



**INSTITUTO POLITÉCNICO NACIONAL**  
**ESCUELA SUPERIOR DE INGENIERÍA**  
**MECÁNICA Y ELÉCTRICA UNIDAD CULHUACAN**

**SECCIÓN DE ESTUDIOS DE POSGRADO E**  
**INVESTIGACIÓN**

**RECONOCIMIENTO DE ROSTROS UTILIZANDO ANÁLISIS DE**  
**COMPONENTES PRINCIPALES Y ECUALIZACIÓN DE**  
**HISTOGRAMA**

**T É S I S**

QUE PARA OBTENER EL GRADO DE:  
**MAESTRO EN CIENCIAS DE LA INGENIERÍA EN MICROELECTRÓNICA**  
P R E S E N T A:  
**KELSEY ALEJANDRA RAMÍREZ GUTIÉRREZ**

**ASESORES:**  
DR. HÉCTOR MANUEL PÉREZ MEANA  
DR. DANIEL CRUZ PÉREZ



**MÉXICO, D. F.**  
**JUNIO 2010**



# INSTITUTO POLITÉCNICO NACIONAL SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

## ACTA DE REVISIÓN DE TESIS

En la Ciudad de México D. F. siendo las 11:00 horas del día 01 del mes de junio del 2010 se reunieron los miembros de la Comisión Revisora de Tesis, designada por el Colegio de Profesores de Estudios de Posgrado e Investigación de SEPI-ESIME-CULH. para examinar la tesis titulada:

"Reconocimiento de Rostros Utilizando Análisis de Componentes Principales y Ecuación de Histograma"

Presentada por el alumno:

<u>Ramírez</u> Apellido paterno	<u>Gutiérrez</u> Apellido materno	<u>Kelsey Alejandra</u> Nombre(s)												
		Con registro: <table border="1" style="display: inline-table;"><tr><td>B</td><td>0</td><td>8</td><td>1</td><td>8</td><td>3</td><td>7</td></tr></table>						B	0	8	1	8	3	7
B	0	8	1	8	3	7								

aspirante de:

Maestría en Ciencias de Ingeniería en Microelectrónica

Después de intercambiar opiniones los miembros de la Comisión manifestaron **APROBAR LA TESIS**, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.

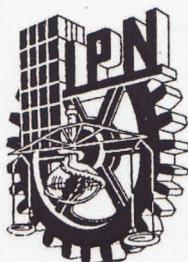
### LA COMISIÓN REVISORA

Directores de tesis

 _____ Dr. Héctor Manuel Pérez Meana	 _____ Dr. Daniel Cruz Pérez
 _____ Dra. Mariko Nakano Miyatake	 _____ Dr. Gabriel Sánchez Pérez
 _____ Dr. Gualberto Aguilar Torres	

**SEPI**  
SECCIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN  
PRESIDENTE DEL COLEGIO DE PROFESORES  
ESIME CUERNAVACA

\_\_\_\_\_  
Dr. Gonzalo Isaac Duchén Sánchez

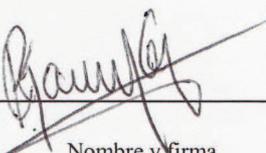


**INSTITUTO POLITÉCNICO NACIONAL**  
**SECRETARÍA DE INVESTIGACIÓN Y POSGRADO**

**CARTA CESIÓN DE DERECHOS**

En la Ciudad de México el día 14 del mes Junio del año 2010, el (la) que suscribe Kelsey Alejandra Ramírez Gutiérrez alumno (a) del Programa de Maestría en Ciencias de Ingeniería de Microelectrónica con número de registro B081837, adscrito a la Sección de Estudios de Posgrado del ESIME Unidad Culhuacan, manifiesta que es autor (a) intelectual del presente trabajo de Tesis bajo la dirección de el Dr. Héctor Manuel Pérez Meana y el Dr. Daniel Cruz Pérez y cede los derechos del trabajo intitulado Reconocimiento de Rostros Utilizando Análisis de Componentes Principales y Ecuación de Histograma, al Instituto Politécnico Nacional para su difusión, con fines académicos y de investigación.

Los usuarios de la información no deben reproducir el contenido textual, gráficas o datos del trabajo sin el permiso expreso del autor y/o director del trabajo. Este puede ser obtenido escribiendo a la siguiente dirección kelseyramirez@gmail.com. Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.

  
Nombre y firma

*A Dios por ser mi refugio y fortaleza en la angustia.*

*A mis padres por su apoyo incondicional.*

*A mis amigos por su compañía.*

## AGRADECIMIENTOS

*El presente trabajo es la culminación de una etapa más en mi vida llena de esfuerzos, retos, aprendizaje y experiencias, con la presencia de muchas personas, que en el transcurso de estos dos años me han acompañado y apoyado.*

*En primer lugar agradezco al Instituto Politécnico Nacional y a la Sección de Estudios de Posgrado e Investigación de la Escuela Superior de Ingeniería Mecánica y Eléctrica Culhuacán por darme el honor de realizar mis estudios de maestría.*

*A mi asesor Dr. Héctor Manuel Pérez Meada, por apoyarme y brindarme su enseñanza en todo momento. Al Dr. Daniel Cruz Pérez y Dr. Gabriel Sánchez Pérez por su apoyo, franqueza y ayuda durante este proyecto. Al M. en C. Jesús Olivares Mercado por apoyarme con sus explicaciones.*

*A mis padres Agustina y Byron, por apoyarme siempre y por ser un ejemplo e inspiración en mi vida.*

*Al CONACYT por brindarme el soporte económico para realizar este trabajo de investigación.*

*A mis amigos y compañeros de SEPI por haberme brindado su apoyo durante todo este tiempo.*

*A la Universidad Nacional de Ingeniería por facilitarme esta gran oportunidad.*

## RESUMEN

Los sistemas biométricos son de gran importancia por sus múltiples aplicaciones, las que abarcan desde aplicaciones comerciales hasta aplicaciones de seguridad, en las cuales se requiere de una alta eficiencia. El reconocimiento de rostro es un área de investigación con muchas aplicaciones desde los años 80. Es probablemente el método biométrico más fácil de entender ya que identifica a las personas por sus rostros, tal como lo realizamos los seres humanos.

La presente tesis propone la ecualización de histograma como una fase en el pre-procesamiento en la que se tratará de llevar los histogramas de los rostros a la uniformidad, este procedimiento se realizará de diferentes maneras utilizando también la Transformada Rápida de Fourier para obtener la fase de los rostros y posteriormente extraer sus características utilizando el Análisis de Componentes Principales. Una vez obtenidos los vectores característicos de cada persona se entrenará la Máquina de Soporte Vectorial que se utilizará como clasificador.

La presente tesis cuenta con un Marco Teórico donde se explican las principales características de los métodos utilizados en el sistema propuesto. Los resultados del sistema se presentan en el capítulo IV.

## ABSTRACT

Biometric systems are of great importance because of its multiple applications, ranging from business applications to security applications, which requires high efficiency. The face recognition is a research area with many applications since the '80s. Face Recognition is probably the easiest biometric method to understand because it identifies people by their faces, as human beings do.

This work proposes the histogram equalization as a phase in the pre-processing which will attempt to bring the histograms of the faces to uniformity, this procedure is carried out in different ways and also use the Fast Fourier Transform to obtain the phase faces and then extract their features using Principal Component Analysis. Once the feature vectors are obtained from each person the Support Vector Machine will be trained to be used as classifier.

This thesis has a theoretical framework that explains the main features of the methods used in the proposed system. The results of the system are presented in Chapter IV.

# CONTENIDO

Dedicatoria	i
Agradecimientos	ii
Resumen	iii
Abstract	iv
Contenido	v
Lista de Figuras	vii
Lista de Tablas	ix
1. Introducción	1
1.1. Justificación	3
1.2. Objetivos	4
1.2.1. Objetivo General	4
1.2.2. Objetivos Específicos	4
1.3. Organización de la Tesis	4
2. Marco Teórico	6
2.1. Introducción	6
2.1.1. Estructura General de un Sistema Biométrico	7
2.2. Aplicaciones de la Biometría	7
2.3. Reconocimiento Facial	8
2.3.1. Enfoques predominantes	8
2.4. Métodos de Extracción de Características	12
2.4.1. Eigenfaces	12
2.4.2. Filtro de Gabor	15
2.4.3. Transformada Discreta de Walsh	17
2.4.4. Análisis de Componentes Principales	20
2.5. Clasificadores	21
2.5.1. Máquinas de Soporte Vectorial	22
2.6. Ecuilización de Histograma	33
2.7. Transformada Rápida de Fourier	40
2.8. Conclusiones	42
3. Sistema Propuesto	44
3.1. Introducción	44
3.2. Estructura del Sistema Propuesto	44
3.3. Pre-procesamiento	45
3.4. Extracción de Características	48
3.5. Entrenamiento SVM	49

3.6. Reconocimiento SVM	50
3.7. Verificación SVM	50
3.8. Conclusiones	51
4. Resultados	52
4.1. Introducción	52
5. Conclusiones Generales	62
5.1. Conclusiones	62
5.2. Trabajos Futuros	63
6. Bibliografía	64
Anexo A	67
Glosario	67
Anexo B	69
Publicaciones	69

## LISTA DE FIGURAS

<b>2.1</b>	Estructura de un sistema biométrico.....	7
<b>2.2</b>	La cota del riesgo real es igual a la suma del riesgo empírico y el intervalo de confianza.....	25
<b>2.3</b>	Hiperplano de separación óptima para el caso bidimensional.....	29
<b>2.4</b>	Variaciones de nivel de gris.....	35
<b>2.5</b>	Una función de transformación de niveles de gris que satisface las condiciones (a) y (b) .....	37
<b>2.6</b>	Imágenes ecualizadas y sus respectivos histogramas.....	40
<b>2.7</b>	Experimento de Oppenheim.....	42
<b>3.1</b>	Sistema de reconocimiento propuesto.....	44
<b>3.2</b>	Pre-procesamiento sin ecualización.....	45
<b>3.3</b>	Pre-procesamiento con ecualización.....	46
<b>3.4</b>	Pre-procesamiento con ecualización por ventana.....	47
<b>3.5</b>	Pre-procesamiento con ecualización por ventana y FFT.....	48
<b>3.6</b>	Diagrama general de entrenamiento de SVM.....	49
<b>3.7</b>	Diagrama de reconocimiento.....	50
<b>3.8</b>	Diagrama de verificación.....	51
<b>4.1</b>	Grupo A de rostros.....	52

<b>4.2</b> Grupo B de rostros.....	<b>52</b>
------------------------------------	-----------

## LISTA DE TABLAS

<b>4.1</b>	Reconocimiento sin ecualización.....	53
<b>4.2</b>	Reconocimiento con ecualización.....	53
<b>4.3</b>	Reconocimiento con ecualización por ventana de 3x3.....	54
<b>4.4</b>	Reconocimiento con ecualización por ventana de 6x6.....	54
<b>4.5</b>	Reconocimiento con ecualización por ventana de 3x3 y FFT.....	54
<b>4.6</b>	Reconocimiento con ecualización por ventana de 6x6 y FFT.....	55
<b>4.7</b>	Verificación sin ecualización.....	55
<b>4.8</b>	Verificación con ecualización.....	56
<b>4.9</b>	Verificación con ecualización por ventana de 3x3.....	56
<b>4.10</b>	Verificación con ecualización por ventana de 6x6.....	57
<b>4.11</b>	Verificación con ecualización por ventana de 3x3 y FFT.....	57
<b>4.12</b>	Verificación con ecualización por ventana de 6x6 y FFT.....	58
<b>4.13</b>	Verificación para control de acceso sin ecualización.....	59
<b>4.14</b>	Verificación para control de acceso con ecualización.....	59
<b>4.15</b>	Verificación para control de acceso con ecualización por ventana de 3x3.....	59
<b>4.16</b>	Verificación para control de acceso con ecualización por ventana de 6x6.....	60

<b>4.17</b> Verificación para control de acceso con ecualización por ventana de 3x3 y FFT.....	60
<b>4.18</b> Verificación para control de acceso con ecualización por ventana de 6x6 y FFT.....	61

# CAPÍTULO 1

## INTRODUCCIÓN

El reconocimiento de rostro ha recibido recientemente una considerable atención [1], [2] debido a que juega un papel importante en muchas áreas de aplicación, tales como la interacción hombre-máquina, la autenticación y la vigilancia. [3]

La biometría consiste en un conjunto de métodos automáticos para el reconocimiento o verificación de identidad de las personas que utilizan físicas o de comportamiento de esa persona. Esta tecnología se basa en el hecho de que cada persona es única y tiene características distintivas que pueden ser utilizados para identificar.

El reconocimiento de rostro es un área activa de investigación con muchas aplicaciones desde los años 80. Es probablemente el método más fácil de entender biométricos porque identificar a las personas por sus caras es el proceso de reconocimiento utilizado por el cerebro humano y el cual no tiene una explicación concreta.

Es esencial tener un sistema confiable de seguridad en oficinas, bancos, empresas, tiendas, etc. Los sistemas de reconocimiento y/o verificación biométrica es una buena alternativa para el desarrollo de estos sistemas. [4]

En las últimas dos décadas, el problema del reconocimiento de rostros ha atraído considerable atención de varias disciplinas y ha sido testigo de un impresionante crecimiento en la investigación básica y aplicada, desarrollo de productos y aplicaciones. Los sistemas de reconocimiento

de rostro ya se han desplegado en los puertos de entrada en los aeropuertos internacionales en Australia y Portugal. [5]

En la actualidad, la seguridad es un tema que tiene gran importancia a nivel mundial. [6] El fraude de identidad se está convirtiendo en un gran problema para los gobiernos alrededor del mundo. Los sistemas biométricos han surgido como una respuesta real para la solución de estos problemas.

## 1.1 JUSTIFICACIÓN

Se desea desarrollar un sistema que sea capaz de reconocer personas a partir de sus características faciales y compararlas con una base de datos para identificar finalmente si se trata de una persona registrada o no. Así mismo el algoritmo deberá ser capaz de verificar si una persona es quien dice ser. El porcentaje de acierto debe de ser alto para que el sistema sea lo más confiable posible.

## 1.2 OBJETIVOS

### 1.2.1 OBJETIVO GENERAL

- Desarrollar un sistema para reconocimiento y verificación de rostros con un porcentaje de acierto superior al 90% y que sea además robusto a variaciones en la oclusión e iluminación.

### 1.2.2 OBJETIVOS ESPECÍFICOS

- Mejorar la calidad de la imagen mediante la ecualización de Histograma.
- Extraer las características de los rostros utilizando Análisis de Componentes Principales.
- Reconocer y Verificar los rostros utilizando la Máquina de Soporte Vectorial.
- Analizar el uso del sistema utilizando la verificación para control de acceso.
- Evaluar los resultados obtenidos.

### 1.3 ORGANIZACIÓN DE LA TESIS

- **Capítulo 1:** Se presenta un panorama general del desarrollo de la tesis, la justificación y los objetivos.
- **Capítulo 2:** Se presenta una introducción al concepto de biometría y reconocimiento de rostros, así como la explicación de los métodos utilizados en este trabajo.
- **Capítulo 3:** Se presenta la explicación del sistema propuesto; los diferentes tipos de pre-procesamiento y como se realiza el reconocimiento y verificación.
- **Capítulo 4:** Se presentan los resultados obtenidos utilizando cada variación de pre-procesamiento.
- **Capítulo 5:** Se presentan las conclusiones a las que se llegaron a lo largo de la realización del trabajo, y los posibles trabajos futuros.

## CAPÍTULO 2

### MARCO TEÓRICO

#### 2.1 INTRODUCCIÓN

Biometría es un sistema automatizado de reconocimiento humano basado en las características físicas y comportamiento de las personas. Es el mismo sistema que utiliza el cerebro humano para reconocer y distinguir una persona de la otra.

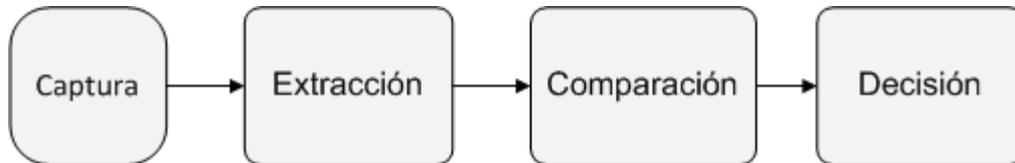
Las tecnologías biométricas, por lo tanto, se refieren a las partes físicas del cuerpo humano o los rasgos personales de los seres humanos. El término "automático" significa esencialmente que la tecnología biométrica debe reconocer o verificar una característica humana rápida y automáticamente, en tiempo real. Las características de una persona son únicas e inalterables lo que hace que estos sistemas proporcionen un alto nivel de seguridad.

Existen básicamente dos tipos de características que se pueden extraer de una persona, las cuales son:

*Física:* son los datos biométricos del ojo (iris y retina), rostro, huella dactilar, mano y la voz.

*Comportamiento:* incluye el ritmo de escritura y la firma.

