_1, \dots, \mathbf{v}_M] = \mathbf{A} \cdot \mathbf{V}.$$

Ecuación 3.12

En la práctica, un pequeño conjunto de M' ($M' < M$) eigen caras es suficiente para la identificación del rostro. Por está razón, solamente M' eigen vectores son significantes en la matriz L , correspondientes a los eigen valores (valor escalar del vector eigen), más largos M' , son seleccionados para el cálculo de eigen caras, esto resulta en mayor compresión. M' es determinada por un umbral θ_λ , del radio de la sumatoria de valores eigen θ_λ :

$$M' = \min_r \left\{ r \mid \frac{\sum_{l=1}^r \lambda_l}{\sum_{l=1}^M \lambda_l} > \theta_\lambda \right\}.$$

Ecuación 3.13

3.5.2 Reconocimiento de Rostros utilizando Eigen Caras

El reconocimiento de rostros basado en eigen caras es un procedimiento basado en dos etapas:

- Etapa de Entrenamiento
- Etapa de Reconocimiento

3.5.2.1 Etapa de Entrenamiento

En la etapa de entrenamiento, la cara de cada individuo conocido, Γ_k , es proyectada en un espacio facial y un vector dimensional M' , Ω_k es obtenido :

$$\Omega_k = U^T(\Gamma_k - \Psi); k = 1, \dots, N_c$$

Ecuación 3.14

Donde N_c es el número de de clases de rostros.

Para describir la clase de rostro en un espacio facial existen dos métodos. El primer método se refiere a una representación media, se calcula el vector de clase promediando los vectores proyectados de cada imagen de entreno. El segundo método, es la representación del punto de ajuste, el cual describe una clase de rostro por medio de un conjunto de vectores proyectados de una imagen individual en todas las imágenes.

El valor del umbral d distancia θ_c , que define la distancia máxima permitida desde una clase de rostro hasta el espacio facial, es definido como el cálculo de la mitad de la distancia más grande entre cualquier par de clases de rostro, como está definido en la siguiente ecuación:

$$\theta_c = \frac{1}{2} \max_{j,k} \{ \|\Omega_j - \Omega_k\| \}; j, k = 1, \dots, N_c.$$

Ecuación 3.15

Calculando el espacio facial y las clases de rostros de todas las imágenes que estarán en la base de imágenes (imágenes de entreno del sistema), concluye la etapa de entrenamiento.

El método de entrenamiento explicado en este sub tema es exclusivo para el reconocimiento por medio de eigen caras. En el capítulo cuatro se exponen técnicas más generales para el entrenamiento y aprendizaje de rostros utilizando Redes Neuronales y Algoritmos genéticos.

3.5.2.2 Etapa de Reconocimiento

En la etapa de reconocimiento, una nueva imagen Γ , es proyectada en el espacio facial calculado, para obtener un vector, Ω :

$$\Omega = U^T \Gamma - \Psi.$$

Ecuación 3.16

La distancia de Ω para cada clase de rostro está definida por:

$$\epsilon_k^2 = \|\Omega - \Omega_k\|^2; k = 1, \dots, N_c.$$

Ecuación 3.17

Con el propósito de discriminar entre imágenes con rostro e imágenes sin rostro, la distancia ϵ entre la imagen original, Γ , y su imagen reconstruida en un espacio de eigen cara, Γ_f , es calculado también:

$$\epsilon^2 = \|\Gamma - \Gamma_f\|^2$$

Ecuación 3.18

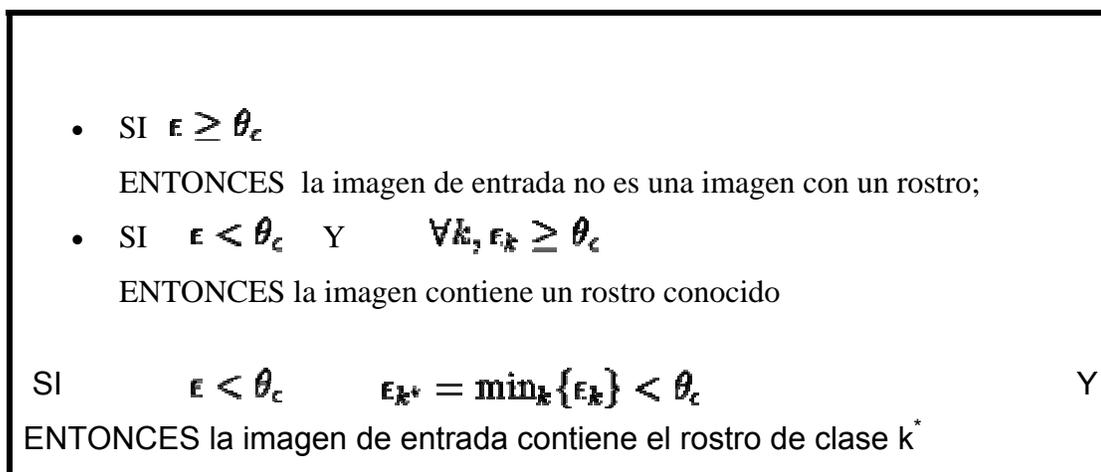
Donde

$$\Gamma_f = U \cdot \Omega + \Psi.$$

Ecuación 3.19

Estas distancias son comparadas con el umbral calculado en la ecuación 3.15 y la imagen de entrada es clasificada, en el algoritmo indicado en la figura trece.

Figura 13. Algoritmo utilizado para el reconocimiento de rostros utilizando eigen caras



4. MÉTODOS DE CLASIFICACIÓN Y APRENDIZAJE, PARA EL RECONOCIMIENTO DE ROSTROS

4.1 Introducción

Cuando queremos aprender algo, primero lo clasificamos, esto permite a los seres humanos interactuar unos con otros, y reconocer características de diferentes cosas en la naturaleza.

Cuando hablamos de aprendizaje de un sistema computacional, pensamos en que primero hay que hacer un reconocimiento y una clasificación de patrones para tomar decisiones correctas.

En la actualidad existen diversos tipos de clasificadores de patrones clasificados en lineales y no lineales.

Como parte de los objetivos de esta investigación, se da más énfasis en las redes neuronales y los algoritmos genéticos, que podrían ser de mayor utilidad para el investigador que pretenda crear un sistema de reconocimiento de rostros robusto.

4.2 Clasificadores lineales

Los clasificadores lineales son utilizados para clasificar patrones por medio de algún método de cálculo matemático.

El objetivo de un clasificador es comparar un patrón de entrada, con un conjunto de patrones prototipo y determinar, a que clase pertenece este patrón.

Como ejemplo de un clasificador lineal, se pueden mencionar los basados en regionalización de espacio abstracto de objetos como el cálculo de eigen caras y los que usan funciones de distancia.

Si bien el cálculo de eigen caras utiliza una clasificación lineal, en si no solo es un clasificador si no que también es una técnica completa de reconocimiento de rostros, por lo que fue expuesta en el capítulo tres. En este capítulo nos limitaremos a explicar otro clasificador lineal, el clasificador de distancia euclidiana.

4.2.1 Clasificador de distancia euclidiana

Cuando se tienen los vectores descriptores del rostro, se procede a entrenar el detector para después realizar la clasificación, y con esto la detección y localización del rostro en una imagen.

El clasificador de distancia euclidiana conocido también como clasificador de distancia mínima ha sido el más usado hasta el momento en tareas de reconocimiento. Supone que las clases de objetos son de naturaleza determinística, es decir, los elementos que componen una clase se representan por un único vector llamado prototipo de la clase. Supone, además, que toda la información necesaria y suficiente para entrenar el clasificador es conocida con antelación. En otras palabras, requiere que todos los prototipos de las clases

estén a la disposición del sistema al comienzo del proceso de entrenamiento del clasificador.

En el diseño de todo clasificador, la selección de características es una parte vital, pues en base a éstas se obtiene el rendimiento. Sin embargo, la elección de buenas características depende de la habilidad, sentido y experiencia del diseñador de un sistema de reconocimiento de rostros, pues no existen reglas universales para llevar a cabo esta tarea.

El clasificador emplea la distancia euclidiana como medida de similitud entre un vector de entrada y un conjunto de vectores prototipo. Sobre la base de esta distancia, el clasificador determina la clase a la cual el patrón de entrada será asignado.

Matemáticamente, este proceso puede ser descrito como sigue:

Sea X es el vector de rasgos o características que describen un patrón de entrada, y Z_1, \dots, Z_n el conjunto de vectores de características que describen a los objetos de las diferentes n clases. El patrón X será clasificado en la clase con índice i si se cumple que exista la distancia mínima:

$$\text{mindist}(X, Z_i), i=1, \dots, n$$

Ecuación 4.1

La distancia Euclidiana entre dos vectores X y Z_i en términos de sus p rasgos se define como:

$$d_E(X, Z_i) = \sqrt{(x_1 - z_1)^2 + \dots + (x_p - z_p)^2}$$

Ecuación 4.2

En forma vectorial está expresada por:

$$d_E(X, Z_i) = \sqrt{(X - Z_i)^T (X - Z_i)}$$

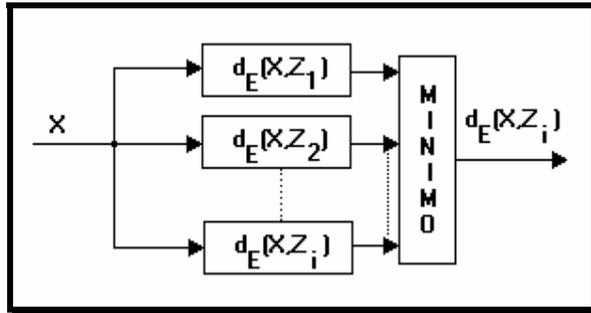
Ecuación 4.3

Por lo tanto, un vector de entrada X es puesto en la clase ω_i si $d_E(X, Z_i)$ es mínima. El funcionamiento de un detector euclídeo consta de los siguientes pasos:

- Diseño del detector, es decir, determinar los prototipos de las clases de interés.
- Dado un vector X a detectar, se calculan las distancias mínimas de X a cada uno de los prototipos Z_1, \dots, Z_n .
- Se asigna el vector X a la clase ω_i si la distancia $d_E(X, Z_i)$ es mínima.

El diagrama de bloques en la figura catorce muestra el proceso para la detección.

Figura 14. Funcionamiento del clasificador euclidiano



Para probar el clasificador se deben tener dos conjuntos de imágenes de prueba: las imágenes empleadas para el entrenamiento del detector y las imágenes propias para la prueba.

En general, los resultados obtenidos con las imágenes de prueba son mejores, debido a que en función de la información contenida en ellas fue diseñado el clasificador.

La prueba de desempeño se lleva a cabo cuando se introducen imágenes con información diferente a la empleada en el diseño del clasificador.

4.2.2 Clasificadores no lineales

Los clasificadores no lineales son con frecuencia utilizados en la clasificación de patrones en un sistema computacional. Como clasificadores no

lineales entendemos los que pueden ser calculados por métodos de simulación o algún método no matemático, por ejemplo evolución, sistemas neuronales, etc.

En este sub tema analizaremos dos clasificadores no lineales, las redes neuronales y los algoritmos genéticos.

4.2.1 Redes neuronales

El cerebro humano es la parte central de nuestro sistema. En él se procesan y toman todas las decisiones con respecto a el cuerpo. En él se realizan las tareas de clasificación y reconocimiento de un rostro u objeto cualquiera.

Durante muchos años se ha mantenido la idea de crear una máquina inteligente utilizando neuronas artificiales. En 1943 McCulloch y Pitts tratan al cerebro como a un organismo computacional, proporcionando los primeros resultados sobre mecanismos simuladores de la mente. Posteriormente, investigadores como Donald Hebb (1949) , Rosenblatt (1957) , Widrow y Hoff (1960) , Minsky y Selfridge (1961) , Block(1962), Minsky y Papert (1969) , entre otros, obtienen diferentes resultados en el transcurso de las siguientes dos décadas.

En 1969 el trabajo de Minsky y Papert pone en duda la eficacia de las redes neuronales para resolver problemas computacionales complejos. La incapacidad del perceptrón para resolver el problema de la función or-exclusiva,

abre un abismo en el tiempo durante el cual la mayoría de los investigadores alejan sus esfuerzos del trabajo con las redes neuronales.

En 1982 resurge el interés por las redes neuronales debido al creciente desarrollo de máquinas cada vez más rápidas e inclusive de procesamiento paralelo; otra razón y quizá la más importante, fue el descubrimiento de nuevas arquitecturas para redes neuronales y de algoritmos de aprendizaje más poderosos.

Las nuevas arquitecturas denominadas *conexionistas* se caracterizan por un gran número de elementos simples que procesan información de forma similar a una neurona biológica, además de un gran número de conexiones con “pesos” entre los elementos de procesamiento, codifican el conocimiento de la red.

El control entre los elementos de este tipo de arquitectura se lleva a cabo en paralelo, lo que permite realizar un aprendizaje automático de representaciones internas.

Como una red neuronal artificial está biológicamente inspirada, es necesario adentrarnos un poco en el funcionamiento de las neuronas cerebrales. En el sub tema 4.2.2.1.1 se exponen conceptos básicos sobre neuronas biológicas.

4.2.2.1.1 Fisiología de una neurona biológica

El sistema nervioso humano desarrolla una gran cantidad de funciones a través de diversas subdivisiones. Los mecanismos de procesamiento del sistema nervioso están constituidos por elementos primitivos conocidos como *células nerviosas* o *neuronas*.

