

**DISEÑO E IMPLEMENTACIÓN DE UN CLASIFICADOR DE
IMÁGENES FACIALES PARA SER UTILIZADO EN UN SISTEMA DE
RECONOCIMIENTO DE ROSTROS EN TIEMPO REAL**

JAIRO ANDRÉS DÍAZ RODRÍGUEZ
EDWARD ELIAS PATIÑO HERRERA

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FÍSICO-MECÁNICAS
ESCUELA DE INGENIERÍA ELÉCTRICA, ELECTRÓNICA Y
TELECOMUNICACIONES
BUCARAMANGA
2010**

**DISEÑO E IMPLEMENTACIÓN DE UN CLASIFICADOR DE
IMÁGENES FACIALES PARA SER UTILIZADO EN UN SISTEMA DE
RECONOCIMIENTO DE ROSTROS EN TIEMPO REAL**

**JAIRO ANDRÉS DÍAZ RODRÍGUEZ
EDWARD ELIAS PATIÑO HERRERA**

**Trabajo de grado presentado como requisito para optar al título
de Ingenieros Electrónicos**

Director

MPE. HENRY ARGUELLO FUENTES

Codirector

Ing. EDWIN OMAR ORTIZ

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FÍSICO-MECÁNICAS
ESCUELA DE INGENIERÍA ELÉCTRICA, ELECTRÓNICA Y DE
TELECOMUNICACIONES
BUCARAMANGA
2010**

A Dios, mi amor eterno y mi sustento. Todo lo que soy se lo debo a Él, y toda mi vida deseo que sea una constante dedicación para corresponder el infinito amor que Él me ha dado.

A mi hermosa esposa, Cindy. Ella, que es la dueña de mi corazón y la razón por la cual he luchado por este logro para poder formar una familia a su lado y que todos nuestros sueños se conviertan en realidad. Te adoro con todo mi corazón.

A mis padres, mis patrocinadores oficiales. Nunca me exigieron llevar una camiseta con sus nombres y aún así fueron los financiadores de este ingeniero. Me faltarán los días para pagarles tanta dedicación. Los amo.

A mi hermanito, que quiero mucho y que sé que llegará a lugares más altos de los que él se imagina.

A mis abuelos, y a toda mi familia.

El fin de todo el discurso es éste: Honra a Dios y cumple sus mandamientos, porque eso es el todo del hombre. Ecc 12:13

Jairo Díaz

A Dios por ser mi creador, mi proveedor, el motor de mi vida y por darme la sabiduría y la fortaleza para que fuera posible alcanzar esta meta.

A mi madre, Yolanda Herrera Anaya por todo el amor que me ha dado y por que con su sacrificio y empeño he podido llegar a donde me encuentro y ser el hombre que hoy soy

A mi hermano, Oscar Javier Patiño Herrera por su compañía en momentos difíciles, su amor fraternal y su apoyo en todos los proyectos que he emprendido

A mis abuelas, por su cariño tan especial, su confianza en mí y toda la ayuda que me han brindado para salir adelante

A mi familia. Ya que estuvieron apoyándome a lo largo de mi carrera y dándome fuerzas para seguir adelante.

Edward Patiño

AGRADECIMIENTOS

A Henry Arguello y Edwin Ortiz por proponernos este reto que hoy damos por feliz término.

A Darío Delgado, por su dirección y apoyo cognoscitivo, por compartirnos su conocimiento dándonos ideas claras, despejando todo interrogante que se nos presentó como obstáculo

Al Grupo de Investigación en Ingeniería Biomédica - GIIB por su apoyo intelectual y técnico.

A la Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones y a la Universidad Industrial de Santander

TABLA DE CONTENIDO

INTRODUCCIÓN	17
1. MARCO TEÓRICO	19
1.1 Reconocimiento de Patrones.....	19
1.2 Clasificadores	20
1.2.1 Aprendizaje Automático	20
1.2.2 Máquinas de soporte vectorial (<i>Support vector machines</i>).....	21
1.3 Clasificador de Rostros.....	23
1.3.1 Procesamiento digital de imágenes: Segmentación.....	23
1.3.2 Procesamiento digital de imágenes: Representación y Descripción ..	23
1.3.3 <i>Active Shape Models (ASM)</i>	23
2. SELECCIÓN DE LAS CARACTERÍSTICAS FACIALES.....	25
2.1 Generalidades.....	25
2.2 Software Interactivo	25
2.3 Resultados.....	26
2.4 Encuesta.....	27
2.5 Discusión	28
3. ALGORITMO DEL CLASIFICADOR	31
3.1 Generalidades.....	31
3.2 Detección del Rostro y ASM	32
3.2.1 STASM.....	32
3.2.2 Algoritmo ASM	33

3.3. Extracción de Características	36
3.3.1. Color de piel	37
3.3.2 Geometría	41
3.3.3 Género	42
3.4 Toma de decisión.....	43
3.4.1 Máquinas de Soporte Vectorial	46
3.4.2 LIB-SVM.....	48
4. RESULTADOS	50
4.1 Base de Datos	50
4.2. Selección SVM.....	51
4.3. Modelos Obtenidos	54
4.4 Comparación con trabajos previos.....	54
5. CONCLUSIONES	56
BIBLIOGRAFÍA.....	57

LISTA DE FIGURAS

Figura 1. Diagrama de Bloques del Proceso de Reconocimiento de Patrones.....	19
Figura 2. Descripción grafica del hiperplano construido por SVM.....	22
Figura 3. ScreenShot del Software Interactivo elaborado para la Encuesta	26
Figura 4. Imagen completa e imagen recortada.....	26
Figura 5. Informe general de resultados	27
Figura 6. Gráfico de barras de resultados de la Encuesta.	28
Figura 7. Ejemplo de Imagen de entrada al clasificador estilo pasaporte	30
Figura 8. Diagrama de bloques del algoritmo clasificador.....	31
Figura 9. Imagen original y resultado después del detector de rostros	33
Figura 10. Rostro con los landmarks utilizados el STASM.....	34
Figura 11. Muestreo alrededor de cada landmark para obtener el perfil.....	35
Figura 12. Niveles de la pirámide gaussiana	35
Figura 13. Entrada. Imagen después del detector. Salida del ASM.....	36
Figura 14. Entrada. LandMarks. ConvexImage. BoundingBox	37
Figura 15. Imagen con sólo el rostro de la persona	37
Figura 16. Redimensionamientos para obtener información de color.	39
Figura 17. Imagen con la información del color.	40
Figura 18. Imágenes de color, normalizada, con umbral y filtrada.....	40
Figura 19. Histograma de la imagen.	41
Figura 20. Rostros forma cuadrada y un cuadrada.....	42
Figura 21. Modelo de puntos (landmarks) asociados a la imagen.	43
Figura 22. Hiperplanos de separación en SVM.....	46
Figura 23. Imágenes de la base de datos.....	50

Figura 24. Barrido de los parámetros C y γ para la clasificación por geometría. ...52
Figura 25. Barrido de los parámetros C y γ para la clasificación por color de piel. 53
Figura 26. Barrido de los parámetros C y γ para la clasificación por género53

LISTA DE TABLAS

Tabla 1. Cantidad de fallos y aciertos	27
Tabla 2. Porcentaje de fallos y aciertos	27
Tabla 3. Resumen de las características extraídas.	44
Tabla 4. Clasificaciones con sus respectivas clases y ejemplos.....	45
Tabla 5. Distribución de imágenes de acuerdo al tipo de clasificación.	51
Tabla 6. Modelos finales obtenidos para el clasificador.....	54
Tabla 7. Comparación con trabajos previos.....	55

RESUMEN

TÍTULO: DISEÑO E IMPLEMENTACIÓN DE UN CLASIFICADOR DE IMÁGENES FACIALES PARA SER UTILIZADO EN UN SISTEMA DE RECONOCIMIENTO DE ROSTROS EN TIEMPO REAL*

AUTORES: JAIRO ANDRÉS DÍAZ RODRÍGUEZ**
EDWARD ELIAS PATIÑO HERRERA**

PALABRAS CLAVES: ASM, Clasificación, Color, Género, Geometría, Rostros, SVM

DESCRIPCIÓN:

En este trabajo se presenta una descripción del diseño de un algoritmo de clasificación de rostros basado en color de piel, género y geometría del rostro. Se hizo un diseño basado inicialmente en extraer 68 *landmarks* (puntos característicos de una imagen) en rostros humanos con la librería STASM 2.4 que utiliza el algoritmo de detección de rostros Viola-Jones y el algoritmo ASM (*Active Shape Models*), para luego, mediante técnicas de procesamiento digital de imágenes, extraer 2 características para la clasificación por color de piel, mediante el método de filtrado e iluminación y el método del histograma; 5 características para la clasificación por género basándose en algunas distancias específicas del rostro; y 2 características para la clasificación de geometría teniendo en cuenta la forma las áreas y las dimensiones de la figura generada por los *landmarks*. Mediante un entrenamiento basado en Máquinas de Soporte Vectorial SVM se determinaron los diferentes modelos para cada clasificación mediante la librería LIB-SVM.