### 2.1.1 ESTRUCTURA GENERAL DE UN SISTEMA BIOMÉTRICO:



**Figura 2.1** Estructura de un sistema biométrico

- 1) Captura: una muestra física o de comportamiento es capturado por el sistema.
- 2) Extracción: se extraen datos únicos de la muestra y se crea una base de datos.
- 3) Comparación: el modelo se compara con una nueva muestra.
- 4) Decisión: Coincide /no coincide; el sistema decide entonces si las características extraídas de la muestra coinciden o no con la nueva.

## 2.2 APLICACIONES DE LA BIOMETRÍA

- Seguridad para el ordenador, softwares y la red informática.
- Comercio electrónico.
- Seguridad para la banca, instituciones financieras.
- Medio de pago por huella digital
- Sistema de información médica.
- Muchas otras aplicaciones que requieren autenticación de usuario.
- Control de Acceso biométrico a centros deportivos.
- Control de Presencia mediante terminal.
- Pago por huella digital.

En la industria de la seguridad, la biometría se considera como la proveedora del más alto nivel de seguridad. Los métodos de

verificación de la identidad de un individuo se desglosan en las siguientes tres etapas:

- *Etapa 1* (nivel más bajo de seguridad) - algo que se posee, como una identificación con foto.
- *Etapa 2* (segundo nivel de seguridad) - algo que usted sabe, como una contraseña para acceder a un ordenador o un número de identificación personal (PIN) para acceder a fondos en un cajero automático.
- *Etapa 3* (nivel más alto de seguridad) - algo que usted hace o algo que usted es, que comprende comportamientos fisiológicas y / o biométricos, incluyendo huellas digitales, voz, firmas, etc.

## 2.3 RECONOCIMIENTO FACIAL

El reconocimiento y verificación facial es uno de los sectores de más rápido crecimiento en la industria de la biometría. Su atractivo radica en el hecho de que es el que más se asemeja a la forma en que los seres humanos nos identificamos entre nosotros. La mayoría de los desarrolladores han tenido dificultad para alcanzar altos niveles de rendimiento. Sin embargo, específicas aplicaciones, como el cribado de las bases de datos del seguro social para duplicados y de los aeropuertos, es probable que aparezcan en el futuro.

### 2.3.1 ENFOQUES PREDOMINANTES

Hay dos enfoques predominantes en el problema de reconocimiento facial:

El geométrico (basado en rasgos) y el fotométrico (basado en lo visual). Conforme al interés investigador en reconocimiento facial continuo, fueron desarrollados muchos algoritmos diferentes, tres de los cuales han sido bien estudiados en la literatura del reconocimiento facial:

Análisis de componentes principales (Principal Components Analysis, PCA), Análisis lineal discriminante (Linear Discriminant Analysis, LDA), y Correspondencia entre agrupaciones de grafos elásticos Elastic Bunch Graph Matching, EBGM).

#### 2.3.1.1 ANÁLISIS DE COMPONENTES PRINCIPALES (PRINCIPAL COMPONENT ANALYSIS, PCA)

PCA, comúnmente referida al uso de Eigenfaces, es la técnica impulsada por Kirby & Sirovich en 1988. Con PCA, el sondeo y la galería de imágenes deben ser del mismo tamaño y deben ser normalizadas previamente para alinear los ojos y bocas de los sujetos en las imágenes. La aproximación de PCA es luego utilizado para reducir la dimensión de los datos por medio de fundamentos de compresión de datos y revela la más efectiva estructura de baja dimensión de los patrones faciales.

Esta reducción en las dimensiones quita información que no es útil [7] y descompone de manera precisa la estructura facial en componentes ortogonales (no correlativos) conocidos como Eigenfaces. Cada imagen facial puede ser representada como una suma ponderada (vector de rasgo) de los eigenfaces, las cuales son almacenadas en un conjunto 1D.

Una imagen de sondeo es comparada con una galería de imágenes midiendo la distancia entre sus respectivos vectores de rasgos. La aproximación PCA típicamente requiere la cara completa de frente para ser presentada cada vez; de otra forma la imagen dará un resultado de bajo rendimiento.

La ventaja primaria de esta técnica es que puede reducir los datos necesarios para identificar el individuo a 1/1000 de los datos

presentados.[8]

### 2.3.1.2 ANÁLISIS LINEAL DISCRIMINANTE (LINEAR DISCRIMINANT ANALYSIS, LDA)

LDA es una aproximación estadística para clasificar muestras de clases desconocidas basadas en ejemplos de entrenamiento con clases conocidas.[9] Esta técnica tiene la intención de maximizar la varianza entre clases (ej. Entre usuarios) y minimizar la varianza de cada clase (Ej. De cada usuario). Hay grandes variaciones entre clases, pero pequeñas en cada clase. Cuando se trata con datos faciales de alta dimensión, esta técnica enfrenta el problema de muestras de tamaño pequeño que surge donde hay un número pequeño de ejemplos de entrenamiento comparados a la dimensionalidad del espacio de muestra.[10]

### 2.3.1.3 CORRESPONDENCIA ENTRE AGRUPACIONES DE GRAFOS ELÁSTICOS ELASTIC BUNCH GRAPH MATCHING, EBGM

EBGM tiene en cuenta que las imágenes faciales reales tienen muchas características no lineales que no son tratadas en los métodos lineales de análisis discutidos previamente, tales como variaciones en la iluminación (Iluminación de exteriores vs. Interior fluorescente), postura (frontal vs. inclinada) y expresión (sonrisa vs. ceño fruncido).

Una ondeleta de transformación Gabor crea una arquitectura de enlace dinámico que proyecta el rostro sobre la planilla elástica. El Jet Gabor es un nodo en la planilla elástica, manifestado por círculos en la imagen debajo. El cual describe el comportamiento de la imagen alrededor de un píxel.

Este es el resultado de una convolución de la imagen con un filtro Gabor, el cual es usado para detectar formas y extraer características utilizando procesamiento de imagen (una convolución expresa la suma de solapamientos de las funciones en la mezcla de funciones entre sí). El reconocimiento está basado en la similitud de la respuesta del filtro Gabor a cada nodo Gabor.

Este método biológicamente basado utilizando filtros Gabor es un proceso ejecutado en la corteza visual de los mamíferos más grandes. La dificultad con este método es el requerimiento de la precisa localización del punto de referencia el cual puede ser algunas veces logrado combinando los métodos PCA y LDA.

## 2.4 MÉTODOS DE EXTRACCIÓN DE CARACTERÍSTICAS

### 2.4.1 EINGENFACES

Para el método de Eigenfaces, el vector formado por todos los píxeles de una imagen se proyecta en un espacio vectorial relativamente pequeño, obteniéndose entonces un vector representativo de menor dimensión. Todos los vectores de las imágenes de un individuo formarán una vecindad dentro de ese espacio vectorial reducido; y todas las imágenes que si quiera contienen un rostro formarán una sección limitada dentro de ese espacio llamado "Espacio de rostros".[11]

Se va a calcular un conjunto de M Eigenfaces a partir de un conjunto base de M imágenes de NxN pixeles.

Sean  $\Gamma_1, \Gamma_2, \Gamma_3 \dots \Gamma_M$  el conjunto vectores  $N^2 \times 1$  de las imágenes base, se calcula el rostro promedio

$$\Psi = (\Gamma_1 + \Gamma_2 + \Gamma_3 + \dots + \Gamma_M)/M \quad (1)$$

Se calcula la diferencia de cada imagen a la imagen promedio  $\Psi$

$$\Phi_i = \Gamma_i - \Psi \quad i = 1, 2, \dots, M \quad (2)$$

Los vectores  $\Phi_i, \Gamma_i$  y  $\Psi_i$  resultarían de longitud NxN,

Suponiendo que  $\lambda_k$  y  $U_k$  son el k-esimo eigenvalor y el k-esimo eigenvector respectivamente, partiremos de la definición de eigenvalores y eigenvectores:

$$U_k^T C U_k = \lambda_k \quad (3)$$

donde C es la matriz de covarianza de las imágenes:

$$C = \frac{1}{M} \sum_{n=1}^M \Phi_n \Phi_n^T = AA^T \quad (4)$$

Donde

$$A = \frac{[\varphi_1^T, \varphi_2^T, \varphi_3^T, \dots, \varphi_M^T]}{\sqrt{M}}$$

Ambos lados de la ecuación (3) se multiplican por la izquierda por el vector  $U_k$

$$U_k^T U_k C U_k = \lambda_k U_k \quad (5)$$

Dado que los vectores  $U_k$  son ortonormales:

$$U_k^T U_k = 1$$

Entonces la ecuación (4) se reduce a

$$C U_k = \lambda_k U_k \quad (6)$$

Sustituyendo (4) en (6)

$$C U_k = A A^T \quad (7)$$

Se multiplica por la izquierda ambos lados de (7) por  $A^T$

$$\begin{aligned} A^T A A^T U_k &= \lambda_k A^T U_k \\ A^T A (A^T U_k) &= \lambda_k (A^T U_k) \end{aligned} \quad (8)$$

Y haciendo

$$Y_k = A^T U_k \quad (9)$$

la ec. (8) resulta

$$A A^T Y_k = \lambda_k Y_k \quad (10)$$

Considerando que  $\lambda_k$  es eigenvalor de  $A A^T$  entonces  $X_k = \alpha_k Y_k$  es eigenvector de la matriz  $A^T A$ . (el tamaño de  $A^T A$  es  $M \times M$ ; significativamente menor que  $N \times N$ )

$$X_k^T X_n = (\alpha_k Y_k)(\alpha_n Y_n)$$

$$\begin{aligned}
X_k^T X_n &= (\alpha_k A^T U_k)^T (\alpha_n A^T U_n) \\
X_k^T X_n &= \alpha_k \alpha_n (A^T U_k)^T (A^T U_n) \\
X_k^T X_n &= \alpha_k \alpha_n (U_k^T A) A^T U_n
\end{aligned} \tag{11}$$

Desde la ec.(7)

$$AA^T U_k = \lambda_k U_k \tag{12}$$

y sustituyendo (12) en (11)

$$X_k^T X_n = \alpha_k \alpha_n U_k^T \lambda_n U_n \tag{13}$$

Cuando  $k=n$ , por ortonormalidad de  $X$ , tenemos

$$X_k^T X_k = \alpha_k^2 \lambda_k = 1 \tag{14}$$

$$\alpha_k = \frac{1}{\sqrt{\lambda_k}} \tag{15}$$

Se multiplica por la izquierda ambos lados de la ec. (9) por  $A$

$$AY_k = AA^T U_k \tag{16}$$

Sustituyendo (12) en (16), obtenemos

$$AY_k = \lambda_k U_k \tag{17}$$

Entonces

$$U_k = \frac{AY_k}{\lambda_k} = \frac{AX_k}{\alpha_k \lambda_k} = \frac{AX_k}{\lambda_k \frac{1}{\sqrt{\lambda_k}}} \tag{18}$$

Una vez calculados los eigenfaces  $U_k$  y los Eigenvalores  $\lambda_k$  para  $k = 1, 2, \dots, M$ , podemos usar todos ellos o solo los de mayor peso y calcular el vector característico  $\omega$  de una imagen a identificar mediante:

$$\Phi = \Gamma - \Psi \tag{19}$$

para luego obtener cada componente de  $\omega$

$$\omega_k = U_k^T - \varphi \quad (20)$$

donde  $k = 1, 2, \dots, M$

Una vez calculados los M datos característicos de  $\omega$  son estos los que se usarán para la verificación de pertinencia a una vecindad. Puede hacerse mediante simples cálculos de distancia o mediante el uso de redes neuronales (detección más precisa de las dimensiones de cada vecindad).

#### 2.4.2 FILTRO DE GABOR

Las funciones de Gabor bidimensionales quedan determinadas por cuatro parámetros, dos que expresan su localización en el dominio espacial  $(x, y)$  y otros dos que expresan la frecuencia espacial de sintonía  $(F)$  y orientación  $(\phi)$  [12], así esta se puede expresar como:

$$h(x, y) = g(x', y') \exp(2\pi j F x') \quad (21)$$

La señal elemental de Gabor bidimensional espacial, está en función de la respuesta Gaussiana bidimensional  $g(x', y')$ , la frecuencia espacial  $(F)$  y la rotación aplicada  $(\phi)$ . La respuesta Gaussiana bidimensional  $g(x', y')$ , puede expresarse mediante la siguiente ecuación:

$$g(x', y') = \left( \frac{1}{2\pi\lambda\sigma^2} \right) \exp \left[ -\frac{(x/\lambda)^2 + y^2}{2\sigma^2} \right] \quad (22)$$

Las funciones de Gabor son funciones que operan en el conjunto de los números complejos, cuya parte real es la función de Gabor simétrica y la parte imaginaria es la función de Gabor asimétrica.

$$(x', y') = (x \cos \phi + y \sin \phi, -x \sin \phi + y \cos \phi) \quad (23)$$

$$h(x, y) = h_e(x, y) - j h_s(x, y) \quad (24)$$

$$h_c(x, y) = g(x', y') \cos(2\pi Fx') \quad (25)$$

$$h_s(x, y) = g(x', y') \text{sen}(2\pi Fx') \quad (26)$$

donde  $h_c(x, y)$  es la señal elemental de Gabor con componentes reales (simetría par) y  $h_s(x, y)$  es la señal elemental de Gabor con componentes imaginarios (simetría impar).

Podemos concebir la información aportada por este par de funciones en cuadratura de fase como un vector bidimensional cuya magnitud informa el contraste de energía en un punto dado y cuya dirección especifica la fase de la energía.

El contraste de energía es llamado también amplitud de la señal el cual, al representarlo en niveles de gris, muestra la respuesta de la imagen al par en cuadratura en función de la posición espacial, que es independiente de la fase. Este contraste de energía presenta una gran similitud con el comportamiento de las células complejas y proporciona una medida de la respuesta del canal, que es independiente del cambio de fase local. Al promediar cada uno de estos contrastes de energía resultantes de cada Banco de Filtros de Gabor, obtenemos los vectores características de la respuesta de la imagen en función de la posición y frecuencia espacial.

La información aportada por este par en cuadratura de fase corresponde al contraste de energía en un punto dado. El contraste de energía  $M(x, y)$  de un par en cuadratura se obtiene mediante la ecuación:

$$M(x, y) = \sqrt{h_c^2 + h_s^2} \quad (27)$$

Esta función  $M(x,y)$  presenta gran similitud con el comportamiento de las células complejas y proporciona una medida de la respuesta del canal, que es independiente del cambio de fase local. Al promediar cada una de estas amplitudes de la señal resultante, obtenemos los vectores característicos de la respuesta de la imagen:

$$M = \frac{\sum_{p=1}^B M_p(x,y)}{B} \quad (28)$$

donde B es el número de bancos de filtros de Gabor. En la sección III se muestran resultados de implementar este tipo de análisis para reconocimiento y/o verificación de rostros.

### 2.3.3 TRANSFORMADA DISCRETA DE WALSH

La Transformada Discreta Walsh (DWT) es una de las técnicas más importantes, así como la Transformada Discreta de Fourier, en el campo de procesamiento de señales [13]. La DWT es un dominio transformado con ciertas propiedades particularmente atractivas, desde el punto de vista de su cálculo. Otro aspecto destacable es que esta transformada es real (no requiere ninguna función exponencial o trigonométrica en su cálculo) y como tal necesita menor capacidad de almacenamiento por píxel que, por ejemplo, la transformada de Fourier que generalmente es compleja.

La DWT es una representación de los datos de una imagen tan válida como la transformada de Fourier, aunque no cuenta con las propiedades de la transformada de Fourier que nos permiten calcular convoluciones y correlaciones de forma eficaz. Sin embargo, proporciona un dominio transformado eficientemente calculado, que puede utilizarse para compresión de imagen.

La DWT trabaja bien para señales digitales debido a la función fundamental llamada la función Walsh. La función Walsh tiene solamente  $\pm 1$ , en general, puede ser generada por el producto de Kronecker de la matriz Hadamard.

La matriz Hadamard  $H_2$  de  $2 \times 2$  es definida por:

$$H_2 = \begin{bmatrix} ++ \\ +- \end{bmatrix} \quad (29)$$

donde los símbolos + y - significan +1 y -1, respectivamente.

Ahora bien, calculando el producto de Kronecker entre dos matrices  $H_2$ , obtenemos una matriz Hadamard  $H_4$  de la siguiente forma:

$$H_4 = H_2 \otimes H_2 = \begin{bmatrix} +H_2 + H_2 \\ +H_2 - H_2 \end{bmatrix} = \begin{bmatrix} ++++ \\ +-+- \\ ++-- \\ -+-+ \end{bmatrix} \quad (30)$$

donde el símbolo  $\otimes$  indica el producto de Kronecker.

La frecuencia característica puede ser dada por la matriz Hadamard. A lo largo de cada renglón de la matriz Hadamard, la frecuencia es expresada por el número de cambios en signo. El número de cambios es llamado "secuencia". La secuencia tiene las características similares a la frecuencia.

Así, la DWT y la inversa DWT se definen de la siguiente forma:

$$V = \frac{1}{N} H_N B \quad (31)$$

$$B = H_N V \quad (32)$$

donde  $B$  es el vector de datos a ser procesados,  $H_N$  es la matriz Hadamard, es decir, la función Walsh ordenada.  $V$  es la DWT de  $B$ , llamado también espectro Walsh.