Una neurona es la unidad celular funcional básica de este sistema que procesa y almacena información. Existe una gran variedad de tipos de neuronas en el cerebro de formas y tamaños diferentes, pero con características comunes en la mayoría de ellas.

La neurona biológica cuenta con cuatro regiones morfológicamente especializadas y son: el cuerpo celular, las dendritas, el axón y los terminales axónicos, cada una de ellas desarrolla una función particular.

El cuerpo de la neurona contiene el núcleo, así como los componentes reproductivos y metabólicos necesarios para las funciones fundamentales de la neurona.

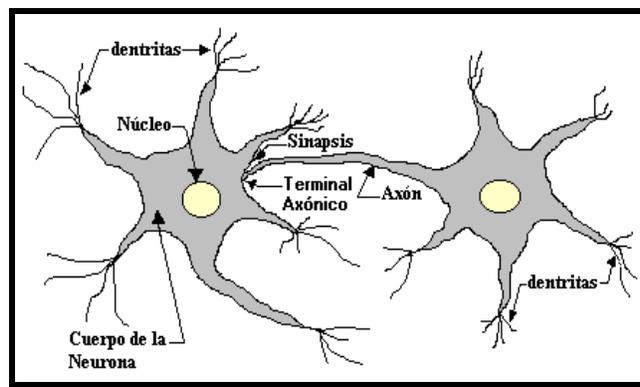
Éste recibe información de otras neuronas y desarrolla funciones integradoras importantes. Del cuerpo de la neurona nacen extensiones conocidas como dendritas que reciben información de otras neuronas, y con una longitud del orden de milímetros.

La estructura cilíndrica denominada axón tiene la función de conducir información codificada en forma de señales nerviosas hacia el terminal axónico.

El axón puede tener una longitud de menos de un milímetro y más de un metro, dividiéndose normalmente en varias ramificaciones cerca de su extremo final.

El proceso típico de transmisión de la información dentro de la neurona comienza cuando las dendritas reciben información de otras neuronas y la envían al cuerpo de la misma; aquí las señales son sumadas y transmitidas a lo largo del axón para transferirlas a otras neuronas.

Figura 15 Neurona y sus partes constitutivas



La transmisión de información entre neuronas se realiza por medio de reacciones químicas y tiene lugar en regiones que reciben el nombre de *sinapsis*, en donde los axones están muy cerca de las dendritas y de las hendiduras sinápticas de otras neuronas; es aquí donde se genera cualquier tipo de representación interna de la información.

4.2.2.1.2 Redes neuronales artificiales

Las redes neuronales son una herramienta matemática que permite la solución de una gran variedad de problemas, en particular los de difícil modelado.

El perceptrón es el modelo básico de una red neuronal. Fue desarrollado por Rosenblatt en 1962. Las redes neuronales artificiales (RNA de aquí en adelante) son redes que tratan de interactuar del mismo modo en que lo hace el sistema nervioso biológico; las neuronas individuales usualmente adaptables son organizadas en forma jerárquica y se interconectan masivamente en paralelo.

En la actualidad se ha incrementado el uso de las RNA, debido a que presentan un gran número de características similares a las del cerebro. No se ha logrado, sin embargo, diseñar una red neuronal que realice procesos tales como la adquisición, el almacenamiento, la representación y el análisis de datos sensoriales complejos, como lo hacen de manera cotidiana las redes neuronales biológicas.

Una de las principales características de las RNA es que son sistemas dinámicos auto adaptables, es decir, que poseen la capacidad de cambiar constantemente para adaptarse a nuevas condiciones. Otra característica muy importante es la de autoorganización, con la cual se logra la generalización,

esto es, la red puede responder en forma adecuada a la presentación de datos a los que no fue expuesta con anterioridad.

La tolerancia a fallos es una característica inherente en una red neuronal. Por una parte, la red puede aprender a reconocer patrones con ruido o incompletos; por otra parte la red puede seguir llevando a cabo su función aun cuando parte de la red haya sido destruida. Lo anterior se debe a que la información se encuentra distribuida en las conexiones entre neuronas, por lo que y existe cierto grado de redundancia en este tipo de almacenamiento.

Una RNA también puede ser usada en sistemas de tiempo real, puesto que la implementación de la red se realiza en paralelo; y el requerimiento de cambio en los pesos de las conexiones es mínimo.

Las RNA son una herramienta computacional de aplicaciones múltiples. Actualmente, las RNA se aplican en la medicina, la industria, el comercio, la biología, la industria militar, entre otros. Existen diversos tipos de redes neuronales y, depende del problema al que se deba dar solución, se selecciona el tipo de red más apropiado.

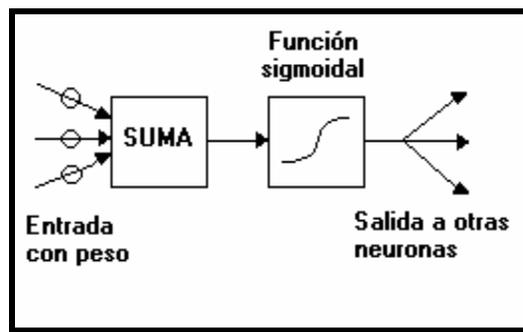
En el área de reconocimiento de patrones no sólo se busca la detección de un objeto o la clase a la que pertenece, sino también se implementan sistemas de percepción artificial que imiten las funciones de los sistemas sensoriales humanos.

4.2.2.1.2.1 Funcionamiento de una RNA

Las partes que constituyen una RNA son básicamente los elementos de procesamiento o neuronas, su estado de activación, sus funciones de salida y la conexión con otras neuronas.

Una RNA está compuesta de elementos de procesamiento denominados unidades de proceso o neuronas a partir de las cuales es posible representar prácticamente los objetos. Las neuronas aceptan señales de entrada de neuronas previas, y son multiplicadas individualmente por un peso y después sumadas para obtener una salida. El resultado es no-lineal y escalado entre los valores 0 y 1, para lo cual se utiliza la función sigmoideal como se explica más adelante. El valor de salida se envía a otras neuronas (Figura dieciséis).

Figura 16. Estructura de una neurona artificial



Un perceptrón, por ejemplo, calcula la suma ponderada de sus entradas y entrega como salida un 1 si dicha suma es mayor que un determinado umbral; o un 0 para el caso contrario.

Los valores de las entradas: y_1, y_2, \dots, y_n , y los pesos de conexión w_1, w_2, \dots, w_n (figura diecisiete), normalmente son números reales positivos o negativos. Si alguna entrada y_i activa al perceptrón el peso w_i es positivo; si la entrada y_i inhibe al perceptrón, el peso w_i es negativo.

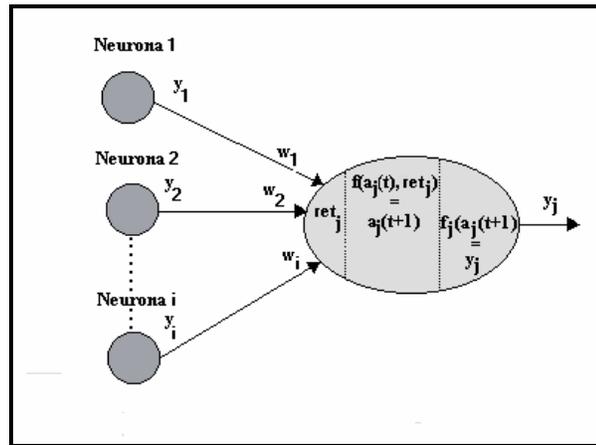
En la figura diecisiete se observa que la i -ésima neurona tiene asociado en cualquier instante un estado de activación $a_i(t)$, así como una función de salida o de transferencia f_j que transforma el estado actual de activación en una señal de salida y_j . La señal de salida se envía por medio de las conexiones unidireccionales a otras neuronas en la red.

Las señales que llegan a la j -ésima neurona se combinan entre sí generando una entrada total representada por la siguiente ecuación:

$$red_j = \sum_i y_i w_{ji}$$

Ecuación 4.4

Figura 17. Funcionamiento de una neurona artificial aislada.



Existen dos formas de actualizar los estados de una neurona. Uno es el modo síncrono en el cual la información que llega a las neuronas es en forma continua y donde los cambios se efectúan simultáneamente como comandado por un reloj interno que indica cuando debe cambiar su estado.

El otro es el modo asíncrono en el que las neuronas verifican su estado continuamente, conforme reciben la información.

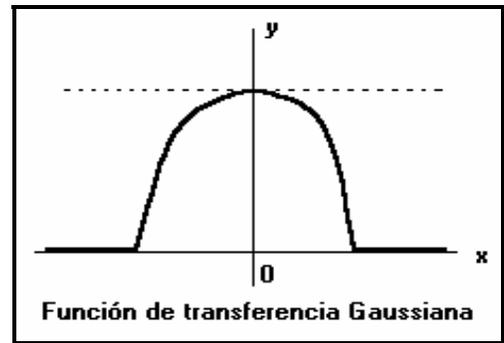
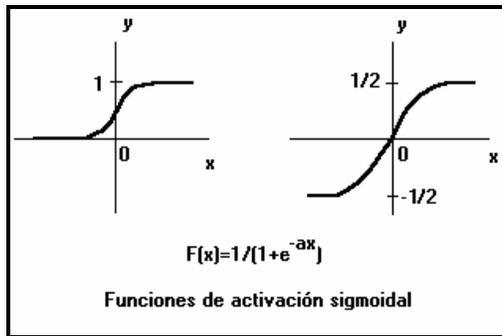
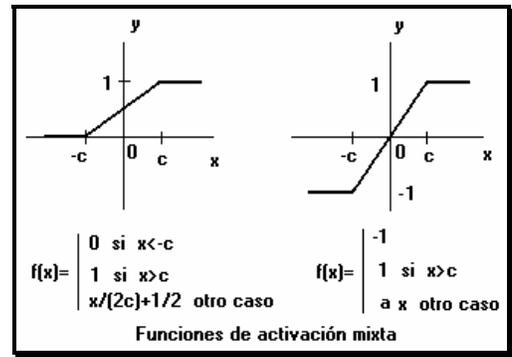
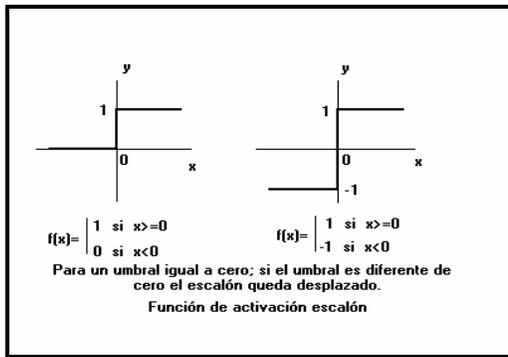
La función de activación determina el nuevo estado de activación $a_j(t+1)$ de la neurona, sobre la base de la entrada total calculada y el estado de activación anterior $a_i(t)$.

Las funciones de activación más comunes son (ver figura dieciocho):

- La función escalón.

- La función lineal y mixta.
- La función sigmoial (continua).
- La función gaussiana.

Figura 18. Funciones típicas de activación de una neurona artificial



Una neurona recibe información del estado de activación de todas las neuronas a las cuales está conectada. Cada sinapsis (conexión) entre neuronas está

ponderada por un peso, por lo que la entrada neta que recibe una neurona es la suma del producto de cada señal individual por el peso que conecta ambas neuronas, lo cual queda expresada por la ecuación 4.4.

4.2.2.1.2.2 Estructura básica de una RNA

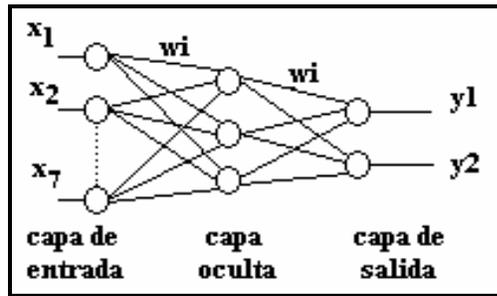
La estructura de una RNA está determinada por el número de capas o niveles, el número de neuronas por nivel, el patrón de conexión y el flujo de información.

En una RNA multicapa la distribución de las neuronas se lleva a cabo formando capas o niveles con un número determinado de éstas y según el lugar en donde se encuentren se clasifican en tres tipos:

- 1. Capa de entrada.** Recibe directamente la información de entrada, la información que será procesada por la red.
- 2. Capa(s) oculta(s).** Se encuentra(n) en la parte interna de la red y no tiene(n) contacto directo con la información de entrada. El número de capas ocultas se determinada manualmente y depende del problema a resolver.
- 3. Capa de salida.** Transfiere hacia el exterior el resultado de la información de entrada procesada por la red.

En la figura diecinueve se muestra una red neuronal con 3 capas: la de entrada con 7 neuronas, la oculta con 3 neuronas, y la de salida con 2 neuronas.

Figura 19. Red neuronal compuesta de 3 capas.



4.2.2.1.2.4 Tipos de RNA

Las RNA se pueden clasificar en redes monocapa (de una capa) y redes multicapa (de 2 o más capas). En el primer tipo, se establecen conexiones laterales entre las neuronas de la misma capa. Estas conexiones pueden ser del tipo auto recurrente. En el segundo tipo, las neuronas se encuentran agrupadas en varias capas. Normalmente, las neuronas de una capa oculta reciben información de la capa anterior más próxima a la entrada de la red, y transmiten la información procesada a las neuronas de una capa posterior más próxima a la salida de la red. Este tipo de conexión es denominada conexión hacia adelante o *feed forward*.

