

**ANÁLISIS Y DISEÑO DE TÉCNICAS DE PROCESAMIENTO DIGITAL DE  
IMÁGENES BASADO EN COLOR E IMPLEMENTADAS EN UN SISTEMA DE  
RECONOCIMIENTO FACIAL**

**BEATRIZ OMAIRA PEDRAZA PICO**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍA DE FÍSICO MECÁNICAS  
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA  
BUCARAMANGA**

**2011**

**ANÁLISIS Y DISEÑO DE TÉCNICAS DE PROCESAMIENTO DIGITAL DE  
IMÁGENES BASADO EN COLOR E IMPLEMENTADAS EN UN SISTEMA DE  
RECONOCIMIENTO FACIAL**

**BEATRIZ OMAIRA PEDRAZA PICO**

**Trabajo de grado para optar al título de  
Ingeniera de Sistemas**

**DIRECTOR**

**MPE. HENRY ARGUELLO FUENTES**

**CODIRECTOR**

**ING. NYDIA PAOLA RONDÓN VILLARREAL**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍA DE FÍSICO MÉCANICAS  
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA  
BUCARAMANGA**

**2011**

## **AGRADECIMIENTOS**

A Dios, por ser el guía de mi vida y por brindarme la fuerza de terminar este proyecto con éxito.

A mis padres, Rodrigo y Omaira por su infinito apoyo.

A mi hermana Laura por sus consejos.

A la Universidad Industrial de Santander, por los conocimientos adquiridos.

Al Msc Fabio Martinez, por compartir su valioso tiempo, sus conocimientos y experiencias para llevar a cabo con éxito este proyecto.

Al Ing. Darío Delgado, por su total disposición para responder cualquier pregunta y por su gran interés en el proyecto.

Al profesor Henry Arguello Fuentes, director del proyecto, por la confianza depositada y por compartir sus conocimientos.

A la Ing. Nydia Paola Rondón, codirectora del proyecto, por su apoyo durante el proyecto.

Al profesor Víctor Eduardo Martínez, por su amistad, su apoyo incondicional y total colaboración.

A mis compañeros del grupo de investigación Hugo, Lina, Andrés, Cindy y Leidy por convertir 8 largas horas de trabajo en 4 horas llenas de alegrías y de experiencias vividas.

A todos mis amigos, por acompañarme en este camino, por alegrar cada uno de los instantes vividos y compartidos. También quiero agradecer a todas las demás personas que de alguna manera fueron partícipes en la realización de este proyecto.

## CONTENIDO

INTRODUCCIÓN.....	13
1. OBJETIVOS .....	15
1.1 OBJETIVO GENERAL.....	15
1.2 OBJETIVOS ESPECÍFICOS .....	15
2. PLANTEAMIENTO DEL PROBLEMA.....	17
3. JUSTIFICACIÓN .....	19
4. ESTADO DEL ARTE .....	20
4.1 HISTORIA.....	20
5. MARCO TEÓRICO .....	26
5.1 FUNDAMENTOS DE LA VISIÓN COMPUTACIONAL .....	26
5.1.1 La Visión Humana.....	26
5.1.2 La Visión Artificial.....	29
5.1.3 Diferencia entre la visión humana y la visión artificial.....	30
5.2 FUNDAMENTOS DEL PROCESAMIENTO DIGITAL DE IMÁGENES .....	32
5.2.1 Imágenes Digitales.....	32
5.2.2 Modelos de Color .....	34
5.2.2.1 Modelo de color RGB .....	35
5.2.2.2 Modelo de color HSV .....	35
5.2.2.3 Conversión del modelo RGB a HSV.....	37
5.2.3 Etapas del Procesamiento Digital de Imágenes .....	38
5.2.3.1 Adquisición.....	39
5.2.3.2 Pre-Procesamiento.....	39
5.2.3.3 Reconocimiento.....	44
5.3 FUNDAMENTOS MATEMÁTICOS.....	46
5.3.1 Eigenvalores y Eigenvectores.....	46

5.3.2	Covarianza.....	46
5.3.2.1	Matriz de Covarianza.....	47
6.	METODOLOGÍA Y DESARROLLO.....	48
6.1	ETAPA I. Adquisición de la Imagen.....	48
6.2	ETAPA II. Detección de Rostros.....	50
6.3	ETAPA III. Pre-Procesamiento.....	52
6.4	ETAPA IV. Reconocimiento de Rostros.....	55
6.4.1	Metodología del PCA en el reconocimiento de rostros.....	57
6.5	ETAPA V: Identificación y Verificación.....	61
7.	RESULTADOS.....	65
8.	CONCLUSIONES Y RECOMENDACIONES.....	70
8.1	CONCLUSIONES.....	70
8.2	RECOMENDACIONES.....	71
9.	REFERENCIA BIBLIOGRAFICA.....	72
10.	BIBLIOGRAFÍA.....	78

## LISTA DE TABLAS

TABLA 1. CATEGORIZACIÓN DE TODAS LAS TÉCNICAS DE RECONOCIMIENTO FACIAL .....	22
TABLA 2 DIFERENCIA ENTRE LA VISIÓN HUMANA Y LA VISIÓN COMPUTACIONAL.....	31
TABLA 3 DESCRIPCIÓN DE LAS BASES DE DATOS IMPLEMENTADAS .....	49
TABLA 4 VECTOR DE DISTANCIA EUCLIDIANA Y DE MAHALANOBIS.....	63
TABLA 5 CLASIFICACIÓN DE LOS ROSTROS EMPLEANDO LA TÉCNICA DE ECUALIZACIÓN DEL HISTOGRAMA PARA LA BASE DE DATOS <i>FACE 94</i> .....	67
TABLA 6 CLASIFICACIÓN DE LOS ROSTROS EMPLEANDO LA TÉCNICA DE ECUALIZACIÓN DEL HISTOGRAMA PARA LA BASE DE DATOS <i>GEORGIA TECH</i> .....	67
TABLA 7 CLASIFICACIÓN DE ROSTROS EMPLEANDO LA PRIMERA TÉCNICA DE LA FIGURA 2167	
TABLA 8 CLASIFICACIÓN DE LOS ROSTROS EMPLEANDO LA SEGUNDA TÉCNICA DE LA FIGURA 21 PARA LA BASE DE DATOS <i>FACE 94</i> .....	68
TABLA 9 CLASIFICACIÓN DE LOS ROSTROS EMPLEANDO LA SEGUNDA TÉCNICA DE LA FIGURA 21 PARA LA BASE DE DATOS <i>GEORGIA TECH</i> .....	68
TABLA 10 CLASIFICACIÓN DE ROSTROS EMPLEANDO LA SEGUNDA TÉCNICA DE LA FIGURA 21. .....	68

## LISTA DE FIGURAS

FIGURA 1. SISTEMA DE RECONOCIMIENTO FACIAL 2D, BASADO EN EL MÉTODO DE EIGENFACES. ....	25
FIGURA 2 ESQUEMA DE FUNCIONAMIENTO DEL OJO HUMANO .....	27
FIGURA 3 EL OJO HUMANO, (A) MUESTRA LAS PARTES DEL OJO HUMANO, (B) MUESTRA LOS FOTORECEPTORES.....	28
FIGURA 4 ESPECTRO ELECTROMAGNÉTICO.....	29
FIGURA 5 PROCESOS DE LA VISIÓN ARTIFICIAL. ....	30
FIGURA 6 CLASES DE IMÁGENES.....	33
FIGURA 7 IMAGEN DIGITAL MUESTREADA Y CUANTIFICADA.....	34
FIGURA 8 ESPACIO DE COLOR DEL MODELO RGB .....	35
FIGURA 9 SATURACIÓN DE LOS COLORES.....	36
FIGURA 10. ESPACIO DE COLOR DEL MODELO HSV .....	37
FIGURA 11 ETAPAS FUNDAMENTALES DEL PROCESAMIENTO DE IMÁGENES.....	38
FIGURA 12 DIGITALIZACIÓN DE UNA IMAGEN.....	39
FIGURA 13. CLASIFICACIÓN DE LA MEJORA DE IMÁGENES.....	40
FIGURA 14. CONTRACCIÓN DE LA IMAGEN .....	41
FIGURA 15 EXPANSIÓN DE LA IMAGEN .....	42
FIGURA 16. ECUALIZACIÓN DE LA IMAGEN .....	43
FIGURA 17 ETAPAS DE UN SISTEMA DE RECONOCIMIENTO FACIAL BASADO EN COLOR. ....	48
FIGURA 18. ESTRUCTURA DEL SISTEMA CLASIFICADOR EN CASCADA .....	50

FIGURA 19 ESTRUCTURA DEL SISTEMA DE DETECCIÓN DE ROSTROS. ....	51
FIGURA 20. REPRESENTACIÓN DE LA IMAGEN INTEGRAL. EL CUADRO DE LA IZQUIERDA BOSQUEJA EL VALOR DE LA IMAGEN INTEGRAL EN EL PUNTO $(X, Y)$ Y EL DE LA DERECHA REPRESENTA LA SUMA DE LOS PÍXELES DENTRO DEL RECTÁNGULO $D$ , EL CUAL ES CALCULADO ASÍ $L4 + L1 - (L2 + L3)$ . ....	52
FIGURA 21 TÉCNICAS DE PRE-PROCESAMIENTO IMPLEMENTADAS. ....	53
FIGURA 22 REPRESENTACIÓN DE LA IMAGEN EN RGB Y SUS RESPECTIVOS PLANOS .....	54
FIGURA 23 REPRESENTACIÓN DE LA IMAGEN EN HSV Y SUS RESPECTIVOS PLANOS.....	54
FIGURA 24 REPRESENTACIÓN DE LA IMAGEN EN ESCALA DE GRISES.....	55
FIGURA 25 ESTRUCTURA DEL SISTEMA DE RECONOCIMIENTO DE ROSTROS. ....	56
FIGURA 26 CONJUNTO DE IMÁGENES DE CARAS DE ENTRENAMIENTO. ....	58
FIGURA 27 ROSTRO PROMEDIO, SEGÚN DEL CONJUNTO DE IMÁGENES DE CARAS DE ENTRENAMIENTO.....	58
FIGURA 28 IMÁGENES DE CARAS AJUSTADAS A LA MEDIA .....	59
FIGURA 29 EIGENFACES OBTENIDOS A PARTIR DEL PRODUCTO ESCALAR DE LOS $uk$ DE LA ECUACIÓN 1.27 Y LOS $\phi$ DE LA ECUACIÓN 1.23 .....	60
FIGURA 30 REPRESENTACIÓN DE LA MATRIZ DE PESOS.....	61

## RESUMEN

**TÍTULO:** ANÁLISIS Y DISEÑO DE TÉCNICAS DE PROCESAMIENTO DIGITAL DE IMÁGENES BASADO EN COLOR E IMPLEMENTADAS EN UN SISTEMA DE RECONOCIMIENTO FACIAL.\*

**AUTOR:** Beatriz Omaira Pedraza Pico. \*\*

**PALABRAS CLAVES:** Análisis de Componentes Principales (PCA), Eigenfaces, AdaBoost, Distancia Euclidiana y de Mahalanobis, Matriz de Covarianza, Procesamiento Digital de Imágenes (PDI).

**DESCRIPCIÓN:** En el presente trabajo se implementó un sistema de reconocimiento de rostros con imágenes bidimensionales (2D) en color basado en el modelo de *Turk y Pentland*, con el fin de determinar la importancia del color en este sistema, por medio de una comparación entre imágenes en escala de grises e imágenes en color.

En la primera parte se describe un módulo de detección de rostros que es capaz de procesar imágenes de forma rápida, ya que involucra varias técnicas de pre-procesamiento en las imágenes detectadas, para el proceso de detección se hace uso de la “imagen integral”, de un clasificador simple, el cual emplea el algoritmo de aprendizaje AdaBoost y una combinación de clasificadores en cascada, obteniendo de esta manera altas tasas de detección y mayor rechazo en imágenes de no caras.

La segunda parte se encarga de reconocer individuos captados en una posición frontal, aplicando un enfoque de *eigenfaces*, para comparar las características principales de los rostros con individuos conocidos contra los rostros con individuos desconocidos. Este sistema propuesto provee la capacidad de reconocer e identificar de manera no supervisada nuevas imágenes de rostros utilizando un sistema clasificador para la identificación del individuo usando la distancia euclidiana y de mahalanobis.

La investigación ha sido realizada con 2 bases de datos internacionalmente reconocidas como lo son la base de datos “Georgia Tech Face Database” y la base de datos “Face 94” las cuales comprenden imágenes con variación de iluminación, expresiones faciales y fondos.

---

\* Trabajo de investigación

\*\* Facultad de Ingeniería Físico Mecánicas. Escuela de Ingeniería de Sistemas e Informática.  
Director: Henry Arguello Fuentes. Codirector: Nydia Paola Rondón Villarreal.

## SUMMARY

**TÍTULO:** TECHNICAL ANALYSIS AND DESIGN OF DIGITAL IMAGE PROCESSING BASED ON COLOR AND IMPLEMENTED IN A FACE RECOGNITION SYSTEM.\*

**AUTHOR:** Beatriz Omaira Pedraza Pico. \*\*

**KEYWORDS:** Principal Component Analysis (PCA), Eigen faces, AdaBoost, Euclidean and Mahalanobis distance, Covariance Matrix, Digital Processing of Images (PDI).

**DESCRIPTION:** In this paper, we implemented a facial recognition system with two-dimensional color images (2D) based on the model of Turk and Pentland, in order to determine the importance of color in this system, through a comparison between gray scale images and color images.

The first part describes a face detection module is able to process images quickly, as it involves several pre-processing techniques on the images detected, for the detection process makes use of the "integral image" of a simple classifier, which uses the AdaBoost learning algorithm and a combination of classifiers in cascade, thereby obtaining high detection rates and more rejection on the images of no faces.

The second part is responsible for recognizing individuals captured in a frontal position, using Eigen faces approach to compare the main features of the faces with known individuals against their faces with unknown individuals. This proposed system provides the ability to recognize and identify non-supervised new images of faces using a filing system for the identification of individuals using the Euclidean and Mahalanobis distance.

The research was conducted with two internationally recognized databases such as database "Georgia Tech Face Database," and the database "Face 94" which include images with varying lighting, facial expressions and backgrounds.

---

\* Researching Work

\*\* Physical-Mechanics Engineering Faculty. Informatics and Systems Engineering School.  
Director: Henry Arguello Fuentes. Co-director: Nydia Paola Rondón Villarreal.

## INTRODUCCIÓN

La detección y reconocimiento del rostro son técnicas de gran relevancia en los últimos años, ya que han sido aplicadas en una amplia variedad de problemas, incluyendo: la identificación biométrica para clasificación de individuos [1], sistemas de seguridad para el reconocimiento y procesamiento digital de imágenes y video para mejorar la calidad y facilitar la búsqueda de información, entre otras.

Debido a que existe gran cantidad de investigaciones y desarrollos algorítmicos para solucionar problemas de detección y reconocimiento facial y que la mayoría de estas investigaciones se han centrado en el uso de imágenes en escala de grises. Este trabajo involucra imágenes en color e implementa un esquema para la detección y reconocimiento facial robusto, eficaz y versátil; con el fin de verificar que el color juega un papel importante y atractivo para el estudio del rostro humano [2].

Además, para la detección de rostros se utilizó el modelo de *Viola Jones* [3], más conocido como el algoritmo de aprendizaje de AdaBoost [4], el cual consiste en detectar y localizar la posición de un número finito de caras en una imagen captada, teniendo en cuenta que el individuo debe ubicarse de forma frontal y la imagen puede diferir en sus colores, expresiones, condiciones de iluminación o resolución. Este modelo se considera especialmente exitoso para cualquier aplicación donde se necesite detectar objetos con variabilidad estructural sin tener que limitarse a sólo caras.

Sin embargo, la etapa de reconocimiento facial se fundamentó en el enfoque de *Eigenfaces* [5], también conocido como PCA<sup>1</sup>, el cual consiste en descomponer un grupo de imágenes de rostros en un pequeño conjunto de vectores denominado *Eigenfaces*, que poseen la información principal del

---

<sup>1</sup> PCA: *Principal Analysis Component*. En español (*Análisis de Componentes Principales*).

grupo inicial de caras de entrenamiento. De esta manera, se logra cambiar el espacio original de la imagen, por otro espacio generado a partir de las imágenes originales haciendo uso de la transformada de *Karhunen-Loeve* también conocida como PCA.