La DWT en dos dimensiones (2D-DWT) es la DWT de una imagen de  $m \times n$  píxeles. La 2D-DWT y la 2D-IDWT se definen de la siguiente manera:

$$F = \frac{1}{MN} H_M f H_N \quad (33)$$

$$f = H_M F H_N \quad (34)$$

donde  $f$  es la matriz de datos a ser procesados y  $F$  es la 2D-DWT de  $f$ .  $F$  es conocida como el espectro de Walsh de dos dimensiones.

Se puede resumir a la DWT con las siguientes características:

- La DWT está definida únicamente para imágenes cuyos tamaños sean potencias enteras de 2.
- Esta transformada tiene como principal aplicación la compresión de imágenes.
- A diferencia de la transformada de Fourier, que está basada en términos trigonométricos, la transformada Walsh se compone de un desarrollo serie de funciones cuyos valores son +1 o -1.
- La transformada directa e inversa son idénticas lo que implica que cualquier algoritmo capaz de implementar una puede implementar ambas.
- La transformada Walsh es real, por lo que requiere menor capacidad de almacenamiento de los datos, que la transformada de Fourier, que es generalmente compleja.

- El cálculo de esta transformada se puede realizar mediante el algoritmo de la FFT, con los términos  $W_N$  igual a 1.

En la siguiente sección se muestran los resultados de la aplicación de la 2D-DWT.

#### 2.4.4 ANÁLISIS DE COMPONENTES PRINCIPALES PCA

Esta técnica fue inicialmente desarrollada por Pearson a finales del siglo XIX y posteriormente estudiadas por Hotelling en los años 30 del siglo XX.

“Para estudiar las relaciones que se presentan entre  $p$  variables correlacionadas (qué miden información común) se puede transformar el conjunto original de variables en otro conjunto de nuevas variables incorreladas entre sí (que no tengan repetición o redundancia en la información) llamado conjunto de componentes principales. Las nuevas variables son combinaciones lineales de las anteriores y se van construyendo según el orden de importancia en cuanto a la variabilidad total que recogen de la muestra.” [14]

Este método busca la proyección según la cual los datos queden mejor representados en términos de mínimos cuadrados. PCA se emplea sobre todo en análisis exploratorio de datos y para construir modelos predictivos. PCA comporta el cálculo de la descomposición en autovalores de la matriz de covarianza, normalmente tras centrar los datos en la media de cada atributo.

Pasos para obtener PCA:

1. Obtener el conjunto de datos a analizar.

$$x^i = [x_1^i \dots x_N^i]^T \quad (35)$$

2. Centrar los datos: Se sustrae el promedio a la matriz original.

$$\bar{x}^i = x^i - m \quad (36)$$

$$\text{Donde } m = \frac{1}{P} \sum_{i=1}^P x^i$$

3. Calcular la matriz de Covarianza

Los vectores obtenidos anteriormente son combinados para obtener una matriz de  $N \times P$  (donde  $P$  es el número de imágenes)

$$\bar{X} = [\bar{x}^1 | \bar{x}^2 | \dots | \bar{x}^P] \quad (37)$$

La matriz de covarianza será:  $\Omega = \overline{\bar{X}\bar{X}^T}$

Esta matriz de covarianza tiene hasta  $P$  Eigenvectores asociados con los Eigenvalores no-cero, asumiendo que  $P < N$ .

4. Cálculo de los Eigenvectores y Eigenvalores

Se encuentran los Eigenvectores ( $V$ ) y Eigenvalores ( $\Lambda$ ) asociados a la matriz de covarianza ( $\Omega$ )

$$\Omega V = \Lambda V \quad (38)$$

5. Ordenar los Eigenvectores

Se ordenan los Eigenvectores  $v_i \in V$  de acuerdo a su Eigenvalor correspondiente  $\lambda_i \in \Lambda$  de mayor a menor. Se conservan solo los Eigenvectores asociados con los Eigenvalores no-cero. Esta matriz de Eigenvectores es el eigenspacio  $V$ , donde cada columna de  $V$  es un eigenvector:

$$V = [v_1 | v_2 | \dots | v_P] \quad (39)$$

## 2.5 CLASIFICADORES

“Aunque existen muchas técnicas alternativas para enfrentar problemas de regresión y clasificación, las máquinas de soporte vectorial han sido

desarrolladas como una herramienta robusta para regresión y clasificación en dominios complejos y ruidosos.” [15]

Las MSV pueden ser usadas para extraer información relevante a partir de conjunto de datos y construir algoritmos de clasificación o de regresión rápidos para datos masivos. Aplicadas al problema de clasificación, mapean los datos a un espacio de características alto-dimensional, donde se puede hallar más fácilmente un hiperplano de separación. Este mapeo puede ser llevado a cabo aplicando el kernel, el cual transforma implícitamente el espacio de entrada en un espacio de características de alta dimensión. El hiperplano de separación es calculado maximizando la distancia de los patrones más cercanos, es decir la maximización del margen. Las MSV pueden ser definidas como un sistema para el entrenamiento eficiente de máquinas de aprendizaje lineal en un espacio de características inducido por un kernel, mientras respeta los principios de la teoría de la generalización y explota la teoría de la optimización.

### 2.5.1 MÁQUINAS DE SOPORTE VECTORIAL

En la teoría del aprendizaje estadístico, el problema de aprender una relación entrada-salida a partir de un conjunto de datos es generalmente visto como el problema de elegir , a partir de un conjunto dado de funciones  $f(x,a)$ ,  $a \in \Lambda$  (donde  $x \in R^n$  es un vector aleatorio tomado independientemente de una función de distribución de probabilidad  $P(x)$ , fija pero desconocida y  $\Lambda$  es un conjunto de parámetros) la que mejor aproxime el valor de la salida “y” para todos los vectores de entrada  $x$ , de acuerdo a una función de distribución condicional  $P(y/x)$ , también fija pero desconocida. [16]

La selección de la función deseada está basada en un conjunto de entrenamiento de  $l$  observaciones independientes e idénticamente distribuidas  $(x_1, y_1), \dots, (x_l, y_l)$  tomados de acuerdo a

$$P(\mathbf{x}, y) = P(\mathbf{x})P(y/\mathbf{x}) \quad (40)$$

Si se considera que el valor esperado de las pérdidas debido a los errores de clasificación o estimación, está dado por el riesgo funcional

$$R(\alpha) = \int L(y, f(x, \alpha)) dP(x, y) \quad (41)$$

Ah  $y, f(x, \alpha)$  es la discrepancia entre las respuestas medidas "y" y la respuesta  $f(x, \alpha)$  proporcionada por la máquina de aprendizaje, la meta sería encontrar la función  $f(x, \alpha_0)$  que minimice el riesgo funcional  $R(\alpha)$  en la situación en donde la única información disponible es el conjunto de entrenamiento.

Para el caso de reconocimiento de patrones de dos clases, la tarea del Aprendizaje a partir de ejemplos puede ser formulado de la siguiente forma:

Se da un conjunto de funciones

$$\{f_\alpha : \alpha \in \Lambda\}, f_\alpha : \mathfrak{R}^n \rightarrow \{-1, +1\} \quad (42)$$

y un conjunto de ejemplos, es decir pares de patrones  $x_i$  y etiquetas  $y_i, (x_i, y_i), \dots, (x_l, y_l), \in \mathfrak{R}^n \times \{-1, +1\}$ , cada uno de ellos generado a partir de una distribución de probabilidad desconocida  $P(x, y)$  que contiene la dependencia subyacente. Lo que se requiere ahora es aprender la función  $f_\alpha$  que proporciona el valor más pequeño posible del error promedio comprometido en los ejemplos independientes tomados aleatoriamente de la misma distribución  $P$ , conocido como el *riesgo*.

Sin embargo, el riesgo -  $R(\alpha)$  - es desconocido, ya que  $P(x,y)$  es desconocida; por lo tanto, se necesita un principio de inducción para la minimización del riesgo. Uno de tales principios es llamado el principio inductivo de la Minimización del Riesgo Empírico (ERM). Sin embargo, esta aproximación directa, que es minimizar el riesgo empírico:

$$R_{emp}(\alpha) = \frac{1}{l} \sum_{i=1}^l \frac{1}{2} |f_{\alpha}(x_i) - y_i| \quad (43)$$

no garantiza un riesgo real pequeño si el número  $l$  de ejemplos de entrenamiento es limitado. En otras palabras, un error más pequeño sobre el conjunto de entrenamiento, no necesariamente implica una habilidad de generalización mayor (es decir un error más pequeño sobre el conjunto independiente de prueba). Para maximizarlo con una cantidad limitada de datos, una novedosa técnica estadística llamada Minimización del Riesgo Estructural ha sido desarrollada. [17]

La teoría de la convergencia uniforme en probabilidad, desarrollada por Vapnik y Chervonenkis (VC) [18], provee cotas sobre la desviación del riesgo empírico con respecto al riesgo esperado. Esta teoría muestra que es crucial restringir la clase de funciones que la máquina de aprendizaje puede implementar a una clase con una capacidad que resulte conveniente para la cantidad de datos de entrenamiento disponibles.

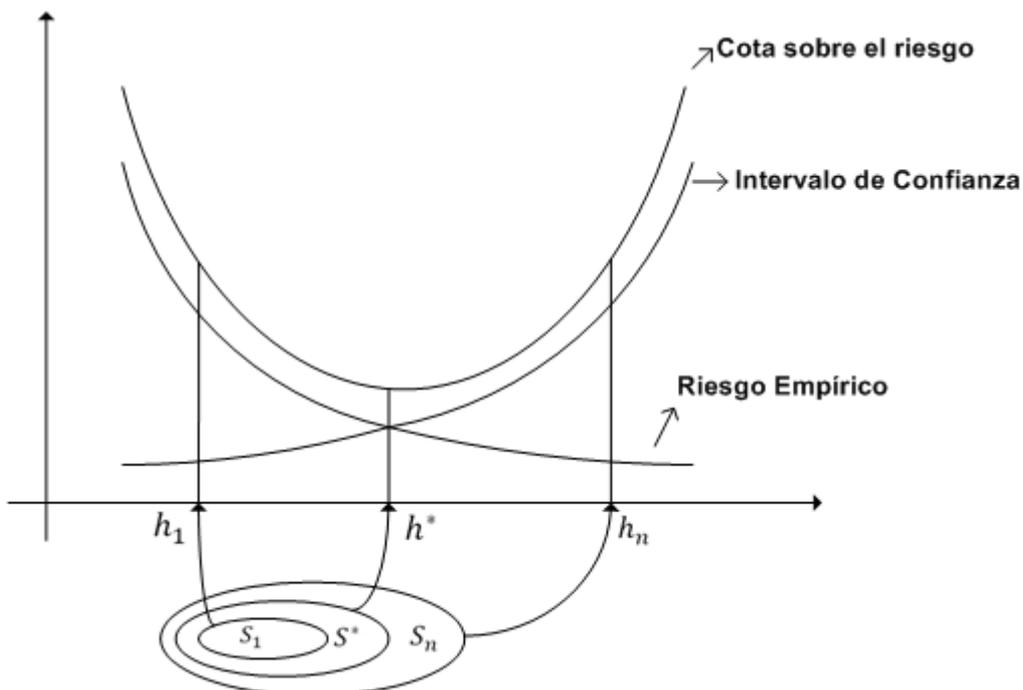
Para  $\alpha \in \Lambda$  y  $l > h$ , una cota VC, la cual se verifica con probabilidad  $1 - \eta$ , tiene la siguiente forma:

$$R(\alpha) \leq R_{emp}(\alpha) + \sqrt{\frac{h(\log \frac{2l}{h} + 1) - \log(\frac{\eta}{4})}{l}} \quad (44)$$

El parámetro  $h$  es llamado la dimensión-VC (Vapnik-Chervonenkis) de un conjunto de funciones y ella describe la capacidad de un conjunto de funciones para representar el conjunto de datos. La dimensión VC es

una medida de la complejidad del clasificador y ella es a menudo proporcional al número de parámetros libre en el clasificador  $f_a$ .

Especialmente cuando  $\frac{l}{h}$  es pequeño, un riesgo empírico pequeño no garantiza un valor pequeño en el riesgo real o estructural. En este caso, con el fin de minimizar el riesgo real  $R(\cdot)$ , se debe minimizar el lado derecho de la desigualdad en la ecuación (44) simultáneamente sobre ambos términos. Para hacer esto, se hace de la dimensión VC una variable controlante. Por lo tanto, el principio general, el cual se denomina "principio inductivo de minimización del riesgo estructural (SRM)", fue motivado para minimizar el riesgo funcional, con respecto a ambos términos: el riesgo empírico y el intervalo de confianza. El término de confianza VC en la ecuación (44) depende de la clase de funciones elegidas, mientras que el riesgo empírico depende de la función particular elegida por el proceso de entrenamiento. El objetivo aquí es encontrar ese subconjunto del conjunto de funciones elegidas, tal que la cota del riesgo para ese subconjunto sea mínima.



**Figura 2.2** La cota del riesgo real es igual a la suma del riesgo empírico y el intervalo de confianza

Esto se hace introduciendo una estructura que divida la clase completa de funciones en subconjuntos anidados (Figura 2.2). SRM consiste en hallar ese subconjunto de funciones que minimiza la cota sobre el riesgo real. Esto se hace simplemente entrenando una serie de máquinas, una para cada subconjunto, donde para un subconjunto dado, la meta de entrenamiento es simplemente minimizar el riesgo empírico. Luego se toma la máquina para la cual la suma del riesgo empírico y la confianza de VC sea mínima.[19]

Aunque existen muchas técnicas alternativas para enfrentar problemas de regresión y clasificación, las máquinas de soporte vectorial han sido desarrolladas como una herramienta robusta para regresión y clasificación en dominios complejos y ruidosos. Las SVM pueden ser usadas para extraer información relevante a partir de conjunto de datos y construir algoritmos de clasificación o de regresión rápidos para datos masivos.

Las SVM, aplicadas al problema de clasificación, mapean los datos a un espacio de características alto-dimensional, donde se puede hallar más fácilmente un hiperplano de separación. Este mapeo puede ser llevado a cabo aplicando el kernel, el cual transforma implícitamente el espacio de entrada en un espacio de características de alta dimensión. El hiperplano de separación es calculado maximizando la distancia de los patrones más cercanos, es decir la maximización del margen. Las SVM pueden ser definidas como un sistema para el entrenamiento eficiente de máquinas de aprendizaje lineal en un espacio de características inducido por un kernel, mientras respeta los principios de la teoría de la generalización y explota la teoría de la optimización.

### 2.5.1.1 CARACTERÍSTICAS

- La teoría de generalización, la cual conduce a una elección de hipótesis basada en principios teóricos; y
- Un kernel de funciones, que introduce no linealidades en el espacio de hipótesis sin requerir explícitamente algoritmos no lineales.

Las máquinas de soporte vectorial (SVM) se definen como un sistema de aprendizaje que usa un espacio de hipótesis de funciones lineales en un espacio de características alto-dimensional, entrenado con un algoritmo de aprendizaje tomado de la teoría de la optimización que implementa un sesgo de aprendizaje derivado de la teoría del aprendizaje estadístico.[20]

El fundamento del tema de las máquinas de soporte (SVM) ha sido desarrollado principalmente por Vapnik y sus colaboradores y los correspondientes métodos SV están ganando popularidad debido a sus muchas características atractivas y el promisorio desempeño empírico. Su formulación incorpora el principio de Minimización del Riesgo Estructural (SRM), el cual se ha demostrado que es superior al más tradicional principio de minimización del riesgo empírico (ERM) empleado por muchas técnicas de modelación.[21] Esta es la diferencia que proporciona a las SVM una habilidad mayor para generalizar, la cual es la meta en aprendizaje estadístico.

### 2.5.1.2 CLASIFICACIÓN POR VECTORES DE SOPORTE

Considere el problema de separar el conjunto de vectores de entrenamiento  $(x_1, y_1), \dots, (x_l, y_l), \in R^n$  que pertenecen a dos clases separadas ( $y_i = \{1, -1\}$ ). En este problema la meta es separar los vectores de entrenamiento en dos clases mediante un hiperplano.

$$(w \cdot x) + b = 0, w \in R^n \text{ y } b \in R \quad (45)$$

Donde  $w$  y  $b$  son parámetros que se inducen a partir de los ejemplos disponibles correspondientes a la función de decisión

$$f(x) = \text{sign}(wx + b) \quad (46)$$

tal que ella se desempeña bien sobre ejemplos no vistos, es decir que generaliza bien.

Para el caso del espacio de entrada bidimensional como se muestra en la Figura 2.3, hay muchos posibles clasificadores lineales que pueden separar los datos; pero hay sólo uno que maximiza el margen (es decir, maximiza la distancia entre el dato más cercano de cada clase). Este clasificador lineal es llamado el hiperplano de separación óptima. Se ha demostrado, además que el hiperplano óptimo, definido como el que tiene el margen máximo de separación entre las dos clases, tiene la capacidad más baja y minimiza la cota sobre el riesgo real.

