

**UNIVERSIDAD MAYOR DE SAN ANDRÉS
FACULTAD DE CIENCIAS PURAS Y NATURALES
CARRERA DE INFORMÁTICA**



TESIS DE GRADO

**“PATRONES DE RECONOCIMIENTO FACIAL MEDIANTE REDES
NEURONALES”**

**PARA OPTAR AL TITULO DE LICENCIATURA EN INFORMÁTICA
MENCIÓN: INGENIERÍA DE SISTEMAS INFORMÁTICOS**

POSTULANTE : Leoncio Gutierrez Mamani
TUTOR : Lic. Eufren Llanque Quispe
REVISOR : Lic. Carlos Mullisaca Choque M,Sc.

**LA PAZ – BOLIVIA
2011**

DEDICATORIA

A la mujer que más admiro en el mundo, mi madre, Manuela por brindarme todo su amor, por su fortaleza y por enseñarme a luchar contra la adversidad.

A mi Padre Abdón Gutiérrez por su amor y sacrificio en mi formación, ejemplo de esfuerzo, trabajo y valores morales que me dieron fuerzas para seguir adelante.

AGRADECIMIENTOS

Gracias a Dios por darme sabiduría y fortaleza para seguir siempre adelante sin desmayar y llevar a feliz término el presente trabajo de investigación.

Gracias al Lic. Eufren Llanque Quispe y al Lic. Carlos Mullisaca por su ayuda y confianza.

Gracias a mis amigos y amigas por todos los momentos compartidos y por hacerme crecer como persona.

Gracias a todos los docentes de la Carrera de Informática que fueron partícipes de mi formación profesional.

Finalmente, agradecer a mis padres por haberme dado la formación, la motivación y la ayuda necesaria para finalizar el trabajo de investigación.

Resumen

En la Tesis se presentan nuevas soluciones al problema del Reconocimiento Facial mediante redes neuronales enmarcadas dentro de la Visión bidimensional. En dos dimensiones se han propuesto dos técnicas de reducción de información, basadas en la utilización de análisis de componentes principales (PCA), el empleo de imágenes de baja resolución, el método de procesamiento que utiliza una plantilla por sujeto y siendo cada plantilla una imagen del sujeto que se quiere verificar.

El problema del reconocimiento de caras se puede dividir en fases, la de pre procesamiento, detección de la cara dentro de la imagen y la de reconocimiento.

La fase de detección se basa principalmente, en la detección de píxeles de piel en la imagen. Posteriormente se realiza una selección de regiones de piel candidatas a ser caras y su validación mediante “mapas” de ojos y boca.

En la fase de reconocimiento se tienen las zonas de las imágenes detectadas como caras. Se utiliza PCA para extraer las características que representan a las imágenes. A continuación, las características sirven para entrenar y simular las redes neuronales. Con las salidas de las redes neuronales se seleccionan las imágenes de la base de datos que se parecen a la cara de la imagen de prueba.

En el trabajo implementado se ve la gran influencia del tipo de imágenes utilizadas para el reconocimiento, siendo mucho mejores los resultados cuando las imágenes cumplen unas ciertas características, como ser las imágenes extraídas en un ambiente controlado.

Abstract

The thesis presents new solutions to the problem of face recognition using neural networks framed within the two-dimensional vision. In two dimensions have been proposed two data reduction techniques based on the use of principal component analysis (PCA), the use of low-resolution images, the processing method that uses a template subject and each image being the subject to check.

The problem of face recognition can be divided into phases, the pre-processing, face detection within the image and recognition.

The detection phase is based mainly on detection of skin pixels in the image. Subsequently, a selection of candidate skin regions to be expensive and its validation by "maps" of the eyes and mouth.

In the recognition phase areas have sensed images as faces. PCA is used to extract features representing images. The following features are used to train and simulating neural networks. With the outputs of neural networks are selected images from the database that resemble the face of the test image.

Implemented at work is the great influence of the type of images used for recognition, with much better results when the images meet a certain characteristics, such as images taken in a controlled environment.

ÍNDICE

	Página
CAPÍTULO I	
MARCO REFERENCIAL	
1.1. Introducción.....	1
1.2. Antecedentes	2
1.3. Planteamiento del problema.....	3
1.4. Objetivos	4
1.4.1. Objetivo General	4
1.4.2. Objetivos Específicos	4
1.5. Hipótesis	4
1.6. Justificación.....	4
1.6.1. Justificación Científica.....	4
1.6.2. Justificación Económica	5
1.6.3. Justificación Social	5
1.6.4. Justificación Técnica	5
1.7. Metodología y herramientas.....	5
1.7.1. Método Científico	5
1.7.2. Herramientas.....	6

1.8. Alcances, límites y aportes.....	7
1.8.1. Alcances.....	7
1.8.2. Límites.....	7
1.8.3. Aportes.....	7

CAPÍTULO II

MARCO TEÓRICO

2.1. Introducción.....	9
2.2. Visión humana y la visión artificial.....	9
2.2.1. Visión ocular.....	9
2.2.2. Visión artificial	10
2.3. Procesamiento digital de imágenes	10
2.3.1. Píxel	11
2.3.2. Imagen digital.....	12
2.3.3. Clasificación de imágenes digitales.....	12
2.3.4. Modelos de colores de imágenes digitales.....	14
2.3.5. Formatos de imágenes.....	14
2.4. Reconocimiento facial a partir imágenes de intensidad 2D.....	16
2.4.1. Técnicas basadas en rasgos locales.....	16
2.4.1.1. Extracción de rasgos locales.....	17

2.4.2. Técnicas holísticas	19
2.4.2.1. Técnicas basadas en análisis de componentes principales	19
2.4.2.2. Técnicas conexionistas	22
2.4.3. Detección de caras en la imagen 2D.....	23
2.5. Redes neuronales artificiales	24
2.5.1. Funcionamiento de una red neuronal.....	25
2.5.2. Ventajas	26
2.5.3. Tipologías de las redes neuronales.....	26
2.5.3.1. Topología de redes neuronales.....	26
2.5.3.2. Aprendizaje	27
2.5.3.3. Tipo de entrada	27
2.5.4. Perceptrón Multicapa	27
2.6. Visual Basic.....	28
2.6.1. Compilador	29
2.6.2. Entorno de desarrollo	29
2.6.3. Versiones	30
2.6.4. Ventajas	30
2.6.5. Inconvenientes	31
2.6.6. Entornos de desarrollo	32

2.6.7. Alternativas multiplataforma o externas a Windows	33
2.7. MySQL	33
2.7.1. Historia del proyecto.....	33
2.7.2. Lenguajes de programación.....	34
2.7.3. Aplicaciones	35
2.7.4. Especificaciones.....	35

CAPÍTULO III

PATRONES DE RECONOCIMIENTO FACIAL MEDIANTE REDES NEURONALES

3.1. Introducción.....	36
3.2. Descripción informal.....	36
3.3. Descripción formal	38
3.3.1. Imágenes	39
3.3.2. Preprocesado de las imágenes.....	40
3.3.2.1. Contracción del histograma.....	41
3.3.2.2. Igualación del histograma.....	42
3.3.2.3. Operador logarítmico.....	43
3.3.3. Detección de rostro	45
3.3.3.1. Detección de píxeles de piel.....	45

3.3.4. Reconocimiento.....	46
3.3.4.1. Algoritmo para el reconocimiento de rostros	49
3.3.4.2. Metodología en el reconocimiento de rostros.....	51
3.3.4.3. Definición de reglas para el reconocimiento.....	59
3.3.4.4. Construcción del software	60
3.4. Evaluación De Resultados	64
3.4.1. Fase de detección.....	64
3.4.1.1. Prueba de detección con selección de candidatos.....	64
3.4.2. Fase de reconocimiento	66
3.4.2.1. Pruebas de identificación de rostros	66

CAPÍTULO IV

CONCLUSIONES Y RECOMENDACIONES

4.1. Conclusiones.....	68
4.2. Recomendaciones	69
BIBLIOGRAFÍA	70

ÍNDICE DE FIGURAS

FIGURA 1:	REPRESENTACIÓN DE UN PÍXEL	11
FIGURA 2:	IMAGEN ORIGINAL Y SU ESTRUCTURA MATRICIAL DE LA IMAGEN.....	12
FIGURA 3:	IMAGEN BINARIA	12
FIGURA 4:	IMAGEN EN ESCALAS DE GRISES	13
FIGURA 5:	IMAGEN EN COLOR.....	13
FIGURA 6:	EJEMPLO DE RED NEURONAL	24
FIGURA 7:	EJEMPLO DE FUNCIONES DE TRANSFERENCIA.....	25
FIGURA 8:	FOTOGRAFÍA ANTES DE LA CONTRACCIÓN DEL HISTOGRAMA.	41
FIGURA 9:	FOTOGRAFÍA DESPUÉS DE CONTRAERSE EL HISTOGRAMA	42
FIGURA 10:	IMAGEN ANTES DE LA ECUALIZACIÓN.....	43
FIGURA 11:	IMAGEN ECUALIZADA	43
FIGURA 12:	IMAGEN ANTES DE APLICARLE EL OPERADOR LOGARÍTMICO.....	44
FIGURA 13:	IMAGEN DESPUÉS DE APLICARLE EL OPERADOR LOGARITMO	45
FIGURA 14:	ALGORITMO PARA EL RECONOCIMIENTO DE ROSTROS HUMANO	50
FIGURA 15:	CONJUNTO DE ENTRENAMIENTO (CE).....	52
FIGURA 16:	ROSTRO PROMEDIO SEGÚN EL CE.	53
FIGURA 17:	Φ I SON RESTA DEL ROSTRO PROM. A CADA IMAGEN DEL CE	53
FIGURA 18:	AUTOCARAS	56
FIGURA 19:	IMAGEN I_a NO PERTENECE AL CE Y SERÁ PROYECTADA AL (ER)	56

FIGURA 20:	PROPORCIONES PARA LA RECONSTRUCCIÓN DE <i>TPROY</i>	57
FIGURA 21:	PROYECCIÓN AL ESPACIO DE ROSTROS.	58
FIGURA 22:	INTERFAZ DE BIENVENIDA	60
FIGURA 23:	INTERFAZ DE ACCESO DE USUARIO AL SISTEMA DE RECONOCIMIENTO	61
FIGURA 24:	INTERFAZ SISTEMA DE RECONOCIMIENTO	61
FIGURA 25:	INTERFAZ DETECCIÓN DE NUEVO USUARIO	62
FIGURA 26:	INTERFAZ LLENA DE DATOS PERSONALES DE NUEVO USUARIO.....	62
FIGURA 27:	INTERFAZ SISTEMA DE RECONOCIMIENTO	63
FIGURA 28:	INTERFAZ APRENDIZAJE DEL SISTEMA DE RECONOCIMIENTO FACIAL	63
FIGURA 29:	IMÁGENES BINARIA DETECCIÓN DE BOCA Y OJOS.....	65
FIGURA 30:	IMAGEN BINARIA DE PÍXELES DE PIEL.....	65
FIGURA 31:	IMÁGENES BINARIAS TRAS LA APERTURA DE RADIO	65

MARCO REFERENCIAL

CAPÍTULO I

1.1. Introducción

El reconocimiento automático de rostros humanos es uno de los problemas que ha supuesto un desafío en informática y que sigue siendo en la actualidad.

Existen muchos problemas en ingeniería que son difíciles de abordar mediante un enfoque algorítmico tradicional. Por ello nació la necesidad de emular algoritmos inteligentes, inspirados en sistemas biológicos. El reconocimiento de patrones en imágenes digitales ha concitado gran interés en las últimas décadas, debido a que ha permitido abordar problemas de inspección visual, de reconocimiento de números y letras manuscritas, como también, el reconocimiento de huellas dactilares, de iris, de rostros, etc. El reconocimiento y clasificación de imágenes heterogéneas ha sido abordado utilizando sistemas dinámicos inspirados en modelos biológicos. En particular, el reconocimiento de rostros humanos por medio de herramientas computacionales, permite varias aplicaciones en donde hoy en día se requiere identificar particularmente a cada individuo. Por ejemplo sistemas de seguridad para laboratorios, bancos, tiendas comerciales y todos aquellos lugares que necesitan identificar a sus propios clientes y funcionarios para dar acceso restringido.

Una de las ventajas de un sistema automático de reconocimiento de rostros humanos basado en visión artificial es no intrusivo: El sujeto no siente invadida su intimidad, no tiene que realizar ninguna acción para identificarse o someterse a análisis.

EL reconocimiento automático de rostros es uno de los problemas que mayores dificultades ofrece en el área de la informática. La variabilidad de condiciones que aparecen a la hora de implementar un sistema, implica que se realicen diversas suposiciones y simplificaciones del problema en sí. Los desafíos que propone la implementación de un sistema automático de reconocimiento de rostros son recogidos por neurofisiólogos, psicólogos cognitivos, psicólogos sociales y científicos de la informática. Una de las metas importantes que persiguen los investigadores es comprender el complejo sistema visual humano, cómo logra almacenar la información de los rostros

humanos y utilizarla para representar y discriminar entre identidades con un alto grado de exactitud.

1.2. Antecedentes

En la vida cotidiana una persona rara vez se maravilla de la capacidad para reconocer un rostro. En general lo hace de golpe, de manera íntegra, sin la necesidad de un análisis consciente de las diversas características faciales. Sin embargo, es sorprendente que se pueda memorizar un número tan considerable de rostros en el transcurso de nuestra vida y que esa capacidad de reconocer rostros se pueda perder por una lesión o accidente cerebral.

En los últimos años, se han publicado numerosos trabajos de investigación sobre sistemas de reconocimiento de rostros humanos siendo un tema de prolífera actividad científica.

Se investiga el desempeño de sistemas de reconocimiento en una población heterogénea basado en la transformada de Karhunen-Loeve, construyendo vectores y valores propios de la imagen del rostro para luego su posterior clasificación. Se concluye que los vectores propios construidos usando la base de datos entera fueron mejores que los creados utilizando grupos de poblaciones separados. Además, se encontró que no hay diferencia significativa entre el desempeño de grupos individuales, que cuando se usa base común [PNB. 1998].

Se ha mostrado los efectos de la frecuencia espacial en el reconocimiento de imágenes de rostros humanos, umbral de razón señal/ruido para el reconocimiento de rostros usando ruido aditivo espacial y sensibilidad relativa al reconocimiento para diferentes frecuencias. Se concluye que en el reconocimiento de imágenes faciales previamente aprendidas, la mayoría de la información es reunida en una banda de frecuencia espacial [NAR. 1999].

Se investiga un método adaptativo, análogo a los algoritmos evolutivos o genéticos, utilizando análisis de componentes principales para disminuir la dimensión del vector y con posibles aplicaciones en codificación, comprensión y clasificación de patrones [LIU. 2000].

1.3. Planteamiento del problema

El problema de reconocimiento facial ha sido abordado, estudiado por una gran cantidad de investigadores en los últimos años, los cuales han desarrollado distintas técnicas. Algunas de las técnicas necesitan una gran cantidad de dinero cuyo costo impide su comercialización a gran escala como ser cámaras de gran resolución sensores láser etc. Para el proyecto solo se requiere cámaras habituales conectadas a una computadora personal.

Con la imagen obtenida de la cámara se analiza la cara del sujeto partiendo de los niveles de gris que se tiene de una imagen. En entorno colaborativo, el sujeto se presenta delante de la cámara, ocupando la mayor parte de la imagen. Se puede reducir la dimensión de la imagen ocupada por la cara sin perder información utilizando el análisis de componentes principales. Otra aproximación estudiada consiste en reducir información partiendo de una pérdida de resolución, reduciendo la imagen a un tamaño de unos pocos píxeles. Una imagen de baja resolución permite a un ser humano verificar la identidad de un sujeto. El objetivo es conseguir que el sistema también sea capaz de conseguirlo.

Una vez fijados los puntos de partida, se deben considerar también los distintos clasificadores que se van a utilizar. En el presente estudio se considera tres tipos: Funciones de base radial, k-vecinos más cercanos y máquinas de soporte vector. La comparación de entre ellos permite definir para cada situación el tipo de clasificador que debemos utilizar.

El objeto de estudio de la tesis es contribuir al estudio de técnicas que permitan realizar de forma adecuada el proceso de verificación facial, técnicas como ser el reconocimiento de rostros humanos basados en una red neuronal.

Formulación Del Problema

¿Con el patrón de reconocimiento facial mediante redes neuronales será capaz de reconocer rostros humanos y proporciona un alto grado de confiabilidad?

1.4. Objetivos

1.4.1. Objetivo General

Determinar por patrones de reconocimiento facial mediante redes neuronales la identificación de rostros humanos, con un alto grado de confiabilidad.

1.4.2. Objetivos Específicos

- Determinar una resolución adecuada para las imágenes de rostros humanos que permitan una máxima clasificación.
- Implementar un modelo neuronal para el reconocimiento de rostros humanos
- Realizar la extracción de características de rostro a partir de imágenes fijas.
- Realizar el llenado de las características de un rostro a la base de datos.
- Comparar distintas formas de reconocimiento en dos dimensiones, partiendo de imágenes de niveles gris.
- Realizar el reconocimiento facial a partir de datos bidimensionales utilizando base de datos.
- Realizar un reconocimiento facial utilizando herramientas de inteligencia artificial (redes neuronales).

1.5. Hipótesis

Con los patrones de reconocimiento facial mediante redes neuronales, se maximiza la tasa de reconocimiento de rostro humano y minimiza la tasa de error, en forma eficaz y eficiente.

1.6. Justificación

1.6.1. Justificación Científica

La biometría es la ciencia que se dedica a la identificación de individuos a partir de una característica anatómica o un rasgo de su comportamiento. Una característica anatómica tiene la cualidad de ser relativamente estable en el tiempo, tal como ser la de un rostro humano. En las tecnologías de la información, la autenticación biométrica se refiere a las

tecnologías para medir y analizar las características físicas y del comportamiento humano con propósito de autenticación, identificación y evitar la suplantación.

1.6.2. Justificación Económica

La identificación por patrones de reconocimiento facial tendrá un costo mínimo y software necesario y la compra de una cámara fotográfica digital.

1.6.3. Justificación Social

La implementación de la tesis depende principalmente del contexto donde se utiliza el reconocimiento. Las aplicaciones de los sistemas pueden ir desde intereses gubernamentales, militares, privados o comerciales, hasta las aplicaciones forenses o policiales.

Un sistema de reconocimiento facial puede ayudar a los investigadores a encontrar a un sospechoso rápidamente. El reconocimiento facial permite buscar e identificar sospechosos incluso con información incompleta de su identidad, incluso sólo con la información que se puede obtener de testigos.