## **1. OBJETIVOS**

### **1.1 OBJETIVO GENERAL**

Implementar técnicas de procesamiento digital de imágenes basadas en color sobre un sistema de reconocimiento de rostros en tiempo real mediante el uso del algoritmo de AdaBoost y de Análisis de Componentes Principales para detectar y reconocer un rostro.

### **1.2 OBJETIVOS ESPECÍFICOS**

- Determinar el estado del arte de las principales técnicas requeridas para procesar imágenes en color a través del Algoritmo de Eigenfaces más conocido como el método de Análisis de Componentes Principales (PCA) para llevar a cabo el reconocimiento de rostros.
- Implementar un sistema clasificador que permita determinar si una imagen dada corresponde a una cara o no.
- Implementar un modelo de detección de caras usando el algoritmo de AdaBoost que utilice para su entrenamiento datos tomados con una cámara digital, con el fin de detectar el rostro humano.
- Estudiar e implementar técnicas de procesamiento digital de imágenes basadas en color a los rostros detectados con el fin de aumentar el porcentaje de reconocimiento.

- Implementar un sistema de reconocimiento de rostros usando el algoritmo basado en la técnica de Análisis de Componentes Principales.
- Verificar el desempeño del sistema desarrollado usando una base de datos internacionalmente reconocida.

## 2. PLANTEAMIENTO DEL PROBLEMA

En los últimos años el desarrollo de técnicas y programas informáticos para el reconocimiento de rostros se ha convertido en un tema principal de investigación, empleando un enfoque de visión por computadora y comprensión de imágenes, donde seguramente este tipo de reconocimiento se ve adaptado en diversas áreas de nuestra sociedad tales como las industrias, las instituciones públicas o privadas; donde dichas técnicas se ven aplicadas a los sistemas de identificación de personas, Sistemas Domóticos, Sistemas de Seguridad, Sistemas de Reconocimiento de usuarios, o Sistemas Biométricos.

La mayoría de los trabajos realizados para resolver el problema de reconocimiento facial se han basado en métodos tecnológicos tradicionales como lo son los Sistemas Biométricos [6], los cuales están soportados en un sistema automático fundamentado en la inteligencia artificial y en el reconocimiento de patrones, que permite reconocer y verificar la identidad de un individuo a partir de sus características morfológicas, convirtiéndose este método en una disciplina habitual en los sistemas de seguridad y en las herramientas criminalísticas. Aunque, para que una característica biométrica sea efectiva en el reconocimiento de rostros, se debe cumplir que la imagen sea invariante, es decir que la imagen no presente cambios de condiciones en la captura para alcanzar a distinguir y diferenciar cada individuo.

La automatización de estos métodos requiere de suficientes recursos computacionales. Entre las dificultades de este método se encuentran las grandes variaciones que presentan las imágenes captadas, como la posición de la cara, las expresiones faciales, la resolución, la iluminación, el color, la presencia de maquillaje en mujeres y de barbas y bigotes en los hombres. Dichas variaciones afectan negativamente a los sistemas biométricos, debido a que las imágenes presentan ciertas alteraciones en el instante de su captura.

Se plantea como solución, implementar un algoritmo de “AdaBoost” y de “Eigenfaces” para la detección y el reconocimiento facial respectivamente, basadas en el color, con el fin de aumentar la precisión de estos sistemas, mejorando el tiempo de ejecución, el espacio en memoria y su tasa de reconocimiento.

### 3. JUSTIFICACIÓN

Cada año evoluciona a pasos grandes la tecnología, ofreciéndole a los seres humanos facilidad de uso y un menor costo en las diferentes líneas de productos que se desarrollan. Un ejemplo de esto son las cámaras digitales, las cámaras web, las cámaras de vídeo, entre otros. Tomando estos dispositivos de captura y enmarcándolos en problemáticas propias de seguridad, se pueden emplear como detectores, identificadores y verificadores de individuos mediante la utilización de la información presente en su rostro.

En el presente trabajo se pretende mostrar la importancia de las técnicas de Análisis de Componentes Principales (PCA), el algoritmo de aprendizaje AdaBoost, la relación entre la identificación y clasificación de rostros, mediante el uso de los sistemas clasificadores y especialmente, la relación que existe al emplear el color en el sistema de reconocimiento facial.

Paul y Viola Jones [3], plantea que se pueden localizar la posición de un número indefinido de caras en una imagen aplicando un sistema clasificador. La fase de detección es considerada una etapa fundamental para cualquier aplicación donde se necesite realizar algún tipo de estudio y análisis de reconocimiento facial, pues es ésta la que permite extraer los rostros que alimentarán al sistema de forma automática.

Luego de la detección de los rostros, se realiza un estudio detallado de la cara detectada, mediante la aplicación de técnicas de procesamiento digital de imágenes basadas en color para el reconocimiento automático de caras, las cuales son técnicas no invasivas. Con esto, se pretende demostrar que el uso de técnicas basadas en color brindan mucha más información que las técnicas basadas en la escala de grises, especialmente cuando la imagen captada presenta variaciones de iluminación.

## 4. ESTADO DEL ARTE

A pesar de que existen diversos métodos de identificación de personas, como lo son las huellas dactilares, el iris del ojo, entre otros, el reconocimiento de rostros es el tema de interés para este trabajo.

Aprovechando la diversidad de dispositivos de captura que existen actualmente es evidente que los retos se encuentran en las aplicaciones que se le pueda dar a dichas imágenes capturadas. Sin embargo, el éxito de los sistemas de reconocimiento facial es muy limitado por las condiciones impuestas en las aplicaciones de la vida real, tales como, la iluminación, la posición del sujeto, entre otros. Por lo cual, se desarrolló un sistema de reconocimiento de rostros subdividido en la fase de detección e identificación del sujeto, lo cual permite plantear un sistema de captura de información (detección) especializado solo en reconocer rostros. Y otro especializado en análisis de información (reconocimiento) especializado en identificar individuos.

A continuación se presenta una reseña de las técnicas de reconocimiento de rostros, explicando brevemente la solución de varios problemas presentados en el sistema de reconocimiento facial.

### 4.1 HISTORIA

Entre los primeros sistemas de reconocimiento de rostros se encuentra el desarrollado por *Claasen y Green* [8] en los años 60, el cual requería de un administrador para localizar los rasgos característicos de un rostro como: los ojos, las orejas, la nariz y la boca, con el fin de calcular las distancias entre los puntos de referencias marcados por el administrador y compararlas luego con las referencias en común de la base de datos.

Para llevar a cabo dicha extracción de rasgos faciales en imágenes 2D se utilizaron técnicas típicas de procesamiento de imágenes como umbralización de la intensidad de los píxeles, detección de bordes, máximos de curvatura local de bordes, análisis de la intensidad del histograma, variaciones en la resolución, iluminación, entre otros, donde autores como *Kelly* [20], propusieron técnicas que permiten analizar las imágenes, extrayendo primero la línea del borde de la cara para posteriormente localizar sus ojos, nariz y boca. Su algoritmo consiste en difuminar la imagen original efectuando promedios locales y trazando las líneas posibles del contorno de la cabeza, proyectando estas líneas en la imagen original para redefinir nuevas líneas continuas que representen mejor el contorno de la cabeza. Una vez obtenidas estas nuevas líneas se calculan las posiciones donde se esperan que estén los rasgos faciales. A su vez, el trabajo de *D. Reisfeld* [26] presenta un operador de búsqueda de puntos en la imagen con un alto grado de simetría, con el fin de ser localizados los rasgos faciales muy rápidamente, mientras que *N.Valentin* [27] propone un método de detección de ojos. Sin embargo, *V.Perlibakas* [28] plantea un modelo de detección automática de rasgos y del contorno exacto de la cara utilizando morfología matemática y los bordes de la imagen.

Por otra parte *Kohonen* [7] en 1982, demostró que usando una red neuronal sencilla podría realizar reconocimiento de rostros con imágenes alineadas y previamente normalizadas. Este tipo de red, se ocupó de computar los rasgos faciales buscando la proximidad de los autovectores de la matriz de autocorrelación de los rostros. Sin embargo, el sistema de *Kohonen* tampoco tuvo un buen éxito en la práctica.

*Kirby y Sirobich* [8], mostraron que eran requeridos menos de 100 valores para cifrar acertadamente la imagen de una cara convenientemente alineada y normalizada, mediante el uso de Análisis de Componentes Principales (PCA), una técnica estándar del álgebra lineal, para la solución del problema del reconocimiento facial.

Gracias a los aportes de Kirby y Sirovich, en 1991 surge la técnica de *Eigenfaces* por Turk y Pentland [9], las cuales están asociadas a los autovectores que contienen las características principales de un conjunto de imágenes previamente evaluadas. En esta técnica, para poder reconocer nuevas imágenes de rostros, se debe calcular el error residual entre la proyección de la nueva imagen con respecto a la proyección del conjunto de imágenes de caras conocidas. Este descubrimiento permitió realizar sistemas de reconocimiento de caras automatizados en tiempo real.

En la Tabla 1 se muestran los principales enfoques de reconocimiento de rostros aplicados hasta la fecha.

**Tabla 1.** Categorización de Todas las técnicas de Reconocimiento facial

<i>Enfoques</i>	<i>Trabajos Representativos por:</i>
<b>Métodos Holísticos</b>	
• <i>Análisis de Componentes Principales (PCA)</i>	
Eigenfaces	<u>Aplicación directa de PCA:</u> Crac & Cameron 1996, [11] Kirby & Sirovich 1990, [8] Turk & Pentland 1991, [9]
Eigenfaces Probabilístico	Moghaddam & Pentland 1997, [12]
Fisherfaces/Subespacio LDA	<u>FLD en eigenspace :</u> Belhumeur et al 1997, [13], Swets & Weng 1996, [14], Zhao et al 1998, [15], He X, S. Yan, Y. Hu y P. Niyogi, [39], Xu, Yang y Jin, 2004 [40], Lu, Plataniotis y Venetsanopoulos, [41]
SVM	<u>Problema de dos-clases de SVM:</u> Phillips 1998, [16]
Feature lines	<u>Punto a distancia y basado en línea:</u> Li & Lu 1999, [17]

ICA	<i>Barlett y Sejnowski, [42]</i> <i>Yuen y Lai, [43]</i> <u>ICA-basado en análisis de características:</u> <i>Barlett et al 1998, [18]</i> <u>ICA-basado en SVM:</u> <i>Déniz, Castrillón y Hernández et al 2003, [44]</i>
• Otras Representaciones	
LDA/FLD	LDA/FLD : <i>Etemad &amp; Chellapa 1997, [19]</i>
Métodos Bayesianos	<i>Moghaddam y Pentland, [37]</i> <i>Xue, Li y Teoh, [38]</i>
Un estudio significativo del color, usando varios algoritmos de eigenfaces	<i>Karimi y Krzyzak, 2007 [46]</i>
<b>Métodos basados en Características</b>	
Métodos Geométricos	<u>Métodos anteriores:</u> <i>Kande 1973 [19], Kelly 1970 [20].</i> <u>Métodos recientes:</u> <i>Cox et al 1996 [10], Manjunath et al 1992 [21]</i>
Arquitectura Dinámica de Enlace	<u>Métodos de HMM.</u> <i>Samaría 1994, [22]</i>
Convolución de Redes Neuronales	<u>Aprendizaje SOM basado en el método de CNN.</u> <i>Lawrence et al 1997, [23]</i>
Métodos basados en extracción de rasgos locales	<u>Usando Templates:</u> <i>Brunelli y Poggio, [34], Yille y Haillinan, [36]</i>
<b>Métodos Híbridos</b>	
Eigenfaces Modular	<u>Eigenfaces and eigenmodules</u> <i>Pentland et al 1994, [24]</i>
Basado en Componentes	<u>Región de cara y componentes:</u> <i>Huang et al 2003, [25].</i> <u>Versus otros enfoques:</u> <i>Heisele, Ho, Wu y Poggio, [35]</i>
Un colo híbrido y características de frecuencias	<i>Liu Z y Liu C, 2008, [45]</i>

Fuente. W. ZHAO, R. chellappa, P. Phillips, A. Rosenfeld, *Face recognition: A literature survey*, 2003, p. 410.

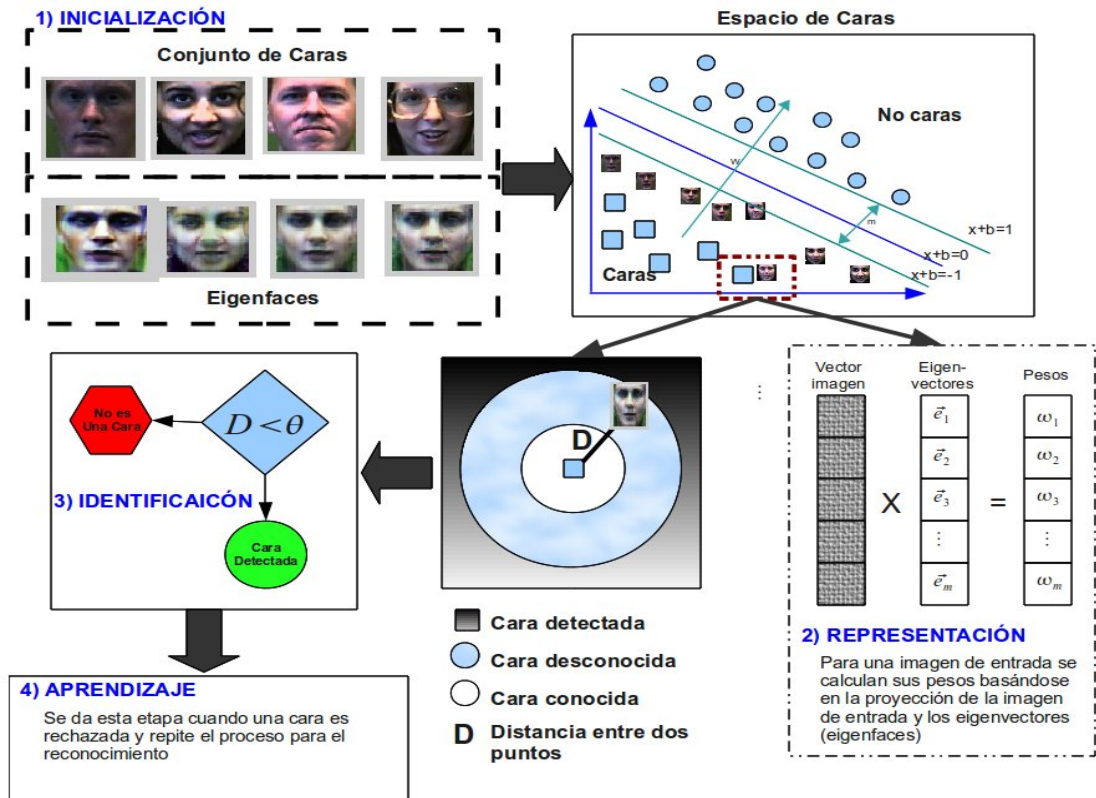
- **Método de eigenfaces**

Es un método estadístico que trata de obtener la información global de la imagen facial realizando una aplicación directa del Análisis de Componentes Principales. Este método parte de una imagen facial  $I(x, y)$ , representada por una matriz bidimensional de  $N \times N$  valores de intensidad de los píxeles, la cual se convierte en un vector de longitud  $N^2$ , donde el espacio vectorial de todas las imágenes de dicho vector para cualquier objeto o escena ya sea de caras, coches, paisajes, entre otros, será siempre de dimensión  $N^2$ . Sin embargo, por ser este un espacio de dimensión tan grande, es necesario encontrar un pequeño subespacio correspondiente a los componentes principales de las imágenes de caras, con el fin de consumir menos recursos en memoria computacional. Por esta razón, se hace uso del PCA, cuya idea principal es encontrar los vectores de la base de este subespacio que describa mejor la distribución del conjunto de imágenes de caras, teniendo como resultado una base de caras con menor dimensión, llamadas *eigenfaces*, de modo que cualquier cara que pertenezca a este subespacio, podrá ser escrita en términos de una combinación lineal de *eigenfaces* siendo los coeficientes de esta combinación lineal, las componentes del vector de características que se representan en el sistema de reconocimiento facial.

Una vez obtenidos los vectores de características se realiza un proceso de clasificación, donde se compara el vector del individuo de entrada, con los de la base de datos de los individuos conocidos mediante el cálculo del valor de alguna función de distancia que proporcione una medida de similitud con el vecino más cercano.

*Turk y Pentland* [9] introdujeron un sistema de reconocimiento facial en 2D, que funciona casi en tiempo real, basándose en el método de *eigenfaces*. Las etapas de este sistema se bosquejan en la Figura 1.