Las RNA multicapa pueden tener conexiones hacia adelante y hacia atrás. En una red con conexiones hacia adelante, toda la información se propaga hacia adelante a través de las capas que conforman la red. En este caso, no se tienen conexiones hacia atrás, laterales o auto recurrentes.

En el caso de una red con conexión hacia adelante-atrás, cuando la red se encuentra en funcionamiento la información es transferido hacia adelante por medio de conexiones tipo *feed forward* y hacia atrás a través de conexiones tipo *feedback*.

4.2.2.1.2.5 Aprendizaje de una RNA

El aprendizaje es un proceso a través del cual la RNA modifica los pesos de las conexiones entre sus neuronas según la información de entrada. Es importante saber cuál es el criterio para modificar los valores asignados a las conexiones, cuando se desea que una RNA aprenda nueva información y determinar así la regla de aprendizaje conveniente para un determinado proceso.

Las reglas de entrenamiento pueden clasificarse en reglas para aprendizaje no supervisado o sin maestro y para aprendizaje supervisado o con guía. La diferencia principal entre ambos tipos de aprendizaje es la existencia o no de un agente externo conocido como supervisor cuya función es controlar el proceso de aprendizaje de la RNA.

Existen varias maneras para llevar a cabo una u otra de estas reglas de aprendizaje.

El aprendizaje *supervisado* puede realizarse a través del:

- Aprendizaje por corrección de error.
- Aprendizaje por refuerzo.
- Aprendizaje estocástico.

En cuanto al aprendizaje *no supervisado* tenemos que éste puede llevarse a cabo por:

- Aprendizaje hebbiano.
- Aprendizaje competitivo y cooperativo.

4.2.2.1.2.6 Asociación de información en una RNA

Las redes neuronales pueden ser consideradas como elementos de una memoria asociativa. Durante el proceso de aprendizaje las redes almacenan patrones de información y presentan un patrón dado a su salida según un estímulo de entrada, es decir, al introducir a la red un estímulo dado la red responde con una salida asociada a ese estímulo.

Existen dos formas de realizar la asociación entre pares de entrada-salidas: la de heteroasociación y la de auto asociación.

En el caso de la heteroasociación, se presentan a la entrada de la red pares ordenados de datos de entrada-salida $(A_1, B_1), (A_2, B_2), \dots, (A_N, B_N)$ para que la red aprenda, de manera que cuando a la entrada se presente la parte A_i , a la salida se manifieste la correspondiente parte B_i .

Para la auto asociación se presentan también a la entrada de la red datos, digamos A_1, A_2, \dots, A_N , con el fin de que la red aprenda, si estos datos forman una secuencia. Al momento del reconocimiento, la red realizará una auto correlación entre un estímulo entrada y los datos almacenados durante el aprendizaje, y a la salida entregará el dato almacenado que más se parezca al estímulo de entrada.

4.2.2.1.2.7 Representación de la información de entrada y salida en una RNA

Una RNA puede también clasificarse por la forma en la que la red representa la información de entrada y salida. En la mayoría de modelos de redes neuronales, tanto los datos de entrada como los datos de salida son valores reales continuos normalizados, siendo su valor absoluto menor que uno; las funciones de activación son continuas del tipo lineal o sigmoideal. Los modelos de redes que admiten valores discretos o binarios a su entrada generan salidas del tipo binario. La función de activación para estos modelos es del tipo escalón.

4.2.2.1.3 La regla delta generalizada

La regla delta generalizada es el algoritmo de aprendizaje de la red, el cual dio lugar al algoritmo de propagación hacia atrás. Este algoritmo utiliza una función de error asociada a la red para establecer el error mínimo. El error del sistema es realimentado hacia atrás para modificar los pesos en un valor proporcional al valor con el que contribuyeron en la salida.

La regla del error mínimo cuadrático medio es un método de supervisión para encontrar el valor de los pesos requeridos en forma de vector, los cuales permiten asociar cada patrón de entrada con su valor de salida correspondiente. Este método se basa en la búsqueda del error mínimo entre la salida deseada y la salida lineal obtenida antes de aplicar la función de activación.

Esta forma de evaluación del error permite que la RNA pueda procesar información analógica, tanto a su entrada como a su salida, por medio de una función de activación continua.

A la regla del error mínimo cuadrático medio también se le conoce como regla delta, porque busca minimizar una delta o diferencia entre el valor obtenido y el valor deseado a la salida de la red, previa a la aplicación de la función de activación lineal. Otros términos que denotan el mismo concepto son: algoritmo de propagación hacia atrás del error, regla de aprendizaje de mínimos cuadrados o gradiente descendente.

A través de la aplicación a la entrada de la red de un valor de pesos descendentes el algoritmo de propagación hacia atrás encuentra un valor mínimo de error. Este valor puede ser local o global. Cada punto de la superficie de la función de error corresponde a un conjunto de valores de los pesos de la red. Por sus características, este algoritmo permite asegurar la bajada por dicha superficie hasta encontrar el valle más cercano, y provocar que el proceso de aprendizaje se detenga en un mínimo local. De acuerdo con esto, lo que se busca con el algoritmo de propagación hacia atrás es minimizar el error, mediante la inicialización de los pesos con valores pequeños y aleatorios (normalmente entre +/- 0.5); de no ser así, el algoritmo puede pasar sobre el punto mínimo sin detenerse en él. No siempre se puede llegar a un mínimo global, pero se puede alcanzar un valor de error mínimo preestablecido.

La selección de un valor adecuado para la velocidad de aprendizaje permite un buen rendimiento de la red. Esta es la velocidad a la que el algoritmo converge. La velocidad de aprendizaje es manipulada por medio del parámetro α , cuyo valor fluctúa normalmente en el rango 0.05 y 0.25. Aunque con un valor dentro de este rango el número de iteraciones tiende a crecer.

Conforme el valor del error va disminuyendo, el valor de η se incrementa. Por tanto, esta es una forma de acelerar la convergencia. Se debe tener, sin embargo, precaución de que α no sea demasiado grande. Esto puede causar que el algoritmo se aleje del valor mínimo real.

Otra forma posible para incrementar la velocidad de convergencia es utilizar un término llamado *momento* η ; esto corresponde a sumar una fracción

del cambio de peso anterior cuando se calcula el valor del cambio del peso actual. Con el momento se tiende a mantener los cambios de peso en la misma dirección.

4.2.2.1.10.1 Funcionamiento de la regla delta generalizada

Para una neurona N_i y su salida y_i , el cambio en el peso de la conexión que une la salida de esta neurona con la neurona U_j para un patrón de aprendizaje x viene dado como:

$$\Delta w_{ji}(t+1) = \alpha \delta_{xj} y_{xi}$$

Ecuación 4.5

Donde α es la constante o tasa de aprendizaje, y

$$\delta_{xj} = (d_{xj} - y_{xj}) f'(net_j)$$

Ecuación 4.6

Donde d_{xj} es la salida deseada en la neurona j para el patrón x y net_j es la entrada neta a la misma neurona. Si la neurona j no se encuentra en la capa de salida, el error δ_{xj} producido viene dado como una función del error existente en las neuronas que reciban como entrada la salida de la neurona j :

$$\delta_{xj} = \left(\sum_k \delta_{xk} w_{kj} \right) f'(net_j)$$

Ecuación 4.7

Donde k cubre a todas las neuronas conectadas a la neurona j . Por lo tanto, el error producido en una neurona oculta es la suma de los errores que se producen en las neuronas conectadas a la salida de dicha neurona oculta multiplicado por el peso de la conexión correspondiente.

4.2.2.1.10.2 Momento en la regla delta generalizada

La regla delta generalizada requiere de gran capacidad de cómputo para realizar cálculos hasta ajustar todos los pesos de la red. Al implementar dicha regla en forma algorítmica se requiere tomar en cuenta una amplitud de paso que viene dada por la tasa de aprendizaje α , esto es:

$$w_j(t+1) = w_{ji}(t) + \alpha \delta_{xj} y_{xi} + \eta (w_{ji}(t) - w_{ji}(t+1)) = \Delta w_{ji}(t+1) = \alpha \delta_{xj} y_{xi} + \eta \Delta w_{ji}(t)$$

Ecuación 4.8

Donde η es un momento que determina el efecto en el instante $t+1$ del cambio de los pesos en el instante t , y que produce la convergencia de la red en un menor número de iteraciones. Si en t y $t+1$ el incremento de un peso es positivo, entonces el descenso por la función de error es mayor. Si en $t+1$ éste

es negativo, el peso que se da en $t+1$ es más pequeño, lo cual es adecuado. Esto significa que ha pasado por un mínimo y que los pesos deben ser menores para poder alcanzarlo.

En términos del momento η , para una neurona de la capa de salida se tiene:

$$\eta(w_{kj}^o(t) - w_{kj}^o(t+1))$$

Ecuación 4.9

Mientras que para una neurona de la capa oculta se tiene:

$$\eta(w_{ji}^h(t) - w_{ji}^h(t+1))$$

Ecuación 4.10

4.2.2.1.11 El algoritmo de propagación hacia atrás

En 1986 Rumelhart, Williams y Hinton desarrollaron el algoritmo conocido como propagación del error hacia atrás *backpropagation*. Se basa en la generalización de la regla delta con conexiones hacia adelante y funciones de activación continuas no decrecientes y derivables. Este algoritmo permite que una RNA de varias capas aprenda la asociación existente entre patrones de entrada y sus clases correspondientes.

El algoritmo de propagación hacia atrás es una regla de aprendizaje que organiza la capa intermedia para conseguir la correspondencia entre estímulos de entrada y patrones de salida.

El funcionamiento de una red entrenada por medio del algoritmo de propagación hacia atrás radica en el aprendizaje de un conjunto predefinido de pares de entrada-salida que se proporcionan como estímulos durante un ciclo conocido como ciclo de propagación-adaptación compuesto de dos fases.

Durante la primera fase se aplica un patrón de entrada a la primera capa de la red y éste se propaga a través de las capas superiores hasta generar un resultado a la salida. Este resultado es comparado con la salida deseada, obteniendo así un error para cada neurona de salida.

El error es transmitido hacia atrás a todas las neuronas de la capa intermedia que contribuya directamente a la capa de salida, asimilando el porcentaje de error aproximado que corresponda a su participación en la salida original para realizar el ajuste de pesos correspondiente. Este proceso se realiza para cada una de las capas hasta que todas las neuronas de la red tengan un error describiendo su aportación relativa al error total.

Con base en los errores obtenidos, los pesos entre neuronas son ajustados; de tal manera que cuando el mismo patrón se presente a la entrada, el error disminuirá. Conforme el error disminuye la salida está más cerca a la salida deseada.

4.2.2.1.12 Algoritmo de aprendizaje

En este sub tema se describe brevemente el proceso seguido para el entrenamiento de una red por medio del algoritmo de propagación del error hacia atrás "backpropagation". Consta de los siguientes pasos:

- Inicialización de los pesos de la red con pequeños valores aleatorios.
- Alimentación de un patrón de entrada $Vx = (v_{x1}, v_{x2}, \dots, v_{xn})$, especificándose la salida $S=(s_1, s_2, \dots, s_m)$ que se desea que debe generar la red.
- Cálculo de la salida actual de la red tomando como base todas las entradas presentadas a la entrada de la red y la salida en cada capa hasta llegar a la capa de salida.
- Se calculan también las entradas netas para las neuronas ocultas que proceden de las neuronas de entrada mediante la ecuación:

$$net_{xj}^h = \sum_{i=1}^N w_{ji}^h v_{xi} + \theta_j^h$$

Ecuación 4.11

En donde h se refiere a la capa oculta "*hidden*", x es el x -ésimo vector de entrenamiento y j es la j -ésima neurona oculta; θ es un factor opcional que actúa como otra entrada.

- Cálculo de la salida de las neuronas de la capa oculta:

$$y_{xj} = f_j^h(\text{net}_{xj}^h)$$

Ecuación 4.12

- Cálculo de los valores netos de las entradas para cada unidad en la capa de salida:

$$\text{net}_{xk}^o = \sum_{j=1}^N w_{kj}^o y_{xj} + \theta_k^o$$

Ecuación 4.13

Donde

$$y_{xk} = f_k^o(\text{net}_{xk}^o)$$

Ecuación 4.14

Y las salidas

$$o_{xk} = f_k^o(\text{net}_{xk}^o)$$

Ecuación 4.15

- Cálculo del error para cada una de las neuronas en la red. Para la neurona

k de la capa de salida se tiene:

$$\delta_{xk}^o = (d_{xk} - y_{xk}) f_k^{o'}(net_{xk}^o)$$

Ecuación 4.16

Donde f es una función de salida derivable y por tanto del tipo sigmoideal para representar los datos de salida en forma binaria:

$$f_k(net_{jk}) = \frac{1}{1 + e^{-net_{jk}}}$$

Ecuación 4.17

La derivada de la función sigmoideal es:

$$f_k^{o'} = f_k^o(1 - f_k^o) = y_{xk}(1 - y_{xk})$$

Ecuación 4.18

Por lo que los valores de error para las neuronas de salida como:

$$\delta_{xk}^o = (d_{xk} - y_{xk}) y_{xk} (1 - y_{xk})$$

Ecuación 4.19

La derivada parcial del error no puede ser evaluada en forma directa cuando la neurona j no pertenece a la capa de salida.

