

# UNIVERSIDAD IBEROAMERICANA



## “RECONOCIMIENTO DE ROSTROS UTILIZANDO ANÁLISIS DE COMPONENTES PRINCIPALES: LIMITACIONES DEL ALGORITMO”

### TESIS

Que para obtener el grado de

**MAESTRO EN SISTEMAS Y PLANEACION**

**P r e s e n t a:**

**CARLOS VILLEGAS QUEZADA**

**Director**

**MTRO. JORGE RIVERA ALBARRAN**

**Asesores:**

**MTRO. PEDRO FERNANDO SOLARES SOTO**

**MTRO. FELIPE ANTONIO TRUJILLO FERNANDEZ**

## Resumen

Una de las tecnologías emergentes que muestra grandes posibilidades de desarrollo en los próximos años, es la Biometría, la cual tiene como propósito la identificación de personas por medio de sus diversas características físicas (voz, huellas digitales, características faciales, etc.). A partir de los ataques terroristas a las dos torres del World Trade Center en New York, el 11 de septiembre de 2001, los sistemas biométricos cobraron un gran auge. Sobre todo, los sistemas de reconocimiento de rostros. Dichos sistemas tendrán el objetivo de detectar posibles sospechosos, a partir de diversas bases de datos con fotografías de personas consideradas terroristas o delincuentes. También, tienen otros múltiples usos de identificación en la vida cotidiana: acceso a cajeros automáticos, a estacionamientos, acceso de personal a oficinas, reconocimiento de personas desaparecidas y fallecidas que no presenten identificación, acceso a computadoras (en lugar de utilizar password), etc.

En la actualidad, los dos paquetes de software comercial más utilizados en el reconocimiento de rostros, utilizan la técnica de Análisis por Componentes Principales o alguna derivación de ella. Asimismo, la gran mayoría de los paquetes computacionales realizados en proyectos de investigación para reconocimiento e identificación de rostros, que utilizan otros métodos estadísticos (regresión, redes neuronales, análisis cluster, etc.), comparten algo en común con el análisis de componentes principales: todos ellos utilizan cálculos y métricas que se llevan a cabo en un espacio denominado Euclidiano o  $L_2$ .

El software basado en técnicas que utilizan la métrica Euclidiana, no ha presentado resultados completamente satisfactorios en el proceso de reconocimiento de rostros. Se considera que dichas limitaciones se presentan por dos motivos principales: el primero, debido al propio algoritmo, en segundo lugar, al tipo de características estadísticas que presentan los datos de entrada al sistema (los rostros). Para poder utilizar adecuadamente métodos estadísticos en espacios  $L_2$ , los datos de entrada deben cumplir los supuestos de linealidad, homoscedasticidad y sobre todo, el de normalidad multivariada. Se postula en el presente trabajo, que los datos numéricos que caracterizan la fotografía digital de un rostro, no cumplen los presupuestos mencionados. Por tanto, dicha violación, contribuye a la obtención de resultados erróneos en el proceso de reconocimiento a pesar de la “robustez” de los métodos de análisis de componentes principales y técnicas similares.

En la tesis se describen los elementos teóricos que subyacen a la técnica de componentes principales y con ellos, se desarrolló un sistema computacional en MATLAB basado en dicha técnica. Asimismo, se realizó un análisis estadístico de las características de las imágenes digitales de rostros (base AT&T y fotografías adicionales del portal del FBI), para comprobar si cumplían los supuestos de normalidad, homoscedasticidad y linealidad.

Con el sistema computacional elaborado, se realizaron diversos experimentos de reconocimiento de rostros utilizando diversas fotografías por sujeto en la base de entrenamiento, con el propósito de observar el comportamiento del sistema. Los resultados muestran que no se cumplen con los supuestos estadísticos mencionados, los cuales se requieren para utilizar adecuadamente técnicas basadas en espacios Euclidianos. Por otra

parte, el método de componentes principales presentó resultados adecuados, pero también, muestra serias dificultades y limitaciones que impiden un reconocimiento de rostros totalmente confiable.

Finalmente, se postuló un nuevo método aplicable al reconocimiento de rostros, que se basa en la utilización de una *familia de funciones polinomiales multivariadas de aproximación manejadas en un Espacio  $L_\infty$* . Este método, al trabajar el problema de aproximación para caracterizar a los rostros en un espacio *n-dimensional* bajo la norma  $L_\infty$  (norma mínimax o Tchebyshev), no requiere el cumplimiento de los supuestos de linealidad, homoscedasticidad y normalidad. Por tanto, se piensa que eliminará o reducirá las limitaciones que presentan los sistemas actuales basados en técnicas de componentes principales o similares al utilizar datos que no satisfacen los supuestos estadísticos mencionados. La técnica que se propone, utilizará el denominado Algoritmo Genético Ecléctico, para encontrar el aproximante que caracterizará a los rostros bajo un enfoque de optimización combinatoria y aproximación multivariada.