La validación de estos algoritmos se realizó mediante una serie de pruebas usando imágenes de rostros humanos de la base de datos experimentales FERET, obteniendo resultados muy buenos para la clasificación de color con una eficiencia corroborada mediante la validación cruzada de los modelos encontrados en las máquinas de soporte vectorial y resultados con la clasificación de color de piel del 90%, 80% para la clasificación por género y marcando un precedente con una eficiencia del 66% respecto a la clasificación por geometría del rostro.

*Tesis de Grado

**Facultad de Ingenierías Físico-Mecánicas. Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones. Director: ARGUELLO Henry

ABSTRACT

TITLE: DESIGN AND IMPLEMENTATION OF A FACE IMAGES CLASSIFICATOR, TO BE USED IN AN REAL TIME FACE RECOGNITION SYSTEM*.

AUTHORS: JAIRO ANDRES DIAZ RODRIGUEZ**
EDWARD ELIAS PATIÑO HERRERA**

KEY WORDS: ASM, Classification, Color, Faces, Gender, Geometry, SVM

DESCRIPTION:

This work shows the description of a Human face classification algorithm design, based on skin color, gender, and geometry. Initially, 68 human face landmarks (Characteristic image points) were extracted with STASM 2.4 library that uses the Viola-Jones face detection algorithm and ASM (Active Shape Models) algorithm to extract, , by means of digital image processing techniques, 2 characteristics for skin-color classification using Filtering and luminance, and histogram methods; 5 characteristics for gender classification taking specific face distances; and 2 characteristics for geometry classification evaluating areas and dimensions of the landmark generated image. With an SVM based training, different models where determined for each classification, using LIB-SVM library.

The validation of these algorithms was done using series of experiments with human face FERET data base, obtaining very good results with a corroborated efficiency by means of the SVM found models cross-validation of 90% for skin color, an 80% acceptable efficiency for gender classification, and giving a start point with 66% efficiency for face geometry classification

*Degree Project

**Physical-Mechanics Engineering Faculty. Electrical, Electronic and Telecommunications engineering School. Director: ARGUELLO Henry

INTRODUCCIÓN

Recientemente la clasificación de imágenes faciales ha despertado la atención tanto de la comunidad científica, como de la industria, debido a su importancia en temas como seguridad y comunicación.

Con la evolución de la interacción “hombre-máquina”, hacerle frente a las crecientes exigencias en cuanto a seguridad, confiabilidad, servicios individualizados y acercamiento a la visión por computadora es todo un reto, debido a que la clasificación de rostros es un trabajo más fácil para los seres humanos que para las máquinas.

Es así que el sistema de visión humano es el mejor ejemplo de sistema que puede realizar dichas tareas de identificación en un alto nivel de precisión, útil para múltiples aplicaciones. Sin embargo, es claro que la capacidad de aprendizaje y memoria humana es limitada, por lo tanto parece natural tratar de entender este sistema para poder mejorarlo con algoritmos computacionales que logren semejarse a la visión humana en el proceso de identificación, pero consiguiendo una capacidad mayor al depender de la propia memoria del sistema clasificador. Esto conlleva a que sea necesario establecer algunas características que normalmente el ser humano identifica rápidamente y de esa manera disminuir el espacio de búsqueda adecuadamente, para mejorar el sistema completo de reconocimiento.

En el presente trabajo se utilizan técnicas de Inteligencia Artificial, así como de Procesamiento Digital de Imágenes para obtener como resultado el diseño e implementación de un sistema clasificador de rostros de acuerdo a características faciales. Dicho trabajo nace como una necesidad del Grupo de Investigación en

Ingeniería Biomédica GIB, para formar parte de un sistema completo de Reconocimiento de rostros, proyecto financiado por la Vicerrectoría de Investigación y Extensión de la Universidad Industrial de Santander UIS.

En el primer capítulo se plantean definiciones y explicaciones relativas a los sistemas de reconocimiento de patrones, algoritmos de clasificación y técnicas de Procesamiento Digital de Imágenes.

El segundo capítulo describe los procesos que se llevaron a cabo para seleccionar las características faciales para la clasificación. Un software desarrollado con este propósito y la descripción de los datos obtenidos dieron como resultado la decisión de realizar el clasificador de Rostros de acuerdo a Color de piel, Género y Geometría del rostro.

El tercer capítulo muestra el diseño e implementación del algoritmo clasificador, formado por 4 etapas básicas: Detección del rostro, Algoritmo “Active Shape Model” (ASM), Extracción de características y Toma de decisión.

El cuarto capítulo explica los resultados obtenidos con imágenes seleccionadas de la base de datos internacional de rostros humanos FERET. Esto con el propósito de establecer un porcentaje de error y verificar el correcto funcionamiento del algoritmo.

Por último se dan algunas conclusiones y recomendaciones para futuros trabajos.

1. MARCO TEÓRICO

1.1 Reconocimiento de Patrones

El objetivo del procesamiento e interpretación de datos sensoriales es lograr una descripción concisa y representativa del universo observado. La información de interés incluye nombres, características detalladas, relacionamientos, modos de comportamiento, etc. que involucran a los elementos del universo (objetos, fenómenos, conceptos) [1].

Estos elementos se perciben como patrones y los procesos que llevan a su comprensión son llamados *procesos perceptuales*. El etiquetado (clasificación, asignación de nombres) de esos elementos es lo que se conoce como *reconocimiento de patrones*. Por lo tanto, el reconocimiento de patrones es una herramienta esencial para la interpretación automática de datos sensoriales

Para poder reconocer los patrones se siguen los siguientes procesos [2] (Fig. 1):

- Adquisición de datos
- Extracción de características
- Toma de decisiones

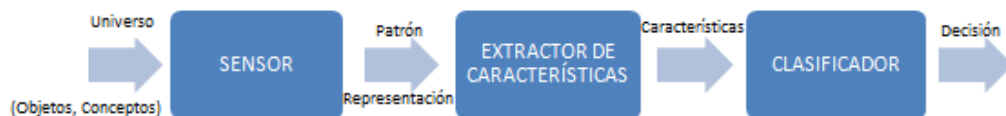


Figura 1. Diagrama de Bloques del proceso de reconocimiento de Patrones

Un sistema completo de reconocimiento de patrones incluye un sensor que recoja fielmente los elementos del universo a ser clasificado, un mecanismo de extracción de características cuyo propósito es extraer la información útil, eliminando la información redundante e irrelevante, y finalmente una etapa de toma de decisiones en la cual se asigna a la categoría apropiada los patrones de clase desconocida a priori.

1.2 Clasificadores

Un clasificador es una función que proporciona una clase etiquetada como salida a partir de un conjunto de características tomadas como entradas. Una manera de construir un clasificador es coger un conjunto de ejemplos etiquetados y tratar de definir una regla que pueda asignar una etiqueta a cualquier otro dato de entrada. Formalmente, un clasificador es una función $h: X \rightarrow Y$ producida por un conjunto de entrenamiento $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, y que mapea cualquier objeto $x \in X$ con su respectivo nivel de clasificación $y \in Y$ definido por un mapeo desconocido $g: X \rightarrow Y$ [3].

1.2.1 Aprendizaje Automático

El Aprendizaje automático es la parte básica que tienen en común los diferentes tipos de clasificadores que existen. La idea básica del aprendizaje consiste en utilizar las percepciones no sólo para actuar, sino también para mejorar la habilidad de un agente para actuar en el futuro.

Existen diferentes tipos de técnicas de aprendizaje, sin embargo, los utilizados en este trabajo son los siguientes [1]:

1.2.1.1 Aprendizaje supervisado

El aprendizaje supervisado consiste en aprender una función, a partir de ejemplos etiquetados anteriormente, que establezca una correspondencia entre las entradas y las salidas deseadas del sistema. No siempre es posible hacer este tipo de entrenamiento ya que se debe disponer de la salida esperada en la función de entrada. El sistema de aprendizaje trata de etiquetar (clasificar) una serie de vectores utilizando una entre varias categorías (clases).

1.2.1.2 Aprendizaje no supervisado

El aprendizaje no supervisado consiste en aprender a partir de patrones de entradas para los que no se especifiquen los valores de sus salidas. El principal problema de esta técnica es la toma de decisiones a la hora de escoger un patrón entre todos los proporcionados. El sistema trata los objetos de entrada como un conjunto de variables aleatorias, construyendo un modelo de densidad para el conjunto de datos [1,4].

1.2.1.3 Aprendizaje semi-supervisado

Actualmente existen técnicas que combinan las dos anteriores, ya que en algunos casos puede resultar muy costoso asignar etiquetas o clases a todos los datos. La finalidad es combinar datos etiquetados y no etiquetados para mejorar la construcción de modelos. Aunque no siempre es útil y existen varios métodos para llevarlo a cabo [1].

1.2.2 Máquinas de soporte vectorial (*Support vector machines*)

Las Máquinas de soporte vectorial (SVM) son técnicas sencillas y de gran aplicación cuando se pretende construir un clasificador utilizando ejemplos. A diferencia de las redes neuronales que intentan construir un modelo a posteriori, las SVM's tratan de obtener la frontera de decisión. Una de sus ventajas es su

facilidad de uso ya que únicamente hay que codificar la geometría de la frontera [5].