1.6.4. Justificación Técnica

Para la implantación de la tesis se cuenta con los requerimientos mínimos hardware como ser una computadora Pentium IV, una cámara fotográfica y una impresora para la optimización del manejo de información, en el desarrollo del sistema se empleará un lenguaje de programación orientado a objetos lo que permitirá la aplicación de una metodología orientada a objetos y así desarrollar un sistema fácil de adaptar y mantener.

1.7. Metodología y herramientas

1.7.1. Método Científico

Por "método científico" se entiende aquellas prácticas utilizadas y ratificadas por la comunidad científica como válidas a la hora de proceder con el fin de exponer y confirmar sus teorías. Las teorías científicas, destinadas a explicar de alguna manera los fenómenos que observamos, pueden apoyarse o no en experimentos que certifiquen su

validez. Sin embargo, hay que dejar claro que el uso de metodologías experimentales, no es necesariamente sinónimo del uso del método científico. Por ello, Francis Bacon definió el método científico de la siguiente manera:

- a. Observación: Observar es aplicar atentamente los sentidos a un objeto o a un fenómeno, para estudiarlos tal como se presentan en realidad, puede ser ocasional o causalmente.
- b. Inducción: La acción y efecto de extraer, a partir de determinadas observaciones o experiencias particulares, el principio particular de cada una de ellas.
- c. Hipótesis: Planteamiento mediante la observación siguiendo las normas establecidas por el método científico.
- d. Probar la hipótesis por experimentación.
- e. Demostración o refutación (antítesis) de la hipótesis.
- f. Tesis o teoría científica (conclusiones).

Así queda definido el método científico tal y como es normalmente entendido, es decir, la representación social dominante del mismo. La definición corresponde únicamente a la visión de la ciencia denominada positivismo en su versión más primitiva. Empero, es evidente que la exigencia de la experimentación es imposible de aplicar a áreas de conocimiento como la vulcanología, la astronomía, la física teórica, etcétera. En tales casos, es suficiente la observación de los fenómenos producidos naturalmente, en los que el método científico se utiliza en los estudios (directos o indirectos) a partir de modelos más pequeños.

1.7.2. Herramientas

La utilización de herramientas en la parte de análisis es: Entrevistas, encuestas las cuales ayudara a profundizar más en los niveles del problema y así concluir con una mejor visualización del mismo.

La metodología del diseño utilizada será orientada a objetos para el proyecto, porque es capaz de disminuir la complejidad y admitir la evolución en el desarrollo del software. “el modelo de proceso de desarrollo orientado a objetos se caracteriza, mediante un conjunto de objetos con atributos y comportamientos específicos. Los objetos son manipulados

mediante una colección de funciones (llamadas, métodos, operaciones o servicios) y se comunican entre ellos mediante un protocolo de mensajes”

En el desarrollo de las herramientas se tiene a Visual .NET que proporciona un entorno dinámico y diseño Visual, con la ventaja, que es un lenguaje Orientado a Objetos, se puede aprovechar los principios fundamentales de: Encapsulación, herencia, polimorfismo, sobrecarga, y reutilización de datos, para el manejo de base de datos se utilizara SQL y MYSQL para así poder utilizar las Base de Datos de forma segura y como servidor.

1.8. Alcances, límites y aportes

1.8.1. Alcances

El presente trabajo de investigación se enfoca al estudio e implementación de un modelo que permita el reconocimiento de rostros humanos en forma organizada e integra. El modelo contara con las siguientes funciones:

- Diseño de un modelo de reconocimiento facial mediante redes neuronales
- Implementación de un modelo de base de datos
- Reconocimiento de sujetos no dispuestos a cooperar.
- Susceptible a problemas de iluminación.
- Sistema vulnerable al reconocimiento de sujetos que se han sometido a operaciones de cirugía plástica (estéticas y de cirugía en general).

1.8.2. Límites

El presente trabajo de investigación se enfoca en la detección de rostros humanos y no toma en cuenta a sujetos que se han sometido a operaciones de cirugía plástica (estéticas y de cirugía en general).

1.8.3. Aportes

El estudio permite integrar el flujo de información concerniente a las imágenes, brindando el reconocimiento y de esa manera se automatice las funciones que se realizan. Brindando una interface de pantallas amigables al usuario y de fácil manejo, para ello se

utiliza el desarrollo de menús de selección. Siendo la implementación del sistema propuesto, que se logra a partir del análisis, diseño, codificación y pruebas de calidad del mismo para finalmente lograr una herramienta capaz de manejar y brindar información precisa e integrada.



MARCO TEÓRICO

CAPITULO II

2.1. Introducción

En el capítulo se van a introducir primeramente una serie de conceptos generales sobre la representación y procesado que hay que tener en cuenta a la hora de trabajar con imágenes. Posteriormente se realiza una descripción de las técnicas existentes para la detección y reconocimiento de caras en imágenes. También se explican algunos conceptos básicos sobre las redes neuronales ya que serán utilizadas en la fase de reconocimiento.

2.2. Visión humana y la visión artificial

Los sistemas computacionales actuales aunque su capacidad, precisión y velocidad resulten impresionantes, son aún lentos y escasos si las comparamos con la inteligencia de un ser humano que es capaz de reconocer en milésimas de segundo. La tarea del reconocimiento para una persona involucra aproximadamente diez trillones de operaciones por segundo que no es alcanzado aún por ningún sistema. Sin embargo, tecnologías futuras como la nanotecnología, los sistemas digitales-biológicos o la computación cuántica pretenden reducir esta brecha.

2.2.1. Visión ocular

Visión ocular es la facultad por la cual a través del ojo se percibe el mundo exterior. Muchos organismos simples tienen receptores luminosos capaces de reaccionar ante determinados movimientos y sombras, pero la verdadera visión supone la formación de imágenes en el cerebro. Los ojos de los distintos organismos proporcionan imágenes de diversa claridad.

La visión está relacionada en especial con la percepción del color, la forma, la distancia y las imágenes en tres dimensiones. En primer lugar, las ondas luminosas inciden sobre la retina del ojo, pero si las ondas son superiores o inferiores a determinados límites no producen impresión visual. El color depende, en parte, de la longitud o longitudes de onda de las ondas luminosas incidentes, que pueden ser simples o compuestas, y en parte del

estado del propio ojo, como ocurre en el daltonismo. La luminosidad aparente de un objeto depende de la amplitud de las ondas luminosas que pasan de él al ojo, y las pequeñas diferencias de luminosidad perceptibles siempre guardan una relación casi constante con la intensidad total del objeto iluminado [AGC.2006].

2.2.2. Visión artificial

Visión es uno de los mecanismos sensoriales de percepción más importantes en el ser humano aunque no es exclusivo. El intento de dotar a las máquinas de un sistema de visión aparece el concepto de visión artificial. La visión artificial es una tarea más difícil a comparación de lo que pueden lograr las personas, sus aplicaciones cada vez son más demandadas en todos los campos del desarrollo humano. La visión artificial involucra tres procesos, procesamiento de la imagen, análisis de imágenes y aplicaciones.

El procesamiento implica la manipulación de las imágenes vistas como señales digitales, para extraer la información deseada. Los métodos de procesamiento de imágenes digitales se fundamenta en dos áreas principales de aplicación: Mejora de la calidad para la interpretación humana, procesamiento de los datos de la escena para la percepción de las máquinas de forma autónoma.

El análisis se encamina a determinar ciertas estructuras elementales tales como contornos o regiones así como las relaciones entre ellas.

Finalmente las aplicaciones tratan de dar solución a los problemas relacionados con ciertas situaciones del mundo real, a saber: reconocimiento, movimiento, reconstrucción 3-D, etc. [AGC. 2006].

2.3. Procesamiento digital de imágenes

El procesamiento digital de imágenes, es una disciplina que desarrolla las bases teóricas y algorítmicas mediante las cuales pueda extraerse información del mundo real, de manera automática a partir de una imagen observada, de un conjunto de imágenes o de una secuencia. Tal información pudiera relacionarse con el reconocimiento de objetos genéricos, descripciones tridimensionales del mismo, posición y orientación de los objetos o la medición de cualquier propiedad espacial tal como la distancia entre dos puntos bien definidos o la sección transversal del objeto.

El espectro de aplicaciones es bastante amplio e incluye desde las aplicaciones industriales, imágenes aéreas y médicas. Por herramientas fundamentales que incluyen el pre procesamiento de la imagen, detección de contornos y segmentación, así como el estudio del movimiento en una secuencia de imágenes.

El uso de la computación en diferentes dominios, ha permitido sofisticar la naturaleza del procesamiento de las imágenes, con el propósito de extraer al máximo, cualquier información disponible en las mismas que pueda ser de utilidad como herramienta. El estudio de técnicas avanzadas de procesamiento digital de imágenes representa un prerrequisito importante con el motivo de desarrollar cualquier investigación en el área, que pueda ser significativa dentro de un contexto global.

El interés por métodos por el procesamiento digital de imágenes se deriva en dos áreas principales de aplicación: Mejoramiento de la información gráfica para una interpretación humana; y procesamiento de la imagen para que pueda almacenarse, transmitirse, y representarse de forma autónoma en los equipos. El objetivo principal en explicar sobre el tema es establecer un puente entre las imágenes digitales y el álgebra lineal y la manipulación para la extracción de datos relevantes [AGC. 2006].

2.3.1. Pixel

Es la abreviatura de las palabras inglesas "*picture element*". Es el menor de los elementos de una imagen al que se puede aplicar individualmente un color o una intensidad o que se puede diferenciar de los otros mediante un determinado procedimiento [GJC. 2009].



Figura 1: Representación de un píxel

Fuente: Gómez Jiménez Carmen Virginia [GJC. 2009].

2.3.2. Imagen digital

Una imagen digital se compone de una agrupación de píxeles, cada uno con un valor de intensidad o brillo asociado. Una imagen digital se representa mediante una matriz bidimensional, de forma que cada elemento de la matriz se corresponde con cada píxel en la imagen (ver figura 2) [GJC. 2009].

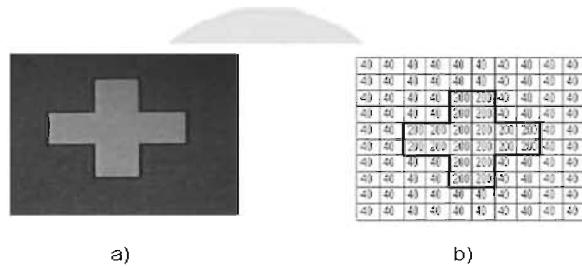


Figura 2: Imagen original y su estructura matricial de la imagen

Fuente: Gómez Jiménez Carmen Virginia [GJC. 2009].

2.3.3. Clasificación de imágenes digitales

Dependiendo del rango de los valores que pueda tomar cada píxel se puede distinguir los siguientes tipos de imágenes [GJC. 2009]:

- **Imágenes binarias:** El rango está formado por los valores negro o blanco [0, 1] únicamente. En la figura 3 se muestra una imagen de tipo binario.

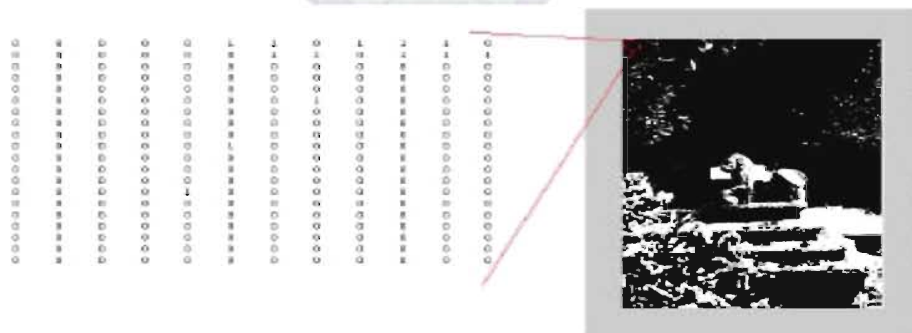


Figura 3: Imagen Binaria

Fuente: Gómez Jiménez Carmen Virginia [GJC. 2009].

- **Imágenes de intensidad:** También conocidas como imágenes en escala de grises, existen hasta 256 niveles de grises, por lo que su rango se encuentra entre [0,255].

24	22	35	31	64	171	108	94	162	104
33	18	33	17	55	105	179	212	125	107
23	10	33	75	88	56	98	92	72	97
34	11	33	69	48	26	22	6	41	26
28	25	47	30	32	20	27	24	30	30
21	37	28	84	38	89	83	188	28	41
22	26	21	27	33	17	32	90	20	34
13	25	35	35	25	29	28	24	42	23
48	34	23	56	52	55	2	37	25	26
23	86	54	33	188	171	117	9	22	37
67	40	33	32	84	36	51	37	30	36
65	109	48	23	2	25	28	42	17	29
42	126	46	7	46	75	58	55	17	35



Figura 4: Imagen en escalas de grises

Fuente: Gómez Jiménez Carmen Virginia [GJC. 2009].

- **Imágenes en color:** Todo color se puede componer a partir de tres componentes básicos. El contenido de cada píxel de la imagen es una terna de valores, un valor por cada componente de color básico.

Azul									
51	27	58	34	67	175	156	85	164	195
41	26	18	23	45	167	178	224	125	269
29	19	46	84	73	494	32	88	75	104
45	32	48	81	59	35	71	32	46	31
41	56								
52	50	23	20	38	32	84	174	192	88
27	35	33	64	12	17	48	189	179	214
13	25	17	9	30	74	47	39	31	84
30	25	12	8	37	68	48	28	33	8
12	52	28	29	47	38	32	23	29	18
49	30	22	28						
50	67	23	31						
51	100	28	35	30	25	30	32	62	165
19	40	48	42	35	22	15	18	35	97
51	26	48	42	25	11	37	74	62	31
60	26	18	36	34	11	57	67	42	19
37	27	78	110	26	21	45	53	27	13
55	29	48	126	15	21	20	27	32	23
52	36	28	39	15	21	14	28	29	11

Verde									
51	27	58	34	67	175	156	85	164	195
41	26	18	23	45	167	178	224	125	269
29	19	46	84	73	494	32	88	75	104
45	32	48	81	59	35	71	32	46	31
41	56								
52	50	23	20	38	32	84	174	192	88
27	35	33	64	12	17	48	189	179	214
13	25	17	9	30	74	47	39	31	84
30	25	12	8	37	68	48	28	33	8
12	52	28	29	47	38	32	23	29	18
49	30	22	28						
50	67	23	31						
51	100	28	35	30	25	30	32	62	165
19	40	48	42	35	22	15	18	35	97
51	26	48	42	25	11	37	74	62	31
60	26	18	36	34	11	57	67	42	19
37	27	78	110	26	21	45	53	27	13
55	29	48	126	15	21	20	27	32	23
52	36	28	39	15	21	14	28	29	11

Rojo									
51	27	58	34	67	175	156	85	164	195
41	26	18	23	45	167	178	224	125	269
29	19	46	84	73	494	32	88	75	104
45	32	48	81	59	35	71	32	46	31
41	56								
52	50	23	20	38	32	84	174	192	88
27	35	33	64	12	17	48	189	179	214
13	25	17	9	30	74	47	39	31	84
30	25	12	8	37	68	48	28	33	8
12	52	28	29	47	38	32	23	29	18
49	30	22	28						
50	67	23	31						
51	100	28	35	30	25	30	32	62	165
19	40	48	42	35	22	15	18	35	97
51	26	48	42	25	11	37	74	62	31
60	26	18	36	34	11	57	67	42	19
37	27	78	110	26	21	45	53	27	13
55	29	48	126	15	21	20	27	32	23
52	36	28	39	15	21	14	28	29	11



Figura 5: Imagen en color

Fuente: Gómez Jiménez Carmen Virginia [GJC. 2009].

2.3.4. Modelos de colores de imágenes digitales

Las computadoras almacenan y manipulan colores según el modelo empleado. Si se usa un byte de memoria para almacenar los valores de cada uno de los colores rojo, verde y azul, pueden representarse una imagen con más de 16 millones de combinaciones cromáticas. Además de determinar el número de colores que puede representar una imagen, los modelos de color afectan al número de canales y al tamaño de archivo de una imagen. Los canales son lo que permite representar la información de los elementos de color de la imagen. En general, al aumentar el número de colores o canales en una imagen, también aumenta el tamaño de la misma [AGC. 2006].

- **Mapas de bits:** Las imágenes con mapas de bits están hechas de un bit de color (blanco o negro) por píxel, y necesitan un espacio mínimo en el disco.
- **Escala de grises:** Las imágenes en escala de grises están compuestas generalmente de 8 bits de información por píxel y usan 256 variaciones de gris para simular gradaciones en color. Aunque es posible añadir canales nuevos a una imagen en escala de grises.
- **Color RVA:** Con el modelo, cada píxel necesita 24 bits de información. Con 8 bits son para representar los 256 niveles de intensidad de color por canal, siendo tres, rojo, verde y azul. Las imágenes RVA (RGB en inglés) se componen de estos canales para presentar imágenes con hasta 16,7 millones de colores en pantalla.
- **Color CMAN:** Se basa en las propiedades de absorción de luz de la tinta impresa en papel. Las imágenes CMAN están formadas por los cuatro colores, Cyan (azul verdoso) Magenta, Amarillo y Negro. Con cuatro canales se pueden representar hasta 2^{32} colores. Reservando cada canal con 8 bits.
- **El modelo HSI:** Se basa en la percepción humana del color. En el modelo, todos los colores se describen según características fundamentales como tono (H), saturación(S) y brillo (I).

2.3.5. Formatos de imágenes

En un gráfico de mapa de bits, cada punto tiene un lugar preciso, definido por su fila y su columna. Algunos de los formatos de gráficos de mapas de bits más comunes son: [AGC. 2006].

- **Formato BMP** (*Windows Bitmap*) es el formato de imagen de mapa de bits estándar de *Microsoft Windows*. Al guardar una imagen con el formato, se puede especificar una profundidad de imagen de 1 a 24 bits. En imágenes de 4 y 8 bits, también puede escoger compresión *Run-Length-Encoding* (RLE); El esquema de compresión no produce pérdidas, es decir, no elimina ningún detalle de la imagen.
- **Formato GIF** (*CompuServe Graphics Interchange Format*) es el formato de archivo utilizado normalmente para mostrar gráficos de color indexado e imágenes en documentos de Lenguaje Marcado como Hipertexto (HTML) en la *World Wide Web* y otros servicios electrónicos. GIF es un formato comprimido, diseñado para reducir al mínimo los tiempos de transferencia de imágenes por las líneas telefónicas.
- **Formato JPEG** (*Joint Photographic Experts Group*) se utiliza normalmente para mostrar fotografías y otras imágenes de tono continuo en documentos de Lenguaje Marcado como Hipertexto (HTML) en *World Wide Web* y otros servicios electrónicos. A diferencia del formato GIF, JPEG conserva toda la información de color de una imagen RVA. JPEG utiliza también un esquema de compresión que reduce de manera muy eficaz el tamaño del archivo mediante la identificación y eliminación de los datos no esenciales para la visualización de la imagen. Al abrir una imagen JPEG se descomprime automáticamente.