# Tabla de Contenido

---

<b>Introducción .....</b>	<b>13</b>
<b>1. Planteamiento del estudio .....</b>	<b>28</b>
1.1. El problema de investigación .....	28
1.2. Delimitación del problema .....	30
1.3. Preguntas de investigación .....	31
1.4. Hipótesis .....	31
1.5. Objetivos .....	32
1.5.1. Objetivo general .....	32
1.5.2. Objetivos particulares .....	32
1.6. Limitaciones del problema .....	33
1.7. Resultados que se pretende lograr con el proyecto .....	34
<b>2. El “Estado del conocimiento” en el reconocimiento de rostros humanos por métodos computacionales .....</b>	<b>36</b>
2.1. Enfoque psicológico y de neurociencias en el reconocimiento de rostros en el ser humano .....	37
2.2. Enfoque computacional .....	41
2.2.1. Segmentación de rostros .....	42
2.2.2. Extracción de características .....	45
2.2.3. Reconocimiento de rostros .....	52
2.2.4. Reconocimiento de rostros a partir de perfiles .....	70
2.2.5. Métodos evolutivos .....	72
<b>3. Análisis de componentes principales .....</b>	<b>74</b>
3.1. Características generales del análisis de componentes principales .....	75
3.2. Planteamiento matemático del análisis de componentes principales .....	77
3.3. Representación de rostros utilizando el análisis de componentes principales .....	85
3.4. Algoritmos para reconocimiento de rostros utilizando análisis de componentes principales .....	87

<b>4.</b>	<b>Análisis estadístico de las imágenes digitales de rostros .....</b>	<b>92</b>
4.1.	Supuestos estadísticos del análisis multivariante tradicional .....	93
4.2.	Análisis estadístico de la imagen de un rostro .....	95
4.3.	Prueba de normalidad univariada y multivariada para fotografías de rostros .....	100
4.3.1.	Prueba de normalidad univariada .....	100
4.3.2.	Prueba de normalidad multivariada .....	106
4.3.3.	Prueba de homoscedasticidad y linealidad .....	109
<b>5.</b>	<b>Análisis del reconocimiento de rostros utilizando técnicas de componentes principales .....</b>	<b>113</b>
5.1.	Características de la base de datos de rostros utilizada .....	114
5.2.	Elaboración de la Base de Entrenamiento .....	118
5.3.	Configuración de los experimentos .....	119
5.4.	Características, Algoritmo y Funcionamiento del sistema de reconocimiento de rostros .....	127
5.4.1.	Características generales del sistema .....	127
5.4.2.	Algoritmo .....	127
5.4.3.	El sistema computacional para reconocimiento de rostros .....	132
5.4.4.	Operación y tipo de resultados que proporciona el sistema .....	135
5.5.	Experimentación .....	144
5.5.1.	Resultados con la base AT&T .....	149
5.5.2.	Experimentos con la base combinada AT&T y la del FBI .....	167
<b>6.</b>	<b>Propuesta de un método para reconocimiento holístico de rostros utilizando aproximación multivariada y algoritmos genéticos eclécticos en un espacio <math>L_\infty</math> .....</b>	<b>178</b>
6.1.	Reconocimiento de rostros como un problema de aproximación multivariada .....	180
6.2.	Reconocimiento de patrones y Aprendizaje .....	182
6.3.	Algoritmos Genéticos .....	186
6.4.	Aproximación multivariada en un Espacio $L_\infty$ como un problema de optimización combinatoria .....	197
6.5.	Método general para el reconocimiento de rostros utilizando algoritmos genéticos eclécticos y aproximantes polinomiales en un espacio $L_\infty$ .....	202
6.5.1.	Muestreo aleatorio de píxeles y configuración de variables .....	203

6.5.2. Reconocimiento de rostros como un problema de aprendizaje y obtención del aproximante multivariado mediante un algoritmo genético ecléctico .....	204
6.5.3. El rostro y sus espectros de señal .....	208
<b>Conclusiones .....</b>	<b>214</b>
<b>Referencias bibliográficas.....</b>	<b>219</b>
<b>Anexos .....</b>	<b>237</b>
Anexo A. Rostros de la Base AT&T .....	238
Anexo B. Resultados del análisis univariado de normalidad para los 400 rostros de la base AT&T .....	259

# Lista de Tablas

---

## TABLA

1.	Estadística descriptiva de los valores a nivel píxel que presenta la Fotografía s0802.pgm .....	96
2.	Estadística descriptiva del “rostro promedio” de las 400 fotografías de la base AT&T .....	99
3.	Resultados de la prueba de Normalidad Univariada (Kolmogorov Smirnov) para algunos de los rostros de la base AT&T .....	102
4.	Prueba de normalidad univariada para el “rostro promedio” de la base AT&T .....	106
5.	Resultados de la prueba de normalidad multivariada utilizando el software PRELIS, a una muestra de la base AT&T .....	108
6.	Resultados posibles considerando la decisión del sistema de reconocimiento en relación a la verdad sobre la fotografía a identificar .....	145
7.	Porcentajes y numero de rostros identificados en cada rubro de acuerdo a diversos Intervalos de confianza (con 40 sujetos y 9 fotografías por cada uno) .....	151
8.	Número de rostros identificados en cada rubro de acuerdo a diversos intervalos de confianza (con 40 sujetos en b.d. y 8 fotografías por cada uno, 53 rostros para identificar) .....	152
9.	Numero de rostros identificados en cada rubro de acuerdo a diversos intervalos de confianza (con 40 sujetos y 5 fotografías por cada uno) .....	152
10.	Numero de rostros y porcentajes de identificación considerando los umbrales $(\bar{X} + 2\sigma)$ y $(\bar{X} + 3\sigma)$ con la menor tasa posible de “falsos positivos” / “falsos negativos” y el mayor porcentaje posible de “positivos verdaderos” para los experimentos realizados con 40 sujetos (con diverso numero de fotografías por sujeto) y 53 fotografías a reconocer .....	154
11.	Número de rostros identificados en cada rubro de acuerdo a diversos intervalos de confianza (con 33 sujetos y 9 fotografías por cada uno) .....	156
12.	Número de rostros identificados en cada rubro de acuerdo a diversos intervalos de confianza (con 33 sujetos y 7 fotografías por cada uno) .....	156