El desarrollo se realiza a partir de valores conocidos y que puedan ser evaluados, esto es:

$$\delta_{xj}^h = v_{xi} (1 - v_{xi}) \sum_k \delta_{xk}^o w_{kj}^o$$

Ecuación 4.20

Donde k representa las neuronas de la capa superior a la neurona j .

De la ecuación anterior se puede ver que el error producido en una neurona de la capa oculta, es proporcional a la suma de los errores conocidos multiplicados cada uno por el peso de la conexión de cada neurona conectada a la salida de la neurona j .

Para actualizar los pesos se utiliza un algoritmo recursivo, tomando las neuronas de la capa de salida y continuando hacia atrás hasta llegar a la capa de entrada.

Los pesos para la capa de salida se calculan como sigue:

$$w_{kj}^o(t+1) = w_{kj}^o(t) + \Delta w_{kj}^o(t+1)$$

Ecuación 4.21

Donde

$$\Delta w_{kj}^o(t+1) = \alpha \delta_{xk}^o y_{xj}$$

Ecuación 4.22

Por otro lado, los pesos para la capa oculta se calcularán como:

$$w_{ji}^h(t+1) = w_{ji}^h(t) + \Delta w_{ji}^h(t+1)$$

Ecuación 4.23

Donde

$$\Delta w_{ji}^h(t+1) = \alpha \delta_{xj}^h v_{xi}$$

Ecuación 4.24

Es posible agregar el momento η descrito con anterioridad para producir la convergencia de la red en un número menor de iteraciones.

El proceso anterior se repetirá hasta que el error sea pequeño, menor o igual a un valor preestablecido para cada uno de los patrones aprendidos.

4.2.2.1.4 Arquitectura de reconocimiento de rostros basada en redes neuronales electrónicas

Como se describió en el capítulo tres, antes de aplicar un método de clasificación y aprendizaje, es necesario obtener una ecuación matemática que represente un patrón.

En este caso los descriptores invariantes, son las herramientas que se utilizan para crear una plantilla que sirva para reconocer un patrón, en el caso de esta investigación, el patrón de un rostro humano.

Cuando se tiene definido un vector de características (propuesto en el capítulo tres), se está listo para diseñar la arquitectura del sistema de reconocimiento de rostros utilizando redes neuronales.

4.2.2.1.4.1 Diseño de la Arquitectura

La arquitectura del sistema de reconocimiento de rostros, está compuesta de varios clasificadores, cada uno compuesto por una red neuronal artificial de tipo ART (*Adaptative Resonance Theory*).

Esta red neuronal está basada en el modelo ART que fue propuesto por G.A. Carpenter y S. Grossberg en 1986 para extender los mecanismos de aprendizaje competitivos de contra propagación y mapas organizados.

La red se conecta a un mapa de memoria (MM). Un mapa de memoria es donde se almacenan las salidas de la red neuronal, que servirán para modificar los pesos en las entradas. Las salidas de los clasificadores se

conectarán a un registro de evidencia y a un conjunto de comparadores (ver figura veinte).

Cada RNA debe ser entrenada, y el mapa de memoria construido. Para la explicación de este caso, usaremos dos fases, la de inicialización y la de indexado (ver figura veintiuno). Durante la primera fase la red neuronal es entrenada y el mapa de memoria construido.

En la siguiente etapa se realiza la detección de características visuales y la codificación de estas.

Figura 20. Arquitectura de la red neuronal

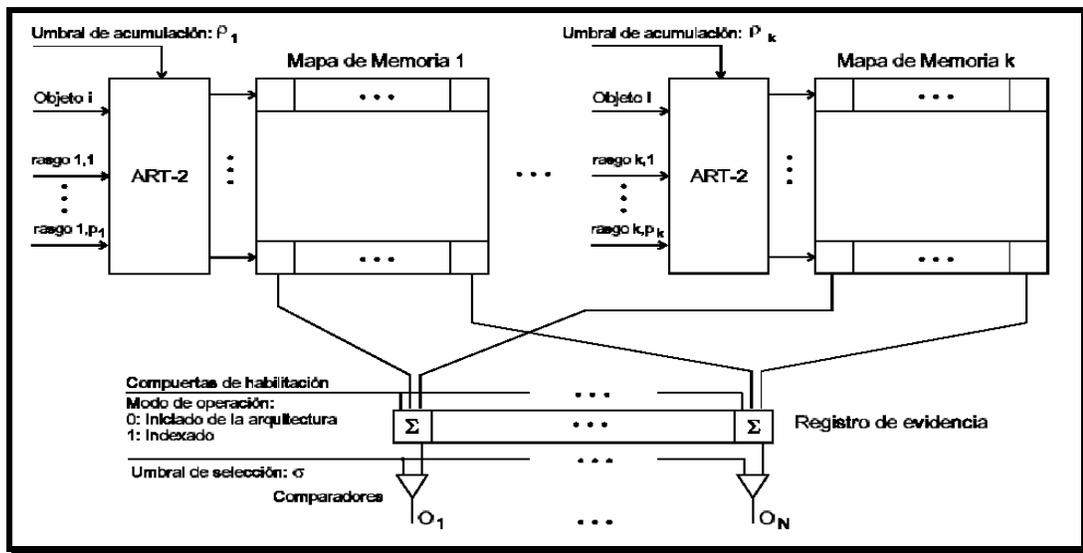
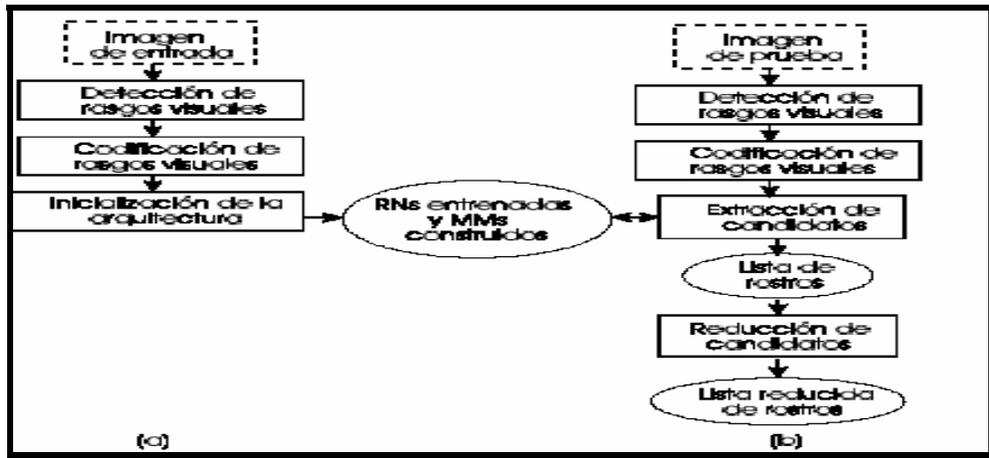


Figura 21. Fases de la arquitectura de una Red Neuronal



4.2.2.1.4.2 Detección de Características Visuales

El funcionamiento de la arquitectura se basa en el hecho de que para reconocer un rostro, es necesario descomponerlo en un conjunto de partes llamadas características visuales (CVs).

Estas CVs permitirán, el reconocimiento eficiente de personas a través de información visual.

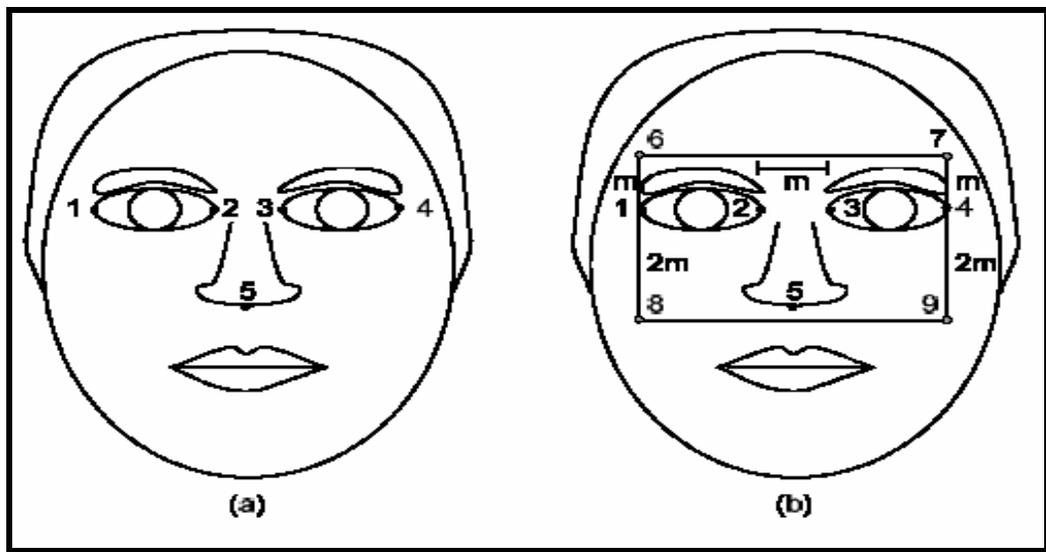
Un conjunto selecto de puntos es primeramente detectado de una imagen del rostro. Estos puntos son los mostrados en la figura veintidós (a).

Cinco puntos son seleccionados manualmente. A partir de estos cinco puntos, otros cuatro puntos son calculados y marcados. Todos estos puntos serán llamados en adelante *puntos de interés*. Se muestran en la figura

veintidós (b) y son obtenidos en términos de la distancia entre los puntos 2 y 3. Con este conjunto de puntos, la siguientes CVs son derivadas:

- ✓ Todos los píxeles encerrados en el rectángulo formado por los puntos 6,7, 8 y 9.
- ✓ El triángulo formado por los puntos 1, 2 y 5.
- ✓ El triángulo formado por los puntos 2, 3 y 5.
- ✓ El triángulo formado por los puntos 3, 4 y 5.
- ✓ El triángulo formado por los puntos 1, 4 y 5.

Figura 22. Puntos de interés a través de los cuales una característica visual es extraída



A partir de los puntos calculados anteriormente, se calculan los descriptores invariantes, que fueron expuestos en el capítulo tres.

4.2.2.1.4.3 Codificación de las características visuales

Las CVs derivadas en la sección anterior son codificadas en términos de los siguientes dos vectores de características:

- Los cinco puntos 1, 2, 3, 4 y 5 mostrados en la figura veintitrés(a) son codificados como el siguiente vector de tres componentes:

$$v_1 = \left[\frac{|M_{125}| + |M_{345}|}{|M_{145}|}, \frac{|M_{135}| + |M_{245}|}{|M_{145}|}, \frac{|M_{235}|}{|M_{145}|} \right]$$

Ecuación 4.25

Donde:

$$|M_{ijk}| = \begin{vmatrix} x_i & x_j & x_k \\ y_i & y_j & y_k \\ 1 & 1 & 1 \end{vmatrix}$$

Los niveles de intensidad encerrados por la ventana formada por los píxeles 6, 7, 8 y 9 (figura veintitrés(b)) son codificados en forma global como un vector de características compuesto por tres de los invariantes a translación, rotación y escala de Hu (consultar capítulo tres): ϕ_1, ϕ_4 , y ϕ_5 .

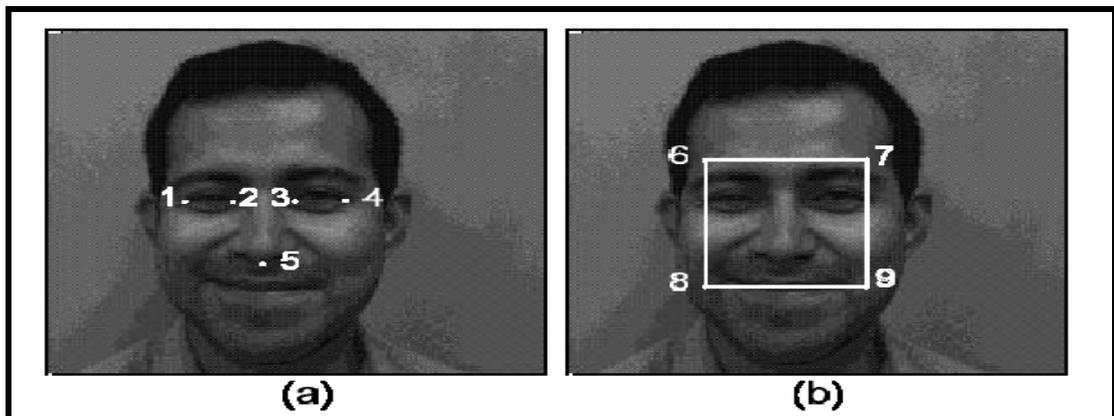
Sea este vector:

$$v_2 = [\phi_1 \quad \phi_4 \quad \phi_5]$$

Ecuación 4.26

Estos dos vectores de atributos (VAs de aquí en adelante) serán enseguida usados para entrenar los dos clasificadores correspondientes.

Figura 23. Configuraciones de puntos para obtener los vectores de atributos o características.



Hasta ahora se definió la arquitectura del sistema de reconocimiento de rostros, en el sub tema 4.1.2.1.2.1 se explican las 2 fases que realiza el sistema.

4.2.2.1.4.4 Fases de la arquitectura

La arquitectura del sistema cuenta de dos fases como se mencionó anteriormente, la fase de inicialización y la fase de indexado.

4.2.2.1.4.4.1 Fase de inicialización

Para llevar a cabo esta fase, un procedimiento compuesto de dos etapas es usado. Cada una de estas etapas es explicada con detalle a continuación.