El esquema de compresión JPEG se conoce como compresión con pérdidas porque elimina datos. Significa que cuando la imagen se ha comprimido y descomprimido, la imagen descomprimida no será idéntica a la original. Un nivel elevado de compresión produce una baja calidad de imagen, un nivel bajo, una calidad mejor. En la mayoría de los casos, al comprimir una imagen utilizando la opción de calidad máxima, se obtiene un resultado muy parecido al original.

La técnica de compresión de JPEG puede reducir el tamaño de un archivo de gráficas hasta en un 96 por ciento.

- **Formato TIFF** (*Tagged-Image File Format*) de uso particular el sistema *Associated Press Picture Desk* utiliza el formato para intercambiar archivos entre aplicaciones y plataformas de ordenadores. El formato TIFF soporta compresión LZW, un método de compresión que no ocasiona pérdidas, ya que no elimina detalles de la imagen para que se puedan leer tanto en Macintosh como en una PC compatible con IBM.

2.4. Reconocimiento facial a partir imágenes de intensidad 2D

Se describe los antecedentes correspondientes al reconocimiento facial automático 2D. Una vez detectada, localizada y segmentada la cara de imagen, las siguientes etapas son la obtención de una representación de la misma y su posterior reconocimiento. De manera general, las técnicas existentes según el tipo de características empleadas se pueden clasificar en dos grandes grupos [CHL. 1995] [ZHA. 2000]:

- **Técnicas basadas en rasgos locales.** Son las que utilizan características que describen la cara parcialmente. Como ejemplo, un sistema podría consistir en extraer rasgos locales, por ejemplo, ojos, boca, nariz, etc. O bien líneas o puntos, a partir de los cuáles se calculan los valores de las características. Pueden consistir en distancias, áreas, ángulos, etc., y sirven para representar a la cara con vistas a la etapa de reconocimiento.
- **Técnicas holísticas.** También se denominan métodos basados en la apariencia o métodos basados en imágenes. Utilizan características de una imagen tales que la describen globalmente. Utilizan la imagen facial completa como patrón, incluido el fondo (aunque es deseable que ocupe la menor superficie posible), y por tanto se evita la etapa de segmentación. Son métodos holísticos los que se basan en el análisis de componentes principales. En ellos se genera una representación de la imagen global de tamaño mucho menor que el de la imagen original). Además existen otros métodos que emplean diferentes tipos de Redes Neuronales para clasificar información global (por ejemplo, imágenes directamente para efectuar el reconocimiento).

Existen múltiples técnicas derivadas de cada una de las mencionadas anteriormente, así como métodos híbridos que combinan ambas. Por ello, en cada una de las secciones, al introducir una técnica, se presentan además posibles derivaciones, y en ocasiones, métodos híbridos existentes que las combinen con otras.

2.4.1. Técnicas basadas en rasgos locales

Hasta los años 90 la mayoría de los sistemas de reconocimiento facial propuestos empleaban rasgos locales extraídos de las imágenes 2D de la cara, una vez localizada en la imagen. Las técnicas holísticas aparecidas con posterioridad, no los han suplantado,

sino que actualmente se siguen investigando ambas técnicas [ZHA. 2000]. En la década de los 80, el número de trabajos que se realizaron fue escaso debido al alto costo del hardware que se necesitaba para implementar aplicaciones en tiempo real [CHL. 1995]. Se comenzó la investigación abordando el problema del reconocimiento a partir de imágenes faciales de perfil, experimentándose sobre un número pequeño de sujetos.

En la década de los 90 se incrementó la investigación sobre el tema de forma vertiginosa, teniendo por objetivo obtener sistemas de reconocimiento facial completamente automáticos. Ello se debió a la gran cantidad de nuevas aplicaciones comerciales que potencialmente adquirieron los sistemas de identificación personal, así como a la reducción en aquellos años del costo del hardware necesario para el tratamiento de imágenes [ZHA. 2000]. Las primeras aplicaciones que se atribuían a sistemas eran las de identificación en las enormes bases de datos de imágenes faciales de las fuerzas de seguridad. Hasta mediados de los 90 la mayor parte de los trabajos, utilizaban imágenes 2D estáticas. En dicha década se comenzó a abordar el problema de reconocimiento utilizando imágenes frontales de los individuos, mayoritariamente [CHL. 1995].

2.4.1.1. Extracción de rasgos locales

La extracción automática de rasgos locales (ojos, boca, etc.) en general y de puntos faciales en particular tiene aplicaciones, no sólo en la obtención de características geométricas locales para reconocimiento facial, sino también en otras como el seguimiento de individuos en secuencias de vídeo, el análisis de expresiones faciales o el ajuste de puntos de control de un modelo 3D de una cabeza a una imagen para obtener un modelo particularizado de un individuo concreto. Por tanto, dicha tarea sigue siendo objeto de investigación en la actualidad y sobre ella continúan apareciendo trabajos recientes. La extracción de rasgos no se realiza todavía con suficiente precisión [ZHA. 2000].

Por otra parte, otro factor por el que la tarea resulta relevante en reconocimiento facial automático, es que los métodos holísticos basados en *eigenfaces* y *fisherfaces* son muy sensibles a variaciones de escala, por lo que necesitarán una etapa de localización de rasgos precisa si se desea una correcta normalización en cuanto a escala de la cara detectada.

Los métodos más importantes de extracción de rasgos faciales de imágenes 2D son:

- a. Los métodos genéricos de procesamiento de imágenes, que realizan búsqueda de bordes, líneas, curvas, etc.
- b. Los métodos basados en el uso de plantillas de los rasgos para localizarlos.
- c. Métodos de correspondencia estructural, que consideran las restricciones geométricas existentes entre los rasgos.

Los dos primeros realizan la búsqueda de los distintos rasgos de forma aislada, por tanto, tienen la desventaja de que cuando la apariencia del rasgo cambia por ejemplo, cuando se tiene un ojo cerrado, hay gafas, o boca abierta, tienen dificultades.

Los del tercer tipo (más recientes), basados en métodos de correspondencia estructural, han demostrado proporcionar mejores resultados.

Técnicas de procesamiento de imágenes. Existen numerosas publicaciones en las que para extraer rasgos locales (puntos, líneas del borde de la cara, ojos, etc.) se han utilizado técnicas típicas de procesamiento de imágenes como umbralización de la intensidad de los píxeles, detección de bordes, máximos de curvatura local de bordes, detección de esquinas, análisis del histograma de intensidad, variaciones en la resolución, etc. Además se han propuesto técnicas encaminadas al procesamiento de imágenes de caras.

- **Métodos basados en el uso plantillas fijas.** Son métodos que consisten en comparar cada región de la imagen con una plantilla de tamaño fijo que contiene únicamente un patrón semejante al rasgo que se desea localizar.
- **Métodos basados en el uso de plantillas deformables.** Son representaciones también de un patrón semejante al rasgo que se desea localizar, variables en cuanto a tamaño y forma, con objeto de adaptarse a todo tipo de variaciones posibles de ese rasgo.
- **Métodos basados en la búsqueda de puntos mediante el uso de wavelets (onditas)** [CAS. 1996]. Las transformadas más utilizadas son las de *Gabor* y *Morlet*. Entre los trabajos enmarcados en la técnica son presentados en [MAN. 1992]. Se utiliza la descomposición en onditas de *Gabor* para extraer puntos de la imagen. Dichos puntos rasgo obtenidos se utilizan para formar un grafo que se almacena en la base de datos. Las caras de test que se presenten al sistema

serán comparadas con las de la base de datos mediante técnicas de emparejamiento de grafos. Se pueden pre procesar en una etapa previa con máscaras de convolución aplicadas en las regiones de interés para resaltar la información importante, y posteriormente aplicar "wavelets".

2.4.2. Técnicas holísticas

Como alternativa al reconocimiento basado en rasgos locales, aparecieron a principios de los 90 técnicas 2D basadas en el cómputo de características holísticas o globales. Dichas técnicas que consideran la información completa de todos los píxeles de la imagen (tanto de la cara como del fondo) como patrón. En ciertas aplicaciones donde el formato de las imágenes es controlado, pueden evitar las dificultosas etapas de localización y segmentación de la cara del fondo y de extracción de características, sin embargo, requieren que las imágenes estén normalizadas en cuanto a escala. Son más apropiadas cuando hay ligeras variaciones en la pose (rotaciones alrededor de las tres direcciones espaciales), en la expresión facial (si es leve) y cuando la imagen está deteriorada. Sin embargo, estos sistemas son más sensibles a las variaciones en la iluminación que los basados en rasgos locales, aunque proporcionan muy buenos resultados cuando las condiciones de adquisición son uniformes en cuanto a escala e iluminación.

Los métodos basados en rasgos son menos sensibles a las variaciones en la iluminación y en la pose en comparación con los métodos holísticos.

Las técnicas propuestas en la literatura pueden clasificarse en dos grupos: las que se basan en el Análisis de Componentes Principales (PCA, del inglés *Principal Component Analysis*) por un lado, y las que enmarcaremos en la categoría de técnicas conexionistas (que utilizan redes neuronales como clasificadores de imágenes de caras).

2.4.2.1. Técnicas basadas en análisis de componentes principales

Se trata de la aplicación directa del análisis de componentes principales (PCA) al reconocimiento facial (lo que se denomina método de eigenfaces). El análisis de componentes principales se utilizó en [KIR. 1990] para procesar imágenes faciales y en [PEN. 1994] como base de una técnica de reconocimiento facial. De la técnica se derivan otras como el método *Fisherfaces* (Análisis Discriminante Lineal de *Fisher*). Ambas

técnicas (*eigenfaces* y *fisherfaces*) se han experimentado en bases de datos de un gran número de imágenes proporcionando buenos resultados cuando las imágenes han sido capturadas bajo condiciones de adquisición uniformes. Por ello, el esfuerzo actualmente se está orientando a la búsqueda de la independencia de las condiciones de la adquisición, para adquirir robustez ante cambios de iluminación, rotación, traslación y escala principalmente [ZHA. 2000].

Método de *eigenfaces*. Se trata de un método estadístico para la obtención de información global de la imagen facial. Es el que realiza una aplicación directa del Análisis de Componentes Principales, en el que se trabaja con las imágenes completas como patrones de entrada. En el método las caras se representan mediante vectores de características, pero no son las intuitivas derivadas de los ojos, boca, etc.. Las imágenes se codifican calculando su variación respecto del resto de la colección de caras, sin hacer ningún juicio sobre sus rasgos locales.

La inspiración de la técnica viene del hecho de que para una imagen facial $I(x,y)$ consistente en una matriz o array bidimensional de $N \times N$ valores de intensidad, si ésta se representa mediante el vector de longitud N formado por todos los valores de sus píxeles alineados, el espacio vectorial de todas las imágenes de longitud N , de cualquier objeto o escena (caras, coches, paisajes, etc.), será de dimensión N . Sin embargo, en el espacio de tan enorme dimensión, sólo un pequeño subespacio es el de imágenes de caras, denominado *espacio de caras* (*eigenspacé*). La idea principal del análisis de componentes principales es encontrar los vectores de la base de subespacio de menor dimensión, es decir, aquéllos que mejor describen la distribución de las imágenes del conjunto de caras.

Los vectores de la base que definen el subespacio de caras, se denominan *eigenfaces* (autocaras, en español). Cualquier cara pertenecerá al subespacio, y por tanto se podrá escribir en términos de una combinación lineal de *eigenfaces* siendo los coeficientes de la combinación lineal, las componentes del vector de características que la representan en el sistema de reconocimiento facial. Puesto que el máximo número de posibles *eigenfaces* o vectores de la base del subespacio es igual al número de caras del conjunto de entrenamiento, los vectores de características que se obtendrán con el método y serán como máximo de longitud (menor que su longitud inicial N).

Método de eigenfaces probabilísticas. Se presenta una mejora de la tasa de reconocimiento de la técnica de Análisis de Componentes Principales, en el caso introduciendo el método bayesiano de clasificación (que utiliza medidas probabilísticas de similitud) en lugar del basado en distancia euclídea.

Método de Fisherfaces. El Análisis Discriminante Lineal (LDA) es una técnica estadística clásica que proporciona la proyección que maximiza el ratio de dispersión entre los datos de las diferentes clases respecto a los de la misma clase.

Las características obtenidas mediante el método son útiles para la clasificación de patrones pues hace que patrones pertenecientes a la misma clase más próximos de unos a otros y las de diferentes clases se alejen entre ellos. El método aplicado al reconocimiento facial se ha denominado método *Fisherfaces*. Puede aplicarse tanto a imágenes de niveles de gris como a vectores de características. Al igual que el método PCA es un método estadístico que se basa en realizar una proyección para obtener una representación. En el caso de PCA, la proyección obtenida reduce la dimensión de los datos, manteniendo la máxima información de las variaciones entre los datos. Aunque LDA aporta mejores resultados de discriminación, sólo cuenta con una matriz de transformación, y cuando se tienen muchas clases y grandes variaciones, una única matriz no es suficiente para efectuar una buena discriminación. Con objeto de dar solución al problema se han presentado trabajos como el de [Kim03] en el que se realiza una partición de las clases (por razas, sexos o caras similares) y usa varias matrices de transformación, y aplica LDA a cada cluster de clases. De manera que utilizando varias matrices de transformación, mejoran los resultados de discriminación.

Métodos de clasificación mediante SVM. (Las *Support Vector Machines* (SVMs))

Son clasificadores que han demostrado tener una gran capacidad de generalización y por tanto ofrecen gran robustez en el reconocimiento de objetos en general y en el reconocimiento facial automático en particular. En el caso de que las clases sean linealmente separables, proporcionan el hiperplano óptimo que separa los patrones de entrenamiento (dicho hiperplano maximiza la suma de distancias entre las clases más cercanas). Los patrones de entrenamiento son proyectados en un espacio de mayor dimensión usando una función kernel. Las funciones más utilizadas como kernel son polinomiales, exponenciales y sigmoidales.

ICA (*Independent Component Analysis*). El Análisis de Componentes Independientes es un método utilizado para calcular vectores de la base de un espacio, que sean estadísticamente independientes. La técnica se ha utilizado para la obtención de características en reconocimiento facial automático, y puede considerarse una generalización de PCA. También trata de obtener un subespacio de proyección previo a la clasificación. ICA se ha aplicado a la separación de señales de audio, y recientemente se ha comenzado a aplicar al reconocimiento facial automático y de expresiones faciales. La técnica PCA consiste en obtener las proyecciones de las muestras en un subespacio de menor dimensión, tal que es óptimo en el sentido del error cuadrático medio, es decir, que al realizar la proyección inversa al espacio original, el error de reconstrucción de los datos comprimidos por el método, es mínimo. De hecho, si se incluyen todos los autovectores con autovalores distintos de cero la codificación se realiza sin pérdida de información. Sin embargo, ICA busca la transformación lineal que minimice la dependencia estadística entre los vectores de la base. Los vectores de la base del subespacio de proyección obtenidos mediante ICA, no son ortogonales, ni están ordenados bajo ningún criterio (a diferencia de los obtenidos mediante PCA).

Es una técnica para extraer variables estadísticamente independientes, de una mezcla de ellas. Tiene como objetivo encontrar una transformación lineal de los datos de entrada usando una base tan independiente (desde el punto de vista estadístico) como sea posible.

Por ello se considera que ICA es una generalización de PCA. PCA trata de obtener una representación de las entradas en variables no correlativas, mientras que ICA trata de obtener una representación en variables estadísticamente independientes. Las imágenes de la base que se obtienen mediante ICA contienen más información local que aquellas obtenidas con PCA, que aportan información global. Sin embargo, el tiempo de entrenamiento es mucho mayor para ICA que para PCA. En el contexto del Reconocimiento Facial PCA proporciona mejores resultados que ICA cuando la métrica de clasificación es la adecuada [BDB, 2000].

2.4.2.2. Técnicas conexionistas

Las redes neuronales artificiales son clasificadoras que se implementan con objeto de imitar al hombre en la tarea de reconocimiento. Se componen de elementos

computacionales que operan en paralelo, llamados nodos o neuronas, que se caracterizan por unos pesos que se van adaptando durante su entrenamiento para mejorar la precisión del reconocimiento. La modificación de los pesos de las conexiones sinápticas entre neuronas constituye el aprendizaje de la red neuronal.

Las redes neuronales son interesantes para ser utilizadas en sistemas de reconocimiento facial debido a: (a) la capacidad de estos sistemas para aprender automáticamente, (b) la propiedad de estos sistemas para continuar funcionando de forma aceptable cuando se utilizan imágenes incompletas, ruidosas o deterioradas, (c) el interés por la búsqueda de modelos que permitan el procesamiento masivamente paralelo y (d) a la similitud de estos sistemas con los modelos neurofisiológicos del cerebro. Las redes neuronales se han usado en el ámbito del reconocimiento facial para resolver varios problemas: clasificación de las imágenes según el sexo del individuo, reconocimiento facial, detección de caras en imágenes y clasificación de expresiones faciales.

Los modelos conexionistas que se aplican al reconocimiento facial pueden operar directamente con representaciones de caras basadas en las propias imágenes, es decir pueden operar directamente sobre la matriz de intensidades de píxeles, sin previa extracción de rasgos locales, incluso con imágenes de caras a muy baja resolución. También se pueden introducir como entradas a las redes neuronales vectores de características geométricas. Por tanto, las redes neuronales cuentan con la posibilidad clasificar tanto datos numéricos de rasgos locales como el conjunto de todos los píxeles de la imagen, es decir pueden trabajar o bien con características locales o bien con características globales [VAL. 1994].

2.4.3. Detección de caras en la imagen 2D

Una etapa necesaria en numerosos sistemas de reconocimiento facial es la de localización de la cara en la imagen (reconociendo si es una cara u otro tipo de patrón) para su posterior segmentación aislándola del resto de objetos. El éxito o fracaso en la etapa es crucial para el posterior reconocimiento: Una cara no puede ser reconocida si no ha sido correctamente localizada en la imagen.

2.5. Redes neuronales artificiales

Las redes neuronales artificiales (*Artificial Neuronal Net, ANN*) son un método de resolver problemas, de forma individual o combinada con otros métodos, para tareas de clasificación, identificación, diagnóstico, optimización o predicción.

Las redes neuronales son modelos que intentan reproducir el comportamiento del sistema nervioso de los animales. Se trata de un sistema de interconexión de neuronas en una red que colabora para producir un estímulo de salida.