13.	Número de rostros identificados en cada rubro de acuerdo a diversos intervalos de confianza (con 33 sujetos y 5 fotografías por cada uno) .....	157
14.	Número de rostros identificados en cada rubro de acuerdo a diversos intervalos de confianza (con 33 sujetos y 3 fotografías por cada uno) .....	157
15.	Número de rostros identificados en cada rubro de acuerdo a diversos intervalos de confianza (con 33 sujetos y 2 fotografías por cada uno) .....	158
16.	Número de rostros identificados en cada rubro de acuerdo a diversos intervalos de confianza (con 33 sujetos y 1 fotografía por cada uno) .....	158
17.	Número de rostros y porcentajes de identificación considerando un umbral de $(\bar{X} + 2\sigma)$ para los seis experimentos realizados con 33 sujetos (con diverso numero de fotografías por sujeto) y 53 fotografías a reconocer .....	160
18.	Número de rostros y porcentajes de identificación considerando una tasa de “falsos positivos” del 0.0% y máximo porcentaje de “positivos verdaderos” para los seis experimentos realizados con 33 sujetos (con diverso numero de fotografías por sujeto) y 53 fotografías a reconocer .....	162
19.	Número de rostros y porcentajes de identificación considerando una tasa de “falsos positivos” del 0.0% y clasificadas por máximo porcentaje de “positivos verdaderos” para los seis experimentos realizados con 33 sujetos (con diverso numero de fotografías por sujeto) y 53 fotografías a reconocer .....	163
20.	Número de rostros y porcentajes de identificación considerando el mejor umbral $(\bar{X} + 2\sigma)$ con la menor tasa posible de “falsos positivos” y el mayor porcentaje posible de “positivos verdaderos” para los seis experimentos realizados con 33 sujetos (con diverso numero de fotografías por sujeto) y 53 fotografías a reconocer .....	165
21.	Número de rostros y porcentajes de identificación considerando el mejor umbral $(\bar{X} + 2\sigma)$ clasificados por mayor porcentaje general de reconocimiento, menor porcentaje de “falsos positivos” y menor numero de fotografías por sujeto para los seis experimentos realizados con 33 sujetos (con diverso número de fotografías por sujeto) y 53 fotografías a reconocer .....	165
22.	Resultados de reconocimiento utilizando la base combinada del FBI con AT&T .....	173
23.	Porcentajes de reconocimiento verdadero para las fotografías de sospechosos del FBI (con respecto a 11 fotografías) .....	174

24	Comparación de resultados considerando un umbral de $\bar{X} + 2\sigma$ entre el experimento 9 y el experimento con la base AT&T / FBI .....	175
25.	Población inicial en un algoritmo genético simple .....	191
26.	Población inicial y selección de cromosomas .....	193
27.	Cruzamiento y nueva población .....	194
28.	Población final .....	195
29.	Matriz de características de píxeles (“m” características x “tm” píxeles) para obtener el aproximante que pueda caracterizar al rostro .....	203

# Lista de Figuras

---

## FIGURA

1.	Sistema de reconocimiento antropométrico de Alphonse Bertillion para el reconocimiento de criminales en Europa, 1877-1884 .....	14
2.	Portada del libro “Finger Prints” publicado por Francis Galton en 1892, primer texto sobre identificación de huellas dactilares .....	15
3.	Reconocimiento de rostros utilizando (a) características geométricas y (b) plantillas .....	55
4.	Reconocimiento utilizando puntos seleccionados del rostro para ser procesados por la técnica de componentes principales .....	56
5.	Reconocimiento utilizando un “clasificador por muestreo” a partir de “n-tuplas” que se obtienen de los valores de píxeles .....	57
6.	Reconocimiento de rostros utilizando el metodo “Eigenfaces” propuesto por Turk y Pentland .....	58
7.	Ejemplo de fotografía de la base AT&T (segunda imagen de la octava persona de la base – s0802.pgm) .....	96
8.	Histograma de los niveles de gris de los pixeles correspondientes a la fotografía s0802.pgm .....	98
9.	Histograma de los niveles de gris correspondientes a los pixeles del “rostro promedio” de la base at&t .....	99
10.	Grafico Q-Q de normalidad univariada para la fotografía s0101.pgm .....	103
11.	Grafico Q-Q de normalidad univariada para la fotografía s0108.pgm .....	103
12.	Nube de puntos en “x” para los pixeles del sujeto 8/foto 8 .....	110
13.	Nube de puntos en “y” para los pixeles del sujeto 8/foto 8 .....	111
14.	Nube de puntos 3-D para los pixeles del sujeto 8/foto 8 .....	112
15.	Configuración de la base de rostros AT&T conteniendo fotografías de 40 sujetos con 10 imágenes por cada persona .....	117
16.	Los primeros 25 rostros de la base AT&T utilizados como “fotografías a identificar” .....	121
17.	Los siguientes 15 rostros de la base AT&T utilizados como “fotografías a identificar”, para dar un total inicial de 40 fotografías a reconocer .....	122