Las Máquinas de Soporte Vectorial (SVM por sus siglas en inglés) son un conjunto de métodos relacionados de aprendizaje supervisado usados para la clasificación. Si se miran los datos de entrada como dos conjuntos de vectores en un espacio n-dimensional, una SVM construirá un hiperplano separando aquel espacio, el cual maximiza el margen entre los dos conjuntos de datos. Para calcular el margen, dos hiperplanos paralelos se construyen, uno en cada lado del hiperplano separador, los cuales son “empujados” hacia los dos conjuntos de datos. Intuitivamente una buena separación es lograda por el hiperplano que tenga la mayor distancia entre ambas clases de datos vecinas, donde se tiene que entre mayor sea el margen, menor el error del clasificador (Fig. 2).

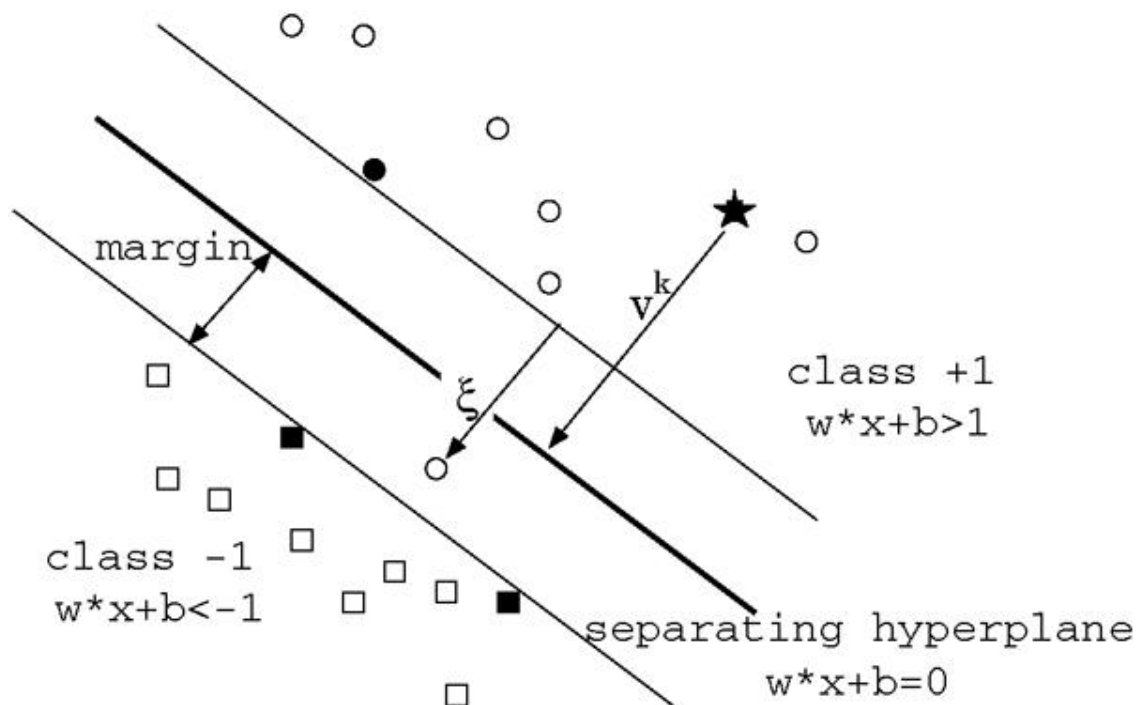


Figura 2. Descripción grafica del hiperplano construido por SVM
Fuente: <http://www.jvrb.org/archiv/760/svm1.jpg>

1.3 Clasificador de Rostros

Un clasificador de rostros toma un conjunto de rostros como entradas y produce como salida una clase etiquetada de la misma. Cada clase es una característica propia del rostro humano: género, color de piel, geometría.

Para lograr un mejor desempeño del clasificador, es necesario utilizar métodos de procesamiento digital de imágenes para extraer las características necesarias que le permitirán determinar el grupo al cual pertenece dicha imagen. Los métodos utilizados en este trabajo son:

1.3.1 Procesamiento digital de imágenes: Segmentación

La segmentación de imágenes es uno de los temas fundamentales más estudiados y útiles en procesamiento y análisis de imágenes, sobre todo porque constituye el primer paso que se usa en la mayoría de tareas de análisis cuantitativo de imágenes. El objetivo de la segmentación es obtener una partición de la imagen en regiones coherentes como paso previo al análisis de su contenido. Por ejemplo, antes de analizar un tumor en una imagen de tomografía volumétrica, es necesario detectarlo y aislarlo del resto de la imagen. Antes de reconocer una cara en una imagen es necesario separarla del resto de la imagen [6,7].

1.3.2 Procesamiento digital de imágenes: Representación y Descripción

Después de la segmentación es necesario representar el conjunto de píxeles obtenidos de una forma adecuada para el procesamiento por computadora. La representación se puede realizar ya sea describiendo las características de contorno tales como longitud, curvatura, simetría y forma o características de región tales como textura, color, área y propiedades topológicas [7].

1.3.3 Active Shape Models (ASM)

Dentro del procesamiento de imágenes también existen los ASM, modelos estadísticos de la forma de los objetos, los cuales iterativamente se deforman hasta encajar en un ejemplo del objeto en una nueva imagen. Las formas son aproximadas por el modelo de distribución de puntos (PDM) para variar solo en las maneras que se ven en un conjunto de entrenamiento con ejemplos etiquetados. La forma de un objeto es representada por un conjunto de puntos (Controlados por el modelo de forma). El algoritmo ASM logra encajar el modelo dentro de una nueva imagen [8]. Este algoritmo se explicará más detalladamente en la sección 2 del capítulo 3.

1.3.3.1 LandMarks

En Morfometría, un *landmark* es un punto en un objeto de forma que las correspondencias entre y dentro de las características del objeto son preservadas [9]. En otras disciplinas, los *landmarks* pueden ser conocidos como vértices, puntos de anclaje, puntos de control, sitios, puntos de perfil, puntos de muestra, nodos, marcadores, etc. Los *landmarks* pueden ser definidos tanto manualmente por expertos como automáticamente por un programa de computador. El tipo de *landmark* que concierne a este trabajo es el *landmark* anatómico. Esta clase de *landmark* es un punto que tiene significado biológico en un organismo. Usualmente los expertos definen puntos anatómicos para asegurar sus correspondencias dentro de individuos de la misma clase. Ejemplos de *landmarks* anatómicos en una figura de un rostro son la esquina del ojo, la punta de la nariz, la mandíbula, etc. *Landmarks* anatómicos determinan partes homólogas de un organismo, las cuales comparten características comunes. Para este trabajo se hacen importantes los *landmarks* ubicados en el rostro humano.

2. SELECCIÓN DE LAS CARACTERÍSTICAS FACIALES

2.1 Generalidades

Con el propósito de diseñar e implementar un sistema clasificador de rostros, se hizo necesario determinar las características faciales que se iban a tener en cuenta para dicha clasificación ya que era vital realizar una elección adecuada que se adaptara a lo que normalmente un ser humano haría y que permitiera acoplarse con algoritmos computacionales.

Por ello, inicialmente se realizó una actividad interactiva a un grupo heterogéneo de personas cuyo objetivo fue tratar de encontrar las características más usadas por la mente humana para la clasificación, basándose en la experiencia en el reconocimiento de rostros. Posteriormente, y en conjunto con la revisión y documentación bibliográfica, se escogieron las características de clasificación más adecuadas y viables para el clasificador.

2.2 Software Interactivo

Para esta tarea se desarrolló un pequeño Software interactivo de reconocimiento de Rostros *Caras.exe* (Figura 3). El programa permite al usuario interactuar en forma de juego, de manera tal que dada una imagen escogida aleatoriamente de una base de datos, el usuario debe tratar de encontrar a la misma persona en otra base de datos en la que también se encuentra, antes de que pase cierta cantidad de tiempo. Al usuario se le permitió la opción de filtrar por género y color de piel para poder comprobar el efecto de dicha clasificación en la rapidez con que la persona encontraba la imagen correspondiente.