Una red neuronal se compone de un conjunto masivamente paralelo de unidades de procesos muy simples, llamadas neuronas, y es en las conexiones entre unidades donde reside la inteligencia de la red. Biológicamente, un cerebro aprende mediante la reorganización de las conexiones sinápticas entre las neuronas que lo componen. De la misma manera, las redes neuronales tienen un gran número de procesadores virtuales interconectados que de forma simplificada simulan la funcionalidad de las neuronas biológicas. En la simulación, la reorganización de las conexiones sinápticas biológicas se modela mediante un mecanismo de pesos, que son ajustados durante la fase de aprendizaje. En una red neuronal entrenada, el conjunto de los pesos determina el conocimiento de esa red y tiene la propiedad de resolver el problema para el que ha sido entrenada [FAL. 2008]

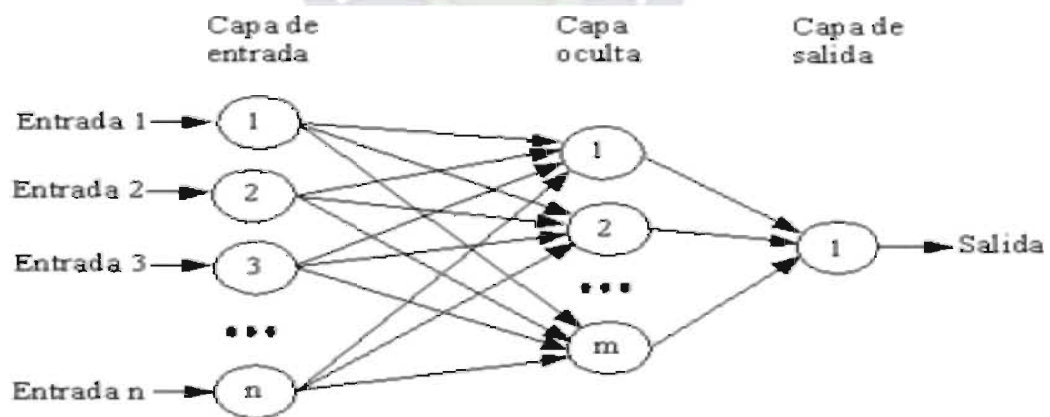


Figura 6: Ejemplo de red neuronal

Fuente: Luis Alberto Flores [FAL. 2008]

Perceptrón simple con n neuronas de entrada, m neuronas en su capa oculta y una neurona a la salida.

2.5.1. Funcionamiento de una red neuronal

Una red neuronal se compone de unidades llamadas neuronas. Cada neurona recibe una serie de entradas a través de interconexiones y emite una salida. Salida viene dada por tres funciones [FAL. 2008]:

a. Una función de propagación (también conocida como función de excitación), que por lo general consiste en el sumatorio de cada entrada multiplicada por el peso de su interconexión (valor neto). Si el peso es positivo, la conexión se denomina excitatoria, si es negativo se denomina inhibitoria.

$$f = \sum_i w_i x_i$$

b. Una función de activación, que modifica a la anterior. Puede no existir, siendo

En el caso la salida la misma función de propagación.

c. Una función de transferencia, que se aplica al valor devuelto por la función de activación. Se utiliza para acotar la salida de la neurona y generalmente viene dada por la interpretación que se quiera dar a dichas salidas. Algunas de las más utilizadas son la función sigmoide (para obtener valores en el intervalo $[0, 1]$) y la tangente hiperbólica (para obtener valores en el intervalo $[-1, 1]$).



Figura 7: Ejemplo de funciones de transferencia

Fuente: Luis Alberto Flores [FAL. 2008]

2.5.2. Ventajas

Las principales ventajas del uso de redes neuronales son [FAL. 2008]:

- **Aprendizaje adaptativo:** Es una de las características más atractivas de las redes neuronales, es la capacidad de aprender a realizar tareas basadas en un entrenamiento o una experiencia inicial. **Auto organización:** la red crea su propia representación de la información en su interior, descargando al usuario.
- **Tolerancia a fallos:** Ya que una red neuronal almacena la información de forma redundante, ésta puede seguir respondiendo aceptablemente aún si los datos de entrada son ruidosos.
- **Flexibilidad:** Una red puede manejar cambios no importantes en la información de entrada, como señales con ruido u otros cambios en la entrada (por ejemplo si la información de entrada es la imagen de un objeto, la respuesta correspondiente no sufre cambios si la imagen cambia un poco su brillo o el objeto cambia ligeramente).

2.5.3. Tipologías de las redes neuronales

2.5.3.1. Topología de redes neuronales

Una primera clasificación de las redes neuronales que se suele hacer es en función del patrón de conexiones que presentan. Así se definen tres tipos básicos de redes:

Dos tipos de redes de propagación hacia delante o acíclicas en la que todas las señales van desde la capa de entrada hacia la salida sin existir ciclos, ni conexiones entre neuronas de la misma capa.

Monocapa: Sólo existe una capa de neuronas de entrada y otra de salida. Por ejemplo el perceptrón y *Adaline*.

Multicapas: Existe un conjunto de capas intermedias entre la capa de entrada y la de salida (capas ocultas). Por ejemplo, el perceptrón multicapa.

Las redes recurrentes que presentan al menos un ciclo cerrado de activación neuronal. Por ejemplo la red de *Elman*, la red de *Hopfield* y la máquina de *Bolzman*.

2.5.3.2. Aprendizaje

Una segunda clasificación que se suele hacer es en función del tipo de aprendizaje de que es capaz (si necesita o no un conjunto de entrenamiento supervisado). Para cada tipo de aprendizaje se encuentran varios modelos propuestos:

- Aprendizaje supervisado: necesitan un conjunto de datos de entrada previamente clasificado o cuya respuesta objetivo se conoce. Ejemplos de tipo de redes son el perceptrón (simple y multicapa), la red Adaline y la memoria asociativa bidireccional.
- Aprendizaje no supervisado o auto organizado: no necesitan de tal conjunto previo. Ejemplos de redes son las memorias asociativas, redes de *Hopfield*, máquina de *Bolzman* y las redes de *Kohonen*.
- Redes híbridas: son un enfoque mixto en el que se utiliza una función de mejora para facilitar la convergencia. Un ejemplo de redes son las funciones de base radial (RBF).
- Aprendizaje reforzado: se sitúa a medio camino entre el supervisado y el auto organizado.

2.5.3.3. Tipo de entrada

Finalmente también se pueden clasificar las redes neuronales según sean capaces de procesar información de distinto tipo en:

- Redes analógicas: Procesan datos de entrada con valores continuos y, habitualmente acotados. Ejemplos son la red de *Hopfield*, de *Kohonen* y redes de aprendizaje competitivo.
- Redes discretas: Procesan datos de entrada de naturaleza discreta; habitualmente valores lógicos booleanos. Ejemplos de tipo de redes son las máquinas de *Bolzman* y *Cauchy* [GJC. 2009].

2.5.4. Perceptrón Multicapa

Un perceptrón multicapa es constituido por un conjunto de neuronas que se comunican enviándose señales entre sí a través de un conjunto de conexiones ponderadas. Es el caso más sencillo de red neuronal y su breve descripción se incluye para presentar los

elementos de una red. Dentro de una red neuronal, podemos distinguir una serie de aspectos

- Un conjunto de unidades de procesamiento (neuronas, células).
- Un estado de activación a_i para cada unidad, que equivale a la salida de la neurona.
- Conexiones entre las neuronas. Generalmente cada conexión se define por un peso W_{jk} el cual determina el efecto de la señal de la neurona j en la neurona k .
- Una regla de propagación, que determina la entrada efectiva U_i de una neurona a partir de las entradas externas.
- Una función de activación $f(\cdot)$ que determina el nuevo nivel de activación, a partir del valor de la entrada efectiva $U_i(t)$ y de la activación actual $a_i(t)$.
- Un método para tratar la información (regla de aprendizaje).
- Un entorno que le proporciona las señales de entrada y con el cual debe operar el sistema.

Morfológicamente, una red de neuronas se forma por múltiples capas de neuronas donde las neuronas de cada capa se conectan con todas las neuronas de la capa siguiente. La conexión se pondera mediante un peso, que es el que determina, junto con la salida, la activación o no de la neurona.

2.6. Visual Basic

Es un lenguaje de programación desarrollado por el alemán Alan Cooper para *Microsoft*. El lenguaje de programación es un dialecto de *BASIC*, con importantes agregados. Su primera versión fue presentada en 1991, con la intención de simplificar la programación utilizando un ambiente de desarrollo completamente gráfico que facilitara la creación de interfaces gráficas y, en cierta medida, también la programación misma. Desde el 2001 *Microsoft* ha propuesto abandonar el desarrollo basado en la API *Win32* y pasar a trabajar sobre un *framework* o marco común de librerías independiente de la versión del sistema operativo, *.NET Framework*, a través de Visual Basic *.NET* (y otros lenguajes como *C Sharp (C#)* de fácil transición de código entre ellos).

Visual Basic (Visual Studio) constituye un IDE (entorno de desarrollo integrado, o, en inglés, *Integrated Development Environment*) que ha sido empaquetado como un programa de aplicación; es decir, consiste en un editor de código (programa donde se escribe el código fuente), un depurador (programa que corrige errores en el código fuente para que pueda ser bien compilado), un compilador (programa que traduce el código fuente a lenguaje de máquina), y un constructor de interfaz gráfica o GUI (es una forma de programar en la que no es necesario escribir el código para la parte gráfica del programa, sino que se puede hacer de forma visual).

2.6.1. Compilador

El compilador de Visual Basic x.0 genera ejecutables que requieren una DLL para que funcionen, en algunos casos llamada MSVBVMxy.DLL (acrónimo de "*MicroSoft Visual Basic Virtual Machine x.y*", siendo x.y la versión) y en otros VBRUNXXX.DLL ("*Visual Basic Runtime X.XX*"), que provee todas las funciones implementadas en el lenguaje. Además existe un gran número de bibliotecas (DLL) que facilitan el acceso a muchas funciones del sistema operativo y la integración con otras aplicaciones.

2.6.2. Entorno de desarrollo

Su entorno de desarrollo es muy similar al de otros lenguajes e IDE's.

Se compone principalmente de su barra de herramientas y menús que se pueden personalizar con prácticamente la completa totalidad de los comandos del IDE a necesidad.

El espacio de trabajo donde se muestran todas las ventanas del proyecto, las vistas de código de módulos y objetos, y los controles con los que se componen las ventanas de nuestra aplicación.

Se pueden agregar todo tipo de controles de terceros, una gran cantidad de ellos de serie con la instalación de Visual Basic 6.0, que vienen embebidos dentro de archivos de extensión *.OCX.

El panel lateral derecho contiene dos vistas principales:

El Explorador de Proyectos, que muestra todos los elementos que componen nuestro proyecto o grupos de proyectos (formularios, interfaz de controles, módulos de código, módulos de clase, etc...)

El Panel de propiedades, donde se muestran todos los atributos de los controles de nuestros formularios o la información de módulos clase y formularios entre otros muchos.

2.6.3. Versiones

Las versiones de Visual Basic para Windows son muy conocidas, pero existe una versión de Microsoft Visual Basic 1.0 para MS-DOS (ediciones Profesional y Estándar) menos difundida y que data de 1992. Era un entorno que, aunque en modo texto, incluía un diseñador de formularios en el que se podían arrastrar y soltar distintos controles [http://es.wikipedia.org/wiki/visual_basic].

La última versión sólo para 16 bits, la 3.0, no incluía una detallada biblioteca de componentes para toda clase de usos. Durante la transición de Windows 3.11 a Windows 95, apareció la versión 4.0, que podía generar programas de 16 y 32 bits a partir de un mismo código fuente, a costa de un gran aumento en el tamaño de los archivos "*runtime*" necesarios. Además, se sustituyen los controles VBX por los nuevos OCX. Con la versión 5.0, estuvo a punto de implementar por primera vez la posibilidad de compilar a código nativo, obteniendo una mejora de rendimiento considerable. Tanto como la posterior 6.0 soportaban características propias de los lenguajes orientados a objetos, aunque careciendo de algunos ítems importantes como la herencia y la sobrecarga. La versión 6.0 continúa utilizándose masivamente y es casi compatible prácticamente al 100% con las últimas versiones de *Windows* como Vista y Windows 7.

Las versiones actuales de *Visual Basic* se basan en la plataforma .NET, que se desligan de las anteriores versiones.

2.6.4. Ventajas

- Es un lenguaje RAD (desarrollo rápido de aplicaciones).
- El código en Visual Basic es fácilmente migrable a otros lenguajes.
- Es un lenguaje muy extendido.
- Fácilmente extensible mediante librerías DLL y componentes ActiveX.

- Posibilidad de añadir soporte para ejecución de scripts, VBScript o JScript, en las aplicaciones mediante Microsoft Script Control.
- Acceso a la API multimedia de DirectX (versiones 7 y 8). También está disponible, de forma no oficial, un componente para trabajar con OpenGL 1.1: VBOpenGL type library [http://es.wikipedia.org/wiki/visual_basic].

2.6.5. Inconvenientes

- No es multiplataforma (sin embargo se pueden usar emuladores e intérpretes para correrlos en otras plataformas).
- Por defecto permite la programación sin declaración de variables, (que puede ser sencillamente corregida escribiendo la frase *Option Explicit* en el encabezado de cada módulo de código, en cuyo caso será menester declarar todas las variables a utilizar, lo que a la postre genera código más estable y seguro).
- No permite programación a bajo nivel ni incrustar secciones de código en ASM (aunque es posible ejecutar código ASM mediante pequeños *hacks*).
- Sólo soporta librerías dinámicas (DLL) que usen la convención de llamadas `stdcall` y componentes y librerías ActiveX.
- Es un lenguaje basado en objetos pero no implementa por completo la filosofía de Orientación a Objetos.
- No permite la sobrecarga de operadores ni métodos.
- No permite nombres de espacio.
- No soporta el puntero a memoria salvo en algunas acciones concretas, como por ejemplo pasar la dirección de memoria de una función como argumento (operador *AddressOf*).
- No soporta tratamiento de procesos como parte del lenguaje.
- No incluye operadores de desplazamiento de bits como parte del lenguaje.
- No permite el manejo de memoria dinámica, punteros, etc. como parte del lenguaje.
- No soporta el tratamiento de excepciones. Su tratamiento de errores se basa en la captura de mensajes y desvío del flujo de ejecución de la forma tradicional del BASIC (*On Error Goto <etiqueta/numero línea>*).

- No controla todos los errores de conversión de tipos dado que en muchas ocasiones hace conversiones al vuelo (sobre todo al usar variables de tipo *Variant*).
- Aunque existen opciones avanzadas en el compilador para desactivar los controladores de desbordamiento de enteros o las comprobaciones de límites en matrices entre otros (presumiblemente para optimizar y lograr algo de rendimiento) no es seguro del todo dado que hay más posibilidades de generar una excepción grave no controlada por el intérprete (y por consiguiente del programador) o un *memory leak* haciendo el programa altamente inestable e impredecible.
- No tiene instrucciones de preprocesamiento.
- El tratamiento de mensajes de Windows es básico e indirecto.
- La gran gama de controles incorporados son, sin embargo en algunos casos, muy generales, lo que lleva a tener que reprogramar nuevos controles para una necesidad concreta de la aplicación. Esto cambia radicalmente en Visual Basic .NET donde es posible reprogramar y mejorar o reutilizar los controles existentes.
- El depurador no es demasiado flexible ni cómodo en ciertas situaciones.
- Los controles personalizados no mejoran la potencia de la API de Windows, y en determinados casos acudir a ésta será el único modo de conseguir el control personalizado deseado.
- No soporta correctamente la programación multihilo haciendo muy inestable su ejecución inclusive desde el propio entorno de desarrollo.
- Su fuerte dependencia de librerías y componentes ActiveX que requieren de privilegios de administrador para poder instalar las aplicaciones (existen opciones de terceras aplicaciones para generar ejecutables independientes que embeben las dependencias en el propio código del ejecutable) [http://es.wikipedia.org/wiki/visual_basic]

2.6.6. Entornos de desarrollo

Existe un único entorno de desarrollo para *Visual Basic*, desarrollado por *Microsoft*: *Microsoft Visual Basic* x.0 para versiones desde la 1.0 hasta la 6.0, (con las diferencias entre las versiones desde la 1.0 (*MS-DOS/Windows 3.1*) hasta la 3.0 (*16 bits, Windows*

3.1/95) y las de la 4.0 (16/32 bits, Windows 3.1/95/NT) hasta la 6.0 (32 bits, Windows 9x/Me/NT/2000/XP/2003 server).

2.6.7. Alternativas multiplataforma o externas a Windows

Existen múltiples alternativas dentro y fuera de *Windows* que intentan imitar al lenguaje y su mecánica de desarrollo. El más conocido y popular es Gambas [http://es.wikipedia.org/wiki/visual_basic].

2.7. MySQL

Es un sistema de gestión de base de datos relacional, multihilo y multiusuario con más de seis millones de instalaciones. MySQL AB desde enero de 2008 una subsidiaria de *Sun Microsystems* y ésta a su vez de *Oracle Corporation* desde abril de 2009 desarrolla MySQL como software libre en un esquema de licenciamiento dual [<http://es.wikipedia.org/wiki/MySQL>].

Por un lado se ofrece bajo la GNU GPL para cualquier uso compatible con licencia, pero para aquellas empresas que quieran incorporarlo en productos privativos deben comprar a la empresa una licencia específica que les permita el uso. Desarrollado en su mayor parte en ANSI C.

Al contrario de proyectos como Apache, donde el software es desarrollado por una comunidad pública y el *copyright* del código está en poder del autor individual, MySQL es patrocinado por una empresa privada, que posee el *copyright* de la mayor parte del código.

Esto es lo que posibilita el esquema de licenciamiento anteriormente mencionado. Además de la venta de licencias privativas, la compañía ofrece soporte y servicios. Para sus operaciones contratan trabajadores alrededor del mundo que colaboran vía Internet. MySQL AB fue fundado por *David Axmark, Allan Larsson y Michael Widenius*.

2.7.1. Historia del proyecto

SQL (Lenguaje de Consulta Estructurado) fue comercializado por primera vez en 1981 por IBM, el cual fue presentado a ANSI y desde entonces ha sido considerado como un

estándar para las bases de datos relacionales. Desde 1986, el estándar SQL ha aparecido en diferentes versiones como por ejemplo: SQL: 92, SQL: 99, SQL: 2003. MySQL es una idea originaria de la empresa *opensource* MySQL AB establecida inicialmente en Suecia en 1995 y cuyos fundadores son *David Axmark, Allan Larsson, y Michael "Monty" Widenius*. El objetivo que persigue la empresa consiste en que MySQL cumpla el estándar SQL, pero sin sacrificar velocidad, fiabilidad o usabilidad. [<http://es.wikipedia.org/wiki/MySQL>].