18.	Trece fotografías adicionales para ser utilizadas como imágenes a identificar .....	123
19.	Modelo simplificado del “espacio de rostros” que se origina al realizar un análisis de componentes principales, ilustrando los cuatro resultados posibles que se obtienen al proyectar la imagen de un rostro en el espacio .....	131
20.	Ejecución del sistema de reconocimiento y Menú principal .....	132
21.	Selección de una fotografía a reconocer .....	135
22.	Reconocimiento de un rostro (verdadero positivo) .....	136
23.	Reconocimiento de un rostro (verdadero positivo). El sujeto a reconocer presenta anteojos y mayor edad .....	138
24.	Reconocimiento de un rostro (verdadero positivo). Las fotografías utilizadas no se encuentran controladas (luz, fondo de imagen, resolución, etc.) .....	138
25.	Reconocimiento “verdadero positivo” de una fotografía controlada de la base AT&T .....	139
26.	Se intenta reconocer una imagen que no corresponde a un rostro real. El sistema reporta adecuadamente “rostro no reconocido” .....	140
27.	Ejemplo de una identificación de tipo “falso negativo” .....	141
28.	Ejemplo de una identificación de tipo “falso positivo” .....	142
29.	Ejemplo de una identificación combinada de tipo “falso negativo” y “falso positivo” .....	143
30.	Reporte de periódico a partir del cual se obtuvieron dos fotografías digitalizadas para experimentación .....	168
31.	Ejemplo de un reporte de delincuentes publicado en internet por el FBI .....	169
32.	Ejemplo de un reporte de terroristas publicado en internet por el FBI .....	170
33.	Rostros que se obtuvieron del portal del FBI y se consideran como parte de la base de entrenamiento .....	171
34.	Rostros de delincuentes y terroristas utilizados como entrada al sistema (rostro a reconocer) .....	172
35.	Cruzamiento de cromosomas en un solo punto .....	188
36.	Óptimos locales y globales en un espacio multidimensional .....	200
37.	Clasificador Mahalanobis de distancia mínima .....	207
38.	Representación vectorial de los valores de gris correspondientes a la imagen de un rostro .....	209
39.	Señales correspondientes a 10 rostros de la base de entrenamiento AT&T. (a) – gradiente, (b) – frecuencia baja, (c) – máxima entropía .....	211

# Introducción

---

En la vida cotidiana rara vez nos maravillamos de nuestra capacidad para reconocer un rostro. Es sorprendente que podamos memorizar un número tan considerable de rostros en el transcurso de nuestra vida. En general lo hacemos de golpe, de manera holística, sin necesidad de una descripción verbal ni de un análisis consciente de las diversas características faciales. Podemos dejar de ver a una persona por algunos años y cuando la encontramos nuevamente, generalmente la reconocemos, a pesar de que haya cambiado de peinado o tenga barba. Sin embargo, el lograr que una computadora pueda reconocer un rostro, es un problema muy complejo.

El reconocimiento de rostros por medios computacionales, forma parte del área denominada **Biometría**. La palabra biometría, se deriva del griego (bio y metria): *βιο* (vida) y *μετρία* (medida)

El propósito de la Biometría, consiste en la elaboración de métodos automatizados para la identificación o verificación de personas mediante el uso de características físicas o de comportamiento. Esta tecnología se basa en la premisa de que cada persona es única y

posee rasgos distintivos que pueden ser utilizados para identificarla. Algunos ejemplos de características físicas son la cara, las huellas digitales, el iris de los ojos y en los últimos años se ha considerado el ADN. Por su parte, algunos ejemplos de características de comportamiento son la voz, la manera de firmar, la forma de caminar.

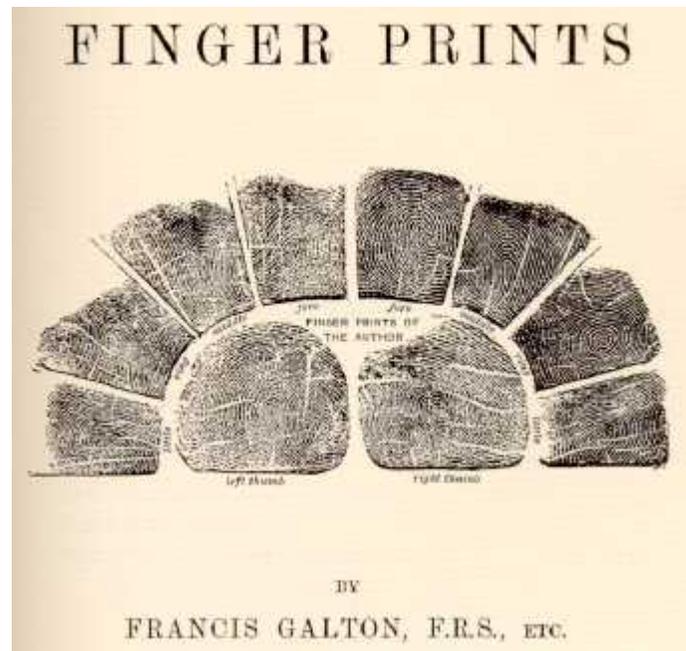
Los antecedentes del uso de medidas físicas del ser humano para propósitos de identificación, se remonta al siglo XIV, en donde los chinos, ya utilizaban ciertas impresiones de los dedos para identificación. En el año de 1877, Alphonse Bertillion (antropólogo y policía de París) está considerado como la primera persona que desarrolló y utilizó de manera sistemática, un método para clasificar e identificar criminales a partir de diversas medidas del cuerpo y la cabeza (figura 1).