Etapas de entrenamiento de la red neuronal y construcción del MM. Durante esta etapa, cada ART2 y su mapa de memoria son, respectivamente, entrenados y construidos.

La figura veinticuatro muestra el algoritmo de la fase de inicialización. Al final de este procedimiento iterativo, la arquitectura habrá sido inicializada, esto es las RNs (Redes Neuronales) y sus MMs (Mapas de Memoria) habrán sido, respectivamente entrenados y construidos.

La arquitectura final tiene la misma estructura que la mostrada en la figura veinte, excepto que el número de clasificadores es dos, uno por cada vector de atributos obtenido.

Figura 24. Inicialización de la arquitectura

```
Algoritmo 1: Inicialización de la arquitectura.  
Entrada: Conjunto de  $N$  imágenes de rostros localizados.  
Salida: ART2s entrenadas y MMs construidos.  
{  
Para cada rostro localizado:  
{  
Detectar puntos iniciales.  
Calcular puntos adicionales.  
Derivar características visuales.  
Codificar características visuales.  
Entrenar el  $k$ -ésimo clasificador usando el  $k$ -ésimo VA.  
}  
}
```

Nótese que cada ART2 tiene como entrada un vector de atributos y un número de salidas igual al número de clases generadas por la ART2, dependiendo del factor de vigilancia ρ usado. Nótese también que cada MM tiene tantas filas como salidas provistas por la ART2 y tantas columnas como rostros usados para entrenar la ART2.

Cada localidad del MM contiene un valor el cual representa el número de veces que un VA se encuentra presente en cada rostro entrenado. Este valor es ponderado dependiendo del número de objetos conteniendo ese VA, entre más objetos contengan ese VA, más pequeño será el peso asignado a ese VA.

El conjunto de compuertas de habilitación determina el modo de operación del sistema (inicialización o indexado). El registro de evidencia almacena los votos ponderados obtenidos por cada rostro durante la etapa de selección de candidatos (ver sección 4.2.1.4.4.2).

Finalmente, el conjunto de comparadores es usado durante la etapa de reducción de candidatos (sección 4.2.1.4.4.2) para seleccionar de entre estos rostros, aquéllos con más votos en términos de un umbral previamente escogido. Normalmente cualquiera de estos rostros corresponderá al rostro presente en la imagen de prueba. Idealmente, la lista reducida debería contener el modelo del rostro presente en la imagen.

Una red neuronal permite, por un lado, agrupar automáticamente patrones para formar cúmulos de una manera auto organizada y, por otro lado, debido a sus propiedades inherentes de indexado. Cada mapa de memoria es construido conforme la red es entrenada.

El número de cúmulos provistos por una ART2 depende del valor escogido para ρ . Si el valor escogido para ρ es muy pequeño, una gran cantidad de cúmulos será generada (vectores de atributos muy similares serán puestos en cúmulos diferentes).

Inversamente, si el valor escogido para ρ es muy grande, pocos cúmulos serán generados (vectores de rasgos muy disimilares serán agrupados en el mismo cúmulo).

La selección de este valor ρ para un clasificador dado ,se debe escoger de forma que el clasificador permita diferenciar lo mejor posible entre VAs de personas diferentes y agrupar Vas de la misma persona.

Para lograr esto, se deberá calcular la diferencia E entre VAs de personas diferentes y la diferencia F entre VAs de la misma persona se seleccionó un valor de umbral inicial ρ_0 , como el promedio de E.

Luego, se comparan los valores de E y F con el valor de ρ_0 . Si $E > \rho_0$, significa que ρ_0 diferenció los VAs de personas diferentes, y se incrementa en 1 un contador CD (previamente inicializado a cero). Si $F > \rho_0$, significa que ρ_0 confunde los VAs de la misma persona, y se incrementa en 1 un contador CC. Este mismo proceso se repite para diferentes valores de ρ_i con $\rho_i = \rho_{i-1} + \Delta$.

Finalmente, se selecciona el valor de umbral óptimo como el máximo de CD-CC, esto es aquel que diferencia mejor entre VAs de personas diferentes y confunde menos entre VAs de la misma persona.

Una vez que los mapas de memoria han sido construidos, estos son ponderados. Para esto se usa el algoritmo de la figura veinticinco.

Figura 25. Algoritmo utilizado para la ponderación de mapas de memoria.

```
Algoritmo 2: Ponderación de mapas de memoria.  
Entrada: K mapas construidos.  
Salida: K mapas ponderados.  
{  
Para mapa de memoria j=1 hasta j=K.  
Para vector de atributos l=1 hasta l=Cj.  
Para i=1 hasta i=N.  
[Mj1,i]=[Ml,i]j/nol  
}  
Donde [Mj1,i] es una de las localidades del j-ésimo mapa de memoria,  
Cj el número de cúmulos generado por el j-ésimo clasificador y nol el  
número de objetos conteniendo ell-ésimo vector de atributos.
```

4.2.2.1.4.4.2 Fase de Indexado

Durante esta fase, los vectores de atributos calculados a partir de las características visuales en una imagen de prueba son usados para recuperar del banco de modelos el conjunto de rostros buscado. Para esto la imagen de prueba es primeramente procesada por medio del conjunto de técnicas descritas en las secciones 4.2.2.1.4.4.2.1 y 4.2.2.1.4.4.2.2.

Cada VA es usado para extraer del mapa de memoria correspondiente los rostros que produjeron dicho VA dando como resultado una lista de rostros candidatos. Esta lista es finalmente reducida por medio de un mecanismo de umbralizado. Estas dos etapas son enseguida descritas en más detalle.

4.2.2.1.4.4.2.1 Selección de candidatos

Durante esta etapa, cada vector de atributos obtenido a partir de las características visuales en una imagen de prueba es presentado al sistema.

Si este vector es lo suficientemente similar a uno de los aprendidos por la red neuronal (en términos de la distancia euclidiana), entonces la salida correspondiente de dicha red será habilitada seleccionando una fila de su mapa de memoria. Los votos ponderados en esa fila serán enviados a un registro de evidencia.

Este proceso se repite para cada VA. Al final se tendrá un conjunto de hipótesis arregladas como un histograma de pesos ponderados (figura veintiséis (a)).

4.2.2.1.4.4.2 Reducción de candidatos

Es claro que al final del proceso anterior se tendrá un histograma conteniendo los pesos ponderados que cada rostro ha recibido durante dicho proceso (para un ejemplo, ver la figura veintiséis (a)). Algunos rostros habrán recibido más votos que otros.

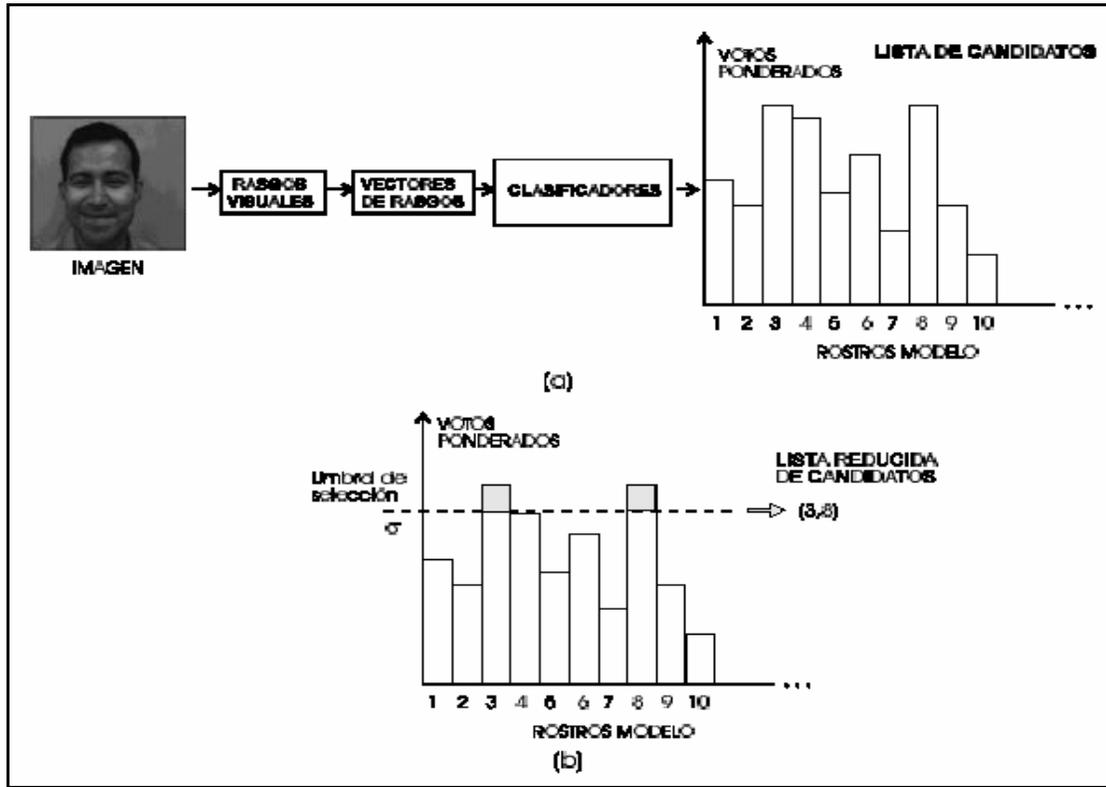
Es necesario decidir que rostros se aproximan lo más al rostro presente en la imagen. Intuitivamente, aquellos rostros con más votos serán los mejores candidatos. Una manera de decidir que objetos se encuentran presentes en la imagen consiste en usar un umbral de selección, digamos σ .

Así, aquellos rostros cuyos votos se encuentren arriba de σ serán seleccionados como los mejores candidatos. Para decidir si un rostro se encuentra presente en la imagen de entrada se usa el algoritmo de la figura veintisiete.

H_i contiene los votos ponderados recibidos por el rostro i durante el proceso de extracción de candidatos y σ es el umbral de selección.

El valor σ se escoge de forma que la lista final de candidatos contenga solamente los dos rostros con los votos más altos. En la figura veintiséis (b), por ejemplo, la lista final contiene los rostros 3 y 8, lo cual significa que el rostro presente en la imagen es ya sea el 3 o el 8.

Figura 26. Fase de selección de candidatos



(a) Fase de selección de candidatos. (b) Fase de reducción de candidatos.

Figura 27. Algoritmo para la detección de rostros.

Procedimiento 3: Lista de posibles rostros en la imagen.
Entrada: Histograma de hipótesis, H_i .
Salida: Lista de los mejores candidatos.
 {Para cada rostro candidato.
 Si ($|H_i| \geq \sigma$) entonces.
 Rostro presente en la imagen.
 Sino rostro no presente en la imagen. }

4.2.2.2 Algoritmos genéticos

4.1.2.2.1 Introducción

Los Algoritmos Genéticos, fueron introducidos por John Holland en 1970 inspirándose en el proceso observado en la evolución natural de los seres vivos.

Los Biólogos han estudiado en profundidad los mecanismos de la evolución, y aunque queda mucho por comprender, varios aspectos se han logrado explicar.

De manera general se puede decir que en la evolución de los seres vivos el problema al que cada individuo se enfrenta cada día es la supervivencia. Para ello cuenta con las habilidades innatas provistas en su material genético. A nivel de los genes, el problema es el de buscar aquellas adaptaciones beneficiosas en un medio hostil y cambiante. Debido en parte a la selección natural, cada especie gana una cierta cantidad de "conocimiento", el cual es incorporado a la información de sus cromosomas.

Así pues, la evolución tiene lugar en los cromosomas, en donde está codificada la información del ser vivo. La información almacenada en el cromosoma varía de unas generaciones a otras. En el proceso de formación de un nuevo individuo, se combina la información cromosómica de los progenitores aunque la forma exacta en que se realiza es aún desconocida.

Aunque muchos aspectos están todavía por discernir, existen unos principios generales ampliamente aceptados por la comunidad científica.

Algunos de estos son:

- La evolución opera en los cromosomas en lugar de en los individuos a los que representan.
- La selección natural es el proceso por el que los cromosomas con "buenas estructuras" se reproducen más a menudo que los demás.
- En el proceso de reproducción tiene lugar la evolución mediante la combinación de los cromosomas de los progenitores. Llamamos Recombinación a este proceso en el que se forma el cromosoma del descendiente. También son de tener en cuenta las mutaciones que pueden alterar dichos códigos.
- La evolución biológica no tiene memoria en el sentido de que en la formación de los cromosomas únicamente se considera la información del período anterior

Los algoritmos genéticos establecen una analogía entre el conjunto de soluciones de un problema y el conjunto de individuos de una población natural, codificando la información de cada solución en un string (vector binario) a modo de cromosoma. John Holland, dice que se pueden encontrar soluciones aproximadas a problemas de gran complejidad computacional mediante un proceso de "evolución simulada".

A tal efecto se introduce una función de evaluación de los cromosomas, que llamaremos calidad (*fitness*) y que está basada en la función objetivo del problema. Igualmente se introduce un mecanismo de selección de manera que los cromosomas con mejor evaluación sean escogidos para "reproducirse" más a menudo que los que la tienen peor.

Los algoritmos desarrollados por Holland inicialmente eran sencillos pero dieron buenos resultados en problemas considerados difíciles. Los algoritmos Genéticos están basados en integrar e implementar eficientemente dos ideas fundamentales: Las representaciones simples como strings binarios de las soluciones del problema y la realización de transformaciones simples para modificar y mejorar estas representaciones.