Michael Widenius en la década de los 90 trató de usar *mSQL* para conectar las tablas usando rutinas de bajo nivel ISAM, sin embargo, *mSQL* no era rápido y flexible para sus necesidades. Lo que llevó a crear una API SQL denominada MySQL para bases de datos muy similar a la de *mSQL* pero más portable.

La procedencia del nombre de MySQL no es clara. Desde hace más de 10 años, las herramientas han mantenido el prefijo My. También, se cree que tiene relación con el nombre de la hija del cofundador *Monty Widenius* quien se llama *My*.

Por otro lado, el nombre del delfín de MySQL es *Sakila* y fue seleccionado por los fundadores de MySQL AB en el concurso "*Name the Dolphin*". El nombre fue enviado por *Ambrose Twebaze*, un desarrollador de software de código abierto africano, derivado del idioma *SiSwate*, el idioma local de Swazilandia y corresponde al nombre de una ciudad en *Arusha*, Tanzania, cerca de Uganda la ciudad origen de *Ambrose*.

2.7.2. Lenguajes de programación

Existen varias APIs que permiten, a aplicaciones escritas en diversos lenguajes de programación, acceder a las bases de datos MySQL, incluyendo C, C++, C#, *Pascal*, *Delphi* (via *dbExpress*), *Eiffel*, *Smalltalk*, *Java* (con una implementación nativa del *driver* de Java), *Lisp*, *Perl*, *PHP*, *Python*, *Ruby*, *Gambas*, *REALbasic* (*Mac* y *Linux*), *(x)Harbour* (*Eagle1*), *FreeBASIC*, y *Tcl*; cada uno utiliza una API específica. También existe una interfaz ODBC, llamado *MyODBC* que permite a cualquier lenguaje de programación que soporte ODBC comunicarse con las bases de datos MySQL. También se puede acceder desde el sistema SAP, lenguaje ABAP [<http://es.wikipedia.org/wiki/MySQL>].

2.7.3. Aplicaciones

MySQL es muy utilizado en aplicaciones web, como Drupal o phpBB, en plataformas (*Linux/Windows-Apache-MySQL-PHP/Perl/Python*), y por herramientas de seguimiento de errores como *Bugzilla*. Su popularidad como aplicación web muy ligada a PHP, que a menudo aparece en combinación con MySQL. MySQL es una base de datos muy rápida en la lectura cuando utiliza el motor no transaccional MyISAM, pero puede provocar problemas de integridad en entornos de alta concurrencia en la modificación. En aplicaciones web hay baja concurrencia en la modificación de datos y en cambio el entorno es intensivo en lectura de datos, lo que hace a MySQL ideal para el tipo de aplicaciones. Sea cual sea el entorno en el que va a utilizar MySQL, es importante adelantar monitoreos sobre el desempeño para detectar y corregir errores tanto de SQL como de programación.

2.7.4. Especificaciones

Plataformas. MySQL funciona sobre múltiples plataformas, incluyendo:

AIX, BSD, FreeBSD, HP-UX, GNU/Linux, Mac OS X, NetBSD , *Novell Netware*, *OpenBSD*, *OS/2 Warp*, QNX, SGI IRIX, *Solaris*, *SunOS*, *SCO OpenServer*, *SCO UnixWare*, *Tru64*, *eBD*, *Windows 95*, *Windows 98*, *Windows NT*, *Windows 2000*, *Windows XP*, *Windows Vista*, *Windows 7* y *Windows Server (2000, 2003 y 2008)*, *OpenVMS*

Características distintivas

Las siguientes características son implementadas únicamente por MySQL:

Múltiples motores de almacenamiento (*MyISAM*, *Merge*, *InnoDB*, *BDB*, *Memory/heap*, *MySQL Cluster*, *Federated*, *Archive*, *CSV*, *Blackhole* y *Example en 5.x*), permitiendo al usuario escoger la que sea más adecuada para cada tabla de la base de datos.

Agrupación de transacciones, reuniendo múltiples transacciones de varias conexiones para incrementar el número de transacciones por segundo.

PATRONES DE RECONOCIMIENTO FACIAL MEDIANTE REDES NEURONALES

CAPITULO III

3.1. Introducción

En el capítulo se expone en forma detallada el método para la identificación de una persona entre un conjunto de imágenes de rostros. Por la complejidad que involucra un sistema completo de reconocimiento, el trabajo trata básicamente de la detección, preprocesado y reconocimiento de rostros.

Previamente a la implementación, es muy importante definir los principales elementos y características de la red neuronal y homologar su modelación matemática a una posible modelación computacional a través de algoritmos y estructuras de datos.

La red neuronal, en el proyecto, se utiliza como un clasificador para reconocimiento de patrones y se logre aprender a reconocer personas por medio de imágenes bidimensionales de rostros.

3.2. Descripción informal

La detección y reconocimiento de rostros han sido los campos de la visión computacional más investigados. El problema ha atraído investigadores de diferentes disciplinas como la psicología, ciencias computacionales e ingeniería habiendo significativos avances. Un sistema completo automatizado de reconocimiento cumple con las etapas siguientes:

- Imagen
- Preprocesado
- Detección
- Reconocimiento
- Clasificación
- Resultados

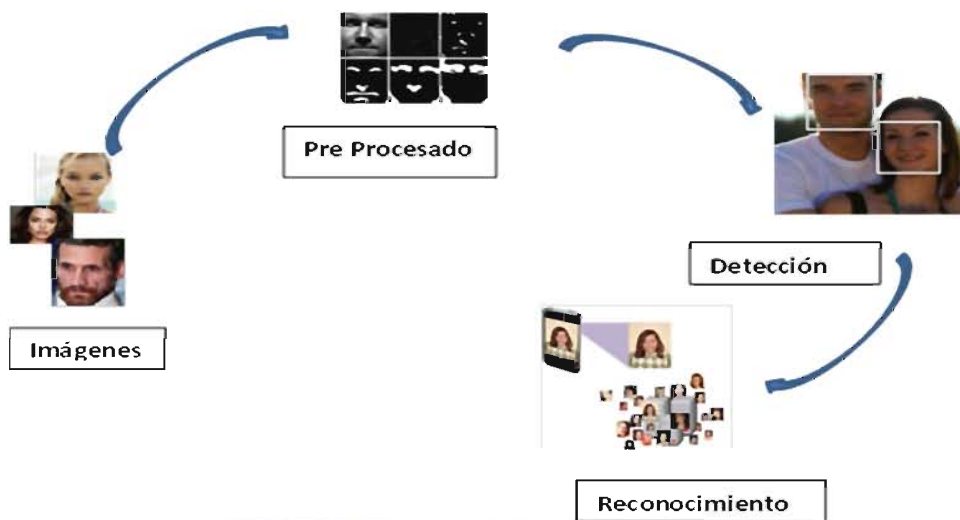


Figura 8: Etapas de reconocimiento facial

Fuente: Elaboración Propia

Cuando existe un conjunto de rostros en una misma fotografía se tiene un problema de detección de rostros y una vez detectados se requiere un proceso de localización. Un rostro ya detectado en una escena requiere operaciones geométricas y un proceso de normalización fotométrica para que sea localizado. Si la imagen de un rostro al compararse con otras imágenes existe una correspondencia entonces el rostro es identificado.

La idea de la detección de objetos es que los ojos humanos pueden detectar objetos sin esfuerzo, mientras las máquinas no pueden hacerlo de forma tan sencilla. Desde la perspectiva de la computación, es como un hombre en la oscuridad trata de encontrar un objeto con la ayuda de sus dedos o un palo.

El problema de detección de objetos puede ser considerado como tareas de segmentación y localización. Entre los métodos generalmente más usados para la detección son aquellos que buscan rasgos constantes para encontrar cierta afinidad o estructura geométrica. Al extraer los bordes de una imagen se puede encontrar ciertas afinidades.

La implementación del sistema de reconocimiento de rostros requiere de ciertos elementos materiales como ser: Cámara de video, Computador. Lo último permite

manipular las imágenes digitales para hacer mejoras en ellas. Como también, poder almacenar la aplicación que realiza el reconocimiento de rostros.

3.3. Descripción formal

Se expone en forma detallada el método para la identificación y reconocimiento facial entre un conjunto de imágenes estáticas tomadas en un ambiente controlado.

El empleo de imágenes estáticas tiene el inconveniente de utilizar proyecciones en dos dimensiones (imágenes) de un objeto que es tridimensional (cara humana). Implica una inevitable pérdida de información. A cambio, se logra una mayor sencillez en los algoritmos necesarios para el reconocimiento facial.

En general, el proceso de reconocimiento de rostros humanos utiliza la misma secuencia de etapas:

- Preprocesado
- Detección
- Reconocimiento (Extracción de características y Clasificación)

El desarrollado de la tesis lógicamente va a seguir también etapas generales. A continuación se muestra un diagrama de bloques en el que se representa de forma más detallada las etapas implementadas.



Figura 9: Etapas generales de desarrollo

Fuente: Elaboración propia

Preprocesado

Recibe como entrada una imagen y devuelve la imagen tras aplicarle un filtro para eliminar ruido.

Detección

Recibe como entrada la imagen preprocesada y devuelve la cara detectada si es que la hay en forma de ventana seleccionada. El algoritmo de detección se basa en el color de la piel.

Reconocimiento

En la fase del reconocimiento de caras se incluye la etapa de extracción de características y la clasificación, las etapas descritas anteriormente corresponderían a la fase de detección de caras. Por tanto se supone que las imágenes con las que se va a trabajar a partir del momento únicamente contienen una cara, aunque también depende del funcionamiento de la fase de detección, ya que puede ocurrir que el funcionamiento no sea tan bueno como se desea.

- **Extracción de características.** Recibe como entrada la cara detectada y devuelve un vector de características. Características que se obtienen mediante la aplicación de PCA a las imágenes.
- **Clasificación.** Recibe como entrada el vector de características de la imagen y devuelve la cara de la base de datos a las que más se parece, para clasificar correctamente, previamente debe haberse sometido a un proceso de entrenamiento. Como clasificador se utiliza una red neuronal.

3.3.1. Imágenes

La “materia prima” del trabajo son las imágenes de los individuos con las que se trabaja. En las que para cada individuo existen fotografías tomadas de frente, así como a color y en blanco y negro, además se ha escogido trabajar con las imágenes de toma frontal, a color y en la escala más reducida, cuya dimensión es 382x256 píxeles.

La elección de dichas imágenes se basa en los siguientes argumentos:

- **Toma frontal:** Las características principales del rostro humano se observan mejor de frente, por lo cual al escoger dicha toma, incluye mayor número de características que usando una imagen lateral. Dado que el objetivo se trata de extraer información de una imagen, por lo que se ha optado trabajar con imágenes frontales.
- **Color:** La elección de trabajar con las imágenes a color se debe básicamente a obtener la máxima información de la imagen tratada por lo que si se incluye la variable, color en nuestro desarrollo, se trabaja con mayor información sobre el individuo. Variable que beneficia también a la hora del preprocesado de la imagen además de proporcionar información en la extracción de características.
- **Dimensiones:** Una de las características es que los resultados obtenidos sean todo lo invariante posible, y para ello se trabaja con una imagen reducida. Con lo que se pretende, asegurar que al obtener las características en el “caso peor”, es decir, con la resolución más adversa, obviamente al tratar con una imagen más detallada los resultados serán mejores. A su vez, otra causa para el criterio de decisión se debe al costo computacional. Al trabajar con un espacio más reducido de puntos en la imagen, el tiempo del preprocesado así como el tamaño será también más reducido y el proceso total más rápido.

Todas las razones han sido usadas para la elección del tipo de imagen, con la necesidad de usar el mismo tipo de imágenes en general. Dado que se trata de implementar un proceso genérico para cualquier tipo de rostro, la necesidad de trabajar con el mismo formato de imagen es un requisito importante, y es la razón principal por la que se utiliza una base de datos.

3.3.2. Preprocesado de las imágenes

La técnica de preprocesado pretende mejorar las propiedades de la imagen para facilitar las operaciones, tales como la etapa de detección y reconocimiento.

La técnica de preprocesado se puede dividir, en general, en función de las pretensiones de sus transformaciones, en algunas de las siguientes facetas:

- Realce o aumento del contraste
- Suavizado o eliminación del ruido

- Detección de bordes

A continuación se expone técnicas de preprocesado.

3.3.2.1. Contracción del histograma

También llamado ecualización del histograma, técnica que mejora el contraste de la imagen al expandir la distribución de los niveles de gris. Siendo dicha expansión lo más suave posible en el sentido que debería de haber el mismo número de píxeles por niveles de gris. Es decir el objetivo es distribuir los niveles de gris de manera uniforme a lo largo de todo el rango de valores de niveles de gris.

La técnica hace una contracción al histograma.

$$\Gamma'(i,j) = \left[\left(\frac{C_{MAX} - C_{MIN}}{\Gamma(i,j)_{MAX} - \Gamma(i,j)_{MIN}} \right) \right] [\Gamma(i,j) - \Gamma(i,j)_{MIN}] + C_{MIN}$$

Ecuación para realizar la contracción del histograma de una imagen.

Donde $\Gamma(i,j)$ es el nivel de gris del píxel en la imagen de entrada Γ ; $\Gamma(i,j)_{MAX}$ es el mayor valor del nivel del gris en la imagen de entrada Γ ; $\Gamma(i,j)_{MIN}$ es el menor valor del nivel de gris en la imagen de entrada Γ ; C_{MAX} y C_{MIN} corresponden al máximo y mínimo de los valores deseados de gris a la compresión del histograma.



Figura 8: Fotografía antes de la contracción del histograma.

Fuente: Olivetti Research Laboratory ORL

Después de aplicar el filtro de contracción del histograma se observa un cambio en la imagen siguiente.



Figura 9: Fotografía después de contraerse el histograma

Fuente: Elaboración propia

3.3.2.2. Igualación del histograma

Se trata de encontrar una función $\Gamma'(\Gamma)$ que realce el contraste general en la imagen original Γ expandiendo la distribución de los niveles de gris. Dicha igualación debe de ser lo más suave posible, debería de haber mismo número de píxeles por niveles de gris. Dada una imagen de dimensiones xy por tener x columnas y y renglones, cada píxel tiene 256 niveles de gris.

Para que exista una mayor distribución en el histograma se calcula primero la probabilidad por cada nivel de gris g que viene dada por la siguiente ecuación.

$$P(g) = \frac{N(g)}{xy}; \quad g = [0, 255]$$

Ecuación para conocer la proporción por el nivel de gris g en una imagen.

Donde $N(g)$ es la cantidad de píxeles por cada nivel de gris, y xy son las dimensiones de la imagen (x columnas, y renglones). Se deduce que:

$$\sum_{g=0}^{g=255} P(g) = 1$$

Ecuación en donde se aclara que la suma de proporciones es igual a uno.

La función de densidad de probabilidad resulta:

$$P_{(i,j)}(X) \cong \sum_{g=0}^X P(g)$$

Función de densidad de probabilidad.



Figura 10: Imagen antes de la ecualización.

Fuente: *Olivetti Research Laboratory ORL*



Figura 11: Imagen ecualizada

Fuente: elaboración propia

3.3.2.3. Operador logarítmico

El rango dinámico de una imagen puede ser comprimido al remplazar el valor de cada píxel con su logaritmo natural. Tiene un efecto para los valores de píxeles que tienen una baja intensidad al ser éstos incrementados. Puede ser útil aplicar el operador logarítmico

en una imagen donde el rango dinámico puede ser muy grande para ser mostrado en pantalla. Puesto que el logaritmo de cero no está definido, se suma un uno al valor de píxel.

Se indica una constante de escala α con un valor máximo de 255, puesto que es el rango de niveles de intensidad que van del cero al de 255 para una imagen con un formato de ocho bits. Se agrega una constante δ de desplazamiento de histograma. Por lo tanto cada píxel es remplazado por la formula siguiente.

$$\Gamma' (x, y) = \alpha \ln(\Gamma (x, y) + 1) + \delta$$

Ecuación para obtener el operador logarítmico de una imagen.

Los nuevos valores de píxeles si son menores a cero ó mayores de 255 se limitan a 0 ó 255 según sea el caso.



Figura 12: Imagen antes de aplicarle el operador logarítmico.

Fuente: *Olivetti Research Laboratory ORL*

Al realizar en la imagen una operación logarítmica se obtiene la siguiente imagen:



Figura 13: Imagen después de aplicarle el operador logaritmo

Fuente: Elaboración propia

3.3.3. Detección de rostro

La fase de detección consiste en la segmentación de la imagen en regiones de interés conforme a algún criterio de homogeneidad. En la tesis se trata de encontrar la región de la imagen en la que aparece una cara humana.

El problema de detección de rostro es considerado como tareas de segmentación y localización. Entre los métodos generalmente más usados para la detección son aquellos que buscan rasgos constantes para encontrar cierta afinidad o estructura geométrica. Al extraer los bordes de una imagen se puede encontrar ciertas afinidades, así como el análisis a bajo nivel, es decir a nivel de píxel, basado en el color de la piel.

El tipo de detección puede tener el problema de que zonas de la imagen que no contengan caras pero con un color parecido al de la piel humana también son marcadas como zonas de piel. Para intentar determinar qué parte de la imagen es realmente una cara se trata de localizar características propias de una cara como los ojos y la boca.

3.3.3.1. Detección de píxeles de piel

Se ha optado por la utilización del espacio de color RGB. Varios investigadores consideran que el color de la piel es descrito por las crominancias y no depende de la luminancia, Lo que haría que el espacio de color RGB quizás no fuera el más adecuado para la detección de piel, ya que mezcla la crominancia y la luminancia. A pesar de todo, se opta por su utilización debido a su gran sencillez.

La detección de píxeles de piel consiste simplemente en comprobar si cada uno de los píxeles de la imagen cumple unas determinadas condiciones. Si las cumple quiere decir que es un píxel de piel humana. Las condiciones que deben cumplirse para que un píxel sea considerado como de piel humana se muestran a continuación [JPF. 2003].

$$R > 95 \text{ AND } G > 40 \text{ AND } B > 20$$

AND

$$\max \{R,G,B\} - \min \{R,G,B\} > 15$$

AND

$$|R - G| > 15$$

AND

$$R > G \text{ AND } R > B$$

El resultado de la detección de piel es una imagen binaria (formada por 0s y 1s) en la que cada píxel de piel detectado en la imagen original es puesto a 1, y el resto son puestos a cero.

La principal ventaja del método es la simplicidad en las reglas de detección de piel, lo que lleva a la construcción de un clasificador rápido. La mayor dificultad es utilizar el espacio de color y las reglas empíricas adecuadas.