**FIGURA 1. SISTEMA DE RECONOCIMIENTO ANTROPOMÉTRICO DE ALPHONSE BERTILLION PARA EL RECONOCIMIENTO DE CRIMINALES EN EUROPA, 1877-1884**

FUENTE: "A brief history of Biometrics", en:  
[http://www.galwayeducationcentre.ie/athenry/a\\_brief\\_history\\_of\\_biometrics.htm](http://www.galwayeducationcentre.ie/athenry/a_brief_history_of_biometrics.htm)

En 1884, con el sistema de Bertillion se tuvo la posibilidad de identificar a 241 criminales. Los cuerpos de policía, tanto de Inglaterra, Francia y los Estados Unidos, adoptaron el sistema. Sin embargo, años después se comprobó que presentaba fallas importantes en el proceso de identificación.



**FIGURA 2. PORTADA DEL LIBRO “FINGER PRINTS” PUBLICADO POR FRANCIS GALTON EN 1892, PRIMER TEXTO SOBRE IDENTIFICACIÓN DE HUELLAS DACTILARES.**  
FUENTE: “from Finger Prints”, en: <http://etext.lib.virginia.edu/railton/wilson/galtonfp.html>

Por su parte, Sir William Herschel realizaba desde principios de 1800, investigaciones sobre el uso de huellas dactilares para tratar de identificar personas. Logró obtener aproximadamente 8000 conjuntos de huellas. Sin embargo fue Sir Francis Galton, antropólogo británico y primo de Charles Darwin, quien logró obtener el primer sistema de clasificación de huellas dactilares completamente operacional. En el año 1892, publicó el primer libro sobre huellas dactilares con el nombre de *Finger Prints* (Galton, 1892). En la

figura 2, se puede apreciar la portada de dicha publicación. Utilizando la investigación de Galton, Edward Richard Henry desarrolló entre 1896 y 1897 el que está considerado como el primer sistema de la policía para identificación por medio de huellas dactilares. Dicho sistema, fue utilizado por Scotland Yard en 1901 y utilizado posteriormente, por diversos cuerpos de policía en el mundo.

Por lo que respecta al desarrollo de sistemas aplicados al reconocimiento de rostros, las primeras aplicaciones se remontan a la década de los años sesentas con una compañía denominada Panoramic Research, Inc. en Palo Alto, California y financiada por el Departamento de Defensa de los Estados Unidos y otras agencias de Inteligencia. Uno de los fundadores de dicha empresa, fue Woodrow Wilson Bledsoe (pionero en el campo del razonamiento automatizado). Bledsoe desarrolló un sistema “semiautomático” para extraer características de la fotografía de un rostro, por medio de señalar en una tableta digitalizadora elementos faciales del rostro (diversos puntos de los ojos, boca, cabeza, etc.). A partir de dichos puntos, se obtenían coordenadas y se comparaban contra una Base de Datos que contenía nombres de personas y un registro de las coordenadas de su correspondiente fotografía. La computadora calculaba las distancias entre los registros de la base de datos y los puntos de la fotografía a identificar. En 1966, el sistema de Bledsoe continuó en desarrollo en el Stanford Research Institute (SRI).

El siguiente trabajo pionero en el área, se debe a Sakai y Fujibayashi (1969), quienes presentan un programa computacional (que no requiere intervención de algún operador humano) para confirmar la existencia o ausencia de un rostro en una imagen.

Por su parte, Kelly (1970) elabora una disertación doctoral sobre reconocimiento de rostros en Stanford. Su técnica permite que una computadora extraiga de manera automática el contorno de la cabeza y el cuerpo de una determinada persona a través de una fotografía; y pueda entonces localizar los ojos, nariz y boca. Sin embargo, el procedimiento requiere de tres imágenes de cada individuo: una imagen del cuerpo, una imagen del “fondo” de la fotografía (sin el cuerpo) y un acercamiento de la cabeza.

El siguiente desarrollo, se realiza en la disertación doctoral de Takeo Kanade (1973) en la Universidad de Kyoto, Japón. Quien reporta los mismos resultados que obtuvo Kelly, pero con la ventaja de utilizar únicamente una sola fotografía del rostro. Asimismo, propone un nuevo esquema más flexible para el análisis de la imagen, logrando reconocer adecuadamente, quince de veinte personas.

Los anteriores, son los trabajos pioneros en el área de reconocimiento de rostros, como se puede apreciar, el tiempo transcurrido entre dichos desarrollos (años 60 y 70) y la actualidad, es de apenas 45 años como máximo. Si lo comparamos contra el trabajo de huellas dactilares en 1892 (a partir del sistema de Galton), a la fecha, se tienen 113 años de desarrollo, por lo cual, los sistemas computacionales aplicados a dicha área presentan éxitos notables y sumamente confiables en la época actual.

De manera general, la mayoría de los sistemas biométricos funcionan de forma muy similar y se puede resumir en dos pasos. El primero, consiste en que la persona debe **registrarse** en el sistema. Durante el proceso de registro, el sistema captura el rasgo característico de la persona, como por ejemplo la huella digital, y lo procesa para crear una representación electrónica denominada “modelo de referencia” o “modelo de entrenamiento”. El modelo de entrenamiento debe ser guardado en una base de datos, una tarjeta inteligente, o en algún

otro lugar del cual será extraído en cualquier ocasión futura para dar origen al segundo paso.