El hiperplano  $(wx) + b = 0$  satisface las condiciones:

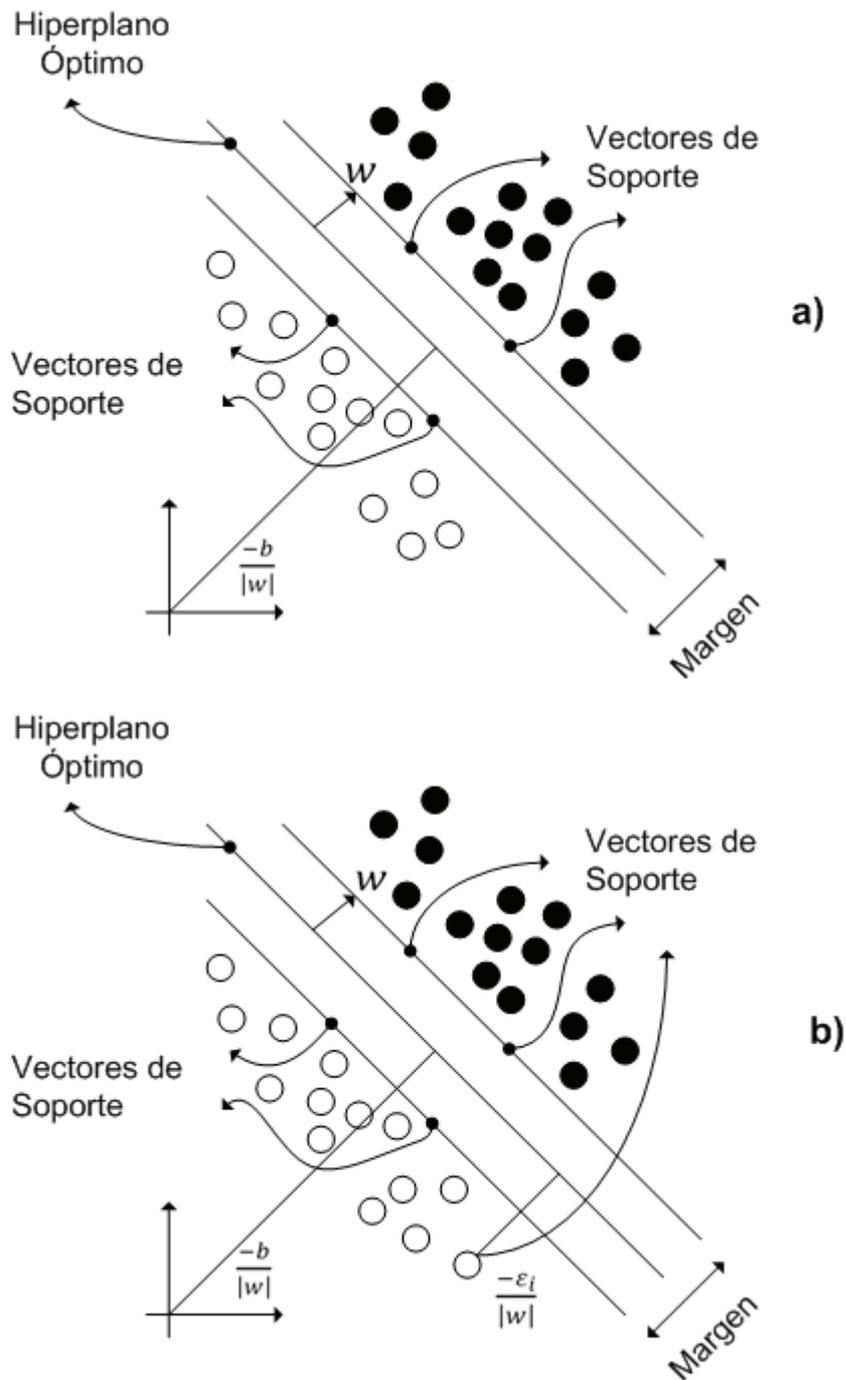
$$(w \cdot x) + b > 0 \text{ si } y_i = 1 \text{ y } (w \cdot x) + b < 0 \text{ si } y_i = -1 \quad (47)$$

Combinando las dos expresiones en la ecuación (14) y escalando  $w$  y  $b$ , con un factor apropiado, una superficie de decisión equivalente se puede formular como aquella que satisfaga la restricción:

$$y_i[(w \cdot x_i) + b] \geq 1, i = 1, 2, \dots, l \quad (48)$$

Se puede demostrar que el hiperplano que separa óptimamente los datos en dos clases es aquel que minimiza el funcional:

$$\Phi(w) = \frac{\|w\|^2}{2} \quad (49)$$



**Figura 2.3** Hiperplano de separación óptima para el caso bidimensional

Por lo tanto, el problema de optimización puede ser reformulado como un problema de optimización no restringida, usando multiplicadores de

Langrange y su solución estaría dada por la identificación de los puntos de silla del funcional de LaGrange [19], como sigue:

$$L(w, b, \alpha) = \frac{\|w\|^2}{2} - \sum_{i=1}^l \alpha_i \{(w \cdot x_i) + b\} y_i - 1 \quad (50)$$

Donde  $\alpha_i$  son los multiplicadores de LaGrange. El Lagrangiano tiene que ser minimizado con respecto a  $w$  y  $b$ , es decir:

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^l \alpha_i y_i = 0 \quad \text{y} \quad (51a)$$

$$\frac{\partial L}{\partial w} = 0 \Rightarrow w - \sum_{i=1}^l \alpha_i x_i y_i = 0 \Rightarrow w = \sum_{i=1}^l \alpha_i y_i x_i \quad (51b)$$

Poniendo las expresiones para  $w_o$  en la ecuación (50) resultará en la siguiente forma dual de la función, que debe ser maximizada con respecto a las restricciones  $\alpha_i \geq 0$

$$W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \quad (52)$$

Encontrar la solución de la ecuación (52) para problemas el mundo real usualmente requerirá la aplicación de técnicas de optimización de programación cuadrática (QP) y métodos numéricos. Una vez se halla la solución en la forma de un vector  $\alpha^0 = (\alpha_1^0, \alpha_2^0, \dots, \alpha_l^0)$ , el hiperplano de separación óptimo estará dado por:

$$w_0 = \sum_{\text{vector de soporte}} y_i \alpha_i^0 x_i \quad (53)$$

y

$$b_0 = -\frac{1}{2} w_0 \cdot [x_r + x_s] \quad (54)$$

Donde  $x_r$  y  $x_s$  son cualesquiera vectores de soporte uno de cada clase. Los clasificadores pueden ser, entonces contruidos como:

$$f(x) = \text{sign}(w_0 \cdot x + b_0) = \text{sign}\left(\sum_{\text{vectores de soporte}} y_i \alpha_i^0 (x_i x) + b_0\right) \quad (55)$$

Solamente los puntos  $x_i$  que tienen multiplicadores de Lagrange  $\alpha_i^0$  diferentes de cero son llamados Vectores de Soporte (VS). Si los

datos son linealmente separables, todos los vectores de soporte estarán sobre el margen y por lo tanto, el número de VS puede ser muy pequeño.

La solución anterior sólo se verifica para datos separables linealmente, y todavía debe ser ligeramente modificada para datos no separables linealmente introduciendo un nuevo conjunto de variables  $\{\xi_i\}$  que mide la cantidad en la cual las restricciones son violadas (Figura 2b). Luego el margen es maximizado, asumiendo una penalización proporcional a la cantidad de la violación de la restricción. Formalmente se resuelve el siguiente problema:

Minimice

$$\Phi(w) = \frac{\|w\|^2}{2} + C \sum \xi_i \quad (56)$$

$$\text{sujeeto a } y_i[(w \cdot x_i) + b] \geq 1 - \xi_i, \quad \xi_i \geq 0 \quad i = 1, \dots, l$$

Donde C es un parámetro elegido a priori y que define el costo de la violación de la restricción. La ecuación (56) proporciona una minimización de la dimensión VC de la máquina de aprendizaje, minimizando por lo tanto, el segundo término en la cota de la ecuación (47). De otra parte, la minimización en la ecuación (56) controla el riesgo empírico, el cual es el primer término en la ecuación (47). Esta aproximación, por lo tanto, constituye una implementación práctica de la Minimización del Riesgo Estructural sobre el conjunto de funciones dado. Con el fin de resolver este problema, el Lagrangiano se construye como sigue:

$$L(w, b, \alpha) = \frac{\|w\|^2}{2} + (\sum_{i=1}^l \xi_i) - \sum_{i=1}^l \alpha_i \{[(w \cdot x) + b]y_i - 1 + \xi_i\} - \sum_{i=1}^l \gamma_i \xi_i \quad (57)$$

Donde  $\alpha_i$  y  $\gamma_i$  están asociados con las restricciones en la ecuación (56) y los valores de  $\alpha_i$  tienen que ser acotados como  $0 \leq \alpha_i \leq C$ . De nuevo,

la solución de este problema se determina por los puntos de silla de este Lagrangiano de forma similar para el caso de datos separables.

En el caso donde una frontera lineal sea definitivamente inapropiada (o cuando la superficie de decisión es no lineal), el SVM puede mapear el vector de entrada  $x$ , en un espacio de características alto-dimensional  $z$ , eligiendo un mapeo no lineal a priori. Entonces la SVM construye el hiperplano de separación óptimo en este espacio más alto-dimensional. En este caso, los problemas de optimización de la ecuación (52) se convierten en :

$$W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j K(x_i \cdot x_j) \quad (58)$$

Donde  $K(x,y)$  es la función kernel que realiza el mapeo no lineal en el espacio de características, y las restricciones permanecen sin cambio. La solución de la ecuación anterior determina los multiplicadores de Lagrange, y un clasificador que implementa en hiperplano de separación óptima en el espacio de característica está dado por,

$$f(x) = \text{sign}(\sum_{\text{vectores de soporte}} y_i \alpha_i^0 (x_i x) + b_0) \quad (59)$$

Consecuentemente, todo lo que se ha derivado para el caso lineal es también aplicable para un caso no lineal usando un kernel conveniente  $K$  en vez del producto punto. Además, usando diferentes funciones kernel, el algoritmo de SV puede construir una variedad de máquinas de aprendizaje, algunas de las cuales parecen ser similares a arquitecturas clásicas. Funciones de base radial, funciones polinomiales y ciertas funciones sigmoideas son entre otras que proporcionan kernels aceptables y los correspondientes mapeos son descritos como sigue:

- *Kernel simple polinomial*:  $K(x, x_i) = ((x \cdot x_i) + 1)^d$ ; donde el grado del polinomio  $d$ , es definido por el usuario.
- *Kernel de Funciones De Base Radial*:  $K(x, x_i) = e^{-\gamma|x-x_i|^2}$ ; donde  $\gamma$  es definido por el usuario.
- *Kernel de Redes Neuronales*:  $K(x, x_i) = \tanh(b(x \cdot x_i) - C)$ ; donde  $b$  y  $c$  son definidos por el usuario.

El otro caso surge cuando los datos están en múltiples clases. Con el fin de obtener una clasificación de  $k$ -clases, se construye un conjunto de clasificadores binarios  $f_1, f_2, \dots, f_k$ , cada uno entrenado para separar una clase del resto, y estos son combinados para llevar a cabo la multclasificación (en un esquema de votación) de acuerdo con la salida máxima, antes de aplicar la función signo. [22]

## 2.6 ECUALIZACIÓN DE HISTOGRAMAS

### 2.6.1 DEFINICIÓN DE HISTOGRAMA

Un histograma es una representación gráfica de una variable en forma de barras, donde la altura o eje vertical es proporcional a la frecuencia de los valores producidos, y la anchura o eje horizontal corresponde a los intervalos o valores de la clasificación.

Un histograma normalizado es aquél en el que la variable representada ha sido escalada para ajustarse a un rango entre 0 y 1.

En un histograma acumulado, al valor correspondiente a cada una de las barras verticales del mismo se le adiciona el valor de la barra inmediatamente anterior.

Por último, un histograma acumulado normalizado es un histograma acumulado escalado para limitarse al mencionado rango. El histograma

de una imagen representa la frecuencia relativa de los niveles de gris de la imagen.

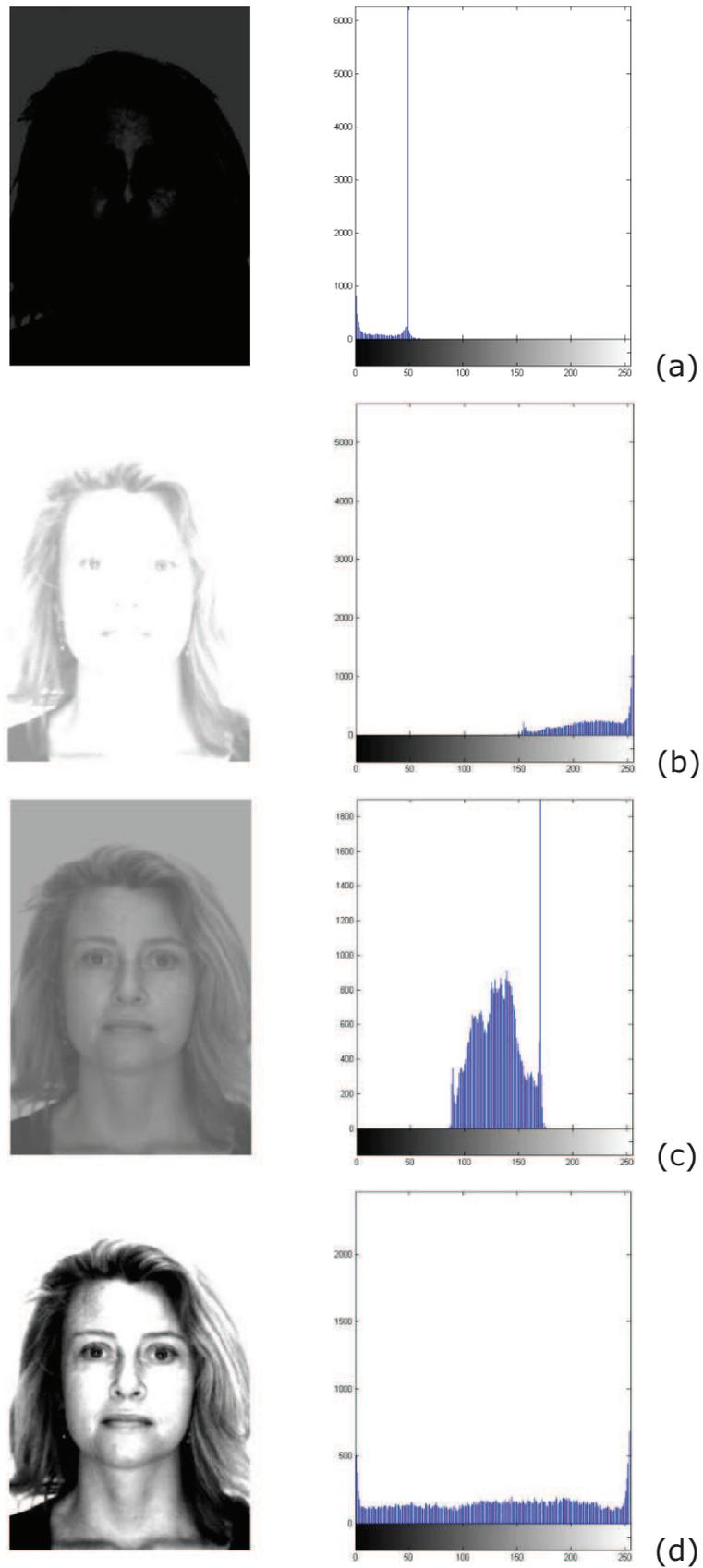
Las técnicas de modificación del histograma de una imagen son útiles para aumentar el contraste de imágenes con histogramas muy concentrados, ya sean imágenes oscuras como claras.

En la Figura 2.4 podemos observar el rostro con cuatro variaciones de nivel de gris: oscura, brillante, bajo contraste y alto contraste; con su correspondiente histograma.

La *Ecuación de Histograma* es una forma de manipulación de histograma que reduce automáticamente el contraste en las áreas muy claras o muy oscuras de una imagen. También expande los niveles de gris a lo largo de todo intervalo. Consiste en una transformación no lineal que considera la distribución acumulativa de la imagen original, para generar una imagen resultante cuyo histograma será aproximadamente uniforme.

La opción de modificación, parte del principio que dice que el contraste de una imagen sería optimizado si todos los 256 niveles de intensidad posibles fueran igualmente utilizados o, en otras palabras, todas las barras verticales que componen el histograma fueran de la misma altura. Obviamente esto no es posible debido a la naturaleza discreta de los datos digitales de una imagen.

Sin embargo, se consigue una aproximación al dispersar los picos del histograma de la imagen, dejando intactas las partes más bajas. Este proceso se obtiene a través de una función de transferencia que tiene una alta inclinación siempre que el histograma original presenta un pico y una baja inclinación en el resto del histograma.