3.3.4. Reconocimiento

La tarea del reconocimiento de rostros es la discriminación de las señales de entradas (datos de imágenes) en diferentes clases (personas). Las señales son altamente ruidosas; pero aún al capturar las imágenes la información no es completamente al azar y a pesar de las diferencias, existen patrones recurrentes. Tales patrones, que pueden ser observados en todas las señales como la presencia de algunos objetos (ojos, nariz, boca, etc.). Los rasgos pueden ser extraídos de la imagen original por medios de métodos como el Análisis de Componentes Principales (ACP).

El ACP es una técnica usada para describir conjuntos de imágenes que se registran espacialmente. La herramienta matemática pretende reducir el número de operaciones haciendo para el ordenador abordable el problema de reconocimiento. La idea es aprovechar la redundancia de un conjunto por hallarse constituido por imágenes de naturaleza similar. (Puesto que todas son de rostros humanos).

El enfoque de ACP comienza con un desglose de un conjunto imágenes no entrenadas que son ingresadas al sistema conocido como Conjunto de Entrenamiento (CE) extrayendo los componentes distintivos para generar un espacio de múltiples dimensiones llamado como Espacio de Rostros, a la etapa se le conoce como Entrenamiento.

Una vez realizada el entrenamiento, cualquier imagen puede ser proyectada al espacio. Sin embargo, solamente aquellas imágenes que compartan características similares serán proyectadas de manera exitosa.

Las imágenes que no existen en el CE aún si llegarán a pertenecer a un individuo; esas imágenes usualmente son detectadas a un grado satisfactorio, siempre y cuando no difirieran más allá de lo permitido con las del CE.

El objetivo del ACP es extraer información que sea relevante en la imagen para compararla con respecto a modelos de base de datos codificados de una manera similar. Es decir, se extrae la información relevante contenida en la imagen, capturando las variaciones existentes del CE, sin prejuicios sobre las características presentes, y usar la información para codificarla y compararla con otras imágenes.

En términos matemáticos, con el ACP se representan grandes dimensiones de datos en un espacio menor con vectores ortogonales significativos. Se trata a las imágenes como vectores, y forman un espacio de multidimensional. El conjunto de características que definen las variaciones se conocen como autovectores. Cada imagen ubicada contribuye de alguna manera a cada autovector, de tal forma que se muestra los autovectores como un rostro fantasmal llamado Rostro Característico, autocaras o *eigenfaces*.

El número de autocaras es igual al número de imágenes del CE y su uso es motivado porque con el ACP se puede representar imágenes a partir de reconstrucciones aproximadas, almacenando una pequeña colección de pesos relativos para cada rostro del CE y las autocaras correspondientes.

Entonces la imagen reconstruida es una aproximación de la imagen original. La imagen original en un CE puede ser reconstruida, si se suman todas las autocaras en la proporción correcta.

Cada autocara representa solamente ciertos rasgos de un rostro, que pueden o no ser presentados en la imagen original. Si el rasgo en la imagen original es distinto al resto de las imágenes del CE, en la “suma” de las autocaras deberá tener mayor presencia. Si por el contrario, el rasgo particular no es (o casi no) presentado en la imagen original entonces la autocara correspondiente deberá contribuir a una menor parte de la suma de las autocaras.

La imagen reconstruida original es igual a la suma de todas las autocaras, con cada autocara teniendo un peso específico. Los pesos indican la cantidad por lo que el rostro en cuestión difiere de los rostros típicamente representados por autocaras. Es decir los pesos determinan el grado que un rasgo específico presente en la imagen original.

Es posible también usar solamente una parte de las autocaras; los rostros pueden aproximarse usando solamente las “mejores” autocaras, considerando solamente las que describan mejor las peculiaridades distintivas de la imagen original. Se puede asegurar que las pérdidas debido a la omisión de algunas pueden ser minimizadas ya que se escogerían solamente los rasgos más importantes o distintivos. La omisión de autocaras es necesariamente debido a la escasez de recursos computacionales pero puede significar en un decremento en la eficiencia del reconocimiento si los rostros son muy semejantes como son en el caso de los gemelos.

Al ser posible no solamente extraer el rostro desde las autocaras dados a un conjunto de pesos, también lo es extraer los pesos de autocaras y el rostro a ser reconocido. Por lo tanto, usando los pesos puede determinar lo siguiente:

- El determinar si la imagen en cuestión es un rostro del todo. En caso de que los pesos de la imagen difieran demasiado de los pesos de las imágenes de rostros entonces probablemente la imagen no es un rostro.
- Los rostros similares poseen rasgos similares en grados similares (pesos). Si se extraen pesos de todas las imágenes disponibles, las imágenes podrían ser

agrupadas. Todas las imágenes que tienen pesos similares probablemente son rostros similares.

Hasta el momento, se puede resumir que en el método de ACP, el reconocimiento de rostros se realiza correlacionando una imagen con respecto a todas las imágenes del Conjunto de Entrenamiento. La imagen del conjunto que proporcione la mayor correlación será la correspondiente a la identidad de la persona que se desea averiguar.

3.3.4.1. Algoritmo para el reconocimiento de rostros

El algoritmo para el reconocimiento de rostros usando autocaras es básicamente descrito en la figura 14, las imágenes originales de rostros del Conjunto de Entrenamiento son adquiridas y transformadas a un conjunto de autocaras E , mismo conjunto que definirá el Espacio de Rostros. Después, los pesos de características son calculados para cada imagen del conjunto de entrenamiento y se almacenan dichos pesos en el conjunto W .

En la obtención de una imagen desconocida X , los pesos son calculados para esa imagen particular y almacenados en el vector Wx . Después, Wx es comparado con los pesos de las imágenes de rostros del conjunto de entrenamiento W , es decir, que se proyectará la nueva imagen en cada una de las autocaras del conjunto E . Una manera de hacer, es considerar cada vector de pesos como un punto en un espacio multidimensional y calcular la distancia euclidiana D entre el vector de pesos W y el vector de pesos de la imagen desconocida Wx . Si la distancia excede un valor umbral predefinido θ entonces el vector de pesos de la imagen desconocida Wx es muy alejado de los pesos de las imágenes. El caso, la imagen no se clasifica. En caso contrario, el vector de pesos Wx es almacenado para una clasificación. La determinación del umbral se define empíricamente.

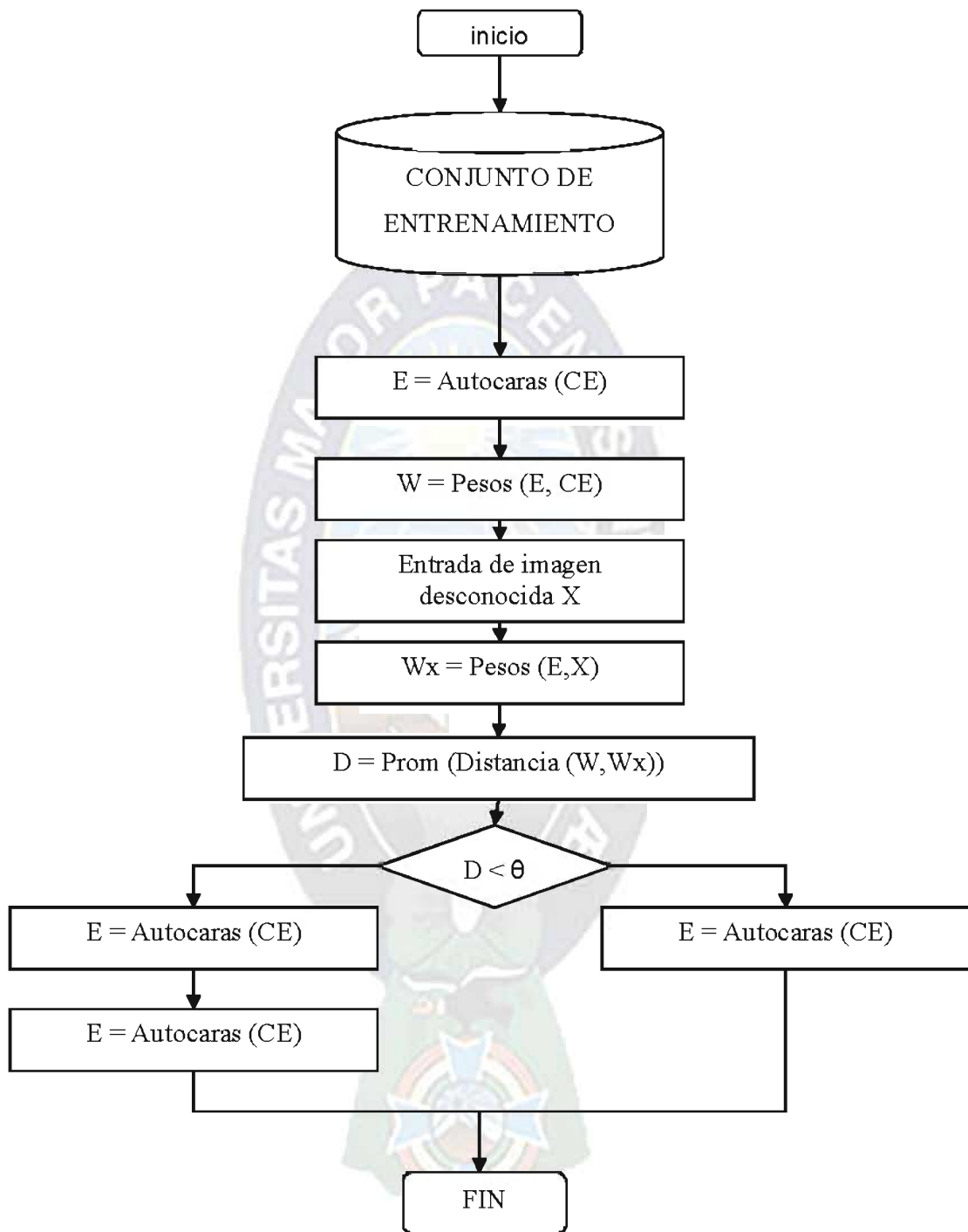


Figura 14: Algoritmo para el reconocimiento de rostros humano

Fuente: Elaboración propia

La aproximación de autocaras para el reconocimiento de rostros involucra varias operaciones englobadas en dos fases. En la fase de inicialización las operaciones son:

- Adquirir un conjunto de imágenes para su entrenamiento.
- Calcular las autocaras del conjunto de entrenamiento, manteniendo solamente las mejores M imágenes con los autovalores más altos. Las M imágenes se define como un “Espacio de Rostros”. A medida que se experimenta el sistema con nuevos rostros las autocaras pueden ser actualizadas.
- Calcular la distribución correspondiente en un espacio de M dimensiones de pesos para cada individuo conocido (imagen de entrenamiento) proyectando su imagen de rostros en el espacio de rostros.

Habiendo inicializando el sistema, los siguientes pasos se hacen para el reconocimiento de nuevas imágenes de rostros.

- Dada una imagen para ser reconocida, se calculan un conjunto de pesos de las M autocaras al proyectar los pesos a cada una de las autocaras.
- Determinar si la imagen es un rostro verificando si dicha imagen representa en un grado suficiente al Espacio de Rostros.
- Si es un rostro, clasificar los pesos de patrones ya sea como una persona conocida o desconocida.
- Actualizar las autocaras y/o los pesos de los patrones (opcional).
- Calcular las características de los peso de patrones de la nueva imagen, e incorporar a los rostros conocidos (opcional).

3.3.4.2. Metodología en el reconocimiento de rostros

Teniéndose un conjunto de entrenamiento (CE) Γ , los vectores $\Gamma_1, \Gamma_2, \dots, \Gamma_M$ donde cada vector Γ_i permite mostrar el rostro de una persona en 256 tonos de gris. Las imágenes con un tamaño de x columnas por y renglones generarán vectores de dimensión N (xy). Es recomendable antes de generar los vectores aplicar un filtro como una ecualización, para mejorar la imagen.

Los vectores del CE definen un espacio vectorial denominado como “Espacio de Rostros”. Como el conjunto es de un tipo de imágenes con una configuración general similar (porque son todas las imágenes rostros humanos), no se distribuirán los vectores de

manera aleatoria en el espacio, por lo que se pueden describir en un subespacio dimensional menor. La idea principal del ACP es encontrar el vector que mejor explique la distribución de los rostros dentro del espacio.

Considérese un CE con 3 imágenes diferentes de 6 personas (M=18). Cada imagen tiene un tamaño de 92 × 112 píxeles, por lo tanto cada vector Γ_i tiene una dimensión de 10.304 (N).

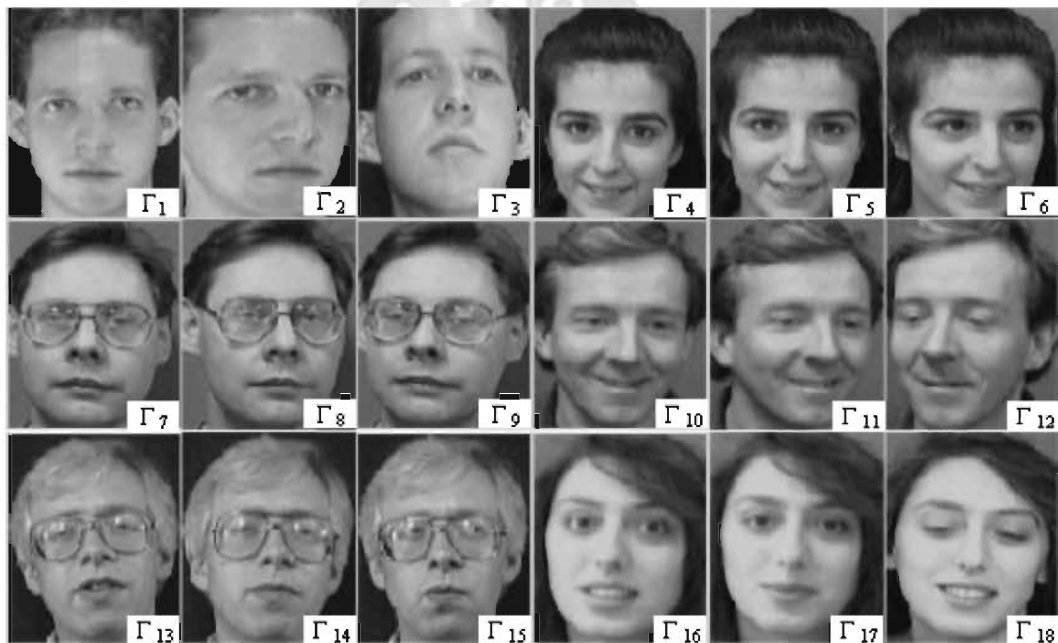


Figura 15: Conjunto de Entrenamiento (CE).

Fuente: Olivetti Research Laboratory ORL

El conjunto de entrenamiento se utilizará para mostrar como el ACP funciona y como operan las ecuaciones. Entonces con el CE se calcula el Rostro Promedio Ψ .

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

EC 3-1 Ecuación para obtener el Rostro Promedio.

Para el CE el vector ψ se muestra así.



Figura 16: Rostro Promedio según el CE.

Fuente: Elaboración propia

Cada rostro del CE difiere en algún grado de éste Rostro Promedio ψ .

$$\Phi = \Gamma - \Psi$$

EC 3-2 Ecuación para calcular las diferencias al Rostro Promedio (Φ).

Los vectores Φ_i visualmente se observan de la siguiente manera:



Figura 17: Φ_i son resta del Rostro Prom. a cada imagen del CE .

Fuente: Elaboración propia

Se buscan los M autovectores u_k ortonormales que mejor describan la distribución de datos, de manera que cualquiera de las imágenes pueda reconstruirse de la siguiente forma (EC 3-3):

$$\Gamma_j = \Psi + \sum_{k=1}^M u_k \omega_{jk}$$

EC 3-3 Ecuación para proyectar una imagen Γ_j al Espacio de Rostros.

Los coeficientes w_{jk} son determinados por las proyecciones sobre los autovectores u_k :

$$w_{jk} = \Phi_j^T u_k$$

EC 3-4 Ecuación para conocer la matriz de pesos W .

Lo que se busca es encontrar las direcciones de los vectores u_k donde las desviaciones se encuentren concentradas. En la siguiente ecuación provee una medida de desviación en dirección al vector u_k .

$$\lambda_k = \frac{1}{M} \sum_{j=1}^M (u_k^T \Phi_j)^2$$

EC 3-5 Medidas de desviación de los autovectores u_k (autovalores de la matriz de covarianzas).

Ahora, se introduce el concepto de Matriz de Covarianzas:

$$C = \frac{1}{M} \sum_{i=1}^M \Phi_i \Phi_i^T$$

EC 3-6 Cálculo de la matriz de covarianzas.

Por lo tanto los vectores u_k y los escalares λ_k son los autovectores y autovalores de la matriz de covarianzas C . Pero con (EC 3-6) la matriz C tendría dimensiones de N^2 y el cálculo de los autovectores u_k sería una tarea muy difícil. Considerando que el tamaño del CE, es un número menor a la dimensión del espacio de las imágenes ($M < N$) se puede construir una matriz con una menor dimensión. Ahora, considérese la siguiente matriz (EC 3-7):

$$A = [\Phi_1 \Phi_2 \dots \Phi_M]$$

EC 3-7 Definición de la matriz para el cálculo de Matriz de Covarianzas reducida A .

Entonces el cálculo de los autovectores v_k de $A^T A$ sería:

$$A^T A v_k = \lambda_k v_k$$

EC 3-8 Autovectores v_k de la matriz $A^T A$.

Pre multiplicando ambos lados de la ecuación (EC 3-8) por A entonces se tienen que $A v_k$ son los autovectores de $C = A A^T$. Entonces la ecuación (EC 3-6) puede reescribirse:

$$C = \frac{1}{M} A A^T$$

EC 3-9 Cálculo de Matriz de Covarianzas reducida (dimensiones MN).

Por lo anterior, los autovectores u_k de la matriz C (EC 3-5) se pueden obtener a partir de los autovectores v_k de la matriz reducida $A A^T$ de la siguiente forma:

$$u_k = A v_k$$

EC 3-10 Cálculo de los autovectores u_k

Entonces la matriz C con (EC 3-9) las dimensiones serán menores, obtener los autovectores v_k a través de un proceso iterativo como el método de potencias y la ortonormalización de los v_k con un método como el de *Householder*, permite que se obtengan los autovectores u_k (EC 3-10).

Para observar correctamente los autovectores u_k , (EC 3-10) los vectores Φ_i (EC 3-2) y el Rostro Promedio ψ (EC 3-1) se normalizan los vectores a un rango entre 0 y 255, (solo para propósitos de visualización) para que se representen en imágenes de gris (EC 3-11).

$$V = 255 - \left(\frac{255 (MAX(V) - V)}{|MAX(V) - MIN(V)|} \right)$$

EC 3-11 Procedimiento para representar correctamente un vector V .