El segundo paso depende si la función del sistema biométrico consiste en *verificar* la identidad de la persona o *reconocer* a la persona. En el caso de verificación, la persona le informa al sistema cual es su identidad ya sea presentando una tarjeta de identificación, dando una clave o proporcionando su huella digital o imagen del rostro. El sistema captura el rasgo característico de la persona y lo procesa para crear una representación también electrónica llamada “modelo en vivo”. Por último, el sistema compara el modelo en vivo con el modelo de referencia de la persona. Si ambos modelos son idénticos o la diferencia es menor que un determinado umbral, se considera una verificación adecuada. A este proceso se le conoce también como comparación uno-a-uno (one-to-one).

Cuando la función del sistema es de reconocimiento, la persona no le informa al sistema cual es su identidad. El sistema tan solo captura el rasgo característico de la persona y lo procesa para crear el modelo en vivo. Luego el sistema procede a comparar los modelos respectivos para determinar la identidad de la persona en la base de datos o en diversas bases de datos. Es un proceso denominado uno-a-muchos (one-to-many).

Los sistemas biométricos se han considerado elementos clave para cuestiones de seguridad, desde hace varias décadas. En enero del año 2000, la revista Technology Review publicada por el MIT, propuso al área de la biometría como una de las “diez tecnologías emergentes que cambiarán al mundo”. Y de acuerdo a la empresa Internacional Biometric Group (una de las empresas líderes en el desarrollo de sistemas biométricos), se consideró que las

ventas de sistemas biométricos crecerían de \$399 millones de dólares en el año 2000, a \$1.9 billones de dólares para el año 2005.

A partir de los ataques terroristas a las dos torres del World Trade Center en New York, el 11 de septiembre de 2001, los sistemas biométricos cobraron un gran auge. Sobre todo, los sistemas de reconocimiento de rostros. La propuesta de las instituciones de seguridad de los Estados Unidos, consistía en proporcionar a los aeropuertos y lugares de concentración masiva (estadios, auditorios, centros comerciales, etc.), sistemas de reconocimiento de rostros. Dichos sistemas tendrían el objetivo de detectar posibles sospechosos (delincuentes y sobre todo terroristas), a partir de diversas bases de datos con fotografías de personas consideradas terroristas o ligados de alguna forma a ellos.

Existen diversos procedimientos para elaborar sistemas computacionales de reconocimiento de rostros, los cuales se detallan en el siguiente capítulo. Sin embargo, los dos algoritmos que más eficiencia has mostrado y con los cuales se desarrollaron los dos sistemas comerciales aplicados al reconocimiento de rostros y que se encuentran en uso por diversas agencias de seguridad, tanto en Estados Unidos como en otros países, son:

- **Algoritmo Eigenfaces**, utilizado en el sistema **FaceNet/FaceFinder**<sup>1</sup> de la compañía Viisaje.
- **Algoritmo de Análisis de Características Locales (LFA)** utilizado en el sistema **FaceIt**<sup>2</sup> de la compañía Visionics.

---

<sup>1</sup> FaceNet/FaceFinder son marcas registradas de Viisaje. <http://www.viisage.com>

<sup>2</sup> FaceIt es marca registrada de Visionics Corp/Identix. <http://www.identix.com>

Ambos algoritmos se basan en la técnica denominada **Componentes Principales**, desarrollada hace décadas y utilizada originalmente en las áreas de Psicología, Ciencias Sociales y Educación.

El primer sistema comercial exitoso, fue **FaceNet**, el cual se desarrolló a partir de la técnica de análisis de componentes principales. Originalmente fueron Sirovich y Kirby (1987), quienes plantearon la posibilidad de caracterizar un rostro humano por medio de componentes principales y denominaron al modelo con el nombre de “eigenpictures”. Posteriormente, Turk y Pentland (1991) investigadores del MIT, modificaron y hicieron operativa la propuesta de los eigenpictures para desarrollar el algoritmo denominado *Eigenfaces*. Dicho algoritmo fue patentado por el MIT y posteriormente lo adquirió la empresa Viisaje para desarrollar comercialmente el sistema FaceNet/FaceFinder.

Por lo que se refiere al algoritmo LFA, fue desarrollado por Joseph Atick, Paul Griffin y Norman Redlich (1996), basándose también en la técnica de componentes principales. El algoritmo LFA, intenta reducir la información de la imagen digital de un rostro, y obtener únicamente de 12 a 40 elementos que caracterizan al rostro. Posteriormente, el Dr. Atick fundó la empresa Visionics y comercializó su sistema de reconocimiento de rostros con el nombre de **FaceIt**.

Ambos sistemas presentan cierto éxito, sus tasas de eficiencia reportadas por las compañías respectivas, indican un porcentaje de reconocimiento entre el 90% y el 95%. Sin embargo, en un estudio específico para evaluar al sistema FaceIt (Heo, Abidi, Paik y Abidi, 2003), reportan porcentajes de reconocimiento eficiente de rostros únicamente, entre el 79.8% y un

máximo de 95.2%, dependiendo de las características de las fotografías de los rostros: expresión, edad, iluminación, pose, etc. Incluso, en algunos experimentos con fotografías del sujeto tomadas después de varios meses y con variaciones en la resolución y detalles de las fotografías, la eficiencia únicamente fue del 32%.