Para llevar a la práctica el esquema anterior y concretarlo en un algoritmo, hay que especificar los siguientes elementos:

- Representación del cromosoma.
- Población inicial.
- Medida de evaluación.
- Criterio de selección / eliminación de cromosomas.
- Técnicas de cruzamiento.
- Técnicas de mutación.

4.2.2.2.1 Representación del cromosoma

En los primeros trabajos las soluciones se representaban por strings binarios, es decir, listas de 1s y 0s. Este tipo de representaciones ha sido ampliamente utilizado, incluso en problemas en donde no es muy natural.

En 1985, De Jong introduce las siguientes preguntas:

- ¿Qué se debe hacer cuando los elementos del espacio de búsqueda se representan de modo natural por estructuras complejas como vectores, árboles o grafos?.
- ¿Se debe intentar linealizar en un string o trabajar directamente con estas estructuras?.

En la actualidad se pueden distinguir formas de representar un cromosoma:

- Limitarse a usar strings binarios .
- Utilizar otro tipo de configuraciones más complejas como árboles o grafos.

Hemos de notar que las operaciones genéticas dependen del tipo de representación, por lo que la elección de una condiciona a la otra.

La ventaja de limitarse a usar strings binarios, es que permite definir fácilmente operaciones de recombinación, además los resultados sobre convergencia están probados para el caso de strings binarios.

Sin embargo en algunos problemas puede ser poco natural y eficiente el utilizarlas.

4.2.2.2.3 Población inicial

La población inicial suele ser generada aleatoriamente. Sin embargo, últimamente se están utilizando métodos heurísticos para generar soluciones iniciales de buena calidad. En este caso, es importante garantizar la diversidad estructural de estas soluciones para tener una "representación" de la mayor parte de población posible o al menos evitar la convergencia prematura.

4.2.2.2.4 Medida de evaluación

Respecto a la evaluación de los cromosomas, se suele utilizar la calidad como medida de la bondad según el valor de la función objetivo en el que se puede añadir un factor de penalización para controlar la infactibilidad.

Este factor puede ser estático o ajustarse dinámicamente de la siguiente forma:

$$\text{Calidad} = \text{Valor Objetivo Normalizado} - \text{Penalización} * \text{Medida Infactibilidad}$$

Donde el Valor Objetivo normalizado, será el valor obtenido por la función objetivo, luego el factor de penalización será un porcentaje de la Medida de Infactibilidad (que el valor se ajuste a la necesidad de resolver el problema).

4.2.2.2.5 Técnicas de selección

Para aplicar los operadores genéticos tendremos que seleccionar un subconjunto de la población. Algunas de las técnicas que se disponen son:

- **Selección directa:** toma elementos de acuerdo a un criterio objetivo, como son «los x mejores», «los x peores»... los del tipo «el cuarto individuo a partir del último escogido» son empleados con mucha frecuencia cuando se quieren seleccionar dos individuos distintos, y se selecciona el primero por un método aleatorio o estadístico.
- **Selección aleatoria:** puede ser realizado por *selección equiprobable* o *selección estocástica*.
 - **Selección equiprobable:** todos tienen la misma probabilidad de ser escogidos. Por ejemplo, en nuestro algoritmo la madre en el cruce es escogida con probabilidad equiprobable.
 - **Selección estocástica:** la probabilidad de que un individuo sea escogido depende de una heurística. En este algoritmo tanto el padre como quién va a emigrar y quién va a morir son decididos por un método estocástico. Los distintos procedimientos estocásticos son:
 - **Selección por sorteo:** cada individuo de la población tiene asignado un rango proporcional o inversamente proporcional a su adaptación. Se escoge un número aleatorio dentro del rango global, y el escogido es aquel que tenga dicho número dentro de su rango. La probabilidad de ser escogido es proporcional/ inversamente proporcional al grado de adaptación del individuo.
 - **Selección por escaños:** se divide el rango del número aleatorio en un número predeterminado de escaños. Los escaños se reparten de acuerdo con la ley d'Hont, tomando como «puntuación» para repartir los escaños el grado de

adaptación. Observamos que es más probable escoger un elemento de baja probabilidad por este método que en el de selección por sorteo.

- **Selección por restos estocásticos:** igual que el método de selección de escaños, sólo que los escaños no asignados directamente, es decir, aquellos en que se aplica directamente la ley d'Hont, se asignan de forma aleatoria. La probabilidad de escoger un elemento de muy baja probabilidad es más alta que en el de selección por escaños.
- **Por ruleta:** definimos un rango con las características de la selección por sorteo. El número al azar será un número aleatorio forzosamente menor que el tamaño del rango. El elemento escogido será aquel en cuyo rango esté el número resultante de sumar el número aleatorio con el resultado total que sirvió para escoger el elemento anterior. El comportamiento es similar al de una ruleta, donde se define un avance cada tirada a partir de la posición actual. Tiene la ventaja de que no es posible escoger dos veces consecutivas el mismo elemento, y que puede ser forzado a que sea alta la probabilidad de que no sean elementos próximos en la población -esto último no es una ventaja de por sí; salvo que algunos de los otros operadores genéticos emplee un método de selección directa basado en la posición relativa de los individuos de la población. En la bibliografía más antigua se emplea este término para

definir lo que aquí hemos definido como selección por sorteo.

- **Por torneo:** escoge un subconjunto de individuos de acuerdo con una de las técnicas anteriores ,habitualmente, aleatoria o estocástica y de entre ellos selecciona el más adecuado por otra técnica habitualmente, determinística de tipo “el mejor” o “el peor”. Esta técnica tiene la ventaja de que permite un cierto grado de elitismo -el mejor nunca va a morir, y los mejores tienen más probabilidad de reproducirse y de emigrar que los peores- pero sin producir una convergencia genética prematura, si la población es, al menos, un orden de magnitud superior al del número de elementos involucrados en el torneo. En caso de que la diferencia sea menor no hemos observado mucha diferencia entre emplear el torneo o no. La selección por torneo ha sido la técnica empleada en nuestro algoritmo para decidir tanto el padre, en unión con el criterio «el mejor», como quién va a emigrar, en unión con el criterio “el mejor” y quién va a morir en unión con el criterio “el peor”.

4.2.2.2.6 Técnicas de Cruzamiento

Existen gran cantidad de técnicas de cruce. Las técnicas básicas son:

- **Cruce básico:** se selecciona un punto al azar de la cadena. La parte anterior del punto es copiada del genoma del padre y la posterior del de la madre.
- **Cruce multipunto:** igual que el cruce básico, sólo que estableciendo más de un punto de cruce.
- **Cruce segmentado:** existe una probabilidad de que un cromosoma sea punto de un cruce. Conforme se va formando la nueva cadena del descendiente, para cada gen, se verifica si ahí se va producir un cruce.
- **Cruce uniforme:** para cada gen de la cadena del descendiente existe una probabilidad de que el gen pertenezca al padre, y otra de que pertenezca a la madre.
- **Cruces para permutación:** Existe una familia de cruces específicas para los problemas de permutación, siendo algunos de ellos:
 - **Cruce de mapeamiento parcial:** toma una subsecuencia del genoma del padre y procura preservar el orden absoluto de los fenotipos (características únicas del genoma) es decir, orden y posición en el genoma del resto del genoma lo más parecido posible de la madre.
 - **Cruce de orden:** toma una subsecuencia del genoma del padre y procura preservar el orden relativo de los fenotipos del resto del genoma lo más parecido posible de la madre.
 - **Cruce de ciclo:** Tomamos el primer gen del genoma del padre, poniéndolo en la primera posición del hijo, y el primer gen del genoma de la madre, poniéndolo dentro del genoma del hijo en la

posición que ocupe en el genoma del padre. El fenotipo que está en la posición que ocupa el gen del genoma del padre igual al primer gen del genoma de la madre se va a colocar en la posición que ocupe en el genoma del padre, y así hasta rellenar el genoma del hijo.

Los tres anteriores son conocidos también como PMX, OX, y CX, la X es por cambio (*exchange*) y PM es por *Partial Mapping*, la O por *Order*, y la C por *Cycle*.

Es una buena idea que, tanto la codificación como la técnica de cruce, se hagan de manera que las características buenas se hereden; o, al menos, no sea mucho peor que el peor de los padres. En problemas en los que, por ejemplo, la adaptación es función de los pares de genes colaterales, el resultante del cruce uniforme tiene una adaptación completamente aleatoria.

4.2.2.2.7 Técnicas de Mutación

Se define mutación como una variación de las informaciones contenidas en el código genético, habitualmente, un cambio de un gen a otro producido por algún factor exterior al algoritmo genético. En Biología se definen dos tipos de mutaciones: las generativas, que se heredan y las somáticas, que no se heredan. En los algoritmos genéticos sólo nos serán interesantes las mutaciones generativas. Más, ¿por qué puede interesar que se incorpore este mecanismo aleatorio?

Algunas de las razones que pueden motivar a incorporar mutaciones en un algoritmo genético son:

- **Desbloqueo del algoritmo.** Si el algoritmo se bloqueó en un mínimo parcial, una mutación puede sacarlo al incorporar nuevos fenotipos de otras zonas del espacio.
- **Acabar con poblaciones degeneradas.** Puede ocurrir que, bien por haber un casi-mínimo, bien porque en pasos iniciales apareció un individuo demasiado bueno que acabó con la diversidad genética, la población tenga los mismos fenotipos. A priori se pueden plantear algunas soluciones, como el escalamiento de la función de adaptación; más, si ya se ha llegado a una población degenerada, es preciso que las mutaciones introduzcan nuevos genomas. Como se analiza en el operador de cruce, esto se hace implícitamente en cada cruce.
- **Incrementar el número de saltos evolutivos.** Los saltos evolutivos, la aparición de un fenotipo especialmente valioso, son muy poco probables en un algoritmo genético *puro* para un problema genérico. La mutación permite explorar nuevos sub espacios de soluciones, por lo que, si el sub espacio es bueno en términos de adaptación, se producirá un salto evolutivo después de la mutación que se expandirá de forma exponencial por la población.
- **Enriquecer la diversidad genética.** Es un caso más *suave* que el de una población degenerada -por ejemplo, que la población tenga una diversidad genética pobre-, la mutación es un mecanismo de prevención de las poblaciones degeneradas.

Las técnicas más utilizadas de mutación existentes son:

- **Mutación de bit:** existe una única probabilidad de que se produzca una mutación de algún bit. De producirse, el algoritmo toma aleatoriamente un bit, y lo invierte.
- **Mutación multibit:** cada bit tiene una probabilidad de mutarse o no, que es calculada en cada pasada del operador de mutación multibit.
- **Mutación de gen:** igual que la mutación de bit, solamente que, en vez de cambiar un bit, cambia un gen completo. Puede sumar un valor aleatorio, un valor constante, o introducir un gen aleatorio nuevo.
- **Mutación multigen:** igual que la mutación de multibit, solamente que, en vez de cambiar un conjunto de bits, cambia un conjunto de genes. Puede sumar un valor aleatorio, un valor constante, o introducir un gen aleatorio nuevo. Esta mutación es la que se produce implícitamente en nuestra implementación de cruce.
- **Mutación de intercambio:** existe una probabilidad de que se produzca una mutación. De producirse, toma dos bits/genes aleatoriamente y los intercambia.
- **Mutación de barajado:** existe una probabilidad de que se produzca una mutación. De producirse, toma dos bits o dos genes aleatoriamente y baraja de forma aleatoria los bits o genes, según se hubiera escogido de los dos.

Se debe notar que el papel que juega la mutación es el de introducir un factor de diversificación ya que, en ocasiones, la convergencia del procedimiento a buenas soluciones puede ser prematura y quedarse atrapado en óptimos locales.

4.2.2.2.8 Aplicación de un algoritmo genético en la detección de rostros

Como punto final de este capítulo, se expone una técnica que se basa en descriptores invariantes (capítulo tres) y algoritmos genéticos para realizar la detección de un rostro.

Los pasos a realizar, dada una imagen de entrada que posiblemente contiene un rostro, son:

- Dividir la imagen en pequeñas celdas de un tamaño de 5x5 píxeles.
- Calcular los momentos ordinarios para cada celda (ver capítulo tres).
- De los momentos ordinarios calculados, calcular los invariantes de Hu (capítulo tres).
- Calcular una función de costo agrupando las celdas individuales usando un algoritmo genético. La función de costo corresponde a la descripción invariante del rostro humano, en términos de las características detectadas en la imagen (nariz, ojos y boca).

Ahora, se debe tomar en cuenta las siguientes suposiciones:

- El número de rostros humanos en la imagen es desconocido.
- El tamaño de los rostros es desconocido.
- La localización y la orientación del rostro es desconocida.
- Pequeñas rotaciones y cambios en la expresión facial son permitidas.

La etapa de el cálculo de descriptores invariantes fue expuesta en el capítulo tres. En esta sección se trata cálculo del algoritmo de aprendizaje y de búsqueda, un algoritmo genético.

El vector de característica de interés $VCI = [\Phi_1 \ \Phi_2 \ \Phi_3 \ \Phi_4]$, es el que corresponderá al el rostro que se desea verificar en la imagen, el problema es reducido a buscar un conjunto de celdas cuyos descriptores invariantes de región $[\Phi_1 \ \Phi_2 \ \Phi_3 \ \Phi_4]$ obtenidos de la combinación de los momentos estándar de cada celda, se ajusten mejor al vector de interés VCI.