$MAX(V)$ es el mayor valor del vector V y $MIN(V)$ es el menor. La longitud del Rango de valores del vector V es $|MAX(V) - MIN(V)|$. Cuando se transforman los autovectores u_k , en imágenes se obtiene según en el campo de reconocimiento de rostros "autocaras", "Rostros Característicos" ó "eigenfaces". Al transformar en imágenes los autovectores u_k del CE son vistos de la siguiente manera:

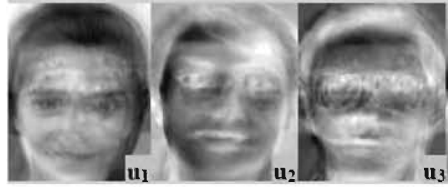


Figura 18: Autocaras

Autocaras obtenidas a partir de la conversión de los autovectores u_k (EC 3-10) con el CE.

Fuente: Elaboración propia

Después los autovectores u_k son “pesados” (EC 3-4) para reconstruir las imágenes proyectadas al espacio de rostros (ER) (EC 3-3) construyendo primero la matriz W en (EC 3-12).

$$W = \begin{bmatrix} u_1^T \Phi_1 & u_1^T \Phi_2 & \dots & u_1^T \Phi_M \\ u_2^T \Phi_1 & u_2^T \Phi_2 & \dots & u_2^T \Phi_M \\ \vdots & \vdots & \ddots & \vdots \\ u_M^T \Phi_1 & u_M^T \Phi_2 & \dots & u_M^T \Phi_M \end{bmatrix}$$

EC 3-12 Definición de la matriz de pesos W del Espacio de Rostros.

Supóngase que se quiere proyectar la siguiente imagen Γ_a al Espacio de Rostros:



Figura 19: Imagen Γ_a no pertenece al CE y será proyectada al (ER)

Fuente: Olivetti Research Laboratory ORL

Cuando se proyectan los vectores Γ_i que pertenecen al CE sobre el mismo Espacio de Rostros la reconstrucción de la imagen es casi perfecta. Sin embargo, al hacerlo con la

imagen (figura 19) diferencias son esperadas. Al vectorizar la imagen Γ_a y restarle el Rostro Promedio ψ de manera similar a (EC 3-2) se obtiene el vector Φ_a . Y si desea ver el resultado se le aplica (EC 3-11)

Ahora si se proyecta Γ_a al Espacio de Rostros, primero se calculan de las proporciones, (EC 3-10) cada autovector u_k aportará en diferente grado para la reconstrucción de la imagen Γ_{PROY} (Proyección al espacio de rostros) y conocer el grado en que se proyecta la imagen Γ_a al Espacio de Rostros según (EC 3-3). Primero se genera el vector ω_a con (EC 3-4) de la siguiente forma:

$$\begin{aligned} \omega_{1a} &= \Phi_a^T u_1 = -652.37 \\ \vdots & \quad \quad \quad \vdots \\ \omega_{18a} &= \Phi_a^T u_{18} = -154.22 \end{aligned}$$

EJ 3-1 Ejemplo de cómo generar el vector de pesos ω_a según (EC 3-4). Con el vector Φ_a y los autovectores u_k mostrados como autocaras en (figura 18).



Figura 20: Proporciones para la reconstrucción de Γ_{PROY} .

Las imágenes son observadas usando la ecuación (EC 3-11)

Fuente: Elaboración propia

Ahora, se aplica la suma en (EC 3-3) comenzando por el Rostro Promedio (figura 16) a las imágenes (figura 20).



Figura 21: Proyección al Espacio de Rostros.

Por una suma acumulativa de las imágenes

Fuente: Elaboración propia

Se puede observar que al reconstruir la imagen tiende a semejarse a Γ_8 . Para sustentar tal afirmación se encuentran las distancias euclidianas entre cada vector columna de la matriz \mathbf{W} y vector \mathbf{w}_a . La menor distancia euclidiana corresponde a la imagen más semejante a la imagen reconstruida.

$$\mathbf{w}_a = \begin{bmatrix} \mathbf{u}_1^T \Phi_a \\ \mathbf{u}_2^T \Phi_a \\ \vdots \\ \mathbf{u}_M^T \Phi_a \end{bmatrix} \quad \mathbf{M} = \begin{bmatrix} \mathbf{u}_1^T \Phi_1 & \mathbf{u}_1^T \Phi_2 & \mathbf{u}_1^T \Phi_3 & \dots & \mathbf{u}_1^T \Phi_M \\ \mathbf{u}_2^T \Phi_1 & \mathbf{u}_2^T \Phi_2 & \mathbf{u}_2^T \Phi_3 & \dots & \mathbf{u}_2^T \Phi_M \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \mathbf{u}_M^T \Phi_1 & \mathbf{u}_M^T \Phi_2 & \mathbf{u}_M^T \Phi_3 & \dots & \mathbf{u}_M^T \Phi_M \end{bmatrix}$$

$$\mathbf{W} = [\mathbf{w}_1 \quad \mathbf{w}_2 \quad \dots \quad \mathbf{w}_M]$$

$$\begin{aligned} \mathbf{e}_1 = \mathbf{w}_a - \mathbf{w}_1 & \quad |\mathbf{e}_1| = \sqrt{\mathbf{e}_1^T \mathbf{e}_1} \\ \mathbf{e}_2 = \mathbf{w}_a - \mathbf{w}_2 & \quad |\mathbf{e}_2| = \sqrt{\mathbf{e}_2^T \mathbf{e}_2} \\ \vdots & \quad \vdots \\ \mathbf{e}_M = \mathbf{w}_a - \mathbf{w}_M & \quad |\mathbf{e}_M| = \sqrt{\mathbf{e}_M^T \mathbf{e}_M} \end{aligned}$$

$$\mathbf{E} = [|\mathbf{e}_1| \quad |\mathbf{e}_2| \quad |\mathbf{e}_3| \quad \dots \quad |\mathbf{e}_M|]$$

$$\text{MIN}(\mathbf{E}) = \mathbf{e}_\alpha$$

$$\Gamma_{\text{PROY}} \Rightarrow \mathbf{c}$$

α Puntero de la imagen más semejante

EC 3-13 Procedimiento para encontrar la imagen más semejante al CE.

Entonces con (EC 3-13) se calculan todas las distancias euclidianas formándose el siguiente vector de distancias euclidianas E .

α	E	Orden	α	E	Orden
01	3725	5	10	3442	4
02	5871	17	11	3743	6
03	5569	14	12	4433	9
04	6245	18	13	4531	10
05	5710	16	14	4421	8
06	5569	15	15	4590	11
07	2878	2	16	5055	16
08	2108	1 ^{er}	17	5299	13
09	3365	3	18	4042	7

TB 3-1 Vector de distancias euclidianas E , estableciendo la imagen Γ_8 ($\alpha=8$) del CE (figura 15) más semejante a Γ_{PROY} (figura 21) y Γ_4 como la menos semejante, las cantidades fueron redondeadas.

Por tener la menor distancia euclidiana, la imagen del CE Γ_8 ($\alpha=8$) (figura 15) es el más semejante a Γ_a . (Figura 19) Le siguen en semejanza las imágenes Γ_7 y Γ_9 (Γ_8 , Γ_7 y Γ_9 son imágenes de la misma persona). Ahora si la distancia euclidiana no excede del valor umbral predefinido θ entonces la persona se considerará como identificada por el Sistema de Reconocimientos de Rostros.

3.3.4.3. Definición de reglas para el reconocimiento

Es complejo definir un umbral, sobre todo viendo que una misma persona puede cambiar radicalmente con un maquillaje bien elaborado puede transformar a cualquier persona.

Si la capacidad de transformación del aspecto de una persona es sorprendente, no es menos sorprendente el grado de semejanza que pueden alcanzar dos personas, por ejemplo en el caso de los gemelos idénticos.

Es diferente el esfuerzo que las personas toman para reconocer a las personas. La gente cuando convive con alguien empieza a reconocer a la persona en todas sus facetas y el conocimiento tiene una plasticidad suficiente como para seguir considerando como un mismo. También con el tiempo se detectan sutilezas para considerar como individuos

diferentes a las personas. En ocasiones las sutilezas son tan poco obvias para los demás como el tono de voz o hasta el temperamento.

Así los patrones de reconocimiento de rostros pueden solo ofrecer grados de semejanza, ya que tienen una problemática que no se presenta en los sistemas de reconocimiento de huellas digitales o el iris. La probabilidad de que las huellas digitales de dos personas sean idénticas es mínima, inclusive en gemelos idénticos no se presenta tal eventualidad. Como en otros países del mundo, las huellas digitales son reconocidas legalmente como sustituto de la firma escrita, indispensable para imponer obligación en un contrato o documento, en los casos en que la persona involucrada no pueda firmar.

3.3.4.4. Construcción del software

Se presenta el software de patrones de reconocimiento facial mediante redes neuronales.

Para poder ejecutar el programa de reconocimiento facial han de cumplirse ciertos requisitos técnicos, tanto a nivel de hardware como de software.

En la Figura se puede ver la interfaz de usuario de bienvenida.



Figura 22: Interfaz de bienvenida

Para el usuario del sistema de reconocimiento facial

Fuente: Elaboración propia

El funcionamiento del sistema de reconocimiento facial diseñado es bastante intuitivo, sobre todo en lo referente a la interfaz gráfica se ha creado de forma simple con la ayuda de visual.Net. Para ingresar al sistema se debe apretar el botón "INGRESAR", pero previamente se debe tener un nombre de usuario y una clave de acceso dada por el administrador del sistema de reconocimiento facial.



Figura 23: Interfaz de acceso de usuario al sistema de reconocimiento

Fuente: Elaboración propia

Una vez superada la interfaz de seguridad, se puede comenzar hacer uso del software. La interfaz de usuario para el software es mostrada en la siguiente figura donde se pueden ver las diferentes opciones.



Figura 24: Interfaz sistema de reconocimiento

Fuente: Elaboración propia

Nuevo Usuario

Si se elige la opción “Nuevo Usuario”, el sistema muestra por pantalla una interfaz en la imagen que se muestra a continuación se puede ver como la interfaz consta de varios botones como ser: “Agregar imagen”, “negativo”, “escala de grises”, “Contracción de Histograma”, “Igualación de Histograma”, “Operador Logarítmico”, “Detección de rostro” y “Cancelar” y espacios en los que aparecen las imágenes.

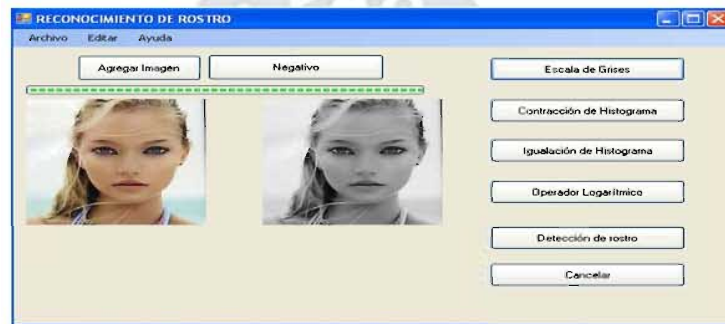


Figura 25: Interfaz detección de nuevo usuario

Fuente: Elaboración propia

Una vez que la cara ha sido detectada, el proceso de “Nuevo Usuario” despliega un mensaje “Cara detectada por favor introduzca datos de nuevo usuario” se elige aceptar y se despliega la siguiente pantalla para introducir datos personales.



Figura 26: Interfaz llena de datos personales de nuevo usuario

Fuente: Elaboración propia

Reconocimiento

Si se pulsa el botón de “Reconocimiento” se accede a la interfaz para introducir la imagen que va a ser reconocida, pasa por el proceso de detección de rostro, si existe un cara en la imagen el usuario debe pulsar el botón “Buscar”. Al hacerlo se pone en marcha el sistema de reconocimiento, si el proceso es satisfactorio, en el recuadro derecho aparece la imagen de la base de datos, si no el sistema devuelve un mensaje como resultado.

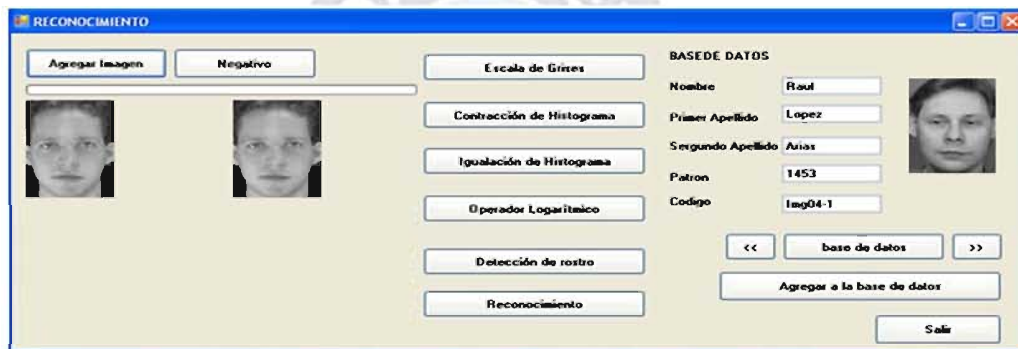


Figura 27: Interfaz sistema de reconocimiento

Fuente: Elaboración propia



Figura 28: Interfaz aprendizaje del sistema de reconocimiento facial

Fuente: Elaboración propia

El software permite ingresar los parámetros del aprendizaje y reconocimiento.

3.4. Evaluación De Resultados

Se analizan los resultados obtenidos en la evaluación del sistema tanto en la fase de pre procesamiento de imágenes, detección y reconocimiento.

Para la fase de detección se analizan los resultados que se obtienen cuando se utiliza solo la parte de selección de regiones candidatas a caras, y cuando se le añade el método alternativo de selección de caras dentro de la imagen

Para la fase de reconocimiento se realizan varias pruebas en las que se variaran algunos parámetros del sistema, para intentar llegar a tener una configuración óptima

3.4.1. Fase de detección

Se detalla reconocimiento de caras que es dedicada a la detección del rostro dentro de una imagen.

Para analizar los resultados obtenidos se impone la condición de que para que se considere que una cara ha sido detectada correctamente deben aparecer los dos ojos y la boca.

En la fase de detección se evalúan dos opciones. Una basada en la selección de regiones de piel candidatas a ser caras y su validación mediante mapas de ojos y boca.

3.4.1.1. Prueba de detección con selección de candidatos

Realizando la prueba de detección se observa que puede haber varias situaciones. Puede ocurrir que se detecte correctamente la cara dentro de la imagen, también puede ser que se devuelva como cara parte de la imagen en la que aparece la cara pero no la cara completa o que se devuelva como cara parte de la imagen que no lo es. Por último, puede que no se detecte ninguna cara dentro de la imagen.

Para tratar de mejorar los resultados obtenidos, se ha decidido mantener la detección mediante la selección de candidatos y validación por mapas de ojos y boca ya que proporciona buenos resultados. Además el método de detección es más general y puede aplicarse a imágenes en las que aparecen varias caras.

Debido a que todas las imágenes contienen una cara, por tanto es fácil que se cumpla el requisito mínimo de encontrar una región de piel de área mayor y considerando los ejes de la elipse mayores se devuelve una mayor región de la imagen como resultado.

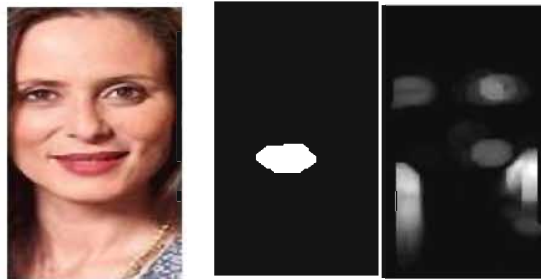


Figura 29: Imágenes binaria detección de boca y ojos

Fuente: Elaboración propia



Figura 30: Imagen binaria de pixeles de piel

Fuente: Elaboración propia

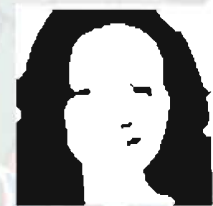


Figura 31: Imágenes binarias tras la apertura de radio

Fuente: Elaboración propia

3.4.2. Fase de reconocimiento

Se ha procurado explicar en detalle el método de Análisis de Componentes Principales (ACP) en su implementación en un Sistema para el Reconocimiento de Rostros. En lugar de explicar el código, por las reglas de sintaxis y semántica propios del lenguaje de programación se prefirió describir matemáticamente el funcionamiento del ACP y se mostró su operación con ejemplos.

Para validar la hipótesis de la tesis, se implementó el ACP en un programa escrito en el lenguaje de programación de Visual.net. Resultados experimentales se hicieron según a un conjunto de imágenes que se encuentran ampliamente disponibles en Internet, la base de datos de imágenes tomadas por el Laboratorio de Investigaciones de la *Olivetti Research Laboratory*.

3.4.2.1. Pruebas de identificación de rostros

Para validar el sistema implementado se hacen una serie de pruebas. Por ejemplo del conjunto de imágenes se escogen 3 imágenes de cada persona.

Ahora considérese que se busca proyectar cada imagen al Espacio de Rostros dado el CE. Es natural que la distancia Euclidiana de las imágenes que son ya miembros del CE será de cero. Para la siguiente tabla, se muestran los resultados de proyectar las imágenes al Espacio de Rostros,

En la tercera columna, indica la distancia euclidiana de la imagen más semejante del CE a la imagen de entrada. En todas las pruebas sólo una dio un resultado negativo, la imagen Img40-2.

Imagen de entrada	Imagen más semejante	Distancia euclidiana
Img01-1	Img01-1	0
Img01-2	Img01-3	3352
Img01-3	Img01-1	3225
Img02-1	Img02-1	0
Img02-2	Img02-1	2111

Img02-3	Img02-2	2677
Img03-3	Img03-1	3542
Img04-1	Img04-2	3522
Img04-2	Img04-1	2453
Img04-3	Img04-3	0
.....
Img40-2	Img05-1	2594

En la prueba de la imagen img40-2 da un resultado negativo el sistema al determinar que img05-1 es la más similar.

Resultados son diferentes si se aumenta o disminuyen el número de imágenes por persona para el CE.



CONCLUSIONES Y RECOMENDACIONES

CAPITULO IV

4.1. Conclusiones

Gracias al constante avance de la tecnología, día a día las labores que tradicionalmente es realizada por seres humanos ahora son realizadas por sistemas automatizados. Una de las actividades que pueden automatizarse y que ha cobrado gran importancia, es la capacidad de establecer la identidad de los individuos.