Se considera que las deficiencias que presentan estos sistemas, entre otras cosas, se deben a la utilización de algoritmos que se basan en la técnica de componentes principales. Esta técnica presupone el cumplimiento de ciertas características estadísticas por parte de los datos que utiliza. Específicamente, uno de los principales supuestos que deben cumplir los datos, es el de ajustarse a una distribución de probabilidad normal multivariada. Asimismo, las técnicas de análisis de componentes principales, LFA y otras técnicas similares que utilizan la gran mayoría de las aplicaciones de reconocimiento de rostros, se manejan en un espacio Euclideo ( $L_2$ ).

Dichos métodos son considerados robustos y permiten obtener resultados con cierta confiabilidad, a pesar de que los datos no cumplan estrictamente con los supuestos estadísticos requeridos. Sin embargo, es indudable, que entre más se alejen los datos manipulados de los supuestos estadísticos que se requieren para utilizar análisis de componentes principales y otras técnicas similares, los resultados cada vez serán menos confiables. Por otra parte, las técnicas mencionadas se basan también, en la utilización de métricas en espacios lineales normados o Euclidianos, lo cual está directamente relacionado con el requerimiento de los supuestos estadísticos y el manejo de métricas en espacios  $L_1$ ,  $L_2$  y  $L_p$ .

Los métodos tradicionales generalmente estadísticos (regresión, análisis factorial, análisis de componentes principales, etc.), utilizados en la mayoría de los sistemas actuales de reconocimiento de rostros presentan ciertas limitaciones al tener que trabajar bajo la norma  $L_2$  (Dean, 1988; Jonson, 1991) o bajo algún subconjunto de ella. Entre las limitaciones, se requiere que las distribuciones de los datos deben satisfacer las características de normalidad, homoscedasticidad e independencia.

El autor del presente trabajo ha postulado un potencial método aplicable al reconocimiento de rostros, que se basa en la utilización de una **familia de funciones polinomiales multivariadas de aproximación manejadas en un Espacio  $L_\infty$**  (Villegas, 1999). La propuesta anterior, se basa en la idea intuitiva, acerca de que los datos digitales que configuran la fotografía de un rostro a nivel de píxel (elemento básico para cualquier sistema de reconocimiento de rostros), no presentan una distribución normal multivariada. Por tanto, como se ha mencionado, la utilización de técnicas basadas en análisis de componentes principales y en general cualquier técnica tradicional de análisis multivariado, tiene grandes posibilidades de encontrar resultados erróneos y/o presentar deficiencias en el análisis, lo cual provocará disminuciones en la eficiencia de los algoritmos como sucede con las aplicaciones comerciales actuales.

El método propuesto considera, que al trabajar el problema de aproximación para caracterizar a los rostros como un sistema que encuentre el mejor aproximante en un espacio  $n$ -dimensional bajo la norma  $L_\infty$  (norma mínimax o Tchebyshev), se eliminarán (o

por lo menos se reducirán) las limitaciones que presentan los sistemas actuales al utilizar técnicas de componentes principales o similares y utilizar datos que no satisfacen presumiblemente, el supuesto de normalidad.

Por tanto, un elemento indispensable para conocer si tiene caso utilizar técnicas en un espacio  $L_\infty$ , es verificar si efectivamente los datos que representan a un determinado rostro y a un conjunto de rostros (base de datos), presentan respectivamente una distribución normal univariada y multivariada. Asimismo, se desearía conocer el comportamiento de un sistema de reconocimiento de rostros basado en análisis de componentes principales, para experimentar diversas alternativas con respecto a las características de las fotografías, el número adecuado de fotografías por sujeto que se deben tener en la base de entrenamiento y algunos otros elementos que incidan en la eficiencia del sistema.

Con el propósito de verificar lo anterior, se planteó el presente trabajo de tesis, que pretende realizar un análisis sobre las limitaciones de la técnica denominada componentes principales en su aplicación al reconocimiento de rostros. Así como proponer un nuevo método que se considera, eliminará algunas de las limitaciones que se presentan en las técnicas basadas en componentes principales y en general, en los sistemas tradicionales que utilizan la métrica  $L_2$ .

Como se ha mencionado, la tecnología en reconocimiento de rostros, se espera que tenga grandes incrementos en ventas en las próximas décadas. De los 399 millones de dólares por ventas de sistemas biométricos en el año 2000 (como se mencionó en párrafos anteriores),

\$34.4 millones en ventas correspondió a sistemas de reconocimiento de rostros. Sin embargo, lo anterior se ha incrementado sustancialmente en los últimos años. Para el año 2007, se pronostican ventas por \$429.1 millones de dólares.

Por tanto, los sistemas de reconocimiento de rostros es un nicho de mercado que no se debe perder de vista y es importante continuar con investigaciones que contribuyan a su desarrollo. A continuación, se presenta una panorámica de las temáticas que se presentan en la tesis.

## **Descripción general de la tesis**

En el capítulo 1 se realizó el planteamiento general de la investigación, presentando las principales preguntas de investigación, la hipótesis, los objetivos de la tesis y los resultados que se pretenden obtener con el proyecto.

En el capítulo 2, se presenta el “Estado del Conocimiento” en el área de reconocimiento de rostros. Se investigaron las principales aplicaciones, técnicas y algoritmos que se han considerado relevantes desde el año 1968 al año 2004.