**Figura 2.4** Variaciones de nivel de gris: (a) oscura, (b) brillante, (c) bajo contraste, (d) alto contraste

Consideremos las funciones continuas, donde la variable  $r$  representa los niveles de grises de la imagen que será optimizada. Inicialmente suponemos que  $r$  se ha normalizado al intervalo  $[0, 1]$ , cuando  $r = 0$  representa al negro y  $r = 1$  representa al blanco. Luego consideramos una formulación discreta y permitir que los valores de los píxeles estén en el intervalo  $[0, L - 1]$ .

Para cualquier  $r$  que cumpla las condiciones antes mencionadas, la función de transformación está dada por:

$$s = T(r) ; 0 \leq r \leq 1 \quad (60)$$

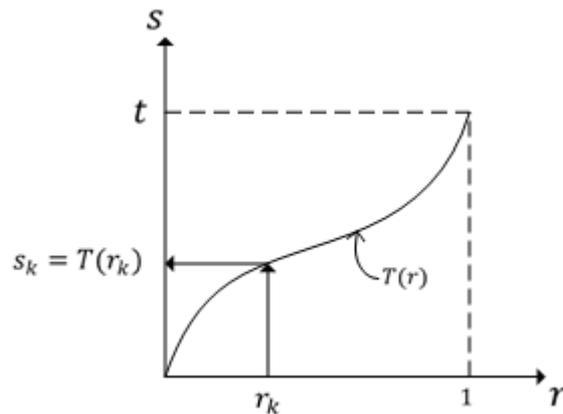
Lo que produce un nivel de  $s$  para cada valor pixel  $r$  en la imagen original. Asumimos que la función de transformación  $T(r)$ , satisface las siguientes condiciones:

- a)  $T(r)$  es un solo valor que incrementa monotónicamente en el intervalo  $0 \leq r \leq 1$
- b)  $0 \leq T(r) \leq 1$  para  $0 \leq r \leq 1$

La condición (a) en la que  $T(r)$  debe ser un solo valor es necesaria para garantizar que la transformación inversa exista, y la condición de monotonidad conserva el orden creciente de negro a blanco en la transformación de la imagen de salida. La función que no sea monotónicamente creciente podría resultar en que por lo menos una sección del rango de intensidad se invierta, lo que produciría algunos niveles de grises invertidos en la imagen de salida. Mientras esto podría ser un efecto deseado en algunos casos, no es lo que estamos buscando en la presente discusión. Finalmente, la condición (b) garantiza que los niveles de gris estarán en el mismo rango que los niveles de entrada. La Figura 2.5 da un ejemplo de la función de transformación que satisface estas dos condiciones. La transformación inversa de  $s$  a  $r$  está dada por:

$$r = T^{-1}(s) ; 0 \leq r \leq 1 \quad (61)$$

Incluso si  $T(r)$  satisface las condiciones (a) y (b), es posible que la inversa correspondiente no sea un único valor.



**Figura 2.5** Una función de transformación de niveles de gris que satisface las condiciones (a) y (b)

Los niveles de gris en una imagen pueden ser vistos como variables aleatorias en el intervalo  $[0,1]$ . Uno de los descriptores más fundamentales de una variable aleatoria es la Función de Densidad de Probabilidad (PDF).  $p_r(r)$  y  $p_s(s)$  denotan la función de densidad de probabilidad de las variables aleatorias  $r$  y  $s$ , respectivamente. Donde los subíndices de  $p$  se utilizan para indicar que  $p_r$  y  $p_s$  son diferentes funciones. Un resultado básico de la teoría de probabilidad elemental es que, si  $p_r(r)$  y  $T(r)$  son conocidos y  $T^{-1}(s)$ , satisface la condición (a), entonces la función de densidad de probabilidad  $p_s(s)$  de la variable transformada  $s$  puede obtenerse mediante la siguiente fórmula:

$$p_s(s) = p_r(r) \left| \frac{dr}{ds} \right| \quad (62)$$

Por lo tanto, la función de densidad de probabilidad de la variable transformada  $s$ , se determina por la PDF de escala de grises de la imagen de entrada y por la función transformación elegida.

Una función de transformación de particular importancia en el procesamiento de imágenes tiene la forma:

$$s = T(r) = \int_0^r p_r(w)dw \quad (63)$$

donde  $w$  es una variable ficticia de integración. Del lado derecho de la ecuación (63) se reconoce como la función de distribución acumulativa (CDF) de variable aleatoria  $r$ . Dado que las funciones de densidad de probabilidad son siempre positivas, y recordando que la integral de una función es el área bajo la función, se deduce que esta función de transformación es de un solo valor y monótonamente creciente, y, por tanto, satisface la condición (a). Del mismo modo, la integral de una función de densidad de probabilidad para variables en el rango  $[0, 1]$  también está en el rango  $[0, 1]$ , por lo que la condición (b) se satisface también.

Dada la función de transformación  $T(r)$ , encontramos  $p_s(s)$  mediante la aplicación de la ecuación (62). Sabemos por cálculo básico (regla de Leibniz) que la derivada de una integral definida con respecto a su límite superior no es más que el integrando evaluado en ese límite. En otras palabras,

$$\begin{aligned} \frac{ds}{dr} &= \frac{dT(r)}{dr} \\ &= \frac{d \int_0^r p_r(w)dw}{dr} \\ &= p_r(r) \end{aligned} \quad (64)$$

Sustituyendo este resultado para  $ds/dr$  en la ecuación (62), y teniendo en cuenta que todos los valores de probabilidad son positivos,

$$p_s(s) = p_r(r) \left| \frac{dr}{ds} \right|$$

$$\begin{aligned}
&= p_r(r) \left| \frac{1}{p_r(r)} \right| \\
&= 1 \quad ; \quad 0 \leq r \leq 1 \quad (65)
\end{aligned}$$

Debido a que  $p_s = (s)$  es una función de densidad de probabilidad, debe ser cero fuera del intervalo  $[0, 1]$  en este caso su integral sobre todos los valores de  $s$  debe de ser 1. Se reconoce la forma de  $p_s(s)$  dada en la ecuación (65) como una función de probabilidad de densidad uniforme. En pocas palabras, hemos demostrado que la realización de la función de transformación dada en la ecuación (63) obtiene una variable aleatoria  $s$  caracterizada por una función de densidad de probabilidad uniforme. Es importante tener en cuenta de la ecuación (63) que  $T(r)$  depende de  $p_r(r)$  pero, como se indica en la ecuación (65), el resultado  $p_s(s)$  siempre es uniforme, independiente de la forma de  $p_r(r)$ .

Para valores discretos se utilizan probabilidades y sumas en lugar de funciones de densidad de probabilidad e integrales. La probabilidad de ocurrencia de los niveles de grises  $r_k$  en una imagen se aproxima por

$$p_r(r_k) = n_k/n \quad k = 0, 1, 2, \dots, L - 1 \quad (66)$$

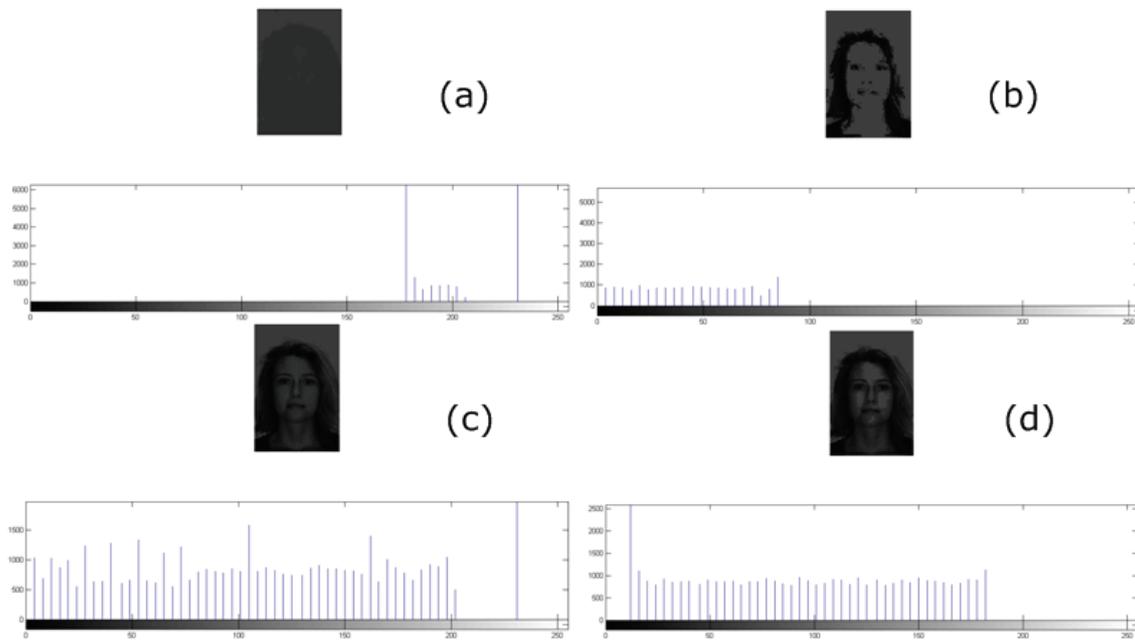
Donde,  $n$  es el número total de píxeles en la imagen,  $n_k$  es el número de píxeles que tienen nivel de gris  $r_k$ , y  $L$  es el total número de posibles niveles de gris en la imagen. La versión discreta de la función de transformación dada en la ecuación (30) es

$$s_k = T(r_k) = \sum_{j=0}^k p_r(r_j) = \sum_{j=0}^k \frac{(n_j)}{n} \quad k = 0, 1, 2, \dots, L - 1 \quad (67)$$

Por lo tanto, una imagen procesada de salida se obtiene mediante el mapeo de cada pixel con nivel  $r_k$  en la imagen de entrada con su pixel correspondiente del nivel  $s_k$  en la imagen de salida, ver ecuación (67). Como se indicó anteriormente, la gráfica de  $p_r(r_k)$  versus  $r_k$  se llama histograma. La transformación (mapeo) dada en la ecuación (67) se

denomina ecualización de histograma o linearización de histograma. [23]

La Figura 2.6 muestra las cuatro imágenes de la Figura 2.5 una vez que se les ha realizado la ecualización de histograma.



**Figura 2.6.** Imágenes ecualizadas y sus respectivos histogramas: (a) oscura, (b) brillante, (c) bajo contraste, (d) alto contraste

## 2.7 TRANSFORMADA RÁPIDA DE FOURIER

La Transformada de Fourier Discreta (DFT) es una forma específica de análisis de Fourier para convertir una función (a menudo en el tiempo o el dominio espacial), a otro dominio de la frecuencia. La DFT es ampliamente utilizada en el procesamiento de señal y campos relacionados con el análisis de frecuencias contenidas en una señal de la muestra, para resolver ecuaciones diferenciales parciales y realizar otras operaciones, como convoluciones.

La Transformada Rápida de Fourier (FFT) es una aplicación eficiente de la DFT y se utiliza, además en otros campos, en el procesamiento digital de imágenes. La Transformada Rápida de Fourier se aplica para convertir una imagen de la imagen (espacial) de dominio al dominio de la frecuencia.

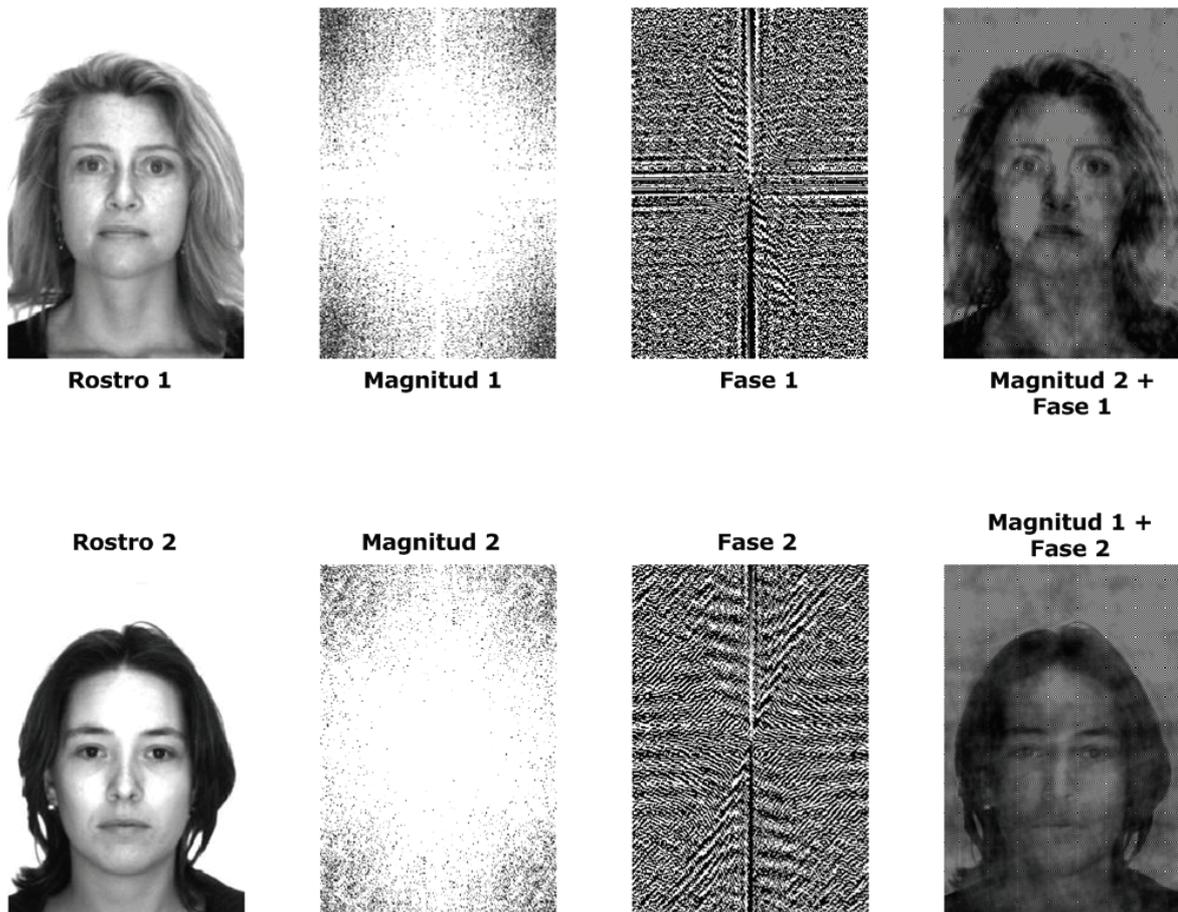
La Transformada de Fourier descompone una imagen en sus componentes reales e imaginarios que es una representación de la imagen en el dominio de la frecuencia. La señal de entrada es una imagen y luego el número de frecuencias en el dominio de la frecuencia es igual al número de píxeles de la imagen o el dominio espacial. La transformada inversa de nuevo transforma las frecuencias a la imagen en el dominio espacial.

El 2D FFT viene dada por la siguiente ecuación:

$$F(x) = \sum_{n=0}^{N-1} f(n) e^{-j2\pi(x\frac{n}{N})} \quad (68)$$

Oppenheim [24, 25] han demostrado que la información de fase de una imagen conserva la mayor parte de la inteligibilidad de una imagen. Su investigación también muestra que obteniendo sólo el espectro de fase de una imagen, uno puede reconstruir la imagen original hasta un factor de escala, por lo tanto, la información de fase es la más importante en la representación de una señal de 2D en el dominio de Fourier.[26]

Esto se demuestra en el siguiente experimento que se muestra en la Figura 2.7 en donde obtenemos la magnitud y la fase de cada rostro y luego intercambiamos las magnitudes y las fases, de donde podemos observar que se predomina la información de la fase de cada rostro.



**Figura 2.7** Experimento de Oppenheim

## 2.8 CONCLUSIONES

El reconocimiento de rostros es actualmente un campo ampliamente estudiado y que cuenta con muchas áreas de investigación. Una serie de procesos son realizados para realizar el reconocimiento, es importante que estos procesos sean eficientes, para la extracción de características el Análisis de Componentes Principales se presenta como una herramienta robusta capaz de identificar patrones en los datos, y expresar los datos de tal forma que se destaquen sus similitudes y diferencias, además de ser capaz de comprimir estos datos sin mucha pérdida de información. La Máquina de Soporte Vectorial es un clasificador de patrones binarios que pueden ser usadas para extraer

información relevante de conjunto de datos y construir algoritmos de clasificación o de regresión rápidos.