El proceso de agrupamiento de celdas se realiza utilizando un algoritmo genético, el cual permite recorrer el espacio de representación a usar una función de confianza $C(f_k)$. La figura veintiocho muestra una imagen dividida en celdas fijas.

Figura 28. Objeto de interés dividido en celdas



En el sub tema 4.2.2.2.8.1, se explican las principales características de el algoritmo genético a utilizar en la detección del rostro.

4.2.2.2.8.1 Características del algoritmo genético en la aplicación

- Un cromosoma equivale a una región rectangular definida por dos puntos $p_1(x_1, y_1)$ y $p_2(x_2, y_2)$. Esta región es un múltiplo exacto de las celdas de 5 por 5.
- La población inicial contiene solamente celdas individuales.
- Las nuevas generaciones son obtenidas recombinando nuevas y viejas poblaciones.
- Las probabilidades de cruce y mutación están fijas al 70% y al 5% respectivamente, para la fase de exploración.
- Para la fase de explotación, se considera una probabilidad de cruce del 90% y de mutación del 1%. En la fase de explotación se aplica, adicionalmente, una optimización local.
- El tamaño de la población es fijado a 100.
- Se utilizan múltiples puntos de cruce, en total cuatro puntos fijos.
- La aptitud de un cromosoma se obtiene con un valor de confianza que involucra los invariantes de Hu de objeto de interés (VCI) y los correspondientes invariantes de la región que se evalúa, así como un factor de forma, apt , mostrado en la ecuación 4.27.

$$apt = \frac{|\phi_i - \phi_d|}{|\phi_i + \phi_d|} * f_f$$

Ecuación 4.27

Donde

apt es el valor de aptitud del cromosoma.

Φ_i es el valor de los invariantes del objeto de interés.

Φ_d es el valor de los invariantes de la región de la imagen especificada por un cromosoma.

f_f Factor de la forma definido por la ecuación 4.28.

- El algoritmo genético procesa un número arbitrario de generaciones, definido en el momento de la simulación.
- La función de costo a minimizar se define en la ecuación 4.29.

$$f_f = \max\left(\frac{E_M}{E_m}, \frac{E_m}{E_M}\right)$$

Ecuación 4.28

Donde, E_m es el tamaño del eje principal menor de la región y E_M el tamaño del eje Principal mayor de la región.

$$C(f_k) = \frac{\phi_i - \phi_d}{\phi_i + \phi_d}$$

Ecuación 4.29

Donde Φ_i , Φ_d y f son, respectivamente los invariantes buscados, los invariantes de región y la función de costo a optimizar.

4.2.2.2.8.2 Descripción del algoritmo genético

Luego de la definición del algoritmo genético, se muestran los pasos que conforman el algoritmo. A continuación un listado de estos:

El algoritmo genético inicia generando una población inicial que corresponde a regiones de tamaño igual a una celda.

La siguiente población se obtiene mediante la cruce y/o mutación de los cromosomas de la población inicial.

La siguiente población contendrá cromosomas correspondientes a regiones con más de una celda.

Este proceso se repite hasta que se haya obtenido la mejor minimización de la función de costo.

La solución (la detección del rostro), se da cuando se obtiene la mejor minimización de la función de costo descrita en el sub tema 4.2.2.2.8.3.

Debido al alcance de esta investigación, se exponen aquí solamente las bases necesarias para que el lector se adentre en el tema de los algoritmos

genéticos y su uso en el aprendizaje y búsqueda de soluciones por medio de una computadora.

5. APLICACIÓN DE UN SISTEMA DE RECONOCIMIENTO DE ROSTROS

5.1 Sistema de seguridad para el aeropuerto internacional La Aurora implementando un sistema de reconocimiento de rostros

La aplicación de un sistema de reconocimiento de rostros en un aeropuerto, tiene el objetivo de ubicar personas, las cuales legalmente tienen prohibido abandonar el país, y actualmente lo hacen portando identificaciones falsas, o personas que son clasificadas como indeseables para el país, como terroristas, mafiosos o personas peligrosas buscadas en otros países.

La necesidad de la seguridad en un aeropuerto es obvia. No es suficiente el chequeo de los pasajeros antes de abordar un avión. Los aeropuertos también funcionan como puntos de inmigración, por lo que la seguridad es básica.

Un sistema de reconocimiento de rostros puede contribuir mucho a la seguridad de un aeropuerto, se pueden mencionar algunas funciones que se pueden implementar:

- Prevención para que personas no autorizadas ingresen a áreas de servicio del aeropuerto.
- Identificación de sospechosos, tan pronto ingresan al aeropuerto.
- Prevenir la inmigración ilegal.

- Proveer un registro de las personas que ingresan y que dejan el país.

5.1.1 Componentes del Sistema

Para el sistema de detección de rostros, se propone utilizar el producto *C-VIS Face Recognition* (descrito en la sección 5.2.1), que implementa algoritmos de redes neuronales. Y son requeridos los siguientes elementos:

- Cámaras IP con sensor de movimiento.
- *C-VIS FACE Snap Recorder*.
- *C-VIS FACE Check*.
- *C-VIS FACE Track*.

5.1.2 Captura de imágenes

Las cámaras IP, se utilizaran para obtener las imágenes de personas que circulan en todas las áreas del aeropuerto. Estas cámaras se distribuirán en las entradas principales, en las puertas de abordaje, puestos de Migración y en áreas en las que solo personal autorizado debe ingresar.

Las cámaras IP, se conectan a través de una red Ethernet, por lo que es muy conveniente para esta aplicación, ya que no se requiere de ningún cableado adicional.

La cámara es de tipo VGA, y la resolución que tiene es de 640 x 480, tiene integrado un sensor de movimiento, el cual registra imágenes, solo cuando detecta movimiento, a través de un detector de imagen CMOS de 35k

píxeles. Adicionalmente posee la característica de visión nocturna, lo cual es útil en situaciones de poca luz.

Figura 29. Cámara IP con sensor de Movimiento.



5.1.3 Extracción y Almacenamiento de Imágenes Faciales

La extracción de la imagen facial, se realiza procesando la imagen de la cámara, para su identificación y verificación. La aplicación de extracción de rostros se encarga de grabar en la base de datos un rostro varias veces en diferentes ángulos. Si hay varios rostros en el ángulo de visión de la cámara, todos se grabaran. Para realizar la extracción, la aplicación preprocesa la imagen y detecta los rostros en ella, para luego almacenarlos en la base de datos. El software que realiza el proceso de extracción y almacenamiento de imágenes faciales *FACE Snap Recorder*.

5.1.4 Normalización e Identificación de rostros de Personas

La identificación de personas, se da cuando la cámara detecta movimiento, luego *FACE Snap Recorder*, graba el rostro en la base de datos.

Un motor de búsqueda de rostros, se encarga de comparar los rostros que están en la base de datos, con los rostros que se están obteniendo de la cámara, estos rostros que se están obteniendo deben ser normalizados, debido a que el tamaño del rostro y el ángulo es impredecible. La normalización de un rostro la realiza *FACE track* y del reconocimiento de rostros se encarga *FACE Check*.

Una vez el rostro es identificado de un conjunto de sospechosos en la base de datos, se activa una alarma en el sistema que muestra la foto del sospechoso, mostrando toda la información asociada a ese rostro, así como también la localización de esta persona en el aeropuerto.

5.1.5 Sala de control

La sala de control será donde está centralizado el sistema. Cada cámara estará asociada a un monitor, en el cual se podrá observar todo el movimiento de personas en un sector. Una computadora será la encargada de procesar las imágenes y realizar las labores de detección. C-VIS, ofrece una computadora especial dedicada a sus productos de software.

Desde la sala de control, cuando exista un evento, los encargados de seguridad verán la alerta en pantalla, y tomarán las acciones necesarias.

En el caso del departamento de Migración, existirá una cámara que toma a detalle cada persona que pasa por el detector de metales, asegurándose una toma más detallada de la persona, y una búsqueda más específica en la base de datos de rostros de personas arraigadas.

5.1.6 Consideraciones Adicionales

Se deben tomar algunas consideraciones para implementar el sistema de reconocimiento de rostros para el aeropuerto.

Primero se deben hacer los contactos necesarios con organismos internacionales para tener una conexión con la base de datos de personas peligrosas o buscadas.

Se debe entrenar al sistema con todas las imágenes que no provienen de las cámaras. El entrenamiento se da ingresando un conjunto de imágenes (que deberán ser provistas en la base de datos de personas sospechosas) en diferentes ángulos.

Cuando una persona se registre en el gabinete de identificación de la policía nacional, adicionalmente de las huellas digitales, deberá tomarse un conjunto de fotos, que alimentarán a la base de datos de personas arraigadas. Siempre que sea posible, estas imágenes deberán ser actualizadas, por ejemplo luego que el juez dictamine un arraigo, la persona deberá actualizar la imagen de su rostro.

Por ultimo, las cámaras se deberán colocar en puntos estratégicos, evitando los puntos que no son de acceso, ya que en ellos existiría movimiento y se estaría registrando constantemente, y esto consumiría capacidad de procesamiento en el analizador de rostros.

5.2 Software para el reconocimiento de rostros

5.2.1 FaceIT ARGUS

Casa Fabricante: Identix.

Pagina Web: www.identix.com/products/pro_security_bnp_argus.html

Tipo: Sistema de Vigilancia.

Tecnología base: Redes Neuronales.

Descripción: El sistema de identificación Facelt, está basado en un circuito de cerrado televisión que utiliza cámaras de video para la obtención de imágenes para reconocer rostros.

Una de sus principales características es que puede reconocer ente varios cientos de rostros, cosa que un operador normal de vigilancia no podría hacer.

El sistema genera señales audibles y visibles cuando se detecta a algún sospechoso. Se puede ajustar a un nivel de confianza para maximizar la activación correcta de alarmas dependiendo los requerimientos de seguridad.

Esta construido sobre la plataforma de la red biométrica de Identix e incorpora todas las capacidades de un motor de reconocimiento facial.

Facelt ha sido probado efectivamente para detectar individuos en una lista en entornos como aeropuertos, centros comerciales y fronteras en varias partes del mundo.

Tiene la capacidad de ser usado con una sola cámara o con un complejo sistema de circuito cerrado por televisión (CCTV).

Figura 30. El sistema Facelt ha detectado un individuo que está en la lista de sospechosos

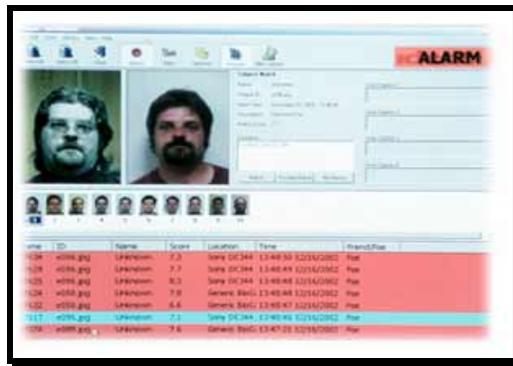
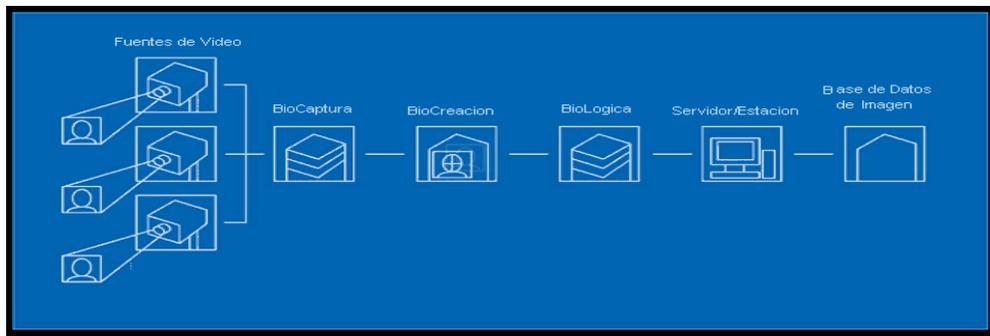


Figura 31. Configuración de una aplicación de reconocimiento de rostros



En la figura treinta y uno se muestra una configuración de ejemplo de una aplicación de reconocimiento de rostro utilizando el sistema de identix Facelt.

El sistema se divide en las siguientes secciones:

- Fuente de Video.
- Bio Captura.
- Bio Creación.
- Bio Lógica.
- Cliente/Servidor.
- Base de datos de imagen.

La fuente de video son el conjunto de cámaras de vídeo, las cuales alimentaran al sistema Facelt para el análisis. La Bio Captura se refiere a la detección automática de un rostro en la señal de video.

En la etapa de Bio creación, se convierte cada rostro en datos digitales o patrones, que se enviarán para ser comparados en contra de una base de datos maestra o lista de observación. En la etapa de Bio Lógica se administra la alimentación de datos de servidor y estación de trabajo.

El sistema cliente servidor realiza la comparación y provee una interfaz de usuario para las respuestas y alarmas y la base de datos de imagen, contiene la lista de observación que contiene todos los rostros de los individuos a los que interesa reconocer. Además contiene varias fuentes de imágenes, incluidas fotografías digitales, fotografías escaneadas, dibujos robot y renderizaciones 3d.

5.2.2 C-VIS Face Recognition

Casa Fabricante: C-VIS

Página Web: <http://www.c-vis.com/htd/home.html>

Tipo: Identificación Biométrica de Rostros.