Figura 1. Sistema de reconocimiento facial 2D, basado en el método de eigenfaces.



Este método ha sido aplicado para otros propósitos como la detección de caras, compresión de imágenes, entre otros. En los últimos años se han realizado numerosas publicaciones derivadas de ese método *C.Thomaz* [29], *H.Kim* [30], *R.Gottumuakkal* [31], *A.Pentland* [24], donde este último presenta un método basado en el análisis de componentes principales en el que se reconocen caras con variaciones en la pose, escala y orientación.

## **5. MARCO TEÓRICO**

### **5.1 FUNDAMENTOS DE LA VISIÓN COMPUTACIONAL**

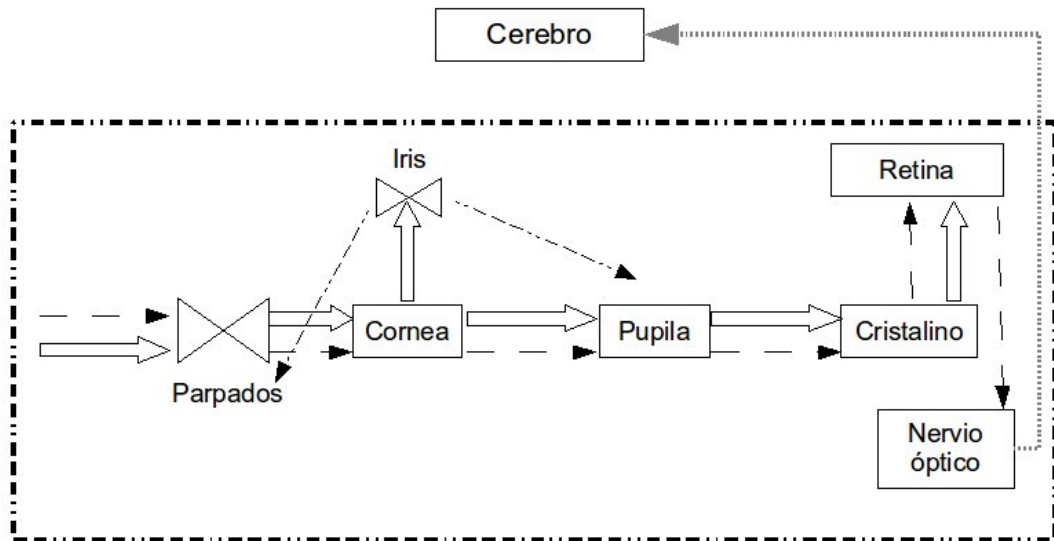
Las técnicas de visión artificial por computadora presentan una multitud de aplicaciones, entre las que se encuentran la identificación de huellas dactilares, reconocimiento de caras, crecimiento y detección de enfermedades, control de tráfico vehicular, monitorización del sistema de vigilancia y muchas otras aplicaciones.

En este capítulo se explicará brevemente lo que es la visión computacional, empezando por entender que es la visión humana, que es la visión computacional y cuál es la diferencia entre estas dos visiones.

#### **5.1.1 La Visión Humana**

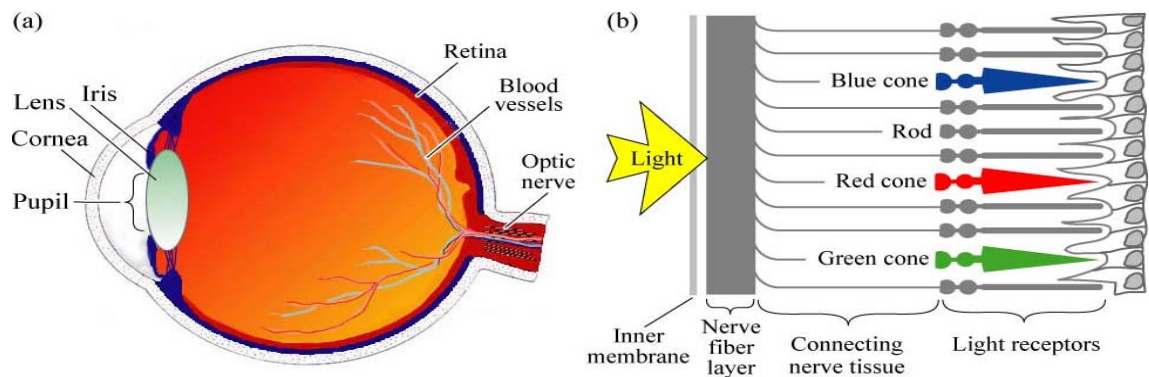
El ojo es el órgano que permite diferenciar colores, luminosidad y distancia de los objetos visualizados. Sin embargo, el órgano que efectúa el proceso de la visión no es el ojo realmente sino el cerebro. La función del ojo es traducir las vibraciones electromagnéticas de la luz en un determinado tipo de impulsos nerviosos que se transmiten al cerebro, como se ilustra en la Figura 2.

**Figura 2** Esquema de funcionamiento del ojo humano



En la Figura 3 se observa claramente las partes que conforman el ojo humano. El glóbulo ocular tiene la forma de una esfera y está rodeado por una membrana llamada esclerótica que por la parte anterior se hace transparente para formar la córnea. Después de la córnea hay un diafragma, el iris, que posee una abertura, la pupila, por la que pasa la luz hacia el interior del ojo, donde el iris es el encargado de regular la intensidad luminosa que recibe el ojo controlando automáticamente el diámetro de la pupila.

**Figura 3** El ojo humano, (a) Muestra las partes del ojo humano, (b) Muestra los fotorreceptores

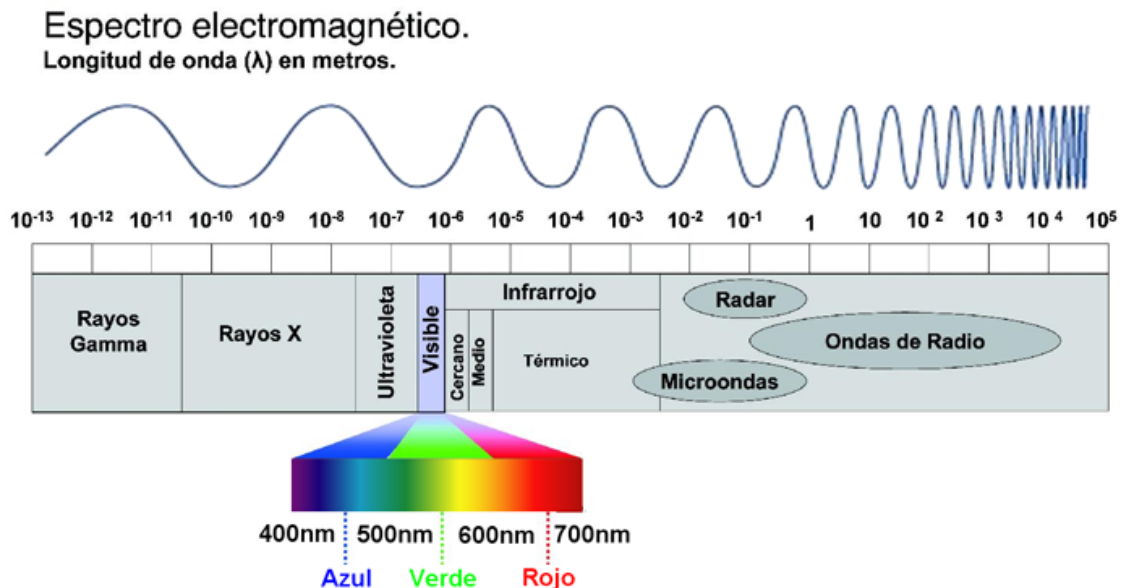


Fuente: Richard A. Moscoso Bullón, *el Blog de la Física*, imagen adaptada de la Enciclopedia Británica, 1994. [Fecha de consulta: 20 Enero 2011]. Disponible en: <http://blog.pucp.edu.pe/category/8396/blogid/3197>.

El cristalino está unido por ligamentos al músculo ciliar y está encargado de enfocar las imágenes sobre la envoltura interna del ojo, es decir sobre la retina. Esta envoltura contiene fibras nerviosas que terminan en unas pequeñas estructuras denominadas conos y bastones, las cuales son muy sensibles a la luz. Además, existe un punto en la retina, llamado fovea, donde alrededor de éste hay una zona que sólo tiene conos para ver el color y sobre ella se forma la imagen del objeto que miramos. Los millones de nervios que van al cerebro se combinan para formar un nervio óptico que sale de la retina por un punto que no contiene células receptoras llamado punto ciego.

La visión es la facultad por la cual a través del ojo se percibe el mundo exterior y está relacionada especialmente con la percepción del color, la forma, la distancia y las imágenes. El color incide sobre la retina del ojo mediante las ondas luminosas, pero, si estas ondas son superiores o inferiores a determinados límites (Ver Figura 4), la retina no producirá visión alguna. El color depende, en parte, de las longitudes de onda de las ondas luminosas incidentes, que pueden ser simples o compuestas. Y la luminosidad aparente de un objeto depende de la amplitud de las ondas luminosas que pasan de él al ojo.

**Figura 4** Espectro Electromagnético



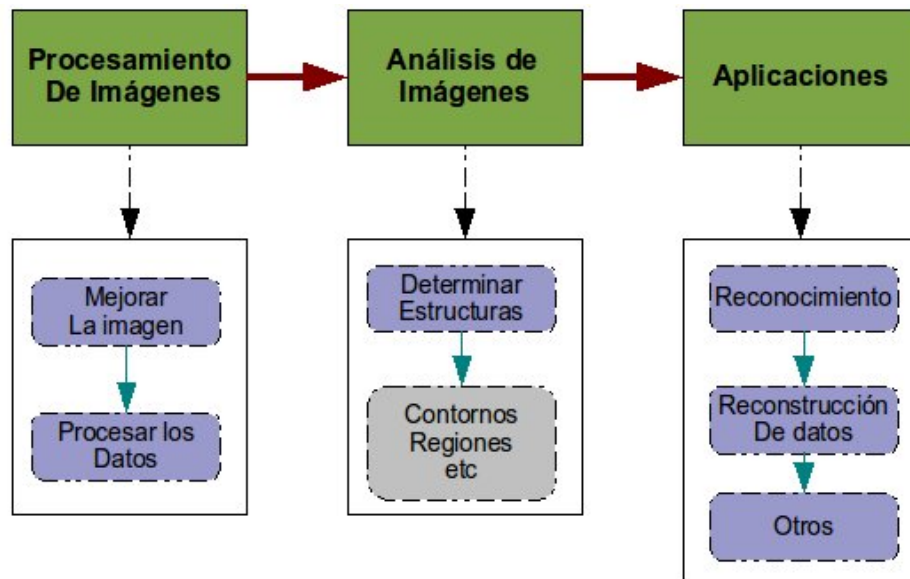
Fuente: Pa'í Olivan, *Poner en orden el espectro electromagnético de radio y TV*, Actualizado el 17 de enero de 2011 [Fecha de consulta: 20 Enero 2011]. Disponible en: <http://paioliva.blogspot.com/2011/01/poner-en-orden-el-espectro.html>

### 5.1.2 La Visión Artificial

La visión es uno de los mecanismos sensoriales de percepción más importantes en el ser humano aunque no es exclusivo.

La visión artificial, es una de las ramas de la inteligencia artificial, y se le conoce también como la visión por computadora o visión técnica, la cual hace referencia al intento de dotar a las máquinas de un sistema de visión aparente. Este tipo de visión tiene como propósito modelar matemáticamente los procesos de percepción visual en los seres vivos y generar programas que permitan simular estas capacidades en la computadora.

**Figura 5** Procesos de la visión artificial.



En la Figura 5 se observan los procesos que debe realizar la visión artificial para simular la visión humana. El primero de ellos implica una manipulación de las imágenes vistas como señales digitales, para extraer la información deseada mediante un exhaustivo análisis en la imagen que se da en el segundo proceso para finalmente ser llevadas a ciertas aplicaciones.

Los métodos de procesamiento digital de imágenes se fundamentan en dos áreas principales: Mejoramiento de la calidad de la imagen para la interpretación humana y procesamiento de los datos de la escena para la percepción de las máquinas de manera autónoma.

### **5.1.3 Diferencia entre la visión humana y la visión artificial**

Los sistemas computacionales actuales son lentos si los comparamos con la inteligencia del ser humano, aunque su capacidad, precisión y velocidad son muy buenas. Por esta razón es necesario mejorar los algoritmos, para poder emular las capacidades humanas en los sistemas de cómputo dentro de los

límites establecidos hasta el momento. A continuación se señala en la tabla 2 las diferencias entre la visión humana y la visión artificial.

**Tabla 2** Diferencia entre la visión humana y la visión computacional.

	<b>VISIÓN HUMANA</b>	<b>VISIÓN COMPUTACIONAL</b>
<b>Flexibilidad</b>	Muy adaptable y flexible en las tareas de visión.	Rígido en la tarea. Requiere cuantificar datos (píxeles).
<b>Habilidad</b>	Puede ser relativamente preciso. Estima de manera subjetiva. Ejemplo: Detectar fruta mala basada por su color y textura	Puede hacer mediciones de dimensiones basándose en entrada de datos predeterminados, basado en el conteo de píxeles.
<b>Color</b>	Subjetivo en color.	Medición de la magnitud de los parámetros cromáticos (RGB)
<b>Sensibilidad</b>	Adaptación en las diferentes condiciones de iluminación, características físicas en detalles de la superficie y distancia al objeto. Limitación en la habilidad a distinguir entre las tonalidades de grises. Varía en función del individuo así como del tiempo, se pueden identificar de 7 a 10 niveles de grises.	Sensitivo al nivel y frecuencia de iluminación también como la naturaleza física de la superficie y distancia al objeto. Habilidad para cuantificar es relativamente grande y ajustado por el sensor, características del entorno y del sistema. Puede identificar fácilmente 256 niveles de grises.
<b>Respuesta</b>	Velocidad en un orden de 0.1 de segundos.	La velocidad de la respuesta depende del sistema computacional y el tipo de problema, usando mejores técnicas y hardware.
<b>2-D y 3-D</b>	Pueden manejar operaciones 3-D y longitudes de onda de color múltiples fácilmente.	Puede manejar operaciones 2-D fácilmente pero limitada en tareas de 3-D se requieren dos cámaras y es muy lento.

<b>Salida de datos</b>	Puede manualmente suministrar a un sistema una descripción de lo observado. El tiempo es mucho y muy impreciso.	Puede automáticamente suministrar datos discretos de entrada precisos al sistema continuamente. Con costos al sistema variables.
<b>Percepción</b>	Percibe iluminación en escala logarítmica. Se afecta la percepción por el área circundante.	Puede percibir la iluminación, ya sea una escala lineal o logarítmica.
<b>Espectro</b>	Limitado a un espectro visual de 300 a 700 milimicrones.	Con la tecnología adecuada puede hacer mediciones en todo el espectro visual.

*Fuente: Encyclopedia Britannica 2003 Ultimate Referente CD's*

## **5.2 FUNDAMENTOS DEL PROCESAMIENTO DIGITAL DE IMÁGENES**

El procesamiento digital de imágenes, es considerado como un área de mayor investigación a nivel mundial y es utilizado en diferentes disciplinas científicas. Una de esas disciplinas es el reconocimiento de rostros [32], [33], en el cual se desarrollaron bases teóricas y algorítmicas para extraer información del mundo real de manera automática, a partir de un conjunto de imágenes observadas.

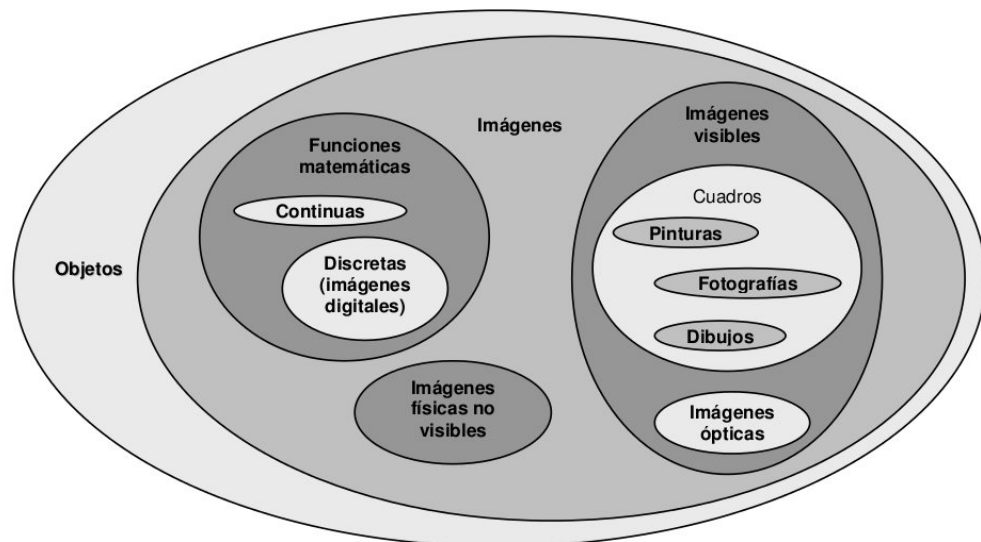
Con el propósito de aclarar los términos y las técnicas usadas para este trabajo de investigación, se explicarán brevemente los conceptos relacionados con este tema, con el fin de establecer un puente entre las imágenes digitales y el álgebra lineal.