Como se ha mencionado, dos de los sistemas comerciales que se consideran más eficientes en la actualidad para reconocimiento de rostros, basan sus algoritmos en el uso de la técnica de componentes principales o alguna derivación de ella. En el capítulo 3, se mencionan las características básicas de los elementos matemáticos que subyacen a la técnica de

componentes principales. Asimismo, se presenta en este capítulo el algoritmo basado en componentes principales para realizar reconocimiento de rostros. A partir del cual, se desarrollará como parte del trabajo de tesis, un programa computacional para reconocimiento de rostros basado en dicha técnica y que se reportará en un capítulo posterior.

Uno de los objetivos principales de la tesis, consiste en realizar un análisis estadístico de las características que presentan los valores numéricos de los píxeles que configuran cualquier fotografía digital de un rostro. Lo anterior se presenta en el capítulo 4, en el cual se utilizó el paquete SPSS<sup>3</sup> para llevar a cabo el análisis respectivo.

Otro elemento esencial del proyecto, consistió en verificar si los píxeles que configuran los datos de un rostro, se ajustan a una distribución normal. Lo anterior se realizó en el capítulo 5, utilizando los paquetes computacionales SPSS y LISREL/PRELIS<sup>4</sup>.

La verificación del supuesto de normalidad multivariada, es un elemento indispensable, los métodos estadísticos tradicionales basan sus algoritmos en el supuesto de normalidad de los datos de entrada, entre otras cosas. El autor del presente trabajo, ha postulado *a priori*, que las fotografías digitales de rostros humanos no se ajustan a una distribución normal multivariada. Si efectivamente no cumplen con una distribución normal, puede ser una de las causas de las fallas que presentan los sistemas comerciales actuales en el eficiente reconocimiento de rostros.

---

<sup>3</sup> SPSS es una marca registrada de SPSS, Inc. (<http://www.spss.com>)

<sup>4</sup> LISREL y PRELIS son marcas registradas de SSI Scientific Software Internacional. (<http://www.ssicentral.com>)

En el capítulo 6, se mencionan las características principales del sistema de reconocimiento de rostros basado en el análisis de componentes principales que se desarrolló utilizando MATLAB<sup>5</sup> y las rutinas proporcionadas por Image Processing Toolbox<sup>6</sup>.

La mayor parte del capítulo, se dedica a los diversos experimentos de reconocimiento de rostros que se llevaron a cabo, utilizando principalmente las fotografías de la base AT&T<sup>7</sup>. Se efectuaron diversos experimentos utilizando 40 sujetos en la base de entrenamiento, en donde cada sujeto puede tener 9, 8 y 5 fotografías por cada persona en la base de datos. Y se utilizaron 52 fotografías, como los posibles sujetos a reconocer. Asimismo, se contemplaron algunas fotografías de objetos, para probar el adecuado reconocimiento del sistema.

Otro conjunto de experimentos se llevó a cabo con 33 sujetos en la base de entrenamiento, utilizando 9, 7, 5, 3, 2 y 1 fotografía por cada persona de la base de entrenamiento. Al igual que en el grupo de experimentos anteriores, se utilizaron 52 fotografías para reconocimiento.

El último conjunto de experimentos, se realizó con una base de entrenamiento que combina fotografías controladas de la base AT&T (como en los casos anteriores) e imágenes de fugitivos y terroristas que se obtuvieron del portal internet del FBI (dichas fotografías no se encuentran controladas en cuanto a luz, fondo de imagen, resolución, etc.). El propósito consiste en verificar la eficiencia del sistema de reconocimiento cuando se trabaja con imágenes que presentan diversas características. Al igual que en los casos anteriores, se

---

<sup>5</sup> MATLAB es marca registrada de MathWorks, Inc. (<http://www.mathworks.com>)

<sup>6</sup> Image processing Toolbox es marca registrada de MathWorks, Inc.

<sup>7</sup> La base de fotografías AT&T fue desarrollada por los Laboratorios de Investigación AT&T – Cambridge junto con Computer Laboratories of Cambridge University, UK. Se ha utilizado en diversas investigaciones de reconocimiento de rostros a nivel internacional.

utilizan 52 fotografías para reconocer. A partir de los experimentos anteriores, se realizan diversas comparaciones y se obtienen algunas limitaciones que presenta el sistema.

Por último, en el capítulo 7 se presenta de manera general, la propuesta de un método alternativo para el reconocimiento de rostros. Dicho método, se basa esencialmente en utilizar un espacio de trabajo en  $L_\infty$  en el cual, a partir de diversas características que presentan los píxeles de la imagen digital de un rostro, se obtendrá un aproximante mediante una familia de polinomios de grado “ $n$ ”.

Dichos polinomios caracterizarán al rostro respectivo. La obtención del aproximante en un espacio  $L_\infty$  se conceptualizará como un problema de optimización combinatoria, el cual se obtendrá utilizando Algoritmos Genéticos y la teoría de aproximación multivariada. Al utilizar un paradigma de cálculo bajo la norma  $L_\infty$ , no se requiere que los datos cumplan los presupuestos estadísticos de normalidad, homoscedasticidad y linealidad, entre otros, que se requieren en el uso de técnicas estadísticas tradicionales. Se considera que la elaboración de un sistema de reconocimiento de rostros utilizando el método propuesto, ayudará a la eficacia y eficiencia de dichos sistemas. La elaboración y programación de tal sistema, se empezará en el corto plazo y es motivo de otra investigación.