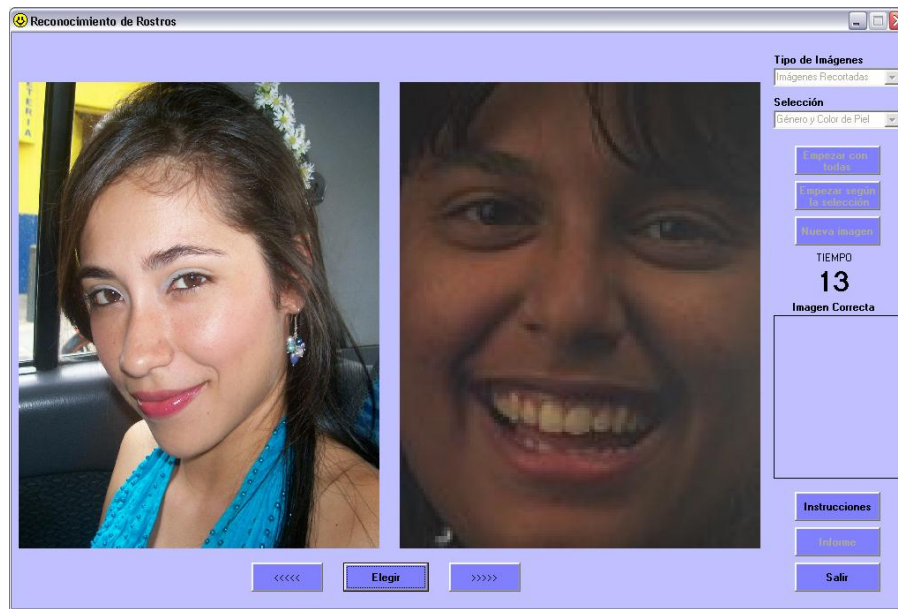


Figura 3. ScreenShot del Software Interactivo elaborado para la Encuesta

Se realizaron alrededor de 400 actividades en un grupo heterogéneo, bajo diferentes características seleccionadas del programa. Los resultados obtenidos se relacionan a continuación.

2.3 Resultados

De acuerdo a lo anterior, se hicieron pruebas con imágenes completas y con imágenes recortadas (Fig. 4). También se realizaron pruebas con imágenes seleccionadas, es decir, clasificándolas de acuerdo a color y género, y con todas las imágenes.



Figura 4. A la izquierda una imagen completa y a la derecha una imagen recortada
Fuente: IMM Face Database [10]

Los resultados se presentan en la figura 5 y las tablas 1 y 2.

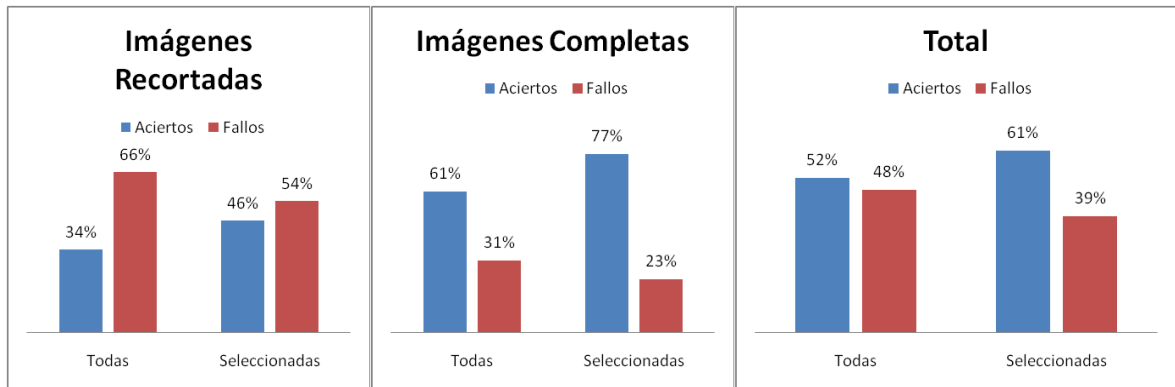


Figura 5. Informe general de resultados

	Recortadas		Completas		Total	
	Aciertos	Fallos	Aciertos	Fallos	Aciertos	Fallos
Todas	36	71	75	33	111	104
Seleccionadas	46	53	65	19	111	72
Total	82	124	140	52	222	176

Tabla 1. Cantidad de fallos y aciertos

	Recortadas		Completas		Total	
	Aciertos	Fallos	Aciertos	Fallos	Aciertos	Fallos
Todas	34%	66%	61%	31%	52%	48%
Seleccionadas	46%	54%	77%	23%	61%	39%
Total	40%	60%	73%	27%	56%	44%

Tabla 2. Porcentaje de fallos y aciertos

2.4 Encuesta

Posterior a la actividad con el software interactivo se realizó una encuesta donde se preguntaba por las tres características faciales que más ayudaron en la actividad para realizar el reconocimiento. Los resultados se observan en la figura 6.

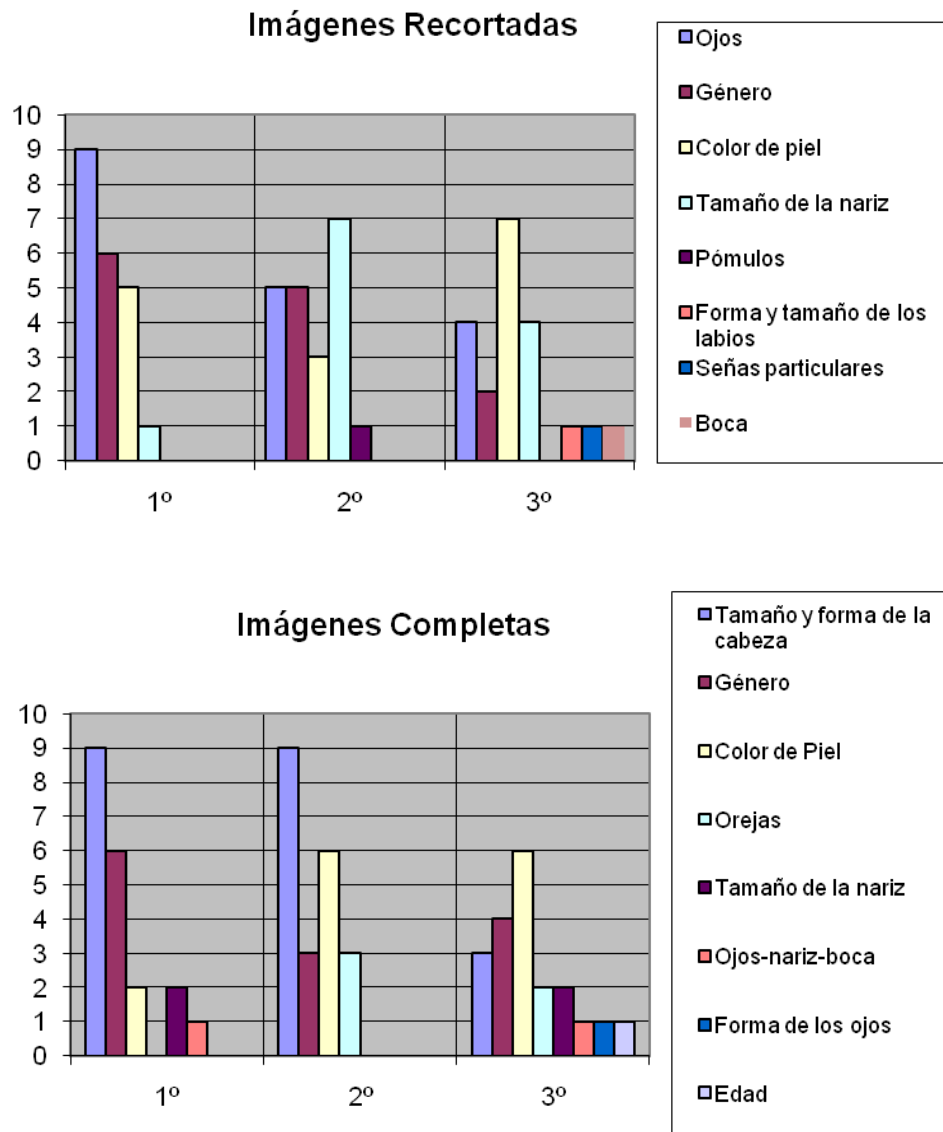


Figura 6. Gráfico de barras de resultados de la Encuesta. En el eje Y la cantidad de personas que escogieron dicha característica. En el eje X la prioridad de la característica en el momento de la escogencia. En las convenciones, las diferentes características escogidas por los encuestados

2.5 Discusión

Como se puede apreciar fácilmente, el porcentaje de aciertos fue mejor al realizar la actividad con imágenes completas debido a la predisposición del ser humano para determinar un reconocimiento basándose en los aspectos externos de la persona (Ropa, peinado, etc.) lo que claramente no está dentro de los alcances de

este trabajo dada la variedad de estilos personales que cualquier persona puede tener en diferentes etapas de su vida. Aunque la clasificación se hizo en general más sencilla con las imágenes recortadas (Muchas de estas tenían la misma ropa en ambas fotos), se vio una clara disminución de la efectividad cuando las fotos comparadas eran la misma persona pero con diferente atuendo o peinado. Estas observaciones llevaron a la conclusión de que era necesario realizar una detección total del rostro para poder llevar a cabo una clasificación adecuada que permitiera un reconocimiento imparcial.

Aún así, fue muy notorio que la efectividad en el reconocimiento mejoró en todos los tipos de prueba cuando se realizó la actividad con imágenes seleccionadas de acuerdo a una clasificación previa dada por el usuario con respecto a color y género. Por otro lado los resultados de la encuesta ayudaron a determinar que las características más observadas en el momento del reconocimiento fueron ojos, tamaño de la cabeza, género y color de piel.

Respecto a la característica de los ojos se tuvieron algunos inconvenientes a la hora de identificar personas con gafas, ya que dicho accesorio no permitía la correcta apreciación de los ojos, causando mayor dificultad en el reconocimiento en comparación con imágenes de personas sin gafas, por lo cual esta característica fue descartada. El tamaño de la cabeza es un aspecto variable si las imágenes tomadas difieren en cercanía y también depende mucho de la configuración del peinado.