### **5.2.1 Imágenes Digitales**

En términos generales, una imagen puede ser considerada como la información descriptiva de un objeto, la cual le proporciona a una persona la

visualización de ese objeto en un campo específico. Las imágenes pueden ser clasificadas en varios tipos, de acuerdo con la forma o el método de generación, tal como se muestra en la Figura 6.

**Figura 6** Clases de imágenes.



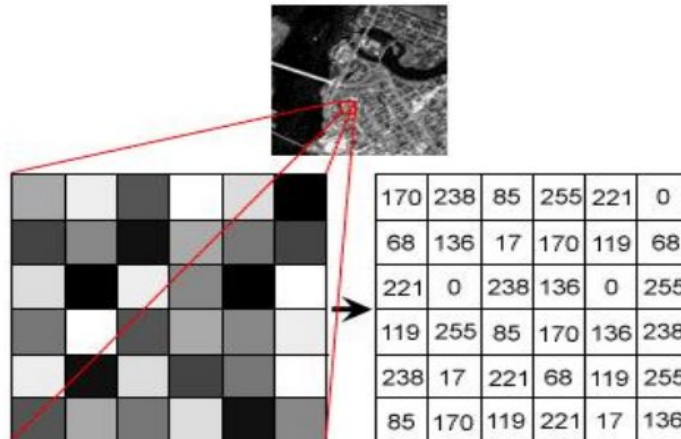
Fuente: CASTELMAN, Kenneth. *Digital Image Processing*. Upper Saddle River. Prentice Hall, 1996, pp. 5.

Las imágenes forman parte de un subconjunto que contienen todos los objetos del universo. Allí, se encuentran las imágenes digitales las cuales se ubican dentro del conjunto de las imágenes multiespectrales que son aquellas que se caracterizan por tener más de un valor en un punto determinado de las imágenes visibles como las pinturas, las fotografías, la televisión, etc. y son procesadas por una computadora.

De acuerdo con lo anterior, una imagen digital se define como una función bidimensional discreta (muestreada y cuantificada) que ha sido captada por medios ópticos, cuyo muestreo se realiza en secuencias de bits que representan los píxeles de la imagen como un patrón cuadrado de

rectángulos igualmente espaciados (Ver Figura 7), para que puedan ser visualizados y cuantificados en intervalos de amplitudes iguales.

**Figura 7** Imagen digital muestreada y cuantificada.



Fuente: Francisco, *La Imagen de Satélite... otra forma de ver el mundo* Actualizado el 18 de Julio de 2019 [Fecha de consulta: 20 Enero 2011]. Disponible en: [http://www.aulati.net/?category\\_name=informacion-geografica](http://www.aulati.net/?category_name=informacion-geografica)

La resolución del bit, mide la cantidad de la información de color disponible para cada píxel de una imagen. Cuanto mayor sea la profundidad del píxel en una imagen, más colores habrá disponibles y más exacta será la representación del color en la imagen digital.

## 5.2.2 Modelos de Color

Una imagen en color es una representación tridimensional en el plano de la imagen, la cual posee tres componentes independientes donde son almacenadas y manipuladas por las computadoras según el modelo del color empleado.

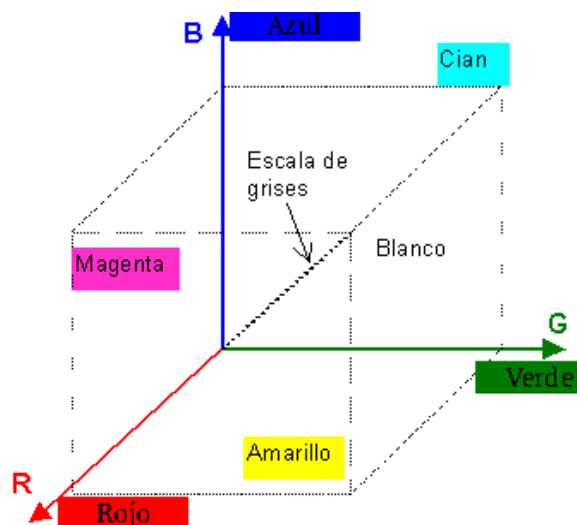
El uso del color en el procesamiento digital de imágenes está motivado por el análisis de las imágenes, ya que el color es un gran descriptor que con frecuencia simplifica la identificación y extracción de objetos de una escena.

A continuación se hará una introducción a los espacios de color utilizados en este trabajo de investigación.

### 5.2.2.1 Modelo de color RGB

Este modelo de color se obtiene por mezcla aditiva de luz, como se observan en la televisión, las pantallas gráficas, las fotografías, etc. Aquí, cada color aparece en sus componentes espectrales primarios, rojo, verde y azul, y está basado en un sistema de coordenadas cartesianas como se observa en el tetraedro de la Figura 8, el cual ilustra el subespacio de color de interés.

Figura 8 Espacio de Color del modelo RGB

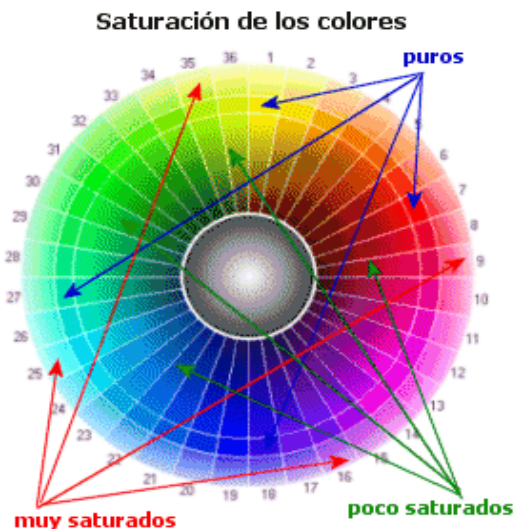


### 5.2.2.2 Modelo de color HSV

El modelo HSV, consiste en una transformación no lineal del espacio de color RGB. En este modelo los colores se clasifican de acuerdo a tres características básicas: Tono, Saturación y Brillo. La componente H, representa el matiz, o sea el tono, el cual hace referencia a la longitud de onda dominante en la luz

emitida o reflejada por un objeto. Para asignar un valor al tono se utiliza una rueda de colores normalizada, en la que los tres colores primarios (rojo, verde y azul) y los tres colores secundarios (cyan, magenta y amarillo) se alternan a lo largo de una circunferencia como se ilustra en la Figura 9.

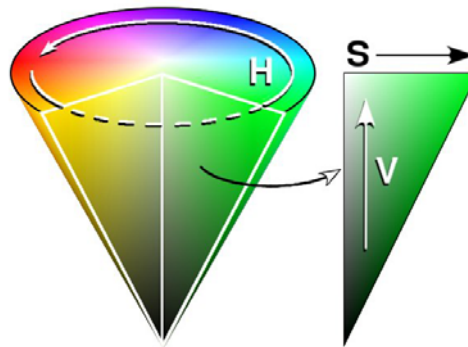
**Figura 9** Saturación de los colores



Fuente: Tomada en <http://nemo152008.blogspot.com/2008/11/saturacion.html> , actualizada el día 18 de Noviembre de 2008, [Fecha de consulta: 20 Enero 2011].

Sin embargo, la componente S o saturación describe la viveza del color, es decir la cantidad del matiz diluido en el color blando. Por eso, cuando se habla de un color muy saturado, lo que realmente se quiere decir, es que dicho color se encuentra con una tonalidad intensa y pura. Por el contrario, un color poco saturado es el que tiene una tonalidad apagada, casi oscura. En la Figura 9, los colores muy saturados se encuentran cerca del borde y los colores poco saturados son los que están cerca del centro del círculo. No obstante, en la Figura 10, la saturación se representa como la distancia a lo largo del radio de la circunferencia del modelo de color HSV. Finalmente, la componente V simboliza el brillo del color que representa la altura en el eje del espacio de la figura 10 que va desde el blanco hasta el negro.

**Figura 10.** Espacio de color del modelo HSV



Fuente: Tomada en <http://www.esacademic.com/dic.nsf/eswiki/445640>, [Fecha de consulta: 20 Enero 2011].

Estas características hacen que el modelo HSV, sea una herramienta ideal para desarrollar algoritmos de procesamiento de imágenes basados en color.

### 5.2.2.3 Conversión del modelo RGB a HSV

Como se ha mencionado anteriormente, el modelo RGB está definido con respecto a un tetraedro unitario, mientras que las componentes de color del modelo HSV están definidas con respecto al triángulo de color mostrado en la Figura 10. Los valores representados en cada uno de las componentes de color de los modelos mencionados anteriormente deben estar previamente normalizados en el rango de [0,1]. La función para la componente H se define en la ecuación (1.1), donde R, representa el valor de los niveles de intensidad del espacio Rojo, G el valor de los niveles de intensidad del espacio Verde, y B, el valor de los niveles de intensidad del espacio Azul.

$$H = \cos^{-1} \left\{ \frac{1/2 [(R - G) + (R - G)]}{[(R - G)^2 + (R - G)^2(G - B)^2]^{1/2}} \right\} \quad (1.1)$$

Para la componente S, se tiene:

$$S = 1 - \frac{3}{R + G + B} [\min (R, G, B)] \quad (1.2)$$

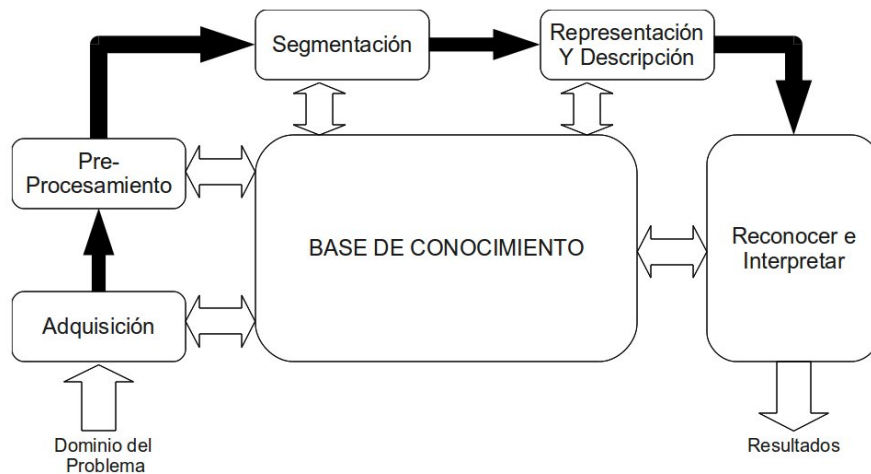
Y para la componente V:

$$V = [\max (R, G, B)] \quad (1.3)$$

### 5.2.3 Etapas del Procesamiento Digital de Imágenes

En esta etapa se bosqueja en la Figura 11 las fases del procesamiento digital de imágenes, las cuales han sido estudiadas a lo largo de este tiempo.

**Figura 11** Etapas fundamentales del procesamiento de imágenes.



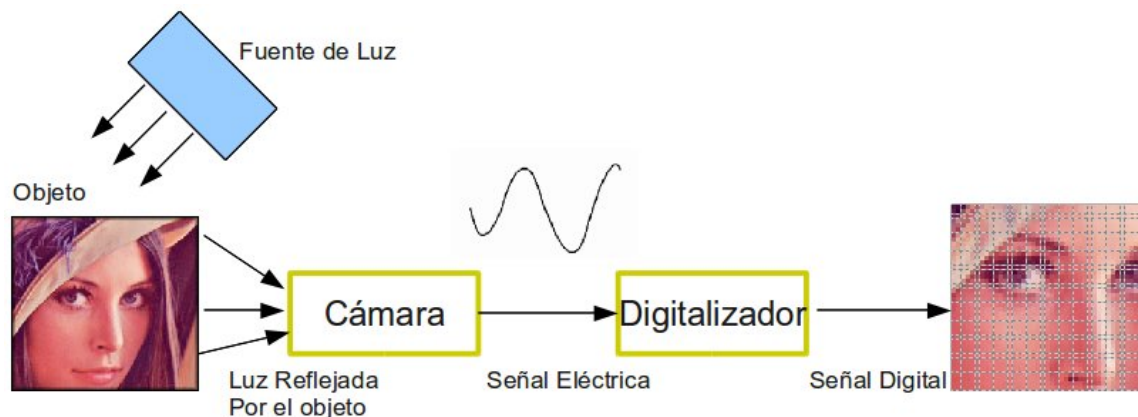
Fuente: GONZÁLEZ, Rafael C, Op, pp. 8.

En este trabajo de investigación, sólo se realizará la etapa de *Adquisición*, *Pre-Procesamiento* y *Reconocimiento* de las imágenes y se explicará cada una de ellas, con el fin de entender la metodología propuesta para cumplir cada uno de los objetivos trazados en esta investigación.

### 5.2.3.1 Adquisición.

Considerado como el primer paso para procesar la imagen de entrada adquirida mediante un dispositivo óptico, la cual produce una señal eléctrica que al ser introducida a un dispositivo llamado digitalizador, convierte esta señal en una señal digital obteniendo como salida una imagen digitalizada (Ver Figura 12)

Figura 12 Digitalización de una Imagen.



Un digitalizador debe tener en cuenta dos criterios importantes al momento de realizar la discretización. En primer lugar el *muestreo*, el cual está relacionado con la dimensión espacial, e indica el tamaño del elemento de la imagen en píxeles, y en segundo lugar la *cuantificación*, hace referencia a la amplitud de la señal, es decir al nivel de intensidad del píxel.

### 5.2.3.2 Pre-Procesamiento.

Es el encargado de la mejora de la información de la imagen. Para ello, existen varios métodos espectrales basados en la información frecuencial y otros en el dominio espacial.

La mejora de imágenes incluye manipulación del contraste, los niveles de grises, reducción de ruido, suavizado o realce de bordes y filtración. La

dificultad en esta etapa es determinar el criterio para el ajuste de la imagen, puesto que la mayoría de las técnicas de mejora son empíricas y requieren procesamientos interactivos para obtener resultados satisfactorios. En la Figura 13 se ilustra la clasificación de los algoritmos de mejora de imágenes.

**Figura 13.** Clasificación de la mejora de imágenes



Fuente: JAIN, Anil K. *Fundamentals of Image Processing*. Englewood Cliffs: Prentice Hall, 1989, pp. 233.

Para el reconocimiento de caras, es necesario aplicar técnicas de mejoramiento de la calidad de la imagen que no permitan perder detalle alguno con el fin de no dificultar la identificación del rostro. Por tal razón, sólo se explicará brevemente el pre-procesamiento de la imagen empleando la manipulación del histograma y utilizando un filtrado de tipo paso alto.

- **Manipulación del Histograma.**

Con esta técnica se mejora la imagen píxel a píxel. Para ello, existen tres técnicas básicas de transformación del histograma, que resultan muy útiles para modificar el brillo y el contraste de las imágenes originales. Estas son contracción, expansión y ecualización del histograma

- **Contracción del histograma.**

Esta técnica disminuye el contraste del histograma, es decir, reduce el rango de los valores de gris. Está definida por:

$$g(i, j) = \left[ \frac{C_{MAX} - C_{MIN}}{f(i, j)_{MAX} - f(i, j)_{MIN}} \right] [f(i, j) - f(i, j)_{MIN}] + C_{MIN} \quad (1.4)$$

donde  $f(i, j)$  es el nivel de gris del píxel  $(i, j)$  de la imagen de entrada  $f$ ,  $f(i, j)_{MAX}$  es el mayor valor del gris en la imagen de entrada  $f$ ,  $f(i, j)_{MIN}$  es el menor valor del nivel de gris en la imagen de entrada  $f$ ,  $C_{MAX}$  y  $C_{MIN}$  es el valor máximo y mínimo deseado en la compresión del histograma.