## CAPÍTULO III

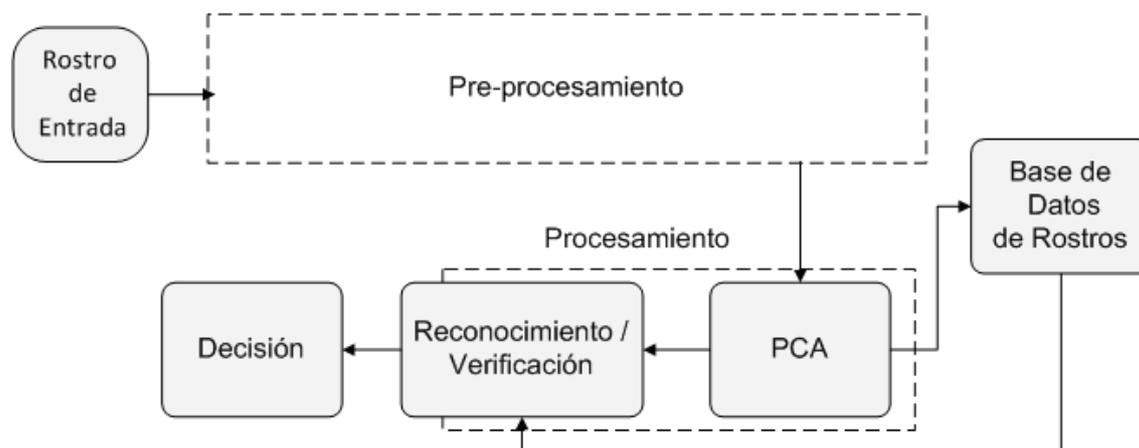
### SISTEMA PROPUESTO

#### 3.1 INTRODUCCIÓN

En este capítulo se presenta el sistema propuesto para el reconocimiento y la verificación de los rostros. El sistema cuenta con las partes básicas de un sistema de reconocimiento: pre-procesamiento, procesamiento y decisión.

#### 3.2 ESTRUCTURA DEL SISTEMA PROPUESTO

En el diagrama que a continuación se presenta, encontramos la estructura que posee el sistema.



**Figura 3.1** Sistema de reconocimiento propuesto

El sistema que se describe en el diagrama anterior es un diagrama general. Ya que posteriormente se definirán variaciones en el pre-procesamiento del rostro, aplicando Ecuilización de Histograma.

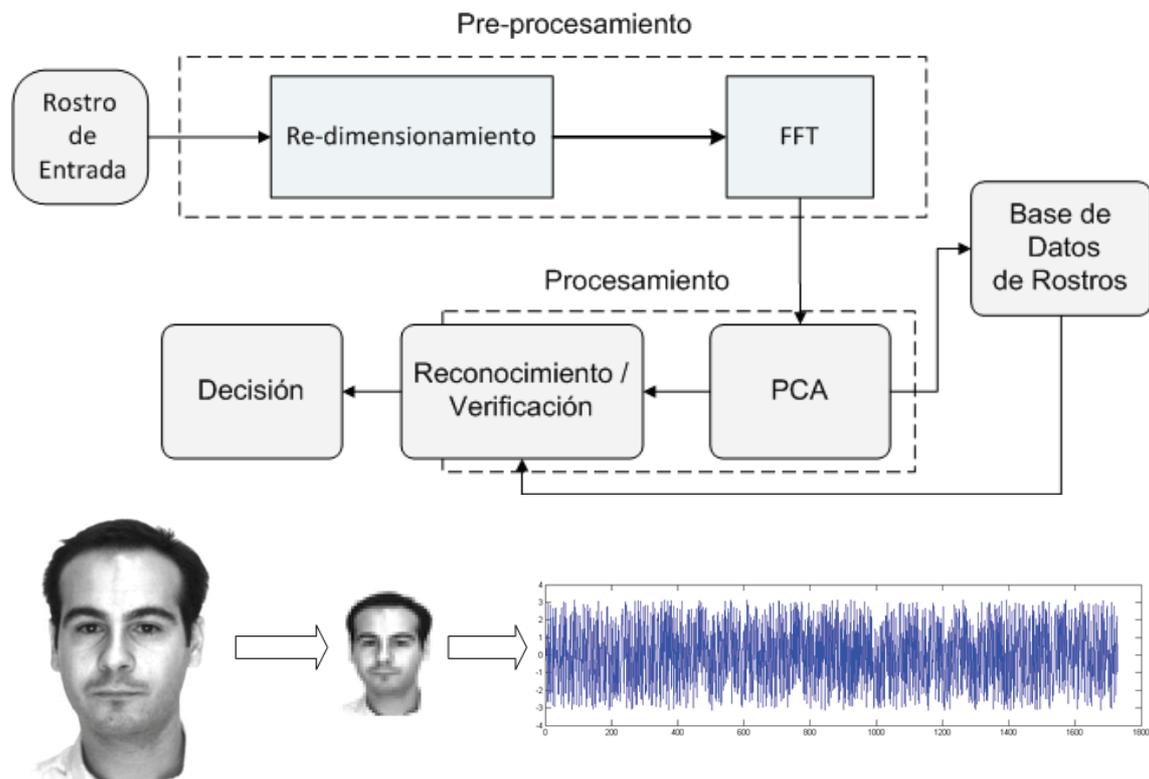
Como entrada del sistema tenemos un rostro que tiene un tamaño de 192x288 píxeles, el cual entra en la etapa de pre-procesamiento en donde se reduce su tamaño a 48x36 píxeles, a este rostro reducido se

le ecualiza el histograma, para ajustar la escala de grises, luego a la imagen reducida y ecualizada se le aplica la Transformada Rápida de Fourier, FFT; para obtener la imagen en el dominio de la fase.

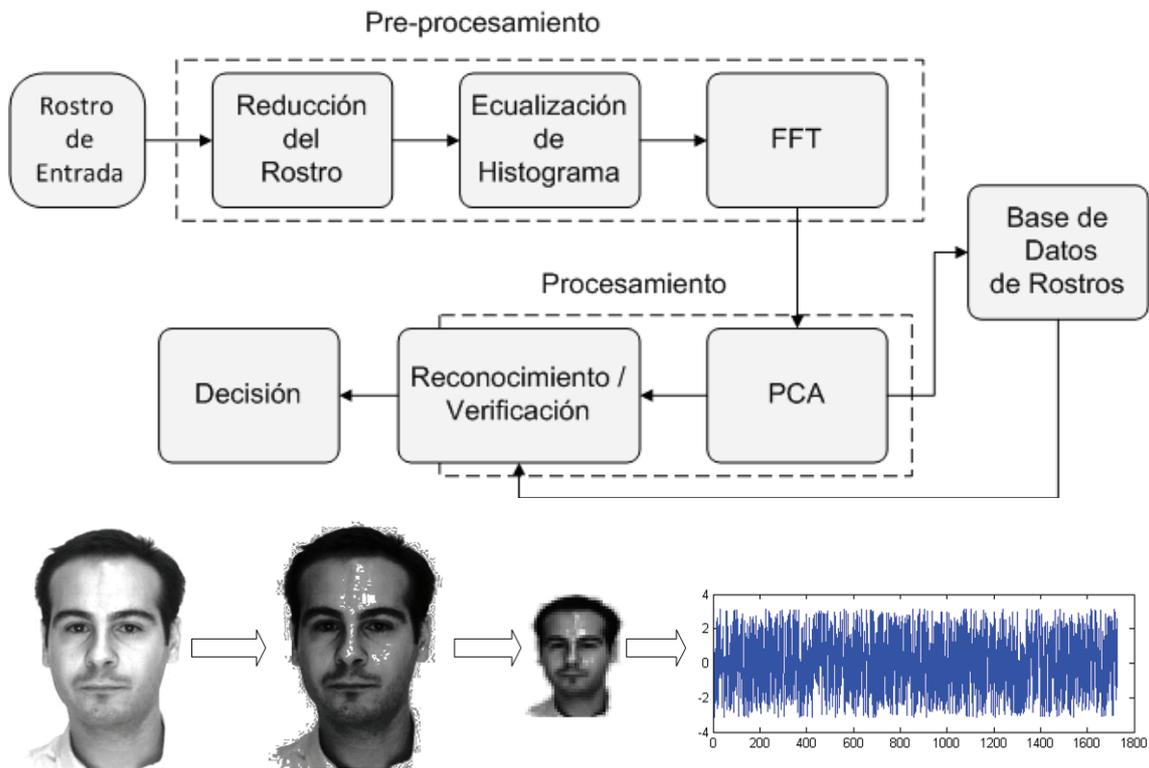
### 3.3 PRE-PROCESAMIENTO

A continuación describiremos cada una de las variaciones en el pre procesamiento del sistema, para esto recurrimos a diagramas de bloque.

Primeramente se realiza el procesamiento sin realizar la ecualización de histograma, con el fin de comparar. Para este caso solo se redimensiona la imagen y aplicamos la transformada rápida de Fourier con el fin de obtener el espectro de fase.



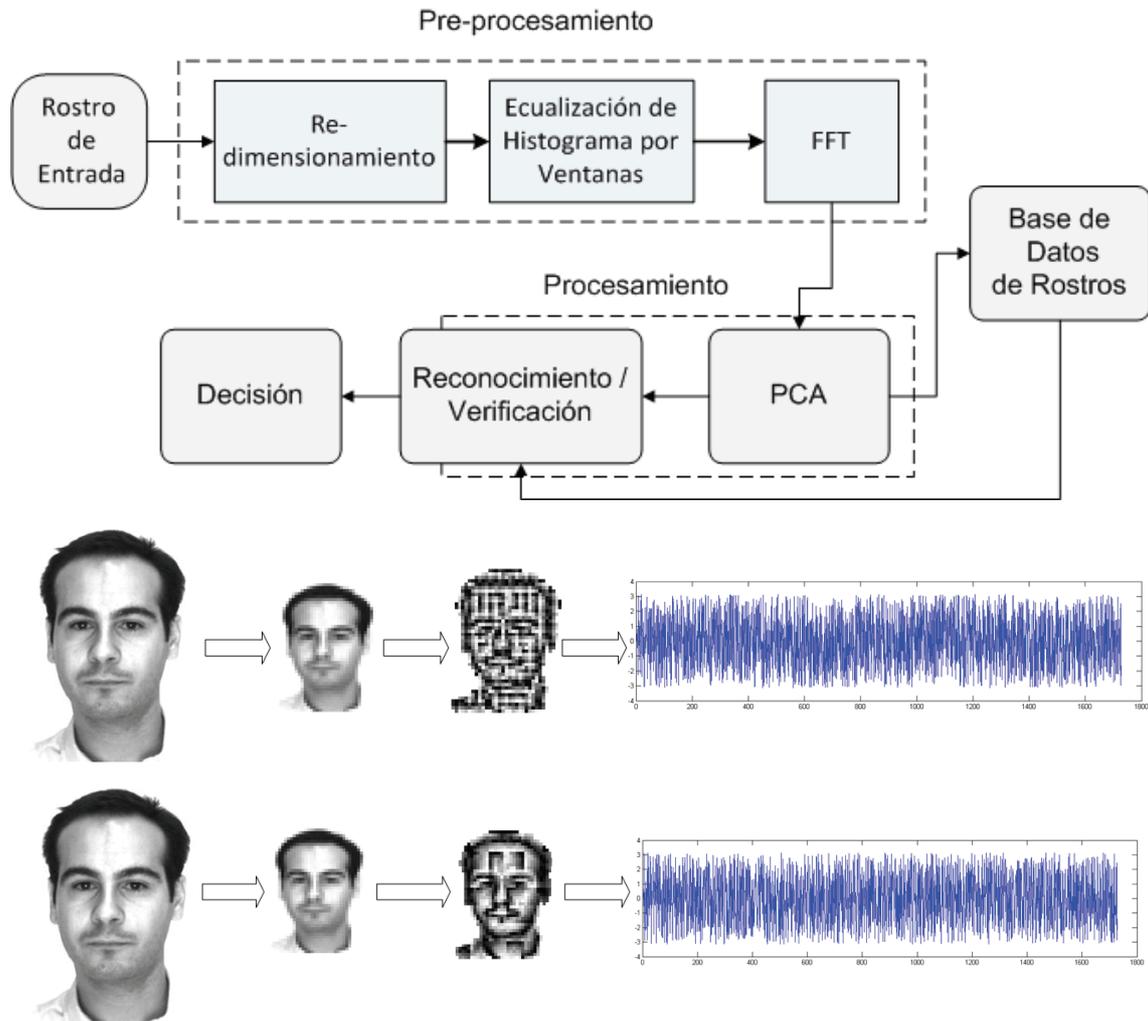
**Figura 3.2** Pre-procesamiento sin ecualización



**Figura 3.3** Pre-procesamiento con ecualización

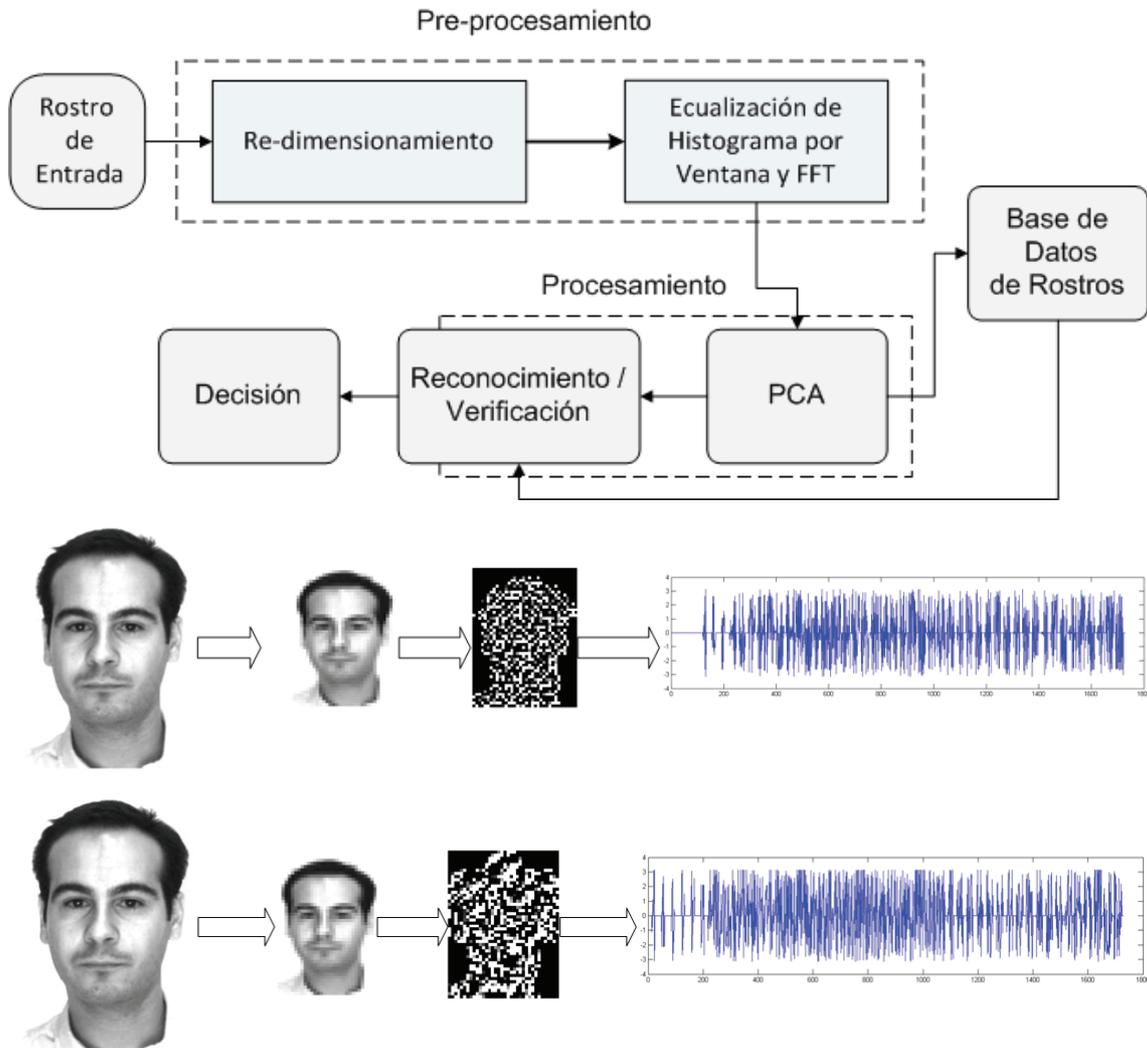
En la Figura 3.3 encontramos la primera variación del sistema propuesto la que consiste en que primero ecualizamos el histograma de la cara y luego redimensionamos la imagen para luego aplicar la Transformada de Fourier y obtener la información relativa a la fase.

En la segunda variación (Figura 3.4) primero redimensionamos la imagen, luego se aplican ventanas que puede ser de  $3 \times 3$  y  $6 \times 6$ , en las cuales se ecualiza el histograma, y luego la imagen se reconstruye para como último paso aplicar la transformada de Fourier y así obtener la fase.



**Figura 3.4** Pre-procesamiento con ecualización por ventana

En la tercera y última variación (Figura 3.5), la imagen original se redimensiona y más tarde se ecualiza el histograma por ventanas de  $3 \times 3$  y  $6 \times 6$  a los cuales se le aplica la transformada rápida de Fourier y como último paso la imagen se reconstruye.