De acuerdo a las consideraciones anteriores las características de **género** y **color de piel** fueron las seleccionadas como óptimas para implementar un clasificador. Por otro lado se encontró una amplia literatura sobre estas características, lo que permitió su elección definitiva.

Posteriormente gracias al ASM utilizado en ambas clasificaciones se dio la facilidad de escoger también la clasificación por la **geometría del rostro**.

Y por último, dado que el efecto de las imágenes recortadas o completas fue notorio, se decidió recibir como entrada imágenes en estilo de pasaporte (Fig. 7).



Figura 7. Ejemplo de Imagen de entrada al clasificador estilo pasaporte
Fuente: The color FERET database [11]

3. ALGORITMO DEL CLASIFICADOR

3.1 Generalidades

El Algoritmo utilizado en el clasificador consta de 4 etapas (Fig. 8). La primera etapa consiste en el procesamiento de la imagen mediante un algoritmo de detección de rostros para hallar exactamente el lugar de la imagen en donde se encuentra la cara y así omitir aspectos de fondo y facilitar la segunda etapa.

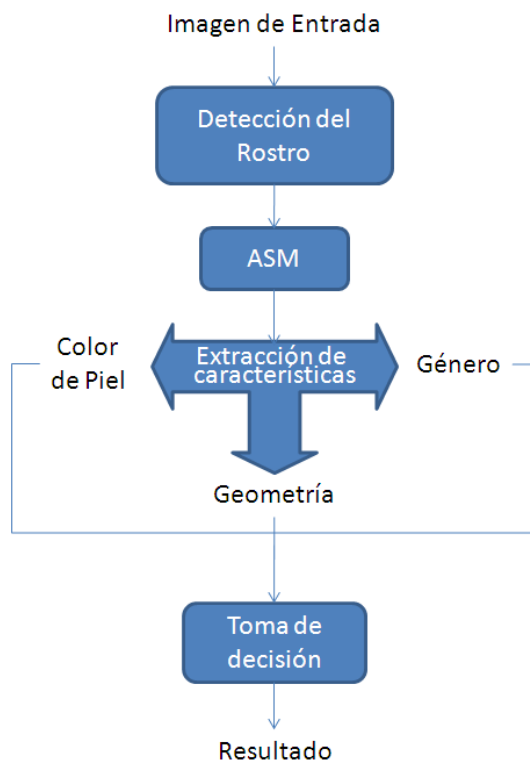


Figura 8. Diagrama de bloques del algoritmo clasificador

La segunda etapa consiste en un algoritmo ASM que permite extraer exclusivamente el rostro eliminando cabello, fondo, etc.; generando los *landmarks* necesarios para una clasificación.

Posteriormente se realiza la etapa de la extracción de las características de acuerdo a cada tipo de clasificador: color de piel, geometría o género. Estas características dependen del objetivo que se quiera lograr y pueden variar desde el rostro mismo, como las distancias entre los *landmarks*.

La última etapa del algoritmo consiste en el reconocimiento de las características y la decisión que debe tomar, esto depende de un entrenamiento previo que debe realizarse en el algoritmo de clasificación para tomar la decisión correcta.

3.2 Detección del Rostro y ASM

Con lo referente a esta etapa se busca determinar la posición exacta del rostro dentro de la foto para posteriormente mediante el algoritmo de *Active Shape Models* obtener el contorno y los principales puntos característicos del mismo.

Estas dos etapas del algoritmo son realizadas mediante el software STASM 2.4 [12] con código abierto bajo GPL2. El software se describe a continuación.

3.2.1 STASM

STASM es una librería en C++ para encontrar características en rostros. Dada una imagen de un rostro se da como resultado la posición de las características faciales (*landmarks*). STASM está diseñado para trabajar con imágenes faciales “estilo pasaporte” con expresión neutral. No funciona bien en rostros con ángulo o con boca abierta u otras expresiones faciales no neutrales. Como todas las técnicas automáticas, no es tan preciso como el cerebro humano y por lo tanto tendrá algunas veces algunas malas colocaciones.

3.2.1.1 Detector de Rostros Viola-Jones

Antes de que la búsqueda de características faciales comience, STASM utiliza el método detección de rostros Viola-Jones [13] contenido en la biblioteca libre de visión artificial OpenCV [14] originalmente desarrollada por Intel publicada bajo licencia BSD, que permite que sea usada libremente para propósitos comerciales y de investigación con las condiciones en ella expresadas. El efecto del detector de rostros puede observarse en la figura 9. Este detector tiene como propósito facilitar la convergencia del algoritmo de ASM

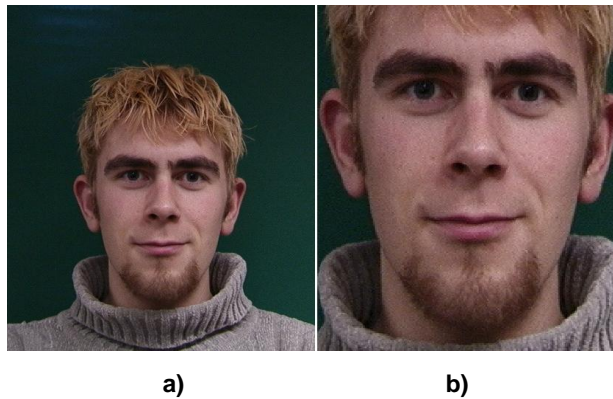


Figura 9. a) Imagen original. b) Resultado después del detector de rostros Viola-Jones.
Fuente: IMM Face Database [10]

3.2.1.2 Active Shape Model

Para la búsqueda de las características faciales STASM utiliza el algoritmo de Active Shape Models desarrollado por Tim Cootes [15].

3.2.2 Algoritmo ASM

3.2.2.1 Entrenamiento

El algoritmo para ASM necesita un entrenamiento previo en donde las imágenes de rostros deben ser manualmente marcadas con los *landmarks*. La ubicación de ellos son los puntos de interés del rostro que van desde el borde de la mandíbula entre lado y lado del rostro, pasando por las cejas los ojos, luego la nariz y por último la boca (Fig. 10). STASM fue entrenado con las bases de datos de rostros

BioID [16], AR [17], y XM2VTS [18]; y 68 *landmarks* [12]. Estas imágenes de entrenamiento se alinean para que todas queden bajo el mismo sistema de coordenadas y orientación.

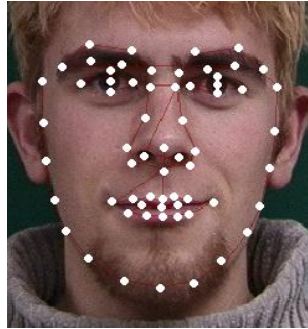


Figura 10. Rostro con los landmarks utilizados el STASM.
Fuente: IMM Face Database [10]

3.2.2.2 Construcción del modelo

Una vez tenida la ubicación de los 68 *landmarks* dentro de todas las imágenes, se elabora un modelo mediante *Principal Component Analysis* (PCA) sobre los datos de entrenamiento. Este modelo obtiene como resultado:

$$x = \bar{x} + Pb \quad (1)$$

En donde \bar{x} es la media de todas las imágenes, P los eigenvectores de la matriz de covarianza generada por PCA y b el set de parámetros que variaran el modelo de acuerdo a la necesidad [15].

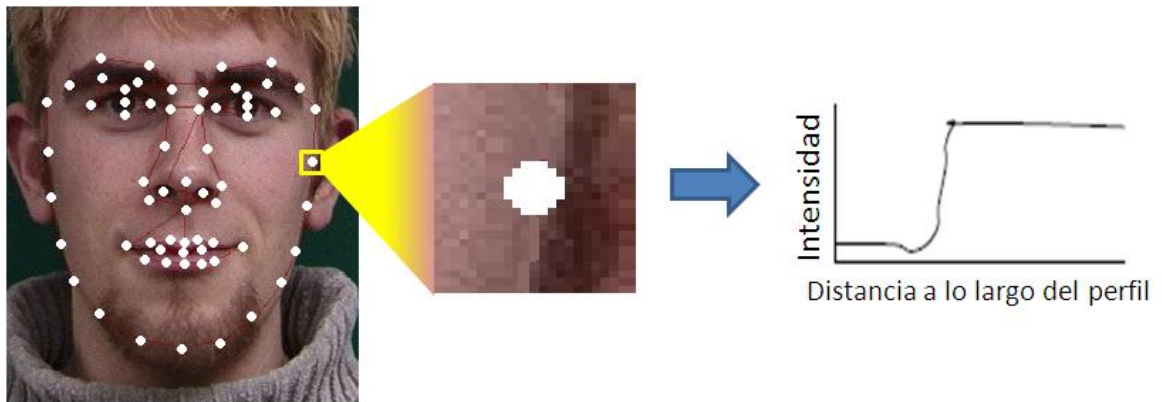


Figura 11. Muestreo alrededor de cada *landmark* para obtener el perfil.
Fuente: IMM Face Database [10] – Autores del Proyecto

Como el rostro humano no es una figura con los bordes bien definidos y algunos *landmarks* (como por ejemplo el que está en el medio de la nariz) no están ni siquiera en uno, se hace necesario también para construir el modelo, muestrear el perfil circundante a cada *landmark* (Fig. 11).