En la Figura 14 se puede observar la mejora de la imagen al aplicar este tipo de técnica, donde la imagen original se muestra al lado izquierdo de la imagen con los valores del histograma de esta imagen concentrados en la parte inferior. Después de aplicar la ecuación (1.4), se obtiene la imagen contraída con su histograma correspondiente en la parte derecha de la imagen.

**Figura 14.** Contracción de la imagen



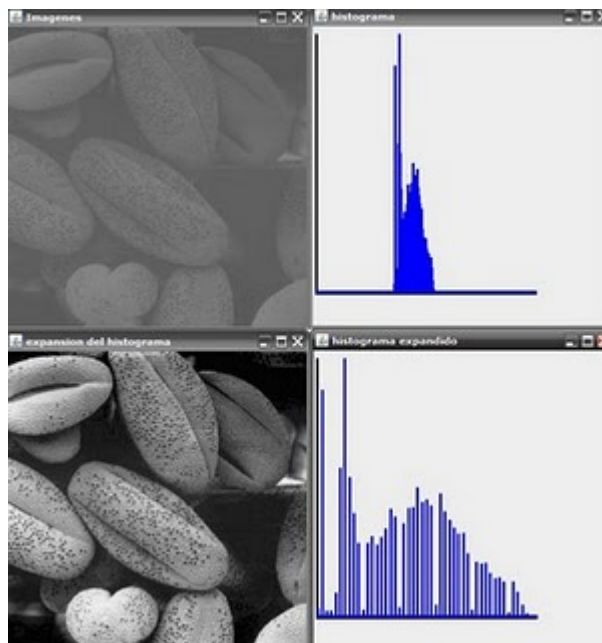
- **Expansión del histograma.**

Es la operación contraria a la contracción del histograma, y se define así:

$$g(i,j) = \left[ \frac{f(i,j) - f(i,j)_{MIN}}{f(i,j)_{MAX} - f(i,j)_{MIN}} \right] [MAX - MIN] + MIN \quad (1.5)$$

donde  $f$  es la imagen de entrada,  $f(i,j)_{MAX}$  es el mayor valor del gris en la imagen de entrada  $f$ ,  $f(i,j)_{MIN}$  es el menor valor del nivel de gris en la imagen de entrada  $f$ ,  $MAX$  y  $MIN$  corresponden al máximo y mínimo valor posible de los niveles de intensidad de la imagen  $f$ . En la Figura 15 se puede observar la mejora de la imagen al aplicar este tipo de técnica y la expansión de su histograma, la cual ocupa el total del rango de niveles de intensidad en el eje horizontal del histograma, abriendo grandes huecos entre niveles de intensidad contiguos de la imagen original.

**Figura 15** Expansión de la imagen



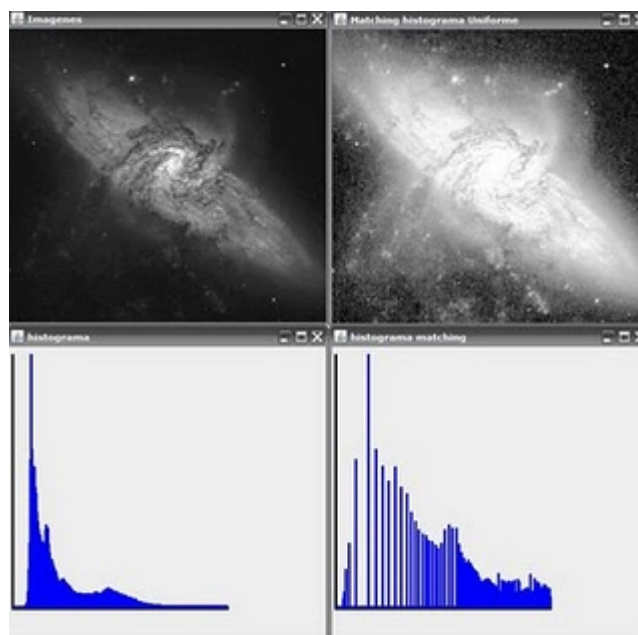
- **Ecuación del Histograma.**

Con esta técnica se trata de encontrar una función  $H$  que realce el contraste general en la imagen original  $f$  expandiendo la distribución de los niveles de gris (Ver ecuación 1.6). Dicha igualación debe ser lo más suave posible, ya que el objetivo de esta ecualización es aumentar y disminuir al mismo tiempo el contraste de la imagen, mediante el incremento y reducción de los niveles de gris que se encuentren más cercanos al valor máximo y mínimo del histograma respectivamente, tal como se ilustra en la Figura 16.

$$H(Q) = \frac{L - 1}{n * m} \sum_{i=0}^Q h_f(i) \quad (1.6)$$

donde  $H(Q)$  es la ecualización del histograma de un imagen  $f$  con dimensión  $n*m$  y  $\forall Q \in \{0, \dots, q - 1\}$  con  $q$  niveles de intensidad de grises en la imagen  $f$ .  $L$  indica el valor máximo que representa a los niveles de gris de la imagen  $f$  y con  $h_f$  como el valor del histograma de la imagen evaluado desde cero hasta  $Q$ .

**Figura 16.** Ecualización de la Imagen



- **Filtración de imágenes**

En este punto, lo que se pretende es manejar un tipo de filtrado que realce en la imagen, con el objetivo de destacar los detalles finos de esta, o también por intensificar aquellas imágenes que han sido difuminadas ya sea por error, o por el método de adquisición donde fueron tomadas.

Para este trabajo de investigación se utilizó el filtro paso alto de tipo *unsharp masking* el cual está basado en una máscara de desenfoque para aumentar el contraste de una imagen, y así mejorar la nitidez de esta. Por lo anterior, este tipo de técnica se considera bastante importante en el procesamiento digital de imágenes y puede ser calculada como la diferencia entre la imagen original y una versión de esta imagen que ha sido pasada por un filtro de paso bajo, es decir:

$$Paso\ Alto = Original - Paso\ Bajo \quad (1.7)$$

$$IF = \alpha * Original - (Original - Paso\ Alto) \quad (1.8)$$

$$IF = (\alpha - 1) * Original + Paso\ Alto \quad (1.9)$$

donde  $\alpha$  es un factor de multiplicación que toma valores entre 1 y mayores que 1.  $IF$  es el resultado que la imagen filtrada se parece más a la imagen original, con un grado relativo de mejora en los bordes el cual depende del valor de  $\alpha$ . En la ecuación 1.8 se indica el proceso de sustracción de una imagen difusa a una original.

### **5.2.3.3 Reconocimiento**

En esta etapa se asignan unos pesos a las imágenes que poseen caras de acuerdo con sus características, permitiendo reconocer e identificar a los individuos, mediante la comparación entre sus pesos y permitiendo identificar el

sujeto que ha sido analizado, haciendo uso de la distancia euclidiana y de mahalanobis.

- **Distancia Euclidiana**

Sea  $(p_1, p_2)$  dos puntos de coordenadas  $(x_1, y_1)$  y  $(x_2, y_2)$  cuya distancia euclidiana está dada por:

$$d_{(p_1, p_2)}^2 = \|(p_1 - p_2)\|^2 \quad (1.10)$$

$$d_{(p_1, p_2)} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (1.11)$$

- **Distancia de Mahalanobis**

Sea  $(v_1, v_2)$  dos puntos cuya distancia de mahalanobis está dada por:

$$d_{EM}(v_1, v_2)^2 = \|(v_1 - v_2)\|_M^2 \quad (1.12)$$

$$d_{EM}(v_1, v_2) = \sqrt{(v_1 - v_2)^T * inv(M) * (v_1 - v_2)} \quad (1.13)$$

donde  $M \in \mathbb{R}^{N \times N}$  es una matriz diagonal cuadrada positiva

## 5.3 FUNDAMENTOS MATEMÁTICOS

### 5.3.1 Eigenvalores y Eigenvectores

Sea  $A$  una matriz de  $n \times n$  con componentes reales. El número  $\lambda$  se llama **eigenvalor** de  $A$  si existe un vector diferente de cero  $\mathbf{v}$  en  $\mathbb{C}^n$  tal que:

$$A\mathbf{v} = \lambda\mathbf{v} \quad (1.14)$$

El vector  $\mathbf{v} \neq \mathbf{0}$  se llama **eigenvector** de  $A$  correspondiente al **eigenvalor**  $\lambda$ . Y sea  $T: V \rightarrow V$  una transformación lineal, es útil encontrar un vector  $\mathbf{v}$  en  $V$  tal que  $T\mathbf{v}$  y  $\mathbf{v}$  son paralelos. Es decir, se busca un vector  $\mathbf{v}$  y un escalar  $\lambda$  tal que:

$$T\mathbf{v} = \lambda\mathbf{v} \quad (1.15)$$

Si  $\mathbf{v} \neq \mathbf{0}$  y  $\lambda$  satisface; entonces se llama un **eigenvalor** de  $T$  y  $\mathbf{v}$  se llama un **eigenvector** de  $T$  correspondiente al eigenvalor  $\lambda$ .

### 5.3.2 Covarianza

Al manejar dos o más variables aleatorias es muy útil describir la forma como varían juntas, y una forma de hacerlo es mediante la covarianza, y se define como el valor esperado de sus desviaciones de la media, para variables continuas:

$$\begin{aligned} \sigma_{XY} &= cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] \\ &= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy (x - \mu_X)(y - \mu_Y) f_{XY} \quad (1.16) \\ &= E(XY) - \mu_X\mu_Y \end{aligned}$$

---

<sup>2</sup> Tomado del libro de GROSSMAN, Stanely. *Álgebra Lineal*. En: McGraw-Hill, Quinta edición. Capítulo 6, pp. 533-546

Para variables discretas se reemplaza la integral por sumatoria, así:

$$\begin{aligned}\sigma_{XY} = cov(X, Y) &= E[(X - \mu_X)(Y - \mu_Y)] \\ &= \sum_{R_X} \sum_{R_Y} (X - \mu_X)(Y - \mu_Y) f_{XY}\end{aligned}\quad (1.17)$$

donde  $R_X$  y  $R_Y$  son los rangos de  $X$  y  $Y$  respectivamente, y se puede decir que de  $-\infty$  a  $\infty$ ;  $f_{XY}$  es la función de probabilidad conjunta de  $X$  y  $Y$ .

Se puede simplificar a la esperanza del producto menos el producto de las esperanzas de las dos variables:  $\sigma_{XY} = E[(X * Y) - E(X) * E(Y)]$  (1.18)

La covarianza es positiva si al aumentar  $X$  generalmente aumenta  $Y$ , es negativa si al aumentar  $X$  generalmente disminuye  $Y$ , y es cero si son independientes, es decir, si al aumentar  $X$ ,  $Y$  puede tanto disminuir como aumentar indistintamente<sup>3</sup>.

### 5.3.2.1 Matriz de Covarianza

Es una matriz que contiene la covarianza entre los elementos de un vector. Entonces, sea  $X$  un vector aleatorio, con media  $\mu$  y matriz de covarianza denotada como  $\Sigma$  y definida así:

$$\sum_{i,j} E [(X_i - \mu_i)(X_j - \mu_j)] \quad (1.19)$$

donde  $\mu_i = E(X_i)$ .

---

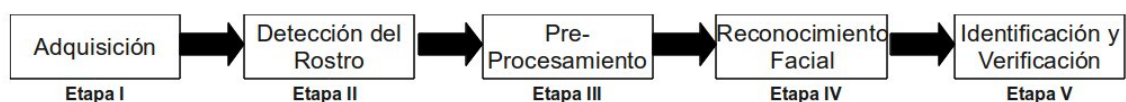
<sup>3</sup> Tomado de los apuntes del Prof. SARMIENTO, Enrique. Estadística Descriptiva, *Capítulo 1*

## 6. METODOLOGÍA Y DESARROLLO

El presente trabajo de investigación está enmarcado en el proyecto No 5537 de la Vicerrectoría de Investigación y Extensión (VIE), en el cual se utilizó una metodología basada en Extreme Programming (XP) para el diseño y desarrollo de un sistema de reconocimiento de rostros en tiempo real sobre un DSP, utilizando imágenes en escala de grises, pero en vista de la importancia del color en las imágenes, se hizo necesario analizar e implementar técnicas de procesamiento digital de imágenes basadas en color sobre un sistema de reconocimiento facial.

Por tanto, este trabajo de investigación se dividió en cinco etapas (Ver Figura 17) adquisición de la imagen, detección del rostro, pre-procesamiento de la imagen, reconocimiento de rostros e identificación y clasificación del individuo, proporcionando de tal forma una simplicidad en el diseño para facilitar los cambios requeridos en el diseño de los algoritmos en ambas fases, durante el desarrollo de estos.

**Figura 17** Etapas de un sistema de reconocimiento facial basado en color.



### 6.1 ETAPA I. Adquisición de la Imagen.

Para la adquisición de la imagen se tuvieron en cuenta dos tipos de base de datos internacionalmente reconocidas *Faces 94* y *Georgia Tech Face Database*. En la siguiente tabla se encuentran las características principales que presentan las imágenes en cada una de las bases de datos mencionadas.

**Tabla 3** Descripción de las bases de datos implementadas

<b>Face 94 Database</b>	<b>Georgia Tech Face Database</b>
<ul style="list-style-type: none"><li>✓ Contiene imágenes de 153 personas captadas a una distancia fija de la cámara en una sola sesión de fotografía.</li> <li>✓ La resolución de la imagen es de 180x200 píxeles y formato JPG.</li> <li>✓ Las imágenes fueron captadas con fondos uniformes, sin cambios de iluminación y con variación en las expresiones faciales de los individuos.</li></ul>	<ul style="list-style-type: none"><li>✓ Contiene imágenes de 50 personas captadas en tres sesiones diferentes contenidas entre las fechas del 06/01/99 al 11/15/99.</li> <li>✓ La resolución de la imagen es de 640x480 píxeles y formato JPG.</li> <li>✓ Las imágenes fueron captadas con fondos diferentes, con variaciones tanto en escala, orientación y expresiones faciales de los individuos y con diferentes condiciones de iluminación en la imagen.</li></ul>

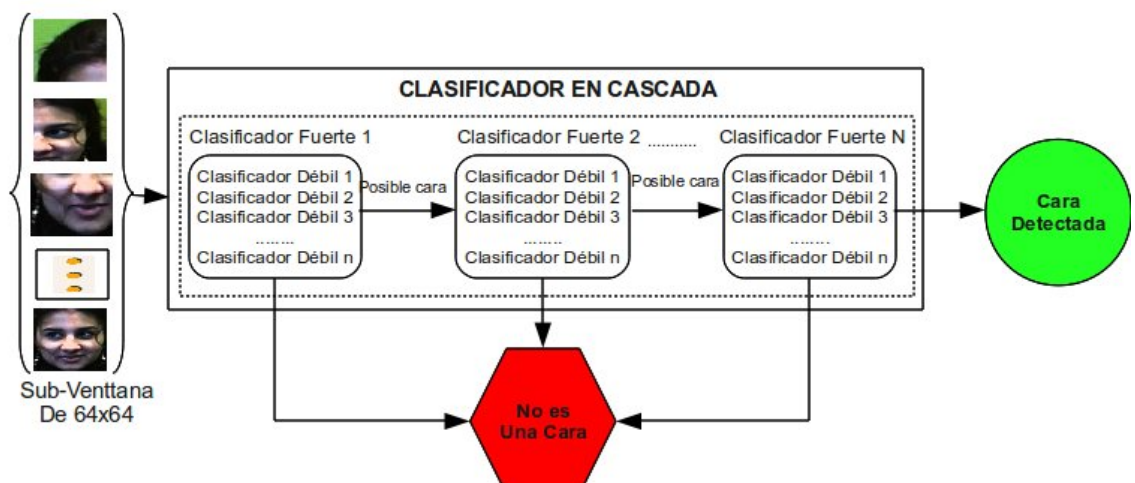
Para este trabajo de investigación, se optó por elegir sólo imágenes frontales sin variación en la orientación del rostro, pero con pequeños cambios en las expresiones faciales, y con o sin, condiciones de iluminación dependiendo del tipo de base de datos a trabajar, estas condiciones se cumplen tanto para la fase de Entrenamiento como de Prueba. En la primera fase se seleccionaron 18 individuos al azar con 4 imágenes por persona, mientras que para la fase de prueba se eligieron 23 sujetos distintos con 4 imágenes para cada uno de ellos.