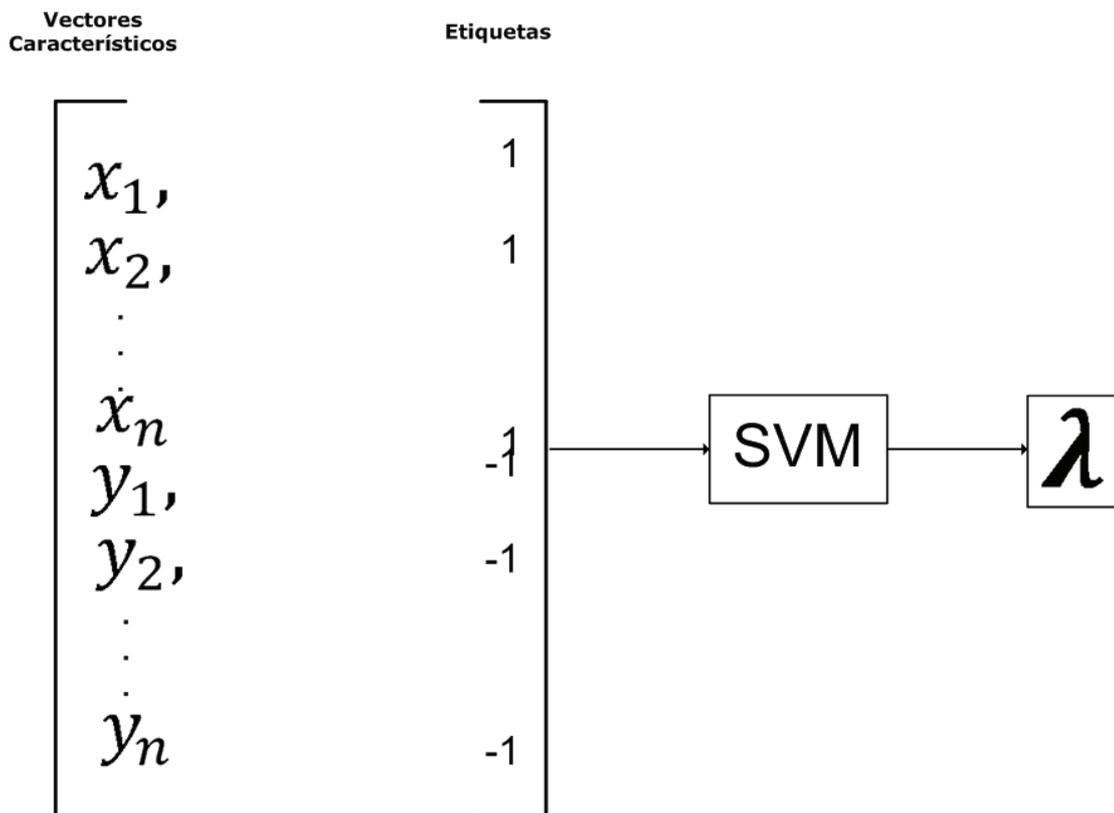
**Figura 3.5** Pre-procesamiento con ecualización por ventana y FFT

### 3.4 EXTRACCIÓN DE CARACTERÍSTICAS

Cuando la fase de pre-procesamiento se termina, el siguiente paso es obtener una matriz de componentes principales, la cual contiene los vectores característicos de los rostros. Para obtener esta matriz creamos dos grupos de diez rostros cada uno, el primer grupo contiene rostros con oclusión, y el segundo rostros sin oclusión. Con cada uno de estos grupos se obtiene una matriz diferente de Componentes Principales. Luego, estas matrices se utilizan para entrenar la SVM, de donde se obtiene un modelo que luego será utilizado en la fase de reconocimiento y verificación.

### 3.5 ENTRENAMIENTO SVM

La Figura 3.6 muestra un esquema general de cómo se entrena el clasificador SVM



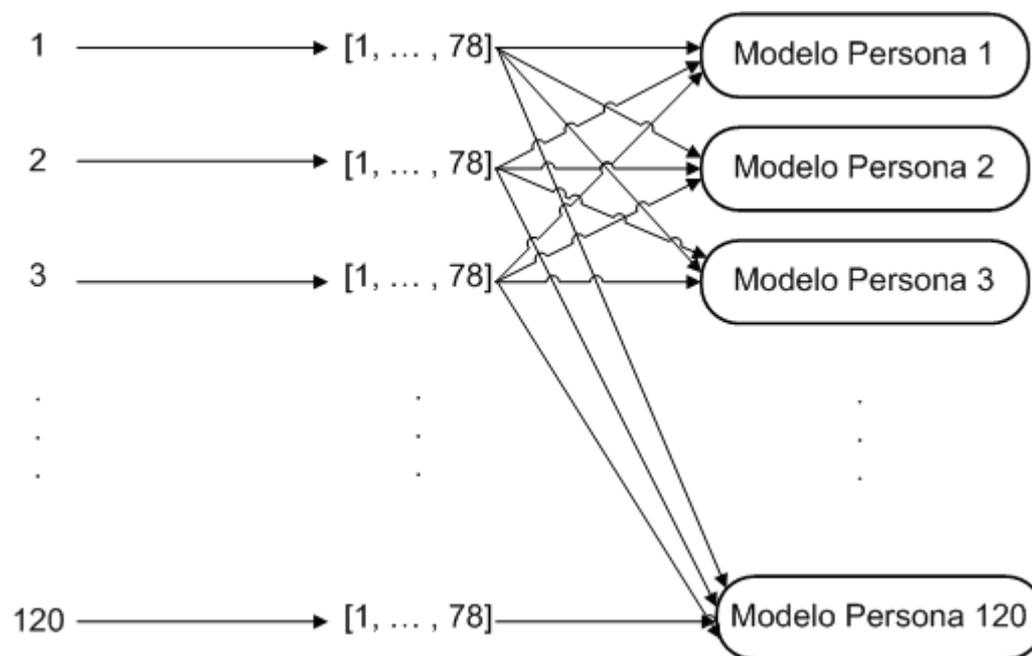
**Figura 3.6** Diagrama general de entrenamiento de SVM

donde  $x_1, x_2, x_3, \dots, x_n, y_1, y_2, y_3, \dots, y_n$ , son los vectores característicos de una persona que se utilizarán para el entrenamiento del SVM,  $n$  es el número de muestras a utilizar por persona, y las etiquetas son para diferenciar una clase de otra en este caso son dos clases ya que se utiliza un modelo de SVM una contra todas (one vs all) en donde 1 es el rostro correcto y -1 es un rostro intruso, posteriormente la SVM es entrenada y obtendremos  $\lambda$  que será el modelo final que utilizaremos para la tarea de reconocimiento o verificación.

Este clasificador nos proporcionara un modelo que está compuesto del número de clases (en este caso 2), el número de vectores de soporte y sus valores, con esto será suficiente para poder hacer satisfactoriamente la tarea de clasificación y verificación.

### 3.6 RECONOCIMIENTO SVM

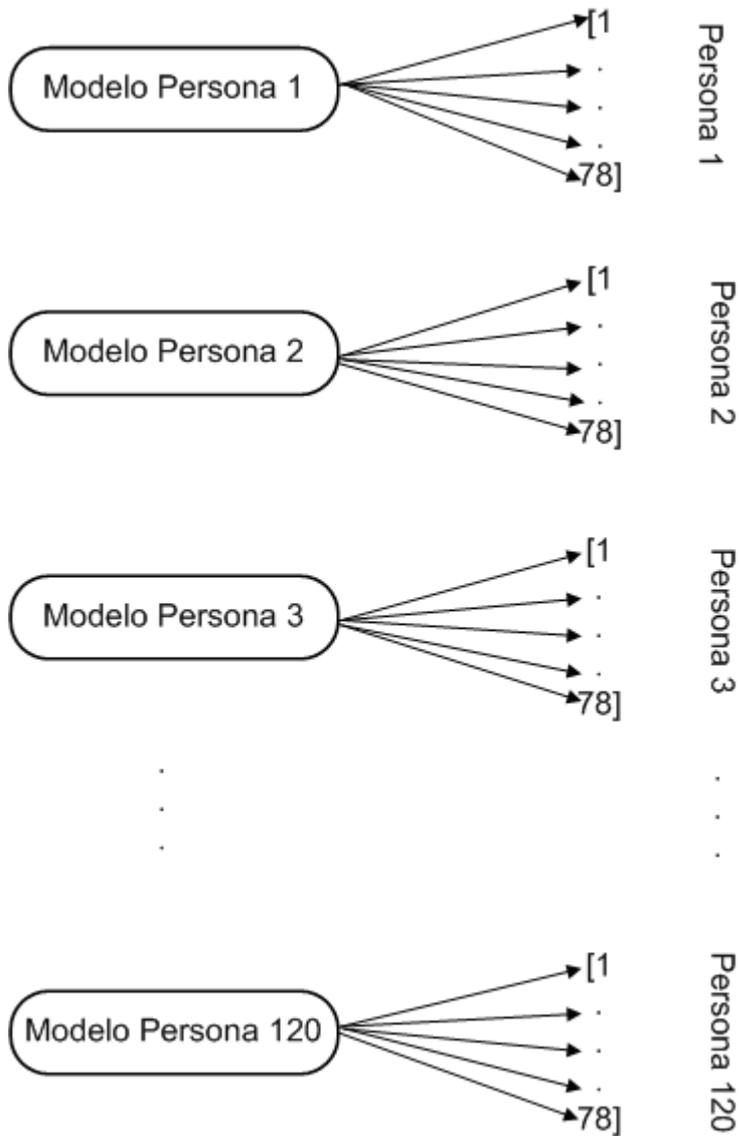
En la fase de Reconocimiento lo que se realiza es comparar cada rostro de cada persona que desea ser reconocido con cada modelo  $\lambda$  obtenido anteriormente.



**Figura 3.7** Diagrama de reconocimiento

### 3.7 VERIFICACIÓN SVM

En la fase de Verificación el procedimiento es diferente, lo que se realiza es comparar cada modelo de cada persona con los rostros de todas las personas, lo que se utiliza en control de acceso para comprobar que la persona es quien dice ser.



**Figura 3.8** Diagrama de verificación

### 3.8 CONCLUSIONES

Se propone un sistema de reconocimiento de rostros que utiliza la Ecuación de Histograma para mejorar el contraste en las imágenes y la iluminación, se utiliza el Análisis de Componentes Principales para la extracción de características y la Máquina de Soporte vectorial como clasificador.

## CAPÍTULO IV

### RESULTADOS

#### 4.1 INTRODUCCIÓN

Para realizar las pruebas se utilizó la base "AR Face Database", esta base de datos cuenta con 120 personas, cada persona cuenta con 78 variaciones de iluminación, expresiones faciales y oclusión parcial del rostro.

Primeramente lo que se realizó fue tomar dos grupos de rostros, que denominaremos Grupo A (Figura 4.1) y Grupo B (Figura 4.2). El Grupo A consta con 10 rostros de cada persona con variaciones de iluminación y el Grupo B toma 10 rostros de cada persona. Estos grupos son utilizados para obtener el modelo de cada individuo que luego será utilizado en la fase de reconocimiento y verificación.



**Figura 4.1** Grupo A de rostros



**Figura 4.2** Grupo B de rostros

Se realizan dos tipos de reconocimiento y de verificación, los que categorizaremos como Tipo I y Tipo II, en el de Tipo I incluimos los 10

rostros utilizados para obtener el modelo y en el Tipo II estos rostros son excluidos.

Con el fin de comparar los resultados al utilizar ecualización, se realizaron pruebas de reconocimiento y verificación sin ecualización, los resultados obtenidos en esta fase se muestran en la siguiente tabla.

**Tabla 4.1**

Resultados de reconocimiento sin ecualización

	<b>Sin Ecualización</b>	
	<i>Grupo A</i>	<i>Grupo B</i>
<i>Tipo I</i>	80.86%	96.41%
<i>Tipo II</i>	78.05%	95.88%

Las pruebas realizadas con Ecualización se muestran a continuación:

**Tabla 4.2**

Resultados de reconocimiento con ecualización

	<b>Con ecualización</b>	
	<i>Grupo A</i>	<i>Grupo B</i>
<i>Tipo I</i>	80.21%	96.27%
<i>Tipo II</i>	77.3%	95.72%

Podemos observar que en ambos casos con y sin ecualización el mejor porcentaje de reconocimiento es el que se obtiene utilizando rostros con oclusión para la obtención del modelo e incluyendo en el reconocimiento estos rostros.

En la siguiente tabla observamos los resultados obtenidos aplicando ventanas:

**Tabla 4.3**

Resultados de reconocimiento con ecualización por ventana de 3x3

	<b>Ventana 3x3</b>	
	<i>Grupo A</i>	<i>Grupo B</i>
<i>Tipo I</i>	81.58%	96.58%
<i>Tipo II</i>	78.87%	96.07%

**Tabla 4.4**

Resultados de reconocimiento con ecualización por ventana de 6x6

	<b>Ventana 6x6</b>	
	<i>Grupo A</i>	<i>Grupo B</i>
<i>Tipo I</i>	81.04%	95.95%
<i>Tipo II</i>	78.25%	95.35%

De igual forma que en los casos anteriores, el porcentaje más alto de reconocimiento es obtenido utilizando rostros con oclusión para la obtención del modelo e incluyendo en el reconocimiento estos rostros.

A continuación se presentan los resultados para ventanas utilizando la transformada de Fourier.

**Tabla 4.5**

Resultados de reconocimiento con ecualización por ventana de 3x3 y FFT

	<b>Ventana 3x3 FFT</b>	
	<i>Grupo A</i>	<i>Grupo B</i>
<i>Tipo I</i>	85.67%	97.57%
<i>Tipo II</i>	83.56%	97.75%

**Tabla 4.6**

Resultados de reconocimiento con ecualización por ventana de 6x6 y FFT

	<b>Ventana 6x6 FFT</b>	
	<i>Grupo A</i>	<i>Grupo B</i>
<i>Tipo I</i>	84.4%	97.37%
<i>Tipo II</i>	82.10%	96.98%

Aquí podemos observar que los mejores resultados se presentan para la misma combinación, excepto para el caso de la ventana de 3x3 en donde el porcentaje más alto fue en los casos que se utilizó rostros con oclusión para la obtención del modelo y en el reconocimiento se omitieron estos rostros.

Las pruebas de Verificación se realizaron bajo las mismas condiciones. En estas pruebas los resultados que tomamos son los porcentajes de Falsa Aceptación, de Falso Rechazo el Error Global.

El porcentaje de mayor interés para nosotros es el de Falsa Aceptación por su aplicación en el control de acceso.

Los resultados se presentan en las siguientes tablas.

**Tabla 4.7**

Resultados de verificación sin ecualización

		<b>Sin Ecualización</b>		
		<b>% Falsa Aceptación</b>	<b>% Falso Rechazo</b>	<b>% Error Global</b>
<i>Tipo I</i>	<i>Grupo A</i>	0.13	5.38	0.18
	<i>Grupo B</i>	0.02	28.29	0.26
<i>Tipo II</i>	<i>Grupo A</i>	0.15	6.17	0.2
	<i>Grupo B</i>	0.03	32.45	0.3

**Tabla 4.8**

Resultados de verificación con ecualización

		<b>Con Ecualización</b>		
		<b>% Falsa Aceptación</b>	<b>% Falso Rechazo</b>	<b>% Error Global</b>
<i>Tipo I</i>	<i>Grupo A</i>	0.007	14.72	0.12
	<i>Grupo B</i>	0.005	33.87	0.28
<i>Tipo II</i>	<i>Grupo A</i>	0.008	16.88	0.14
	<i>Grupo B</i>	0.0067	38.86	0.33

En ambos casos sin y con ecualización observamos que el menor porcentaje obtenido es en caso donde para la obtención del modelo se utilizaron rostros con oclusión y los mismos fueron incluidos para la verificación.

En las siguientes tablas se observan los resultados al aplicar ventanas para la ecualización.

**Tabla 4.9**

Resultados de verificación con ecualización por ventana de 3x3

		<b>Ventana 3x3</b>		
		<b>% Falsa Aceptación</b>	<b>% Falso Rechazo</b>	<b>% Error Global</b>
<i>Tipo I</i>	<i>Grupo A</i>	0.029	37.61	0.31
	<i>Grupo B</i>	0.003	9.51	0.10
<i>Tipo II</i>	<i>Grupo A</i>	0.003	37.12	0.31
	<i>Grupo B</i>	0.03	10.91	0.12

En el caso de la ventana de 3x3 el porcentaje más bajo que se obtiene es en el que para la obtención del modelo se utilizaron rostros con oclusión y los mismos fueron incluidos para la verificación.

**Tabla 4.10**

Resultados de verificación con ecualización por ventana de 6x6

		<b>Ventana 6x6</b>		
		<b>% Falsa Aceptación</b>	<b>% Falso Rechazo</b>	<b>% Error Global</b>
<i>Tipo I</i>	<i>Grupo A</i>	0.02	28.89	0.26
	<i>Grupo B</i>	0.02	11.34	0.11
<i>Tipo II</i>	<i>Grupo A</i>	0.02	33.14	0.3
	<i>Grupo B</i>	0.02	13.01	0.13

En la ventana de 6x6 el porcentaje se mantiene constante para todos los casos.

A continuación se presentan los resultados para la verificación utilizando ventanas con la transformada de Fourier.