Para mejorar la eficiencia y la robustez del algoritmo, se implementa en estilo multiresolución, siguiendo el modelo de pirámides gaussianas. Esto consiste en hacer varios niveles de la imagen, mediante el método “smoothing and sub-sampling” que va suavizando la imagen dependiendo del nivel (Disminuyendo a la mitad la cantidad de pixeles en cada dimensión) y volviendo a muestrear (Fig. 12). En cada uno de esos niveles se vuelve a tomar el perfil circundante a cada punto característico.



Figura 12. Niveles de la pirámide gaussiana
Fuente: IMM Face Database [10] - Autores del Proyecto

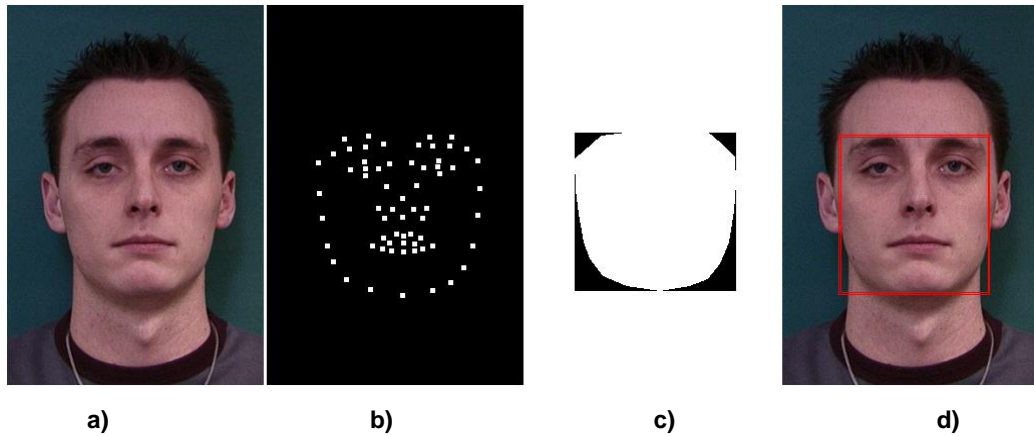


Figura 14. a) Imagen de entrada. b) Imagen de LandMarks. c) ConvexImage. d) BoundingBox
Fuente: IMM Face Database [10] - Autores del Proyecto

3.3.1. Color de piel

Con la Imagen Convexa y las coordenadas de la BoundingBox, se obtiene una imagen donde sólo se encuentra el rostro sin cabello y sin fondo, sólo el rostro como tal (Fig. 15).

La determinación del color se realiza mediante dos métodos diferentes.



Figura 15. Imagen con sólo el rostro de la persona
Fuente: IMM Face Database [10] - Autores del Proyecto

3.3.1.1 Método de filtrado y luminosidad

En este método [19] es necesario eliminar cualquier parte del rostro que impida la correcta determinación del color de la piel y que sea un obstáculo para el algoritmo; fosas nasales, ojos, labios, cejas, barba, entre otras cosas, deben eliminarse.

Para filtrar todos estos colores indeseados, se hace una umbralización de la imagen mediante la función *graythresh* que usa el método Otsu [20], el cual busca el umbral que minimiza la varianza entre las clases umbralizadas negra y blanca. Sin embargo, dada la cantidad de negro en la imagen es necesario un procedimiento previo para que el umbral no se vea afectado.

En la matriz que representa la imagen, se eliminan las entradas "0" para eliminar el fondo negro, esto redimensiona la matriz dentro de un vector de tamaño igual a la cantidad de pixeles que no son negros, en donde los componentes R,G,B quedan de igual tamaño pero concatenados dentro del vector. Luego se redimensiona dicho vector en una matriz de 3 columnas en las que en cada columna quedan los vectores R, G, B, por último, esta matriz se redimensiona en una matriz de 3 dimensiones para que vuelva a estar en el formato de imagen RGB (Fig.16).

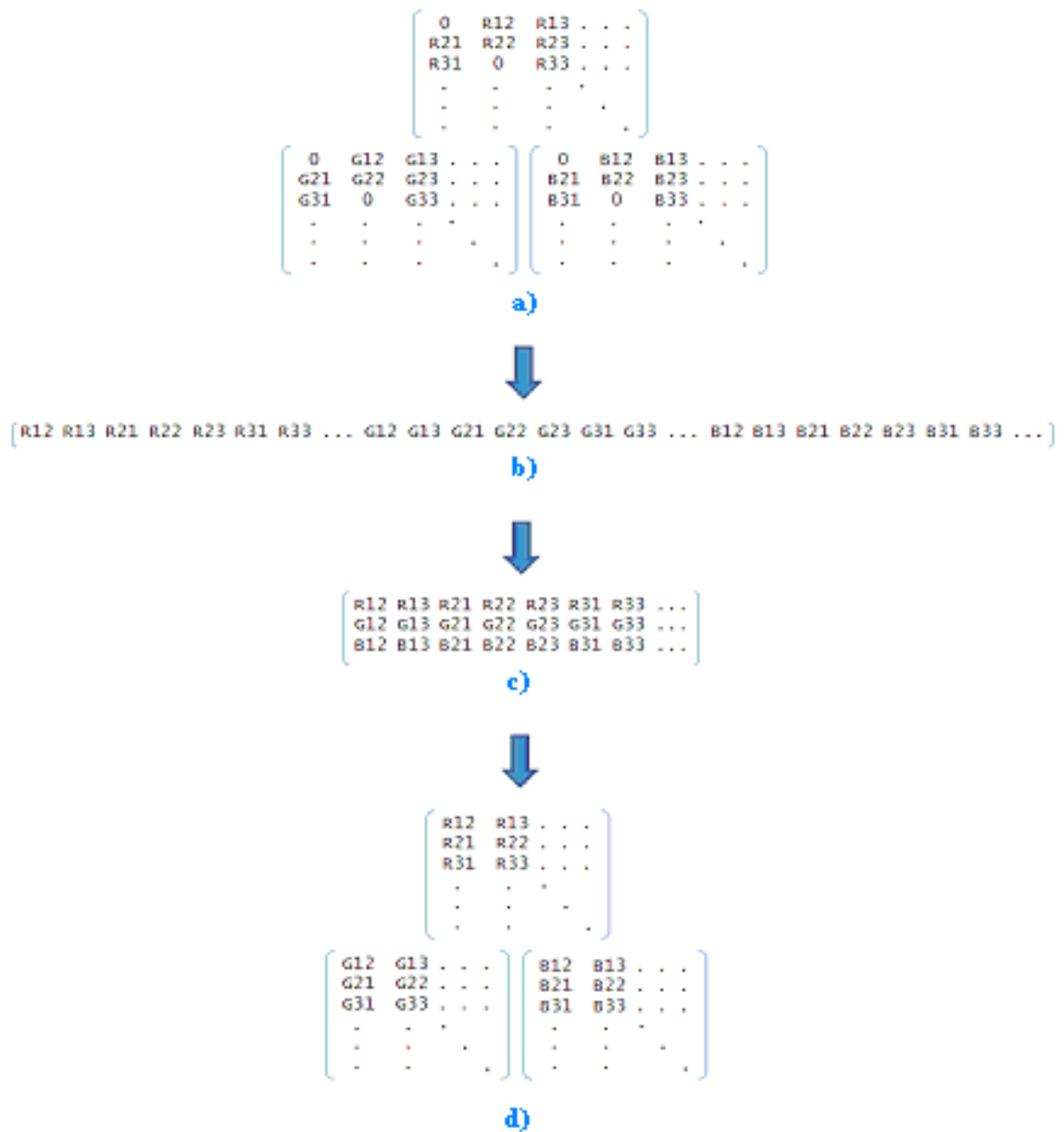


Figura 16. a) Imagen RGB con entradas iguales a cero. b) Vector sin entradas cero con la información de la imagen. c) Matriz donde cada fila contiene la información de los canales R, G, B respectivamente. d) Matriz RGB resultante con la información de color.

Dado que lo único que interesa es la cantidad de color y que cada valor en R, esté en la misma ubicación que sus correspondientes en G y B, la distorsión de la

imagen es irrelevante. La imagen resultante mantiene la información del color, mas no la forma (Fig.16).

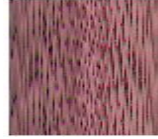


Figura 17. Imagen con la información del color.

Una vez obtenida únicamente la información de color, se realiza filtrado mediante umbralización para quitar las zonas donde hay información no necesaria, previamente hecha una normalización de la imagen (Fig.18)

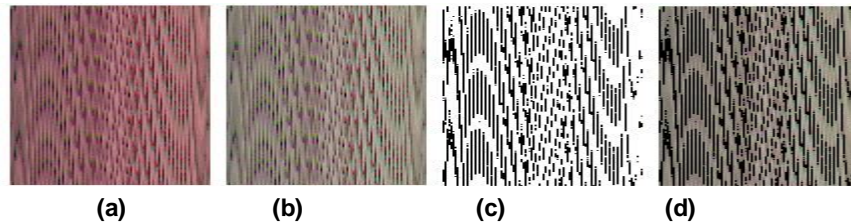


Figura 18. (a) Imagen de color. (b) Imagen Normalizada. (c) Máscara de umbral. (d) Imagen sin colores indeseables.