## 6.2 ETAPA II. Detección de Rostros

La detección de rostros es un proceso que localiza caras en unas imágenes de entrada con diferentes condiciones de iluminación, calidad, variación en su escala y orientación del rostro en la imagen. Esta etapa es considerada como el primer paso fundamental dentro de los sistemas de reconocimiento de rostros, la cual permite optimizar el tiempo de cómputo, tamaño e información de la muestra, para localizar, identificar y clasificar el individuo correctamente.

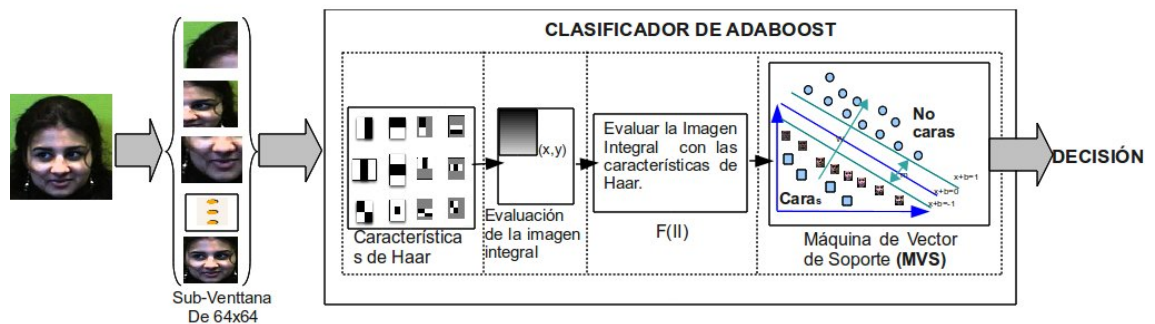
Para llevar a cabo esta fase, se diseñó un método de detección de caras basado en el modelo de *Viola Jones* [3] y enfocado en una arquitectura en cascada de N clasificadores de tipo AdaBoost. La idea principal de este clasificador es ir rechazando los falsos negativos progresivamente como se ilustra en la Figura 18, de tal forma que en las primeras etapas se eliminan un gran número de estos falsos negativos, logrando que las etapas finales se encarguen de clasificar correctamente los casos más difíciles permitiendo así, aumentar el rendimiento de discriminación entre dos posibles categorías de un conjunto de imágenes de Caras y No-Caras.

Figura 18. Estructura del sistema clasificador en cascada



De esta manera, el esquema algorítmico propuesto en la Figura 19, está dado por un conjunto de sub-imágenes de 64x64 píxeles evaluadas en cada una de las etapas de la cascada previamente seleccionadas y combinadas durante el entrenamiento, es decir que cada sub-imagen está evaluada en un conjunto de características de tipo Haar sobre la imagen integral, donde cada una de estas características está asociada a una máquina de aprendizaje de tipo vectorial, para finalmente obtener una acertada decisión.

**Figura 19** Estructura del sistema de detección de rostros.

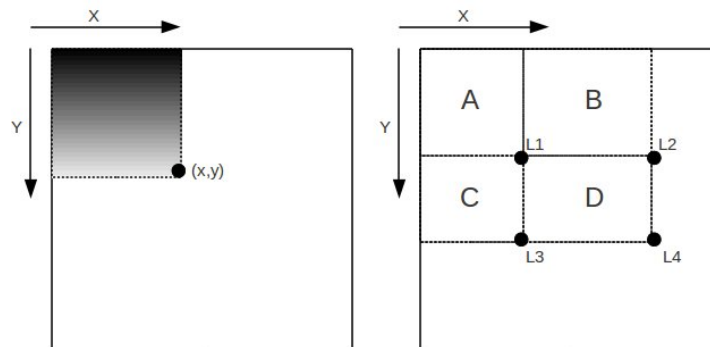


La imagen integral se define según la ecuación (1.20), como la sumatoria de los píxeles por encima y a la izquierda de la imagen ubicada en las coordenadas (x,y) tal como se muestra en la Figura 20.

$$ii(x,y) = \sum_{x=1}^n \sum_{y=1}^m i(x,y) \quad (1.20)$$

donde  $i(x,y)$  es la imagen de entrada  $\forall x = \{1,2,3, \dots, n\}$  y  $\forall y = \{1,2,3, \dots, m\}$  donde  $n * m$  representa la dimensión de la imagen  $i$ .

**Figura 20.** Representación de la imagen integral. El cuadro de la izquierda bosqueja el valor de la imagen integral en el punto  $(x,y)$  y el de la derecha representa la suma de los píxeles dentro del rectángulo  $D$ , el cual es calculado así  $L4 + L1 - (L2 + L3)$ .

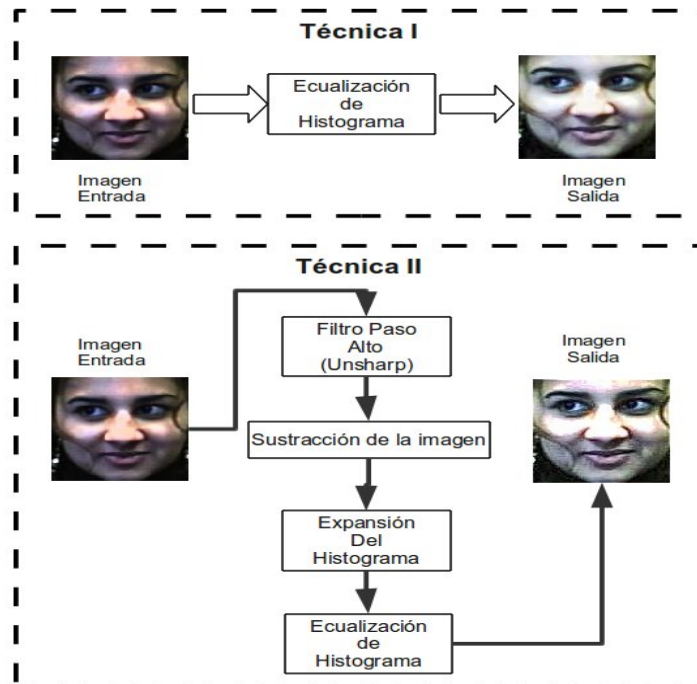


Fuente: VIOLA P. y Jones M. *Robust real-time face detection.*, **En:** Int. J. Comput. Vision. Vol. 57. pp. 137-154, Mayo 2004

### 6.3 ETAPA III. Pre-Procesamiento

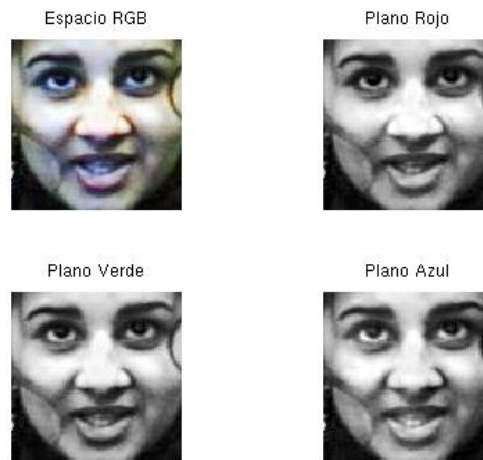
Una vez adquirida y detectada la cara del individuo, el siguiente paso es el pre-procesamiento de esta imagen, para suavizar, realzar los bordes y extraer las características de la cara detectada por medio de la información del color, con el fin de obtener un buen reconocimiento y clasificación de los rostros. En el presente trabajo se implementó un esquema de dos técnicas de pre-procesamiento tal como se ilustra en la Figura 21. La primera hace uso de la técnica de ecualización del histograma y la segunda emplea una combinación de técnicas de tipo espacial como el filtro *unsharp* y sobre píxeles como la expansión del histograma, cuyas técnicas han sido definidas en el capítulo (6.2.3.2).

**Figura 21** Técnicas de Pre-Procesamiento implementadas.



En esta etapa es importante procesar la imagen en los modelos de color RGB, HSV y en escala de grises, aplicando para cada uno de estos modelos el diseño de la figura 21, con el objetivo de comprobar la importancia del uso del color en el sistema de reconocimiento facial. La figura 22 muestra una cara que se obtiene de la etapa anterior y que por defecto se encuentra en el espacio de color RGB.

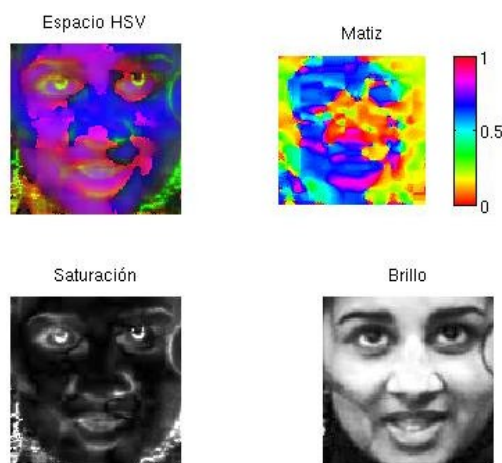
**Figura 22** Representación de la imagen en RGB y sus respectivos planos



- **Conversión al modelo HSV**

Debido a que en una imagen de color RGB no es posible separar la información cromática de la información lumínica, se estableció este tipo de modelo para extraer de forma adecuada la información del color del rostro detectado teniendo en cuenta la variación de iluminación con que fue captada la imagen. Convirtiendo los planos de RGB, en tres planos con diferente información tal como se observa en la Figura 23.

**Figura 23** Representación de la imagen en HSV y sus respectivos planos



- **Conversión al Modelo Gris**

Se hizo necesario trabajar con imágenes en grises, para poder verificar el rendimiento del sistema de reconocimiento de rostros con respecto a las imágenes en color, convirtiendo los planos de RGB, en un solo plano tal como se observa en la Figura 24.

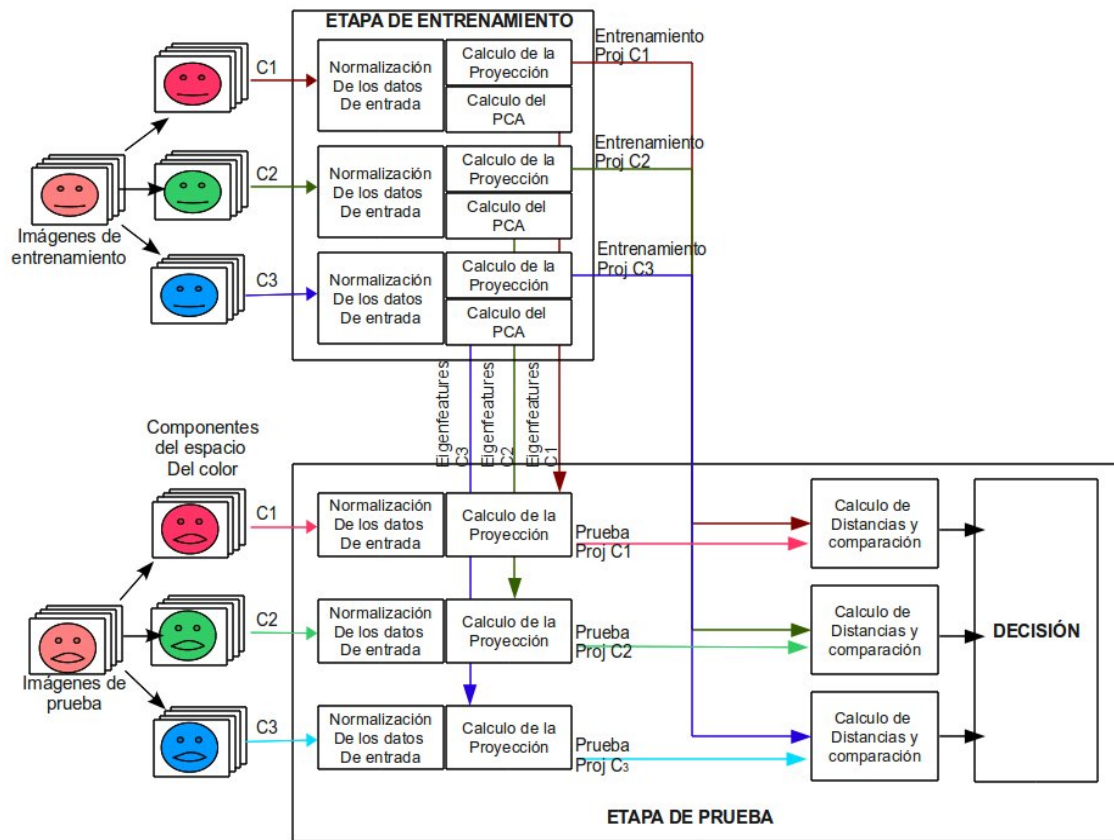
**Figura 24** Representación de la imagen en escala de grises



#### **6.4 ETAPA IV. Reconocimiento de Rostros**

Para llevar a cabo esta etapa, según la Figura 25, se vio en la necesidad de dividir este proceso en dos fases: Entrenamiento y Prueba.

Figura 25 Estructura del sistema de reconocimiento de rostros.



Fuente: TORRES L. Reutter J. y Lorente L. *The importance of the color information in face recognition*. **En:** Image Processing. Vol. 3. pp. 627,631, 1999.

Analizando la Figura 25, se observa que el proceso de normalización en la imagen se debe realizar en ambas fases para tener en cuenta ciertas independencias de las propiedades de la imagen como lo son el brillo y el contraste. Por tanto, este proceso de normalización se define así:

$$N(x, y) = \frac{NV}{Max(i) - Min(i)} * (i(x, y) - Min(x, y)) \quad (1.21)$$

donde  $i(x, y)$  representa un píxel en la imagen en escala de grises, con  $Max(i)$ ,  $Min(i)$  los valores máximos y mínimos de niveles de grises en la imagen respectivamente y  $NV$  el número de niveles.

También se observa, que por cada componente de color en la imagen, se emplea el método de Análisis de Componentes Principales, con el fin de describir mejor el conjunto de rostros a entrenar para luego ser proyectados a un espacio de caras y comparadas en la etapa de prueba, de tal forma que facilita su identificación y clasificación.

#### **6.4.1 Metodología del PCA en el reconocimiento de rostros.**

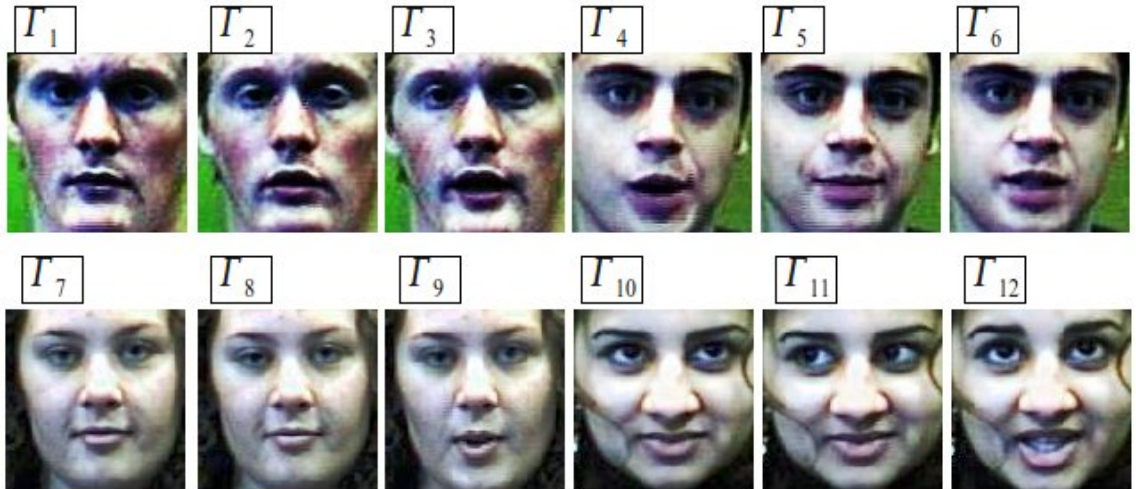
Sea  $I_c$  el conjunto de  $M$  imágenes de entrenamiento con  $\Gamma_1, \Gamma_2, \dots, \Gamma_M$ , donde cada vector  $\Gamma_i$  representa el rostro de una persona ya sea en color o en escala de grises con un tamaño de  $n$  columnas por  $y$  renglones, las cuales generan vectores de dimensión  $N$ .