La motivación general que ha guiado la tesis es profundizar el estudio de las técnicas que permiten realizar un reconocimiento automático de caras e implementar un sistema.

Se ha desarrollado un Sistema de Reconocimiento de Caras que resuelve la detección de la cara dentro de la imagen, la extracción de características y finalmente el reconocimiento del individuo.

En la tesis se han utilizado imágenes estáticas en blanco y negro y a color. Se ha trabajado con las imágenes en el espacio de color RGB (aunque también se utiliza el YcbCr para la creación de mapas y ojos) y con su transformación en escala de grises. Se pretende conseguir una mayor información de las imágenes para poder facilitar el reconocimiento.

En primer lugar las imágenes son sometidas a un preprocesado que trata de mejorar sus características para facilitar así la detección de la cara dentro de la imagen.

A continuación se produce la detección de la cara dentro de la imagen. La fase es determinante a la hora del correcto funcionamiento del sistema, ya que una correcta detección favorece un mejor reconocimiento. Por eso se ha comprobado que en la fase no sólo son importantes el número de imágenes en las que se detecta una cara, sino que también importa la calidad de esa detección.

Los resultados que se obtienen para la fase de detección dependen en gran medida de la correcta detección de los píxeles de piel en la imagen y del proceso de filtrado y agrupamiento posterior. La detección de los píxeles es bastante buena. En cuanto al

filtrado y agrupamiento los mejores resultados se obtienen para distintos parámetros de los filtros morfológicos que se utilizan.

Para la fase de reconocimiento se ha propuesto la extracción de las características que representan las imágenes mediante PCA. El número de autovectores que representan las imágenes viene dado por el número de autovalores. El número no tiene por qué ser el mismo para las distintas pruebas realizadas

Las características extraídas de las imágenes de entrenamiento sirven para entrenar las redes neuronales que forman parte de la etapa de clasificación.

En la tesis se ha comprobado como el tipo de imágenes que se utilizan en el reconocimiento influyen de manera determinante en los resultados que se obtienen, y en la elección de determinados parámetros.

Observando los resultados finales de porcentaje de reconocimiento se ve como el funcionamiento del sistema implementado es bueno sólo cuando las imágenes que se utilizan son tomadas en un ambiente controlado.

4.2. Recomendaciones

Para mejorar el reconocimiento facial mediante redes neuronales utilizando el Análisis de Componentes Principales (ACP) se requiere lo siguiente:

- Combinar con otros sistemas de reconocimiento como huellas digitales o de iris y que dichos sistemas sirvan de confirmación de resultados cuando se tenga un alto grado de certidumbre.
- Implementar herramientas que permitan una aproximación en menor tiempo a los autovectores y autovalores correspondientes a una Matriz de Covarianzas con mayores dimensiones, sin que exija mayores recursos de computacionales.
- Para fortalecer el sistema de reconocimiento, se recomienda que se establezcan controles cuando se dé el caso en que autovectores o autovalores tengan valores casi idénticos. Aunque la repetición de autovectores significaría una imagen duplicada en el CE, autovectores casi idénticos podría darse en

gemelos, generando una incertidumbre que solo sería resuelta con información de sistemas de reconocimiento auxiliares.

BIBLIOGRAFÍA

- [AGC.2006] Anguiano Guerrero Carlos "Análisis de características principales para el Reconocimiento de rostros En imágenes bidimensionales Por computadora" 2006.
- [BDB. 2000] *Kyungim Baek*, Bruce A. Draper, J. Ross Beveridge & Kai She, "PCA vs. ICA: A comparison on the FERET data set", 2000.
- [CAS. 1996] K. Castleman, "Digital Image Processing, 2" Edition, Prentice-Hall, 1996.
- [CHL. 1995] R. Chellapa, C. L. Wilson & Sirohey, "Human and Machine Recognition of Faces: A Survey", *Proceedings of the IEEE*. Vol. 83, n°5, pp.704-740, Mayo 1995.
- [FAL. 2008] Luis Alberto Flores Armenta. "Red neuronal artificial". Instituto Tecnológico de Nogales, Área de Ing. En Sistemas Computacionales; 2008.
- [GJC. 2009] Gómez Jiménez Carmen Virginia "Diseño y Desarrollo de un Sistema de Reconocimiento de Caras" Madrid, Abril 2009
- [GOR. 1991] G. Gordon, "Face recognition based on depth maps and surface curvature", in *SPIE P Geometric Methods in Computer Vision*, vol.1570, pp. 234-247, 1991.
- [HRO. 1997] HERNÁNDEZ SAMPIERI, ROBERTO & OTROS. "Metodología de la investigación". 2da ed. México DF: McGraw Hill. 1997.
- [HUA. 1996] Thomas S. Huang & Li-an Tang, "3-D Face Modeling and Its Applications", *International Journal fo Pattern Recognition and Artificial Intelligence*, Vol. 10, N°5, pp. 491-520, 1996.
- [JPF. 2003] Jure Kovac, Peter Peer & Frank Solina, "Human Skin Colour Clustering for Face Detection" *Faculty of Computer and Information Science, University of juliana*; 2003.
- [KIR. 1990] M. Kirby & L. Sirovich, "Application of the Karhunen-Loève Procedure for the Characterization of Human Faces", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, pp. 103-108, 1990.

- [LIU. 2000] Liu Chengjun & Wechsler Harry "Evolutionary Pursuit and Its Application to Face Recognition". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2000, pp. 670-682.
- [MAN. 1992] B. S. Manjunath, R. Chellappa & C. vonder Malsburg, 'A feature based approach to Face Recognition', in *Proc. IEEE Computer Soc. Conf. on Computer Vision and Pattern Recognition*, pp. 373-378, 1992.
- [MDA. 2004] Moreno Díaz Ana Belén "Reconocimiento Facial Automático mediante Técnicas de Visión Tridimensional" Madrid, 2004.
- [NAR. 1999] Nasanen, Risto. "Spatial Frequency Bandwidth Used in the recognition of Facial Images". Helsinki. Finland. 1999. pp. 3824-3833.
- [PEN. 1994] A. Pentland, B. Moghaddam & T. Stamer, "View-based and Modular Eigenspaces for Face Recognition", *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, 1994.
- [PNB. 1998] Purnell, D W., Nieuwoudt, C y Botha, E. C."Automatic Face Recognition in a Heterogeneous Population". 1998. *Pattern Recognition Letters* 19. pp.
- [REP. 2002] Robert Epstein "Fingerprints meet Doubert. The myth of fingerprint 'science' is Revealed" *Southern California Law Review*. Vol .75:605. 2002.
- [TRU. 1998] Emanuele Truoco & Alessandro Verri, "Introductory Techniques for 3-D Computer Vision", *Prentice-Hall*, 1998.
- [VAL. 1994] D. Valentín, H. Abdi, A. J. O'Toole & G.W. Cottrell, "Connectionist Models of face processing: a survey", *Pattern Recognition*, Vol. 27, N° 9, pp. 1209-1230, 1994.
- [VGH. 1999] Vargas Granzow Hector. "Introducción a la biometría y algunas aplicaciones".1999. 2da. Ed.
- [ZHA. 1999] W. Zhao, "Improving the Robustness of Face Recognition", *Proc. Intl. Conf. on Audio and Video Based Person Authentication*, pp. 78-83, 1999.
- [ZHA. 2000] W. Zhao, "Face Recognition: A Literature Survey", *UMD-CFAR, Technical Report CAR-TR-948*, 2000.

WEBGRAFIA

- <http://es.wikipedia.org/wiki/MySQL>
Fecha y hora de acceso: 25/06/2011 21:50
- http://es.wikipedia.org/wiki/visual_basico
Fecha y hora de acceso: 25/06/2011 21:10
- <http://www.chuidiang.com/ood/metodologia/scrum.php>
Fecha y hora de acceso: 03/04/2011 10:30
- <http://www.buscapdf.com/buscar.php>
Fecha y hora de acceso: 03/04/2011 11:30
- <http://www.librosintinta.com/busca/Reconocimiento+facial/pdf/>
Fecha y hora de acceso: 05/05/2011 09:20
- <http://www.manuales-pdf.com/reconocimiento-facial/2/>
Fecha y hora de acceso: 06/05/2011 10:40
- <http://www.omitsis.com/scrum-como-metodologia-de-desarrollo>
Fecha y hora de acceso: 20/05/2011 11:25
- <http://www.tecnociencia.es/monograficos/biometria/biometria2.html>
Fecha y hora de acceso: 20/05/2011 12:40



ANEXOS

Breve Repaso de Álgebra Lineal

El propósito del anexo es dar un contexto general, que sea suficiente para entender términos de álgebra lineal que se menciona en la tesis.

Álgebra Lineal

Es una rama de las matemáticas que estudia los sistemas de ecuaciones lineales, transformaciones lineales, vectores, espacios vectoriales.

Sistema de Ecuaciones

Es un conjunto de ecuaciones cuyas soluciones comunes se pretende hallar. Las ecuaciones de un sistema suelen tener dos o más incógnitas, por lo que cada una de ellas puede tener infinitas soluciones.

Matriz

Es una tabla de números. Una de las principales aplicaciones de las matrices es la representación de sistemas de ecuaciones con varias incógnitas. Cada fila de la matriz representa una ecuación, siendo los valores de una fila los coeficientes de las distintas variables de la ecuación, en determinado orden. Por ejemplo, para representar en una matriz un sistema de ecuaciones.

$$\left. \begin{array}{l} 2x + 4y + 6z = 18 \\ 4x + 5y + 6z = 24 \\ 3x + z - 2z = 4 \end{array} \right\} = \left[\begin{array}{ccc|c} 2 & 4 & 6 & 18 \\ 4 & 5 & 6 & 24 \\ 3 & 1 & -2 & 4 \end{array} \right]$$

Ejemplo de sistema de ecuaciones lineales.

Los elementos de una matriz general, de tamaño $m \times n$, m renglones, n columnas. Se representa normalmente utilizando doble subíndice; el primer subíndice, i , indica el

número de fila y el segundo, j , el número de columna. Así el elemento a_{23} pertenece a la segunda fila y tercera columna. La matriz general se delimita generalmente entre corchetes o entre paréntesis.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{1n} \\ a_{21} & a_{22} & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{mn} \end{bmatrix}$$

Definición de una matriz.

El tamaño de una matriz es dado por el número de filas y el de columnas. Si $m = n$, es una matriz cuadrada y el número de filas (o columnas) es del mismo orden. Dos matrices $A = (a_{ij})$ y $B = (b_{ij})$, son matrices iguales, si y sólo si son de igual tamaño y si para todo i e j , $a_{ij} = b_{ij}$. La matriz transpuesta de una matriz A es otra matriz A^T en la cual en cada elemento a_{ij} de A , su fila i será la columna i de A^T , y la columna j será la fila j de A^T . Una matriz cuadrada es una matriz simétrica si es igual a su transpuesta.

Una matriz diagonal es una matriz cuadrada donde todos los elementos son 0 salvo los elementos de la diagonal principal. La matriz unidad o matriz de identidad I_m , es una matriz cuadrada de orden m en la cual todos los elementos son cero excepto los de la diagonal principal, que es 1. El orden de la matriz de identidad se puede omitir si se sobrentiende con el resto de la expresión, con lo que I_m se escribe simplemente I .

La adición y la multiplicación de matrices son definidas de manera que ciertos conjuntos de matrices forman sistemas algebraicos. Considérese los elementos de las matrices con números reales cualesquiera. La suma de dos matrices sólo es definida si ambas tienen el mismo tamaño. Si $A = (a_{ij})$ y $B = (b_{ij})$ tienen igual tamaño, entonces la suma $C = A + B$ se define como la matriz (c_{ij}) , en la que $c_{ij} = a_{ij} + b_{ij}$, es decir, para sumar dos matrices de igual tamaño basta con sumar los elementos correspondientes.

En el conjunto de todas las matrices de un determinado tamaño, la adición tiene las propiedades asociativa y conmutativa. Además hay una matriz única 0 tal que para

cualquier matriz A , se cumple $A + 0 = 0 + A = A$ y una matriz única B tal que $A + B = B + A = 0$.

En el producto de dos matrices AB , A y B , es definido sólo si el número de columnas del factor izquierdo, A , es igual al número de filas del factor derecho, B ; si $A = (a_{ij})$ es de tamaño $m \times n$ y $B = (b_{jk})$ es de tamaño $n \times p$, el producto $AB = C = (c_{ik})$ es de tamaño $m \times p$, y c_{ik} está dado por

$$c_{ik} = \sum_{j=1}^n a_{ij} b_{jk}$$

Ecuación para el producto de dos matrices.

La determinante de una matriz es una función $\det(A)$ que asocia un escalar denotado como $|A|$ a una matriz cuadrada A . El significado geométrico de un determinante es un factor escalar. Para conocer, por ejemplo, la determinante de una matriz de tamaño 2×2 es igual $a_{11}a_{22} - a_{21}a_{12}$. De forma general, para encontrar el determinante de una matriz de tamaño $n \times n$ se hace por medio de un desarrollo de cofactores a lo largo del primer renglón.

$$\text{Det}(A) = \sum_{j=1}^n a_{1j} C_{1j}$$

Determinante de una matriz.

$$C_{ij} = (-1)^{i+j} \det(A_{ij})$$

Definición de un cofactor.

Una matriz es inversa de otra matriz, si el producto entre las matrices es la matriz identidad. Es decir, la matriz A^{-1} es una matriz inversa de la matriz A , si $A^{-1}A = AA^{-1} = I$.

Vectores Linealmente Independientes

Dos vectores \mathbf{u} y \mathbf{v} de un mismo espacio vectorial son vectores independientes si no son proporcionales, es decir si uno de ellos no es un múltiplo del otro. Para cualquier escalar k , se tiene $\mathbf{u} \neq k\mathbf{v}$. Es equivalente decir que el sistema (\mathbf{u}, \mathbf{v}) compuesto por los dos vectores es libre. Geométricamente, dos vectores son independientes si no tienen la misma dirección (con sentidos idénticos u opuestos). La definición supone que el vector nulo tiene todas las direcciones.

Tres vectores son independientes si y sólo si no están contenidos en el mismo plano vectorial, o sea si ninguno de ellos es una combinación lineal de los otros dos (en cuyo caso estaría en el plano generado por estos vectores).

La noción se generaliza a un número cualquiera de vectores. Los n vectores son independientes si ninguno de ellos es una combinación lineal de los demás.

Cálculo de los Autovectores y Autovalores de la Matriz

Conceptos de autovectores y autovalores de una matriz. Sea una matriz de $n \times n$ con elementos reales o complejos. El número λ (real o complejo) recibe el nombre de autovalor de \mathbf{A} si existe algún vector diferente de cero \mathbf{x} tal que:

$$\mathbf{Ax} = \mathbf{x} \lambda, \quad \mathbf{x} \neq \mathbf{0}$$

El vector \mathbf{x} es un autovector de \mathbf{A} correspondiente al autovalor λ . Los autovectores también se conocen como vectores propios, vectores característicos o eigenvectores. Los autovalores se conocen como valores propios, valores característicos, valores latentes o eigenvalores. Los autovalores de \mathbf{A} serán sí y solo sí.

$$p(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I}) = 0$$

A la ecuación se conoce como ecuación característica de **A**. Además a $p(\lambda)$ se le llama polinomio característico de **A**. Cuando se desarrolla la ecuación característica se obtiene el polinomio característico.

Concepto de polinomio característico en una matriz de 2 x 2.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad A - \lambda I = \begin{bmatrix} a & b \\ c & d \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix}$$

$$p(\lambda) = \det(A - \lambda I) = (a - \lambda)(d - \lambda) - bc = \lambda^2 - (a + d)\lambda + (ad - bc)$$

Polinomio característico para una matriz de 2 X 2.

Si **A** es de $n \times n$, su polinomio característico será de grado n . Un polinomio de grado n con coeficientes reales o complejos tiene como máximo n raíces distintas. Si se aplica al polinomio característico, se ve que una matriz de $n \times n$ con entradas reales o complejas tiene a lo más n distintos autovalores. Por ejemplo, para calcular los autovectores y autovalores de la siguiente matriz.

$$A = \begin{bmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{bmatrix}$$

$$\det(A - \lambda I) = \begin{vmatrix} 1 - \lambda & -1 & 4 \\ 3 & 2 - \lambda & -1 \\ 2 & 1 & -1 - \lambda \end{vmatrix} = -(\lambda - 1)(\lambda + 2)(\lambda - 3) = 0$$

$$\text{donde } \lambda_1 = 1; \quad \lambda_2 = -2; \quad \lambda_3 = 3$$

Con los autovalores ya conocidos se obtendrá sus autovectores correspondientes.

$$(A - I)v = \begin{pmatrix} 0 & -1 & 4 \\ 3 & 1 & -1 \\ 2 & 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad v_1 = \begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix}$$

$$(A - (-2I))v = \begin{pmatrix} 3 & -1 & 4 \\ 3 & 4 & -1 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$$

$$(A - 3I)v = \begin{pmatrix} -2 & -1 & 4 \\ 3 & -1 & -1 \\ 2 & 1 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad v_3 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

Obtención de los autovectores de A por sus autovalores

Si en sus n autovectores son todos ellos distintos entonces estos vectores son linealmente independientes.

Acrónimos

ACP

El Análisis de Componentes Principales (ACP) es una técnica estadística de síntesis de la información, o reducción de la dimensión (número de variables). Es decir, ante un banco de datos con muchas variables, el objetivo será reducirlas a un menor número perdiendo la menor cantidad de información posible.

BMP

Los archivos con extensión .BMP, en los sistemas operativos Windows, representan la sigla BitMaP, que significa mapa de bits. Los archivos de mapas de bits se componen de direcciones asociadas a códigos de color, uno para cada cuadro en una matriz de píxeles.

CE

Conjunto de Entrenamiento. Conjunto de Aprendizaje.

GIF

Graphics Interchange Format, es un formato gráfico utilizado ampliamente en la World Wide Web, tanto para imágenes como para animaciones.

JPEG

Joint Photographic Experts Group. Es un algoritmo diseñado para comprimir imágenes con 24 bits de profundidad o en escala de grises.

ORL

Base de datos de *Olivetti Research Lab*, la cual contiene varias imágenes diferentes de 40 sujetos distintos.

PÍXEL

El píxel (del inglés *picture element*, o, "elemento de la imagen") es la menor unidad en la que se descompone una imagen digital, ya sea una fotografía, un fotograma de video o un gráfico.

RNA

En inteligencia artificial, referidas habitualmente de forma más sencilla como redes de neuronas o redes neuronales, las redes de neuronas artificiales (RNA) son un paradigma de aprendizaje y procesamiento automático inspirado en la forma en que funciona el sistema nervioso de los animales.