Luego se halla el vector de luminancia descrito en la ecuación (2), y con las entradas entre los percentiles 33 y 98 se evalúa la media. Este procedimiento elimina los brillos y otras impurezas indeseadas [19]. El resultado es la primera característica de la clasificación por color de piel.

$$Y = R + G + B \quad (2)$$

3.3.1.2 Método del histograma

El segundo método que se realiza, ideado por los autores, está basado en el histograma de la imagen. Teniendo como entrada la imagen redimensionada de sólo color, se saca el histograma y se halla el punto donde éste se hace máximo, obteniendo así la segunda característica (Fig.19). Mediante resultados experimentales se observa que la normalización fortalece los resultados para las personas de color blanco pero debilita el resultado para las personas de color

oscuro, por lo cual se prefiere implementar este segundo método sobre la imagen sin normalizar.

Mediante estas dos características extraídas se puede clasificar posteriormente entre los grupos “Moreno” y “No moreno”. Entre más se alejan de 0 las dos características extraídas, el rostro se hace menos “Oscuro”.

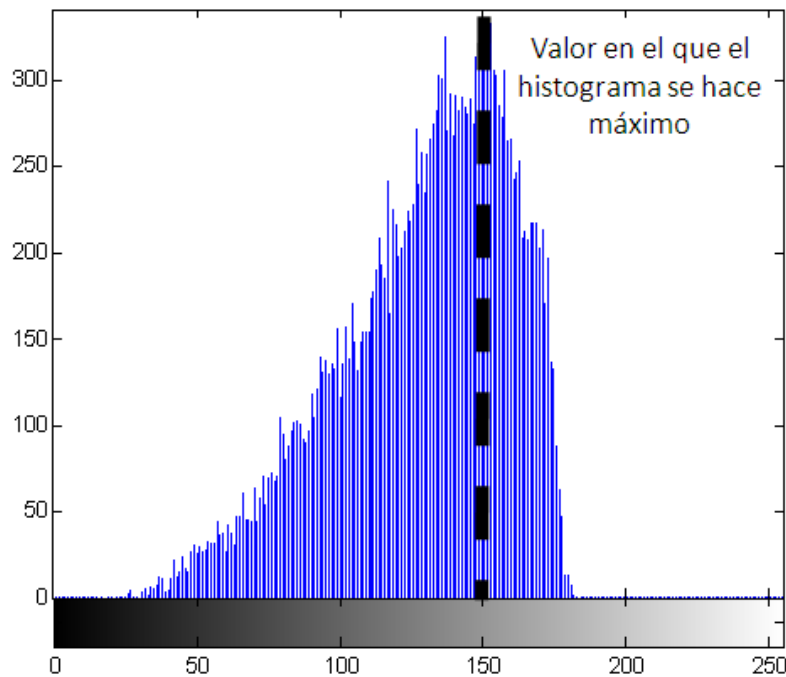


Figura 19. Histograma de la imagen.

3.3.2 Geometría

Mediante la propiedad *convexArea* que halla el área de la figura generada por *convexImage*, y la propiedad *boundingBox* de la función *regionprops* de MATLAB se extraen las dos características escogidas para la clasificación por geometría del rostro que permite clasificar en grupos con “Forma casi cuadrada” y “Forma no cuadrada”.

La primera característica de clasificación se extrae mediante la razón entre el área del área convexa y el área del rectángulo generado por *boundingBox*. La idea es

que entre esta razón se acerque más a la unidad, querrá decir que el área convexa se acerca más al rectángulo. La segunda característica se halla con la razón entre el ancho de la imagen convexa y la altura de la misma. De manera similar entre esta razón se acerque más a la unidad querrá decir que el rostro es menos alargado y por lo tanto más cuadrado.

En la figura 20 se aprecia claramente esta hipótesis. La sombra roja muestra el área convexa mientras que la amarilla el rectángulo *boundingBox*. Nótese la diferencia entre la imagen de la izquierda, más alargada, donde el área roja es mucho menor que la amarilla al compararla con las áreas de la imagen de la derecha donde el área roja se aproxima más al área amarilla.

Cabe destacar que no se encontró literatura disponible sobre este tipo de clasificación.

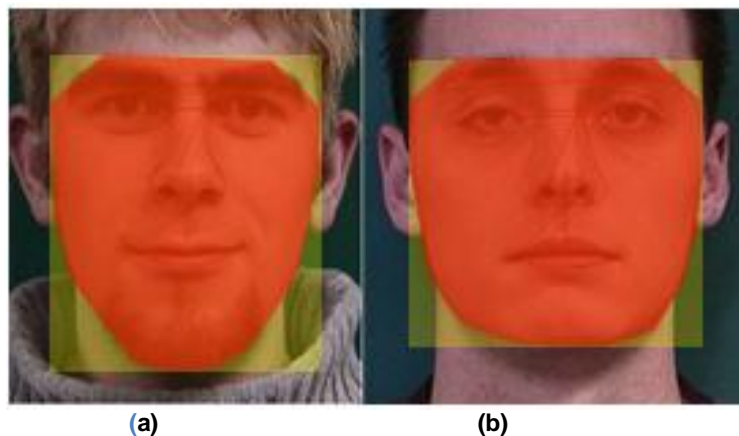


Figura 20. (a) Ejemplo de rostro con "Forma no cuadrada". (b) Rostro con "Forma Cuadrada".
Fuente: The IMM Face Database – Autores del Proyecto

3.3.3 Género

El género se determina mediante la extracción de 5 características geométricas del rostro: grueso de las cejas (2), ancho de la nariz y las distancias de las cejas a los ojos (2), las cuales son las principales características que difieren en gran medida de un hombre a una mujer [21]. La extracción de estas características se hace a través de la librería Stasm antes mencionada, el cual genera 68 landmarks

para cada imagen. Teniendo los landmarks y conociendo el número de cada landmark asociado a cada una de las partes de interés de la imagen - cejas, nariz y ojos (Fig.21), basta con calcular las distancias entre cada landmark para obtener cada una de las características empleadas para la clasificación de género.

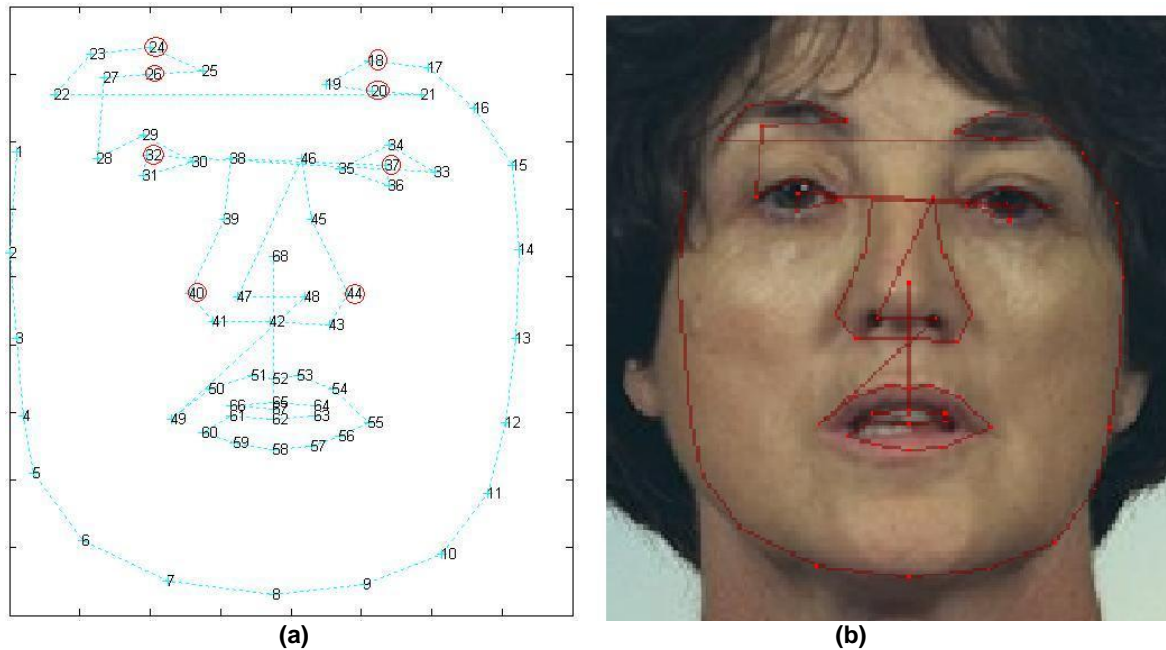


Figura 21. (a) Modelo de puntos (landmarks) asociados a la imagen en (b).
Fuente: FERET Database, Autores del Proyecto

3.4 Toma de decisión

De acuerdo a la extracción anteriormente descrita para la clasificación por color de piel, género y geometría, de cada rostro se obtienen 2 características para clasificación por piel, 5 para clasificación por género y 2 para clasificación por geometría. Esto se ve resumido en la tabla 3.