Estos vectores del  $I_c$  definen un espacio vectorial denominado *Espacio de caras* de tipo uniforme no óptimo, por lo que se recomienda construir un nuevo espacio de caras con menor dimensión, los cuales contienen los componentes principales del  $I_c$ . Es posible elaborar este nuevo espacio debido a que este conjunto  $I_c$  sólo pertenecen imágenes de caras humanas

Para construir el nuevo espacio de caras, es necesario utilizar el método del PCA, el cual consiste en encontrar el vector que mejor exprese la distribución de los rostros dentro de este espacio de caras, por lo que se describirá a continuación paso a paso el desarrollo algorítmico de esta metodología.

Considere que se tiene un conjunto  $\Gamma_c$  (Ver Figura 26) con 3 imágenes diferentes de 4 personas ( $M = 12$ ). Cada imagen tiene un tamaño de  $64 \times 64$  píxeles, por lo tanto cada vector  $\Gamma_i$  tiene una dimensión de 4096 ( $N = 4096$ ).

**Figura 26** Conjunto de Imágenes de caras de entrenamiento.



En seguida, se calcula el rostro promedio  $\Psi$  (ver ecuación 1.22), el cual se muestra en la Figura 27.

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i \quad (1.22)$$

**Figura 27** Rostro Promedio, según del conjunto de imágenes de caras de entrenamiento.

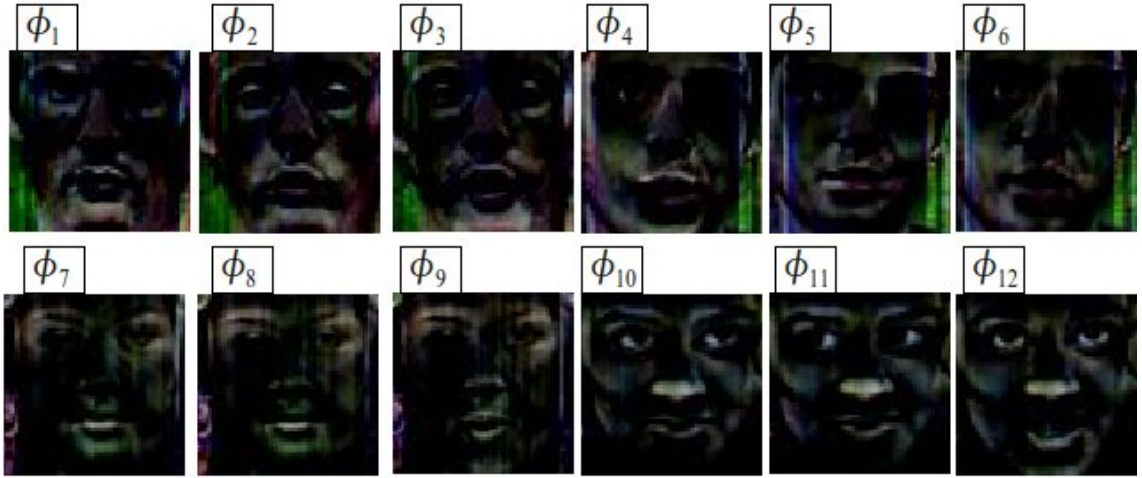


Como cada rostro  $\Gamma_i$  del conjunto  $I_c$  difiere en algún grado del rostro promedio  $\Psi$ , es necesario ajustar cada  $\Gamma_i$  con respecto a  $\Psi$  de la siguiente forma:

$$\phi_i = \Gamma_i - \Psi, \text{ donde } \phi \in \mathbb{R}^{N \times M} \quad (1.23)$$

Estos vectores  $\phi_i$  visualmente se observan de la siguiente manera:

**Figura 28** Imágenes de caras ajustadas a la media



Ahora, se buscan los  $M$  autovectores  $u_k$  ortonormales y los escalares  $\lambda_k$  que mejor describan la distribución de los datos de la matriz de covarianza  $C$ , tal que se satisfaga la ecuación 1.26.

$$C = \frac{1}{M} \sum_{i=1}^M \phi_i \phi_i^t = AA^t, \text{ donde } C \in \mathbb{R}^{M \times M} \text{ y} \quad (1.24)$$

$$A = \frac{1}{\sqrt{M}} \sum_{i=1}^M \phi_i \quad (1.25)$$

$$Cu_k = \lambda_k u_k \quad (1.26)$$

Después, se ordenan de manera descendientes los  $\lambda_k$ , con el fin de hallar los  $M'$  componentes principales definidos así:

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_{M'}$$

$$\vec{e} = [u_1, u_2, u_3, \dots, u_{M'}], \text{ donde } \vec{e} \in \mathbb{R}^{M \times M'} \quad (1.27)$$

Por último, se transforman los autovectores  $u_k$  de la ecuación (1.27) en *Eigenfaces*  $\omega_k$ , definido en la ecuación (1.28) e ilustrado en la Figura 29.

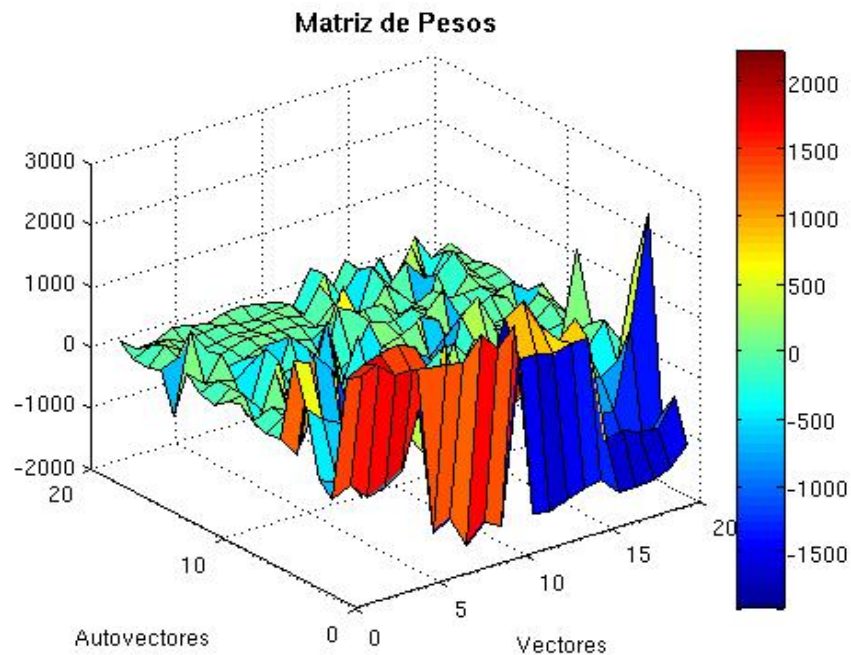
$$\omega_k = u_k * \phi \quad (1.28)$$

**Figura 29** Eigenfaces obtenidos a partir del producto escalar de los  $u_k$  de la ecuación 1.27 y los  $\phi$  de la ecuación 1.23



Posteriormente una matriz de pesos, para que se puedan reconstruir las imágenes proyectadas al Espacio de Caras. Dicha matriz de pesos, está conformada por los autovectores  $u_k$  ortornormalizados definidos en la ecuación (1.27) y los vectores  $\phi$  de la ecuación (1.23)

**Figura 30** Representación de la matriz de pesos



## 6.5 ETAPA V: Identificación y Verificación.

En esta etapa es necesario emplear un criterio de decisión para el reconocimiento e identificación del individuo. Este criterio de decisión, lo determina el argumento mínimo del cálculo de las distancias entre el conjunto de imágenes proyectadas en la fase de entrenamiento y la imagen proyectada de prueba. Con el fin de entender mejor esta etapa, suponga que se tiene una imagen de prueba  $\Gamma_p$ , la cual se desea proyectar al Espacio de Caras para su debido reconocimiento y clasificación.

Por lo tanto, para llevar a cabo la proyección de la imagen  $\Gamma_p$  al Espacio de Caras, es necesario realizar los siguientes pasos:

### Pasos para Proyectar la imagen $\Gamma_p$

1. Vectorizar la imagen  $\Gamma_p$

2. Restarle el Rostro Promedio  $\Psi$  de la Figura 27 que fue calculado en la ecuación (1.22).

$$\Phi = \Gamma_p - \Psi \quad (1.29)$$

3. Finalmente, proyectar la imagen ajustada  $\Phi$  por cada autovector  $u_k$  encontrado en la ecuación (1.27)

$$\Omega = u_k * \Phi \quad (1.30)$$

$$\Phi_f = \sum_{i=1}^M \Omega_i^t * u_k^t \quad (1.31)$$

Al reconstruir la imagen de prueba  $\Gamma_p$ , esta tiende a semejar a una imagen que pertenece a un  $k$  individuo  $\Gamma_k$  del conjunto de entrenamiento de la Figura 26. Para sostener tal afirmación es necesario calcular la distancia euclidiana y de mahalanobis para clasificar correctamente al individuo de la imagen  $\Gamma_p$ , asociándola a una imagen reconstruida semejante.

La distancia euclidiana para la clasificación de los  $k$  individuos se define según ecuación (1.32), mientras que la distancia entre la imagen de prueba ajustada  $\Phi$  y el espacio proyectado  $\Phi_f$  se define en la ecuación (1.33).

$$\varepsilon_k^2 = \|(\Omega - \omega_k)\|^2 \quad (1.32)$$

$$\varepsilon^2 = \|(\Phi - \Phi_f)\|^2 \quad (1.33)$$

Usando la definición de la distancia de Mahalanobis en la ecuación (1.13), se determina la distancia entre la proyección de los datos originales  $\omega_k$  y la proyección de los datos de prueba  $\Omega$  así:

$$d_{EM} = \sqrt{(\omega_k - \Omega)^T * inv(M) * (\omega_k - \Omega)} \quad (1.34)$$

Entonces, con las ecuaciones (1.32) y (1.34) se calculan todas las distancias formándose el siguiente vector de distancias, mostrado en la tabla 4.

**Tabla 4** Vector de distancia Euclidiana y de Mahalanobis

k	Ek	Em	Orden
1	4.5194020e+03	1.7085290e+04	12
2	4.4229152e+03	1.7165149e+04	7
3	4.4827893e+03	1.7266909e+04	8
4	4.4819889e+03	1.6540221e+04	4
5	4.0129436e+03	1.6784090e+04	5
6	4.2888224e+03	1.7942020e+04	9
7	3.5244159e+03	1.5247647e+04	3
8	3.6239903e+03	2.0038107e+04	11
9	3.2489613e+03	1.3990858e+04	2
<b>10</b>	<b>9.4644465e+02</b>	<b>1.1654872e+04</b>	<b>1</b>
11	1.4278559e+03	1.8299936e+04	10
12	2.4501232e+03	1.6935377e+04	6

De la Tabla 4, se puede deducir que la menor distancia tanto euclidiana como de mahalanobis es para  $k = 10$  por lo que la imagen  $\Gamma_p$  es semejante a la imagen  $\Gamma_{10}$  de la Figura 26. Sin embargo, si la distancia de mahalanobis como la euclidiana no exceden un umbral predefinido  $\theta_{Em}, \theta_\varepsilon$  para cada una de las distancias evaluadas respectivamente, se considerará que la persona es identificada correctamente por el sistema de reconocimiento de rostros. Este tipo de consideración se origina a partir de un criterio de decisiones, el cual permite verificar si el individuo fue identificado correctamente.

Entonces, para asegurar que un sujeto es reconocido por el sistema, es necesario utilizar el criterio de decisión para verificar que la distancia de mahalanobis evaluada por el sujeto de entrada no exceda el umbral predefinido  $\theta_{EM}$ ; es decir, suponga, que sea  $k = 1$ , el primer sujeto del conjunto de imágenes de entrenamiento. Por lo tanto, para reconocer el individuo de

entrada con respecto al conjunto de imágenes de entrenamiento (Base de Datos), se debe realizar un proceso iterativo que satisfaga la siguiente condición.

**Mientras** ( $d_{EM} < \theta_{EM}$ ) **hacer**  
 $k = k + 1$   
**Fin mientras**

Sí se satisface la condición anterior se puede asociar dicho individuo de entrada a un  $k$  individuo del conjunto de entrenamiento. Para luego, verificar si realmente ese  $k$  individuo pertenece al conjunto de entrenamiento mediante el uso de la distancia euclidiana evaluada por el sujeto de entrada con la condición de que no exceda el umbral predefinido  $\theta_\varepsilon$ , así:

Asumiendo, que el  $k = \text{argmin}(\varepsilon_k)$  existen cuatro posibilidades de verificación.

1. **Sí** ( $\varepsilon_k \leq \theta_e$ )  $\cap$  ( $\varepsilon \leq \theta_\varepsilon$ )

En este caso la imagen es identificada correctamente

2. **Sí** ( $\varepsilon_k \leq \theta_e$ )  $\cap$  ( $\varepsilon \geq \theta_\varepsilon$ )

En este caso la imagen es rechazada y el individuo evaluado  $\Gamma_p$  se considera un impostor, debido a que la imagen proyectada  $\Phi_f$  está lejos del espacio de caras.

3. **Sí** ( $\varepsilon_k \geq \theta_e$ )  $\cap$  ( $\varepsilon \leq \theta_\varepsilon$ )

En este caso la imagen  $\Gamma_p$  clasifica como un individuo desconocido, ya que la imagen  $\Gamma_p$  evaluada está cerca del espacio de caras, pero lejos de pertenecer a un  $k$  individuo.

4. **Sí** ( $\varepsilon_k \geq \theta_e$ )  $\cap$  ( $\varepsilon \geq \theta_\varepsilon$ )

En este último caso, la imagen  $\Gamma_p$  evaluada no es una cara.

## 7. RESULTADOS

Luego de construido la estructura del sistema de detección y reconocimiento de rostros, se efectuaron pruebas de validación para la clasificación correcta de un determinado individuo. Para estas pruebas se emplearon dos bases de datos internacionales, con la primera base de datos *Face 9*, se emplearon 72 imágenes de caras de entrenamiento, los cuales corresponden a 18 sujetos previamente seleccionados con 4 imágenes por cada individuo. Sin embargo, para la fase de prueba se utilizaron 100 imágenes, comprendidas entre 96 caras y 4 no caras, de las cuales sólo se detectaron 90 imágenes como caras correctamente.

De la segunda base de datos *Georgia Tech Face Database* se utilizaron 18 imágenes de caras de entrenamiento, correspondientes a 6 sujetos previamente seleccionados con 3 imágenes por cada uno de ellos. Además se seleccionaron 9 individuos distintos, donde por cada uno de ellos se escogieron 3 imágenes con diferentes poses para la fase de prueba.

Se realizaron 2 pruebas de reconocimiento de rostros comprendidas así: la primera prueba se desarrolló utilizando la técnica de ecualización del histograma para cada una de las bases de datos mencionadas y la segunda prueba se realizó utilizando la técnica número 2 de la Figura 21 para cada una de las bases de datos mencionadas. A partir de estas pruebas se crearon unas tablas de verdad con el propósito de determinar el número de individuos reconocidos o identificados correcta o incorrectamente. Estableciendo los criterios de especificidad, de sensibilidad y el valor predictivo positivo (VPP) para determinar el rendimiento del clasificador propuesto en el Capítulo (7.5).

La sensibilidad mide la probabilidad de reconocer correctamente a un individuo; la especificidad, determina la probabilidad de clasificar correctamente a un individuo que no pertenece a la base de datos ni al espacio de caras; y el valor

predictivo positivo, establece la probabilidad de clasificar correctamente a un individuo con aquellas personas que han sido catalogadas como imágenes de caras reconocidas. De acuerdo con lo anterior, los criterios para medir el rendimiento de clasificación se definen en las siguientes ecuaciones:

$$\text{Sensibilidad} = \frac{VP}{(VP + FN)} * 100 \quad (1.35)$$

$$\text{Especificidad} = \frac{VN}{(VN + FP)} * 100 \quad (1.36)$$

$$VPP = \frac{VP}{(VP + FP)} * 100 \quad (1.37)$$

donde VP son los verdaderos positivos que miden el proceso de reconocer correctamente a un individuo que pertenece a la base de datos; VN son los verdaderos negativos, el cual mide el proceso de no reconocer a un individuo que no pertenece a la base de datos ni hace parte del espacio de caras; FN mide la cantidad de individuos no reconocidos, sabiendo éste que pertenece al espacio de caras, es decir a una base de datos, por lo cual se les denomina falsos negativos; FP son los falsos positivos, que miden la cantidad de individuos reconocidos pero que no pertenecen al espacio de caras.

En las siguientes tablas, se observan los resultados obtenidos haciendo uso de las técnicas de pre-procesamiento de imágenes planteadas en la Figura 21 para cada una de las bases de datos mencionadas anteriormente.

**Tabla 5** Clasificación de los rostros empleando la técnica de ecualización del histograma para la base de datos *Face 94*.