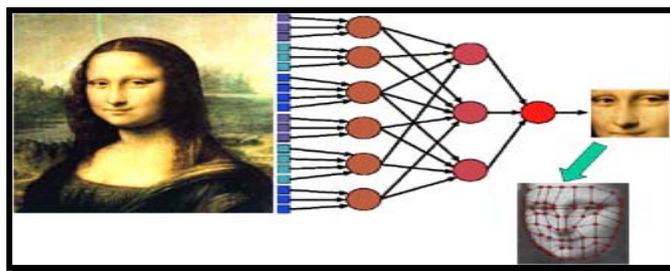
Tecnología base: Redes Neuronales, Computación Neuronal.

Descripción: La tecnología de c-vis, se basa en computación neuronal, que combina las ventajas de elasticidad y redes neuronales.

La computación neuronal provee de métodos de procesamiento de información que son similares a la forma en que se realizan en los sistemas biológicos, como el cerebro humano. Esta tecnología comparte muchas fortalezas clave como por ejemplo robusta tolerancia a fallos y la habilidad de aprender de ejemplos. Las redes elásticas pueden comparar superficies faciales, inclusive si las imágenes no son idénticas, como prácticamente en todas las situaciones en el mundo real. Las redes neuronales son utilizadas para aprender a reconocer similitudes a través de un patrón de reconocimiento.

El reconocimiento de rostros en este sistema es difícil de engañar debido a que este funciona comparando superficies faciales (ver figura treinta y dos), porciones específicas y ángulos definidos en las características faciales, que no pueden fácilmente ser disfrazadas con barbas, lentes o maquillaje.

Figura 32. Clasificación de imágenes de C-Vis.



C-vis, puede ser utilizado en áreas abiertas de bastante tráfico de personas, como por ejemplo:

- Aeropuertos.
- Casinos.
- Estadios.
- Transporte Público.
- Oficinas de gobierno.

C-vis es utilizado como vigilante de acceso biométrico, en comparación de otros sistemas biométricos como tarjetas de identificación, PIN, llaves, etc., que requieren que las personas sean pre identificadas, o los análisis de iris y geometría de mano que son tecnologías que requieren tener algún tipo de intrusión en el cuerpo, los sistemas de reconocimiento de rostros trabajan a distancia, sin ningún contacto físico. Se chequean los rostros en contra de una base de datos de personas conocidas y agrega nuevos rostros cada vez que una persona nueva entra.

Un sistema de monitoreo por video provee un registro de que paso en el área, el reconocimiento de rostros provee un registro de quien estuvo en el área. O sea que el reconocimiento de rostros hace al monitoreo por video, inteligente. Esto ayuda a que el personal que está encargado de el monitoreo por video no tenga que estar vigilando ininterrumpidamente la pantalla.

Ya que se guardan registros en la base de datos, las personas conocidas pueden ser detectadas automáticamente, y las personas no conocidas chequeadas rápidamente, sin que el personal de seguridad tenga

que observar todos los monitores de seguridad todo el tiempo. Los tiempos de reacción son mucho más cortos y el análisis post evento es más rápido.

Las partes con las que cuenta el sistema c-vis son:

- *FACE Snap Recorder.*
- *FACE Check.*
- *FACE Track.*

Face Snap Recorder es una solución de software que se agrega a un sistema de vigilancia y monitoreo. Este extrae las imágenes faciales de un video para identificación y verificación. Los Rostros individuales y sus superficies faciales, son grabados y almacenados en una base de datos que podrá ser vista en una PC. *Face Snap Recorder* reconoce y graba un rostro varias veces en diferentes ángulos, inclusive si varios rostros están en su ángulo de visión. La figura treinta y tres muestra un ejemplo de *Face Snap Recorder*.

Figura 33. Un ejemplo del funcionamiento de *Face Snap Recorder*.



Face Check, es un módulo que sirve para identificar automáticamente a una persona, este modulo es el motor de búsqueda de la base de datos de Face Snap, aunque puede se utilizado por si mismo para otras aplicaciones de reconocimiento, como por ejemplo el chequeo de entrada y salida de personas. La Volkswagen utilizó por primera vez este modulo, para el reconocimiento de un conductor en un automóvil concepto.

Face Check puede ser utilizado en red para identificación en tiempo real. Esto significa que pueden existir puntos de chequeo remotos, como por ejemplo en un lugar donde existan múltiples accesos como en un edificio de oficinas o un estadio.

Face Track, es el módulo que separa los rostros de una imagen de video. Las personas en una imagen de video se pueden ver de diferentes tamaños y sus movimientos son impredecibles. Esto significa que para reconocer un rostro se debe cubrir un área grande para estar seguro que se va a tomar cada rostro, y eso significa también que habrá cantidades inmensas de información de video irrelevante, como torsos y fondo.

Face Track se encarga de obtener imágenes de alta calidad de imágenes complejas.

Para llevar a cabo la tarea de obtención de rostros, *face track* utiliza dos cámaras, en un ángulo amplio. Esta cámara detecta una silueta de un humano que entra en su campo de visión. El software automáticamente localiza el rostro y realiza un zoom sobre él.

CONCLUSIONES

1. La biometría es la ciencia que trata de identificar o autenticar a una persona por medio de una característica física, como puede ser una huella digital, el iris y la retina, la mano y, por supuesto, el rostro.
2. De todos los métodos biométricos, el reconocimiento de rostros es el más difícil de engañar, pero también es un método lento, ya que, la búsqueda requiere de un proceso complejo. Los mejores sistemas biométricos son los que combinan el reconocimiento de huellas, con el reconocimiento de rostros, debido a que el reconocimiento de huellas es menos confiable, pero más rápido, por lo que al hacer una combinación de ellos se obtiene una identificación rápida y segura.
3. Para reconocer un rostro de una imagen digital, es necesario preprocesarla a un formato de imagen que sea más fácil de manejar. El formato PGM conserva las características faciales de la imagen y debido a que cada píxel en la imagen está mapeado a una escala de gris, lo que lo hace conveniente para el cálculo de momentos invariantes.
4. Un rostro se puede describir por medio de un contorno o una región. La descripción por región es mejor debido a que el contorno del rostro en una imagen presenta mucho ruido. Mediante la descripción por región se puede obtener una mejor separación del rostro de la imagen.

5. Para crear una plantilla de un rostro se utilizan los descriptores invariantes de Hu y de Yan en combinación. Se utilizan los dos descriptores debido a que los descriptores de Hu no varían con cambios de tamaño y traslación de la imagen, y los descritos por Yan, no varían con los cambios de contraste y luz en la imagen. Con estos descriptores combinados se construye el vector de característica del rostro.
6. El método de eigen caras describe el rostro de una imagen por medio de eigen vectores. Los eigen vectores son usados en cálculos relacionados con fenómenos de la naturaleza.
7. Los clasificadores, son métodos por medio de los cuales se clasifica un conjunto de patrones.
8. Una red neuronal artificial, intenta imitar algunas funcionalidades de una neurona del cerebro humano. Una red de neuronas artificiales puede ser monocapa, en la que cada neurona de entrada es de salida también y las multicapa, las cuales son más complejas y poseen capas de entrada, ocultas y de salida.
9. En un sistema de reconocimiento de rostros, una red neuronal, clasifica y busca un rostro en una base de datos o un mapa de memoria.
10. Los algoritmos genéticos, están inspirados en el proceso de evolución natural. En algoritmos estos se manejan conceptos de genética como, cromosomas, cruce y mutación. Son utilizados para la solución de

problemas complejos. Una de las claves es la supervivencia del que mejor se adapte a un medio hostil y cambiante. Cada vez que hay evolución un cromosoma agrega información importante a el.

11. Para aplicar un algoritmo genético en el reconocimiento de rostros, es necesario dividir la imagen de entrada en un conjunto de celdas, las cuales serán representadas por un cromosoma.

RECOMENDACIONES

1. Para el reconocimiento de rostros humanos se debe elegir un tamaño estándar de la imagen de entrada para simplificar el procesamiento y detección del rostro en la imagen.
2. Se deberá simplificar lo más posible las imágenes que se utilizaran en el reconocimiento, para ello se debe convertir la imagen a formato PGM, con alguna herramienta existente, la cual realice esta conversión.
3. Cuando se construye una aplicación de reconocimiento de rostros basada tanto en redes neuronales es indispensable realizar un entrenamiento de estas, de una manera exhaustiva, para que los resultados a obtener sean exactos.
4. Si se diseña el reconocimiento de rostros, para una aplicación de vigilancia en lugares públicos, debido a la complejidad del fondo, es recomendable utilizar el método de eigen caras para obtener el rostro en la imagen.
5. Para un sistema biométrico de acceso completo, rápido y seguro, se recomienda combinar la fiabilidad del reconocimiento de rostros, con la rapidez del reconocimiento de huellas dactilares.

6. Se sugiere utilizar redes neuronales electrónicas ART-2, cuando se realicen aplicaciones de reconocimiento de rostros dedicadas y que no se necesite que la computadora realice el trabajo de la búsqueda y reconocimiento, y se dedique solamente al manejo de la base de datos y las estadísticas.
7. Cuando se requiera implementar una solución de reconocimiento de rostros sin utilizar redes neuronales, se recomienda como alternativa el uso de algoritmos genéticos. Debido a que el campo de los algoritmos genéticos no ha sido explotado del todo en el reconocimiento facial, el investigador podrá desarrollar nuevos y mejores métodos para el reconocimiento de rostros.
8. Es aconsejable incluir en los cursos de la carrera de ingeniería en ciencias y sistemas, en el área de ciencias de la computación, algunas técnicas concretas de biometría, de reconocimiento de patrones, algoritmos de aprendizaje y búsqueda.

BIBLIOGRAFÍA

1. Yang G. et al. Human face detection in a complex background, Pattern Recognition Pp 53-63.
2. Rosenblum M., Human Emotion from Motion using a radialbasis function network architecture, Workshop on Motion of Non-rigid and articulated objetcs, Pp 43-49, 1994.
3. Govindaraju V. et al., A computational model for face localization, Proceedings of the IEEE International Conference on Computer Vision, Pp 718-721, 1990.
4. Govindaraju V. et al., Locating human faces in newspaper photographs, Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 1989.
5. Young Ho Kwon, Age classification from facial images, Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition.
6. Poggio R. and T., Face recognition: features versus templates, IEEE Trans. on Pattern Analysis and Machine Intelligence, 1993.

7. R. Pinto Elias, J Sossa, Detección de rostros utilizando descriptores Invariantes y algoritmos genéticos, Universidad del valle México.
8. Pao Y. H., Adaptive pattern recognition and neural networks, Addison Wesley, Mass., 1989.
9. Hu M. K. Visual pattern recognition by moment invariants, IRE Transactions on Information Theory, Pp 79-187, 1962.
10. Gonzalez R.C./Woods, Digital Image Processing, Addison Wesley Publishing company, 1994 .
11. Huang Ren, Detection strategies for face recognition, using learning and evolution. George Mason University, Virginia, 1998.
12. Yan W., D. Mingyue and P. Jiaxiong, Research on using moment invariants in scene matching, Proceedings of the GMICV 95, Pp. 338-341, 1995.
13. Brunelli, Roberto . Template matching. Pattern Recognition. Pp 751-768. 1997.
14. Yacoub, Meziane. Discriminative feature extraction selection applied to face recognition. 1999.

15. Z. Mingfa, S. Hasani, SI Bhattarai, and H Singh. Pattern Recognition with moment invariants on a machine vision system. Pattern Recognition Letters. 1989. Pp 175-180.
16. M. A. Turk and A. P. Pentland. Face recognition using eigenfaces. IEEE Conf. on Computer Vision and Pattern Recognition, Pp 586-591, 1991.
17. H. Murase and S. K. Nayar. Visual learning and recognition of 3-d objects from appearance. International Journal of Computer Vision, Pp 5-24, 1995.
18. Aguilar Jauregui, Maria. Detección y localización de rostros en imágenes bidimensionales. Tesis. Instituto Politécnico Nacional. , Mexico. P.p. 1-16 2000.
19. Tala Ayerdi, Virginia. Redes Neurales aplicadas al pronostico del índice de precios al consumidor. Tesis. Universidad de San Carlos de Guatemala. Guatemala. 1997.
20. G. A. Carpenter and S. Grossberg, The ART of adaptive pattern recognition by a self-organizing neural network, IEEE Computer, p.p.77-88, 1988.
21. Sossa Azuela, Jose, Patricia Rayon. Arquitectura basada en Redes neuronales para el reconocimiento de rostros. Universidad del Valle de México.

22. D. E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Addison Wesley Publishing Company, inc. 1989.

Referencias electrónicas

23. Face Recognition Home Page, <http://www.cs.rug.nl/~peterkr/FACE/face.html>.
24. Sistemas Biométricos. <http://www.udabol.edu.bo/biblioteca/sistemas/sistemas/10redneuro/7r1sisbio19e/sisbiometricos.htm>. Universidad de Aquino Bolivia.
25. Reconocimiento de Rostros utilizando Eigen Faces. www.cim.mcgill.ca/~wsun/sa/project/node9.html. Universidad Mc Gill. California.
26. Redes Neuronales. Biblioteca digital.
http://www.udabol.edu.bo/biblioteca/sistemas/sistemas/10redneuro/2r10rna30n/redes_neuronales/redes_neuronales.html. Universidad Aquino de Bolivia.
27. Marti, Rafael. Algoritmos Genéticos. <http://www.uv.es/~rmarti/genet.html>. Universidad de Valencia, España. Ocerro, David. Algoritmos Genéticos <http://www.orcero.org/irbis/disertación/node189.html>.