**Tabla 4.11**

Resultados de verificación con ecualización por ventana de 3x3 y FFT

		<b>Ventana 3x3 FFT</b>		
		<b>% Falsa Aceptación</b>	<b>% Falso Rechazo</b>	<b>% Error Global</b>
<i>Tipo I</i>	<i>Grupo A</i>	0.72	13.04	0.82
	<i>Grupo B</i>	0.72	2.00	0.73
<i>Tipo II</i>	<i>Grupo A</i>	0.83	14.96	0.95
	<i>Grupo B</i>	0.83	2.3	0.84

En la ventana de 3x3 FFT el menor error se da cuando se utiliza rostros sin oclusión y estos rostros son incluidos en la verificación.

**Tabla 4.12**

Resultados de verificación con ecualización por ventana de 6x6 y FFT

		<b>Ventana 6x6 FFT</b>		
		<b>% Falsa Aceptación</b>	<b>% Falso Rechazo</b>	<b>% Error Global</b>
<i>Tipo I</i>	<i>Grupo A</i>	1.51	11.85	1.59
	<i>Grupo B</i>	1.53	1.14	1.53
<i>Tipo II</i>	<i>Grupo A</i>	1.73	13.60	1.83
	<i>Grupo B</i>	1.76	1.31	1.75

En la ventana de 6x6 FFT el menor error se da cuando se utiliza rostros con oclusión y estos rostros son incluidos en la verificación.

Para control de acceso se realizaron otras pruebas en donde se tomaron solamente los diez rostros con variación de iluminación y para la verificación se utilizaron solamente rostros con variación de iluminación, omitiendo los de oclusión, tomando en cuenta que para control de acceso se establecen restricciones en cuanto al uso de lentes, gorras, bufandas, etc.

En este caso solo tenemos un grupo de 10 rostros para la obtención del modelo y dos tipos de Verificación el tipo I donde incluimos los 10 rostros y el tipo II donde se excluyen.

Los resultados se presentan en las siguientes tablas.

**Tabla 4.13**

Resultados de verificación para control de acceso sin ecualización

	<b>Con Ecualización</b>		
	<b>% Falsa Aceptación</b>	<b>% Falso Rechazo</b>	<b>% Error Global</b>
<i>Tipo I</i>	0.006	5.21	0.05
<i>Tipo II</i>	0.0271	3.6458	0.0573

**Tabla 4.14**

Resultados de verificación para control de acceso con ecualización

	<b>Sin Ecualización</b>		
	<b>% Falsa Aceptación</b>	<b>% Falso Rechazo</b>	<b>% Error Global</b>
<i>Tipo I</i>	0.007	4.26	0.04
<i>Tipo II</i>	0.0092	5.5990	0.0558

Para las pruebas con ecualización y sin ecualización los mejores resultados en Falsa Aceptación son aquellos donde se incluyeron en la verificación los rostros que se utilizaron en la obtención de modelo.

**Tabla 4.15**

Resultados de verificación para control de acceso con ecualización por ventana de 3x3

	<b>Ventana 3x3</b>		
	<b>% Falsa Aceptación</b>	<b>% Falso Rechazo</b>	<b>% Error Global</b>
<i>Tipo I</i>	0.004	6.78	0.06
<i>Tipo II</i>	0.0057	8.9063	0.0799

**Tabla 4.16**

Resultados de verificación para control de acceso con ecualización por ventana de 6x6

	<b>Ventana 6x6</b>		
	<b>% Falsa Aceptación</b>	<b>% Falso Rechazo</b>	<b>% Error Global</b>
<i>Tipo I</i>	0.019	3.37	0.04
<i>Tipo II</i>	0.0256	4.4271	0.0623

Al igual que en las pruebas con ecualización y sin ecualización, en las pruebas con ventanas los mejores resultados en Falsa Aceptación son aquellos donde se incluyeron en la verificación los rostros que se utilizaron en la obtención de modelo.

**Tabla 4.17**

Resultados de verificación para control de acceso con ecualización por ventana de 3x3 y FFT

	<b>Ventana 3x3 FFT</b>		
	<b>% Falsa Aceptación</b>	<b>% Falso Rechazo</b>	<b>% Error Global</b>
<i>Tipo I</i>	0.57	0.45	0.57
<i>Tipo II</i>	0.7596	0.5990	0.7582

**Tabla 4.18**

Resultados de verificación para control de acceso con ecualización por ventana de 6x6 y FFT

	<b>Ventana 6x6 FFT</b>		
	<b>% Falsa Aceptación</b>	<b>% Falso Rechazo</b>	<b>% Error Global</b>
<i>Tipo I</i>	1.15	0.29	1.15
<i>Tipo II</i>	1.5211	0.3906	1.5117

Aplicando Ventanas con la transformada de Fourier, al igual que en los casos anteriores los mejores porcentajes son aquellos donde se incluyeron en la verificación los rostros que se utilizaron en la obtención de modelo, con la diferencia que en estos casos, los errores aumentan con respecto a los anteriores.

## CAPITULO V

### CONCLUSIONES GENERALES

#### 5.1 CONCLUSIONES

Podemos concluir que para el reconocimiento el porcentaje más alto fue utilizando la ecualización por ventana de 3x3 con la transformada de Fourier, el cual fue de 97.75%. Lo que se debe a la obtención de la fase en pequeños bloques.

En el caso de la Verificación el menor porcentaje de error en Falsa Aceptación fue que fue de 0.0029% utilizando la ventana de 3x3 con los rostros sin oclusión e incluyendo todas las variaciones de los rostros.

Para las pruebas realizadas para verificación en control de acceso, los menores porcentajes de error en falsa aceptación fueron para la ventana de 3x3 con un porcentaje de 0.004%

Podemos concluir que el uso de una ventana más grande no presento mejora en cuanto a los porcentajes de verificación y reconocimiento, al igual que el aplicar ventanas con la transformada de Fourier, no representó una gran mejoría, comparados con los de la ventana de 3x3.

Analizando los resultados obtenidos a través de las pruebas realizadas, podemos concluir primeramente que la Máquina de Soporte Vectorial, como clasificador para el reconocimiento y verificación de rostros, es una herramienta que permite obtener resultados muy satisfactorios, debido a su alta habilidad para generalizar.

Podemos concluir que la imagen mejora para la extracción de sus características. El uso de Análisis de Componentes Principales como

extractor de características ayuda a generalizar mejor los vectores haciendo uso de menos datos.

### 5.3 TRABAJOS FUTUROS

Algunos trabajos que se podrían realizar a futuro en base a este trabajo son:

- Utilizar otro tipo de clasificador para comparar los resultados.
- Usar imágenes a color.
- La combinación con otro sistema para crear un sistema biométrico híbrido.
- Probar el sistema con variaciones de rotación, tamaño e inclinación

## BIBLIOGRAFÍA

- [1] Jain A.K., Ross R. and Prabhakar S. "An introduction to biometric recognition", IEEE Trans. On Circuits and Systems for Video Technology, Vol. 14, no. 1, January 2004, pp. 4-20
- [2] Zhao W. Chellappa, R. Phillips P.J. and Rosenfeld A. "Face Recognition: A literature survey" ACM Comput. Surv. Vol. 35, no.4, December 2003, pp. 399-459
- [3] Dao-Qing Dai and Hong Yan Sun Yat-Sen. "Wavelets and Face Recognition", University and City, University of Hong Kong
- [4] J. Olivares-Mercado, K. Hotta, H. Takahashi, M. Nakano-Miyatake, K. Toscano-Medina, H. Perez-Meana, "Improving the Eigenphase Method for Face Recognition", IEICE Electronic Express , vol. 6, no. 15, pp. 1112-1117, 2009.
- [5] R. Plata-Garduño y M. Nakano-Miyatake. "Reconocimiento Automático de Rostros: Método de Eigenfaces"
- [6] Sandra María Villa Palacios, "Sistema de Reconocimiento de Rostros"; Universidad Peruana de Ciencias Aplicadas (UPC)
- [7] D. Bolme, R. Beveridge, M. Teixeira, and B. Draper, "The CSU Face Identification Evaluation System: Its Purpose, Features and Structure", International Conference on Vision Systems, Graz, Austria, April 1-3, 2003. (Springer-Verlag) 304-311.
- [8] "Eigenface Recognition"  
<http://et.wcu.edu/aids/BioWebPages/eigenfaces.htm>
- [9] D. Bolme, R. Beveridge, M. Teixeira, and B. Draper, "The CSU Face Identification Evaluation System: Its Purpose, Features and Structure" International Conference on Vision Systems, Graz, Austria, April 1-3, 2003. (Springer-Verlag) 304-311.

- [10] Lu, K.N. Plataniotis, and A.N. Venetsanopoulos, "Regularized Discriminant Analysis for the Small Sample Size Problem in Face Recognition," Pattern Recognition Letters, December 2003, Vol. 24, Issue 16: 3079-3087.
- [11] R. Plata-Garduño y M. Nakano-Miyatake. "Reconocimiento Automático de Rostros: Método de Eigenfaces"
- [12] Dunn, D., Higgins, W. E., "Optimal Gabor Filters for Texture Segmentation," IEEE Trans. Image Proc., Vol. 4, No. 7, Jul. 1995.
- [13] Shanks, J. L., "Computation of the Fast Walsh-Fourier Transform," IEEE Trans. Comput., Vol. 18, No. 5, pp. 457-459, 1969.
- [14] Tema 3: Análisis de Componentes Principales <http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/AMult/tema3am.pdf>
- [15] Pervys Rengifo Rengifo. "Máquinas De Soporte Vectorial", Facultad De Matemáticas E Ingenierías, Ingeniería De Sistemas. Fundación Universitaria Konrad Lorenz
- [16] Vapnik, V. "The Nature of Statistical Learning Theory", Springer, New York, 1995
- [17] Vapnik, V. "Statitistical Learning Theory", Wiley, New York, 1998
- [18] Vapnik, V., & Chervonenkis Theory of Pattern Recognition [inRussian], Nauka, Moscow 1974.
- [19] Burges, C.J.C., "A Tutorial on Support Vector Machines for Pattern Recognition", Data Mining and Knowledge Discovery, 1998, pp. 121-167
- [20] N. Cristianini and J. Shawe-Taylor," An introduction to support vector machines", Cambridge University Press, 2000

- [21] Osuna, E., Freud, R. and Girosi, F. "An improved training algorithm for support vector machines". In Proc. Of IEEE Workshop on Neural Network for Signal Processing VII, New York, 1997, pp. 276-285
- [22] Minoux M., "Mathematical Programming: Theory and Algorithms". John Wiley and Sons, 1986
- [23] Scholkopf, B "Support Vector Learning" 1997, R. Oldenbourg, Munich
- [24] A. V. Oppenheim M. H. Hayes, J. S. Lim. Signal reconstruction from phase or magnitude. *IEEE Trans. Acoust., Signal Processing*, 28:672-680, 1980.
- [25] J. S. Lim A. V. Oppenheim. The importance of phase in signals. *Proc. IEEE*, 69(5):529-541, 1981.
- [26] Jesus Olivares Mercado "Reconocimiento y Verificación de Identidad por medio del Rostro" Escuela Superior de Ingeniería Mecánica y Eléctrica "Unidad Culhuacan" Instituto Politécnico Nacional, 2008

## ANEXO A

### GLOSARIO

**Algoritmo:** Secuencia limitada de instrucciones o pasos que indica a un sistema computarizado cómo resolver un problema en especial. Un sistema biométrico utiliza múltiples algoritmos; por ejemplo, para el procesamiento de imágenes, la generación de plantillas, comparaciones, etc.

**Característica:** Características matemáticas distintivas derivadas de una muestra biométrica, utilizadas para generar una referencia.

**Extracción:** Proceso de conversión de una muestra biométrica capturada en datos biométricos para que puedan ser comparados con una referencia.

**Ecualización de Histograma:** La ecualización del histograma de una imagen es una transformación que pretende obtener para una imagen un histograma con una distribución uniforme. Es decir, que exista el mismo número de píxeles para cada nivel de gris del histograma de una imagen monocroma.

**Histograma:** Es una representación gráfica de una variable en forma de barras, donde la superficie de cada barra es proporcional a la frecuencia de los valores representados. En el eje vertical se representan las frecuencias, y en el eje horizontal los valores de las variables, normalmente señalando las marcas de clase, es decir, la mitad del intervalo en el que están agrupados los datos.

**Modelo:** Representación utilizada para caracterizar a un individuo.

**Oclusión:** se refiere al estado de algo comúnmente abierto y que se encuentra obliterado o cerrado.

**Pixel:** es la menor unidad homogénea en color que forma parte de una imagen digital.

**Reconocimiento:** Término general utilizado en la descripción de sistemas biométricos (por ejemplo, reconocimiento de rostro o reconocimiento de iris) en relación con su función principal. El término "reconocimiento" no implica necesariamente verificación, identificación de grupo cerrado ni identificación de grupo abierto (lista de vigilancia).

**Reconocimiento de rostro:** Modalidad biométrica que utiliza una imagen de la estructura física visible del rostro de una persona para fines de reconocimiento.

**Rostro:** La cara (también rostro y faz) es la parte frontal de la cabeza, en los humanos la parte anteroinferior, desde las cejas a la barbilla, están incluidos cejas, ojos, nariz, mejilla, boca, labios, dientes, piel y barbilla.

**Tasa de falsa aceptación:** Estadística utilizada para medir el rendimiento biométrico durante la tarea de verificación. Porcentaje de veces que un sistema produce una falsa aceptación, lo cual ocurre cuando un individuo es erróneamente vinculado con la información biométrica existente de otra persona. Por ejemplo: Frank declara ser John, y el sistema verifica esta declaración

**Tasa de falso rechazo:** Estadística utilizada para medir el rendimiento biométrico durante la tarea de verificación. Porcentaje de veces que el sistema produce un falso rechazo. Ocurre un falso rechazo cuando un individuo no es vinculado con su propia plantilla biométrica existente. Por ejemplo: John declara ser John, pero el sistema incorrectamente niega dicha declaración.

**Verificación:** Estadística utilizada para medir el rendimiento biométrico durante la tarea de verificación. Tasa en la cual los usuarios finales legítimos son correctamente verificados.

## ANEXO B

### PUBLICACIONES

*"Face Recognition and Verification using Histogram Equalization"*

Kelsey Ramírez Gutiérrez, Daniel Cruz Pérez, Héctor Pérez Meana

WSEAS Conferences in Iwate Prefectural University, Iwate, Japan, October 4-6, 2010 (Aceptado)

**Dear Author (of accepted paper in our ISI Books, Proceedings and Journals) This is the letter of Acceptance from WSEAS** Inbox | X

from **WSEAS (Letter of acceptance) -- Japan 2010** <support@wseas.org> [hide details](#) Jun 9 (3 days ago) Reply  
 to kramirezg0901@ipn.mx  
 date Wed, Jun 9, 2010 at 5:24 AM  
 subject Dear Author (of accepted paper in our ISI Books, Proceedings and Journals)  
 This is the letter of Acceptance from WSEAS

**Images are not displayed.**

[Display images below](#) - [Always display images from support@wseas.org](#)

Dear Author (of accepted paper in our ISI Books, Proceedings and Journals)

We would like to inform you that your paper (we have just sent all the details to the responsible for the correspondence author, i.e. to the author that had uploaded your paper via the web form) has been accepted for the WSEAS Conference proceedings and our ISI book (+invitation for extended paper for the journal) for the WSEAS Conferences in Iwate Prefectural University, Iwate, Japan, October 4-6, 2010. The Conference is going to be held in the **Iwate Prefectural University, Iwate, Japan** which is the main Sponsor and main Organizer of the 5 Parallel Conferences. Delegates attending these conferences will be invited to submit an extended version of their paper for publication in the International Journal of NAUN and WSEAS after the conference.

#### **REVIEWERS' COMMENTS:**

The reviewers' comments, for the improvement of your paper for final publication in the Proceedings as well as for valuable assistance for the extended journal publication, will be sent to you after your registration, in an attempt to reduce our secretariat cost as well as to ensure your participation & presentation in the conference - discouraging authors that send papers only to receive reviewers' comments). All the papers have been reviewed by 2 or 3 independent reviewers.

Therefore, after your registration (see the attached file and the instructions below) your paper is going to appear in

- a) *Book in ISI*. Please visit: <http://www.wseas.us/indexes/>
- b) *CD-ROM Proceedings*,
- c) *Journal of NAUN or University Press* (you select the journal from the list of NAUN or UniversityPress)
- d) *E-Library of WSEAS* (Over 40000 visits per month)

**Deadline for registration: September 15, 2010 - however, earlier submission of your paper is recommended in order to receive Reviewers' comments in time for improvements.**

*"Face Verification System using Histogram Equalization and Support Vector Machine"*

Kelsey Ramírez Gutiérrez, Daniel Cruz Pérez, Héctor Pérez Meana

The Electronics, Robotics and Automotive Mechanics Conference (CERMA) 2010 (En revisión)

*"Face Recognition Using Principal Components Analysis and Histogram Equalization"*

Kelsey Ramírez-Gutiérrez, Daniel Cruz-Pérez, Héctor Pérez-Meana

2010 Mexican Meeting on Informatics Security (M2IS 2010) (En revisión)