BASE DE DATOS FACE 94			
Espacio Color	Clasificación Real	Clasificación del Rostro por el Sistema	
		Sujeto Reconocido	Sujeto No Reconocido
RGB	Sujeto Reconocido	51	0
	Sujeto No Reconocido	30	14
HSV	Sujeto Reconocido	51	0
	Sujeto No Reconocido	28	16
GRIS	Sujeto Reconocido	51	0
	Sujeto No Reconocido	29	15

**Tabla 6** Clasificación de los rostros empleando la técnica de ecualización del histograma para la base de datos *Georgia Tech*.

BASE DE DATOS GEORGIA TECH			
Espacio Color	Clasificación Real	Clasificación del Rostro por el Sistema	
		Sujeto Reconocido	Sujeto No Reconocido
RGB	Sujeto Reconocido	10	2
	Sujeto No Reconocido	14	1
HSV	Sujeto Reconocido	9	3
	Sujeto No Reconocido	12	3
GRIS	Sujeto Reconocido	9	3
	Sujeto No Reconocido	14	1

**Tabla 7** Clasificación de rostros empleando la primera técnica de la Figura 21

Base de Datos	Espacio de Color	Porcentaje %		
		Sensibilidad	Especificidad	VPP
Face 94	RGB	100	31,82	62,97
	HSV	100	36,36	64,56
	GRIS	100	34,09	63,75
George Tech	RGB	83,33	6,67	41,67
	HSV	75	20	42,86
	GRIS	75	6,67	39,13

**Tabla 8** Clasificación de los rostros empleando la segunda técnica de la Figura 21 para la base de datos *Face 94*.

BASE DE DATOS FACE 94			
Espacio Color	Clasificación Real	Clasificación del Rostro por el Sistema	
		Sujeto Reconocido	Sujeto No Reconocido
RGB	Sujeto Reconocido	51	0
	Sujeto No Reconocido	28	16
HSV	Sujeto Reconocido	51	0
	Sujeto No Reconocido	25	19
GRIS	Sujeto Reconocido	51	0
	Sujeto No Reconocido	25	19

**Tabla 9** Clasificación de los rostros empleando la segunda técnica de la Figura 21 para la base de datos *Georgia Tech*

BASE DE DATOS GEORGIA TECH			
Espacio Color	Clasificación Real	Clasificación del Rostro por el Sistema	
		Sujeto Reconocido	Sujeto No Reconocido
RGB	Sujeto Reconocido	10	2
	Sujeto No Reconocido	14	1
HSV	Sujeto Reconocido	8	4
	Sujeto No Reconocido	10	5
GRIS	Sujeto Reconocido	9	3
	Sujeto No Reconocido	13	2

**Tabla 10** Clasificación de rostros empleando la segunda técnica de la Figura 21.

Base de Datos	Espacio de Color	Porcentaje %		
		Sensibilidad	Especificidad	VPP
Face 94	RGB	100	36,36	64,56
	HSV	100	43,18	67,11
	GRIS	100	43,18	67,11
George Tech	RGB	83,33	6,67	41,67
	HSV	66,67	33,33	44,40
	GRIS	75	13,3	40,91

De los resultados obtenidos de la Tabla 7 y 10, se deduce que para ambas bases de datos, evaluadas cada una en los tres componentes de color, el sistema reconoce siempre una imagen de rostro como rostro. Sin embargo, existe la posibilidad de que un individuo que no pertenece a la base de datos de rostros utilizada, pueda ser clasificado por el sistema como un sujeto reconocido. Los porcentajes de acierto corresponden a la sensibilidad, mientras que la especificidad hace referencia a los casos en los cuales el sistema rechaza correctamente a un sujeto que no pertenece a la base de datos.

## 8. CONCLUSIONES Y RECOMENDACIONES

### 8.1 CONCLUSIONES

Los objetivos planteados al inicio de este proyecto se han cumplido satisfactoriamente, indicando que los espacios de color utilizados, proporcionan una información importante para el sistema de reconocimiento facial, en especial el espacio de color **HSV**, del cual, se empleó únicamente la componente **S** y **V** de este espacio de color, debido a que la componente **H** no proporciona mayor información característica de la imagen, debido a la no linealidad de la información del espacio de color; por la cual se vio en la necesidad de ser excluida en el desarrollo de este proyecto.

La implementación de técnicas de procesamiento digital de imágenes, permite aumentar la capacidad de reconocimiento de caras, debido al tratamiento que se le da a la imagen, la cual permite extraer mejor sus características, sin perder ninguna información de la imagen.

La implementación del algoritmo de AdaBoost, presenta ventajas en el sistema de reconocimiento facial, debido a la reducción del tiempo de clasificación para reconocer a un determinado individuo.

El uso del método de Análisis de Componentes Principales, permite convertir todo un conjunto de imágenes de caras, en un espacio de menor dimensión, ofreciendo grandes ventajas en cuanto al costo computacional y reducción de tiempo en la proyección de los datos.

La rotación de la cara en las imágenes influye directamente en el fallo del clasificador del individuo en el sistema de reconocimiento facial.

## **8.2 RECOMENDACIONES**

Debido a problemas de reconocimiento al realizar la práctica en imágenes con variaciones en sus aspectos físicos y rotación de la cara, se recomienda aplicar técnicas para optimizar los algoritmos de reconocimiento de rostros.

En la etapa de detección se presentaron problemas para la identificación de la región correspondiente al rostro, debido a la raza de los sujetos. Por esto se recomienda realizar la clasificación de los individuos por su raza, mediante la extracción de la información característica del color para agilizar el proceso de reconocimiento.

## 9. REFERENCIA BIBLIOGRAFICA

- [1] HJELMAS Erik. *Biometric Systems: A Face Recognition Approach*, 2000
- [2] TORRES L. Reutter J. y Lorente L. *The importance of the color informacion in face recognition*. **En:** Image Processing. Vol. 3. pp. 627,631, 1999
- [3] VIOLA P. y Jones M. *Robust real-time face detection.*, **En:** Int. J. Comput. Vision. Vol. 57. pp. 137-154, Mayo 2004.
- [4] MASNADI- Shirazi H. *AdaBoost face detection*. **En:** Departament of Electoral and Computer Engineering at the University of California, San Diego
- [5] TURK M. y Pentland A, *Eigenfaces for recognition*. **En:** J. Cognitive Neuroscience. Vol. 3. pp. 71-86, Febrero 1991
- [6] ABBAZIO J. Perez S. Silva D. Tesoriero R. Panna F. y Zack R. *Face biometric systems*. **En:** Proceeding of Strudent-Faculty Research Day, CSIS, Mayo 18<sup>th</sup> 2009.
- [7] KOHONEN T. *Self-organized formation of topologically correct feature maps*. **En:** Biological Cybernetics. Vol. 43. pp. 59-69, 1982
- [8] KIRBY M. Sirovich L. *Application of the karthunen-loeve procedure for the characterization of human faces*. **En:** IEEE Pattern Analysis and Machine Intelligence. Vol. 12. No.1. pp. 103-108, 1990

- [9] TURK M. Pentland P. *Face recognition using eigenfaces*. **En:** IEEE Computer Society Conference on Computer Vision and Pattern Recognition. pp. 586-591, 1991.
- [10] COX. I. Ghoson J. y Yianilos, *Feature-based face recognition using mixture-distance*. **Em:** IEEE Conference on Computer Vision and Pattern Recognition. pp. 209,216, 1996
- [11] CRAC, I. y Cameron, P. *Face Recognition by computer*. **En:** Proceedings British Machine Vision Conference, 1996, pp. 489-507.
- [12] MOGHADDAM B. y Pentland A. *Probabilistic Visual Learning for Object Representation*. **En:** Trans. IEEE Pattern Analysis and Machine Intelligence, Julio 1997, Vol. 19, pp. 696-710
- [13] BELHUMEUR P. Hespanha J. Kriegman D. *Eigenfaces vs. Fisherfaces: recognition using class specific linear projection* **En:** IEEE Transactions on Pattern Analysis and Machine Intelligence, Julio 1997. Vol. 19, pp. 711-720
- [14] SWETS D.L y Weng J. *Discriminant Analysis and Eigenspace Partition Tree for Face and Object Recognition from Views*. **En:** Proc. Intl. Conf. on Automatic Face and Gesture Recognition, 1996, pp. 192-197
- [15] ZHAO W., Chellapa R. y Krishanaswamy. *Discriminant Analysis of Principal Components for face Recognition*. **En:**, Proc. Intl. Conf. Automatic Face and Gesture Recognition, 1998, pp. 336-341
- [16] PHILLIPS P. *Support Vector Machines applied to face recognition*. **En:** Advanced Neural Information Processing Systems, 1998. Vol. 11, pp. 803-809

- [17] LI, S. y Lu, J. *Face recognition using the nearest feature line method*. **En:** IEEE Trans. Neural Network, Vol. 10, pp. 439-443.
- [18] ETEMAD K. y Chellapa R. *Discriminant Analysis for recognition of human face images*. **En:** Journal of the Optical Society of America, 1997. Vol. 14, pp. 1724-1733.
- [19] KANDE T. *Picture processing system by computer complex and recognition of human faces*. **En:** PhD Thesis, Kyoto University, Department of Information Science. Noviembre 1973
- [20] KELLY M. *Visual identification of people by computer*. **En:** Technical Report AI-130, Stanford AI Project, 1973
- [21] MANJUNATH B.S, Chellappa R. y von der Malsburg C. *A feature based approach to Face Recognition*. **En:** Proc. IEEE Computer Soc. Conf. on Computer Vision and Pattern Recognition, 1992, pp. 373-378
- [22] SAMARIA, F. *Face Recognition Using Hidden Markov Models*. **En:** Ph.D. Thesis, University of Cambridge, 1994
- [23] LAWRENCE S. Giles C. Tsoi A. y Back A. *Face Recognition: A Convolutional Neural Network Approach*. **En:** IEEE Transactions on Neural Networks, 1997. Vol. 8, pp. 98-113
- [24] PENTLAND A. Moghaddam B. y Starner T. *View-based and Modular Eigenspaces for Face Recognition* **En:** Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 1994
- [25] HUANG J. Heisele B. y Blanz V. *Component-based face recognition with 3D morphable models*. **En:** Audio and Video based Biometric Person Authentication, pp. 62-69, 2003

- [26] REISFELD D. y Yeshurun, *Robust Detection of Facial Features, by Generalized Symmetry*. **En:** Proc. Of Intl. Conf. on Pattern Recognition, 1992, pp. 117-120
- [27] VALENTIN N. André Melo y Aurélio Campilho, *Location of Eyes in Face Images*. **En:** Proceedings of 12<sup>th</sup> Portuguese Conference on Pattern Recognition (RECPAD), Aveiro, 2002
- [28] PERLIBAKAS V. *Automatical detection of face features and exact face contour*. **En:** Pattern Recognition Letters, 2003, Vol. 24, pp. 2977-2985
- [29] THOMAZ C, D. Gillies y Raul Q. Feitosa, *Using mixture covariance matrices to improve face and facial expression recognitions*. **En:** Pattern Recognition Letters, 2003, Vol. 24, pp. 2159-2165
- [30] KIM H, Daijin Kim, Sung Yang Bang y Sang-Youn Lee, *Face Recognition using the second-order mixture-of-eigenfaces method*. **En:** Pattern Recognition, 2004, Vol. 37, pp. 337-349
- [31] GOTTUMUKKAL R, Vijayan K. Asari, *An improved face recognition technique based on modular PCA approach*. **En:** Pattern Recognition Letters, 2004, Vol. 25, pp. 429-436.
- [32] SWAIN M:J y Ballar D.H. *Color indexing*. **En:** International Journal of Computing Vision, Vol. 7, No. 1, pp. 11-32, 1991.
- [33] HARALICK R. y L:G. Shapiro. *Survey: Image segmentation techniques*. **En:** Computer Vision, Graphics, and Image Processing. Vol. 29, pp. 100-132, 1985.

- [34] BRUNELLI R. y Poggio, *Face Recognition: Features versus Templates*. **En:** IEEE Trans. Pattern Anal. y Mach. Intell, 1993, Vol. 15, pp. 1042-1052
- [35] HEISELE B. Purby Ho, Jane Wu y Tomaso Poggio, *Face recognition: component-based versus global approaches*. **En:** Computer Vision and Image Understanding 91, 2003, pp. 6-21
- [36] YILLE A, D. Cohen y P. Hallinan, *Feature extraction from faces using deformable templates* **En:** Proc. IEEE Comp Soc. Conf. on Comp. Vis. and Patt. Recog, 1989, pp. 104-109
- [37] MOGHADDAM B, Toni Jebara y Alex Pentland, *Bayesian Face Recognition*. **En:** Pattern Recognition, 2000, Vol. 33, pp. 1771-1782
- [38] XUE Z. Stan Z Li y Eam Khwang Teoh. *Bayesian shape model for facial feature extraction and recognition*. **En:** Pattern Recognition, 2003, Vol. 36, pp. 2819-2833.
- [39] HE X, S. Yan, Y. Hu, P. Niyogi, y H.J. Zhang, *Face recognition using laplacianfaces*, **En:** Pattern Analysis and Machine Intelligence, IEEE Transactions, 2005. Vol.27,No. 3, pp. 328-340
- [40] XU Y, Jing-Yu Yang y Zhong Jin, *A novel method for Fisher discriminant analysis*. **En:** Pattern Recognition, 2004, Vol. 37, pp. 381-384
- [41] LU J, K. Plataniotis y A. Venetsanopoulos. *Face Recognition using LDA-based algorithms*. **En:** Neural Network IEEE Transaction. Vol. 14, pp. 195-200, Junio 2003.

- [42] BARLETT, Lades, Sejnowski, *Independent component representation for face recognition* **En:** Proceedings, SPIE Symposium on Electronic Imaging: Science and Technology, 1998, pp. 528-539
- [43] YUEN, P.C y J.H. Lai, *Independent Component Analysis of Face Images*. **En:** IEEE Workshop on Biologically Motivated Computer Vision, Seoul, Springer-Verlag, 2000
- [44] DÉNIZ O, M. Castrillón y M. Hernández, *Face Recognition using Independent Component Analysis and Support Vector Machines*. **En:** Pattern Recognition Letters, 2003, Vol. 24, pp. 2153-2157
- [45] LIU Z, y C. Liu, *A hybrid color and frequency features method for face recognition*. **En:** Image Processing, IEEE Transactions, 2008. Vol. 17, No. 10, pp. 1975-1980
- [46] KARIMI y A. Krzyzak, *A Study on significance of color in face recognition using several eigenfaces algorithms*. **En:** Electrical and Computer Engineering, 2007, pp. 1309-1312

## 10. BIBLIOGRAFÍA

KARIMI y A. Krzyzak, *A Study on significance of color in face recognition using several eigenfaces algorithms*. **En:** Electrical and Computer Engineering, 2007, pp. 1309-1312.

Lorente. L y Torres L, *A global eigen approach for face recognition*, **En:** International Workshop on Very Low Bit-rate Video Coding, Urbana, Illionis, USA, Octubre 8-9, 1998.

Masnadi H. *Adaboost face detection*, **En:** Department of electoral and computer engineering at the university of California, San Diego.

MOON H. y PHILLIPS, *Computational and performance aspects of pca-based face-recognition algorithms*, Vol. 30 No 3, pp. 303-321, 2001.

MORIZET Nicolas, Amiel Frédéric, Dris Hamed Insaf, Ea Thomas. *A Comparative Implementation of PCA face recognition algorithms*. **En:** IEEE, 2007.

SWAIN M:J y Ballar D.H. *Color indexing*. **En:** International Journal of Computing Vision, Vol. 7, No. 1, pp. 11-32, 1991.

TORRES L. Reutter J. y Lorente L. *The importance of the color informacion in face recognition*. **En:** Image Processing. Vol. 3. pp. 627,631, 1999.

TURK M. Pentland P. *Face recognition using eigenfaces*. **En:** IEEE Computer Society Conference on Computer Vision and Pattern Recognition. pp. 586-591, 1991.

TURK M. y Pentland A, *Eigenfaces for recognition*. **En:** J. Cognitive Neuroscience. Vol. 3. pp. 71-86, Febrero 1991.

VIOLA P. y Jones M. *Robust real-time face detection.*, **En:** Int. J. Comput. Vision. Vol. 57. pp. 137-154, Mayo 2004.

YIP, A. y Sinha P. *Role of color in face recognition*, **En:** AIM-2001-035, CEBCL-212, Diciembre 2001.