En esta etapa se hace necesaria la utilización de un algoritmo de clasificación para elegir la mejor separación en cada característica facial escogida. Para esto es importante determinar las clases es las que se quiere clasificar. Para color de piel se separa entre “Moreno” y “No moreno”; género entre “Hombre” y “Mujer”; geometría entre “Forma No Cuadrada” y “Forma Cuadrada”. Ejemplos de dichas separaciones se pueden apreciar en la tabla 4.

Tipo de clasificación	Característica
Piel	Filtrado- Luminosidad Histograma
Género	Grosor ceja izquierda Grosor ceja derecha Tamaño de la nariz Distancia ojo derecho- nariz Distancia ojo izquierdo – nariz
Geometría	ConvexArea/BoundingBox Ancho/largo

Tabla 3. Resumen de las características extraídas.

De acuerdo con esto, se utiliza un algoritmo clasificador para determinar la separación adecuada para cada clasificación. Para dicho propósito al comparar los distintos tipos de clasificación para género, se decide implementar una Máquina de Soporte Vectorial para cada tipo de clasificación [22].





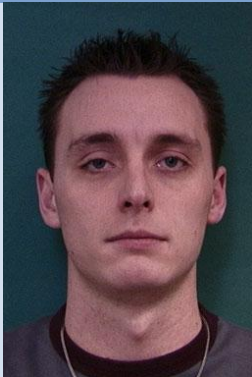

Clasificación	Clases	
	<i>Moreno</i>	<i>No Moreno</i>
Color de Piel		
	(a)	(b)
	<i>Hombre</i>	<i>Mujer</i>
Género		
	(c)	(d)
	<i>Forma Cuadrada</i>	<i>Forma No Cuadrada</i>
Geometría		
	(e)	(f)

Tabla 4. Clasificaciones con sus respectivas clases y ejemplos.
Fuente: The IMM Database, Autores del Proyecto

3.4.1 Máquinas de Soporte Vectorial

Una máquina de soporte vectorial (SVM – *Support Vector Machines*) es un algoritmo de aprendizaje para la clasificación de patrones que requiere de un set de entrenamiento inicial. El principio básico de entrenamiento detrás de las SVM es encontrar el hiperplano lineal más óptimo de manera que el error de clasificación esperado en las muestras de prueba se minimice, de manera que se tenga un buen comportamiento general [23]. La dimensión VC (del inglés *Vapnik-Chervonenkis dimension*) es una medida de la capacidad de los algoritmos de clasificación estadística, definida como la cardinalidad del mayor conjunto de puntos que el algoritmo puede separar [24]. Basado en esto, el SVM utiliza un acercamiento sistemático para encontrar una función lineal con la menor dimensión VC. Para datos que no se puedan separar de manera lineal, el SVM puede mapear la entrada (no-lineal) en un espacio de características de alta dimensión donde se puede encontrar un hiperplano que lo separe. A pesar de que no hay garantía de que siempre exista una solución lineal en el espacio con alta dimensión, en la práctica es factible construir una solución que trabaje bien [23].

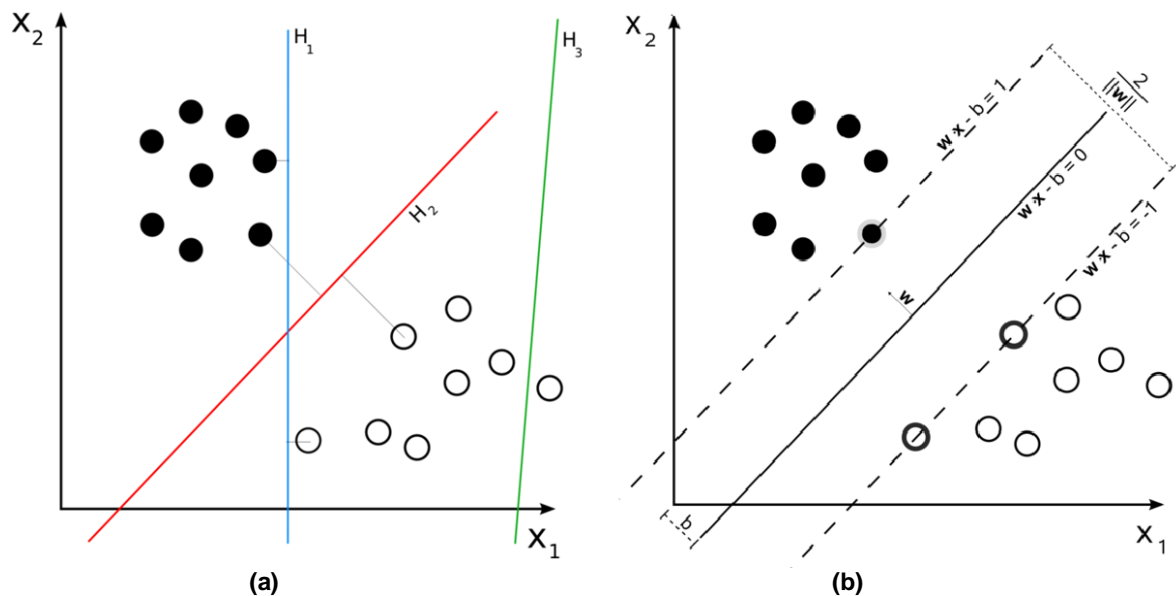


Figura 22. (a) H3 no separa ambas clases, H1 las separa con un pequeño margen mientras que H2 las separa con el máximo margen. (b) Hiperplano con máximo margen, los vectores más cercanos de cada clase son llamados los vectores de Soporte.

Fuente: http://en.wikipedia.org/wiki/Support_vector_machine

Un SVM requiere de un set de parejas de entrenamiento $(x_i, y_i), i = 1, \dots, l$ donde $x_i \in \mathbb{R}^n$ son los vectores de entrenamiento y $y_i \in \{1, -1\}^l$ indica la clase a la que pertenece cada vector de entrenamiento [25]. El objetivo final de un algoritmo SVM es encontrar el hiperplano óptimo que separe en dos grupos con el mayor margen de distancia entre ambos grupos a los vectores x_i , tomando como punto de separación la variable predictora y_i . De esta forma, los puntos del vector que se encuentren situados dentro de los márgenes de una categoría de la variable estarán a un lado del hiperplano y los casos que se encuentren en la otra categoría estarán al otro lado. Si dicho hiperplano existe, es conocido como *hiperplano de máximo margen* (Fig.22).

Dado que todo hiperplano se puede escribir como el conjunto de puntos x que satisfacen la ecuación (3), donde w es el vector normal al plano y el parámetro $\frac{b}{\|w\|}$ determina el offset del hiperplano a partir del origen y a través de w , el problema se reduce a minimizar $\|w\|$, sujeto a que (para todo $i = 1, \dots, l$) se mantenga que $y_i(w \cdot x_i - b) \geq 1$ [26].

$$wx - b = 0 \quad (3)$$

Desafortunadamente los universos a estudiar no se suelen presentar en casos ideales en los que la separación se puede lograr con un hiperplano N-dimensional, sino que para este trabajo el algoritmo SVM debe tratar con curvas no lineales de separación y casos donde los conjuntos de datos no pueden ser completamente separados.

Debido a las limitaciones computacionales de las máquinas de aprendizaje lineal, éstas no pueden ser utilizadas en la mayoría de las aplicaciones del mundo real. La representación por medio de funciones *Kernel* [27] ofrece una solución a este problema, proyectando la información a un espacio de características de mayor

dimensión el cual aumenta la capacidad computacional de las máquinas de aprendizaje lineal. Es decir, se mapea los vectores x_i a un nuevo espacio de características de mayor dimensionalidad (Espacio de *Hilbert*) donde existe un producto punto especial para dicha función *Kernel* que agiliza los cálculos y permite la separabilidad basándose en el mismo problema de optimización:

$$\min_{w,b,\xi} \frac{1}{2} \|w\|^2 + C \sum \xi_i \quad (4)$$

$$\text{dado que } y_i(\phi(w) \cdot \phi(x_i) - b) \geq 1 - \xi_i \quad (5)$$

$$\xi_i \geq 0 \quad (6)$$

Donde el término $\|w\|$ se cambia por $\frac{1}{2} \|w\|^2$ dado que tienen el mismo valor mínimo para agilizar los cálculos computacionales; C es una constante dada por el método para encontrar el mínimo de una función basada en multiplicadores de *Lagrange*; ξ_i es una penalización por el grado del error de clasificación del dato x_i , lo cual se agrega para permitir separar conjuntos de datos que no son completamente separables; y donde $K(w, x_i) = \phi(w) \cdot \phi(x_i)$ es la función *Kernel* [26]. Nótese que esta optimización termina siendo un equilibrio entre un amplio margen y un pequeño error de penalización.

Las funciones *Kernel* más utilizadas son [25]:

- Lineal: $K(x, y) = x \cdot y$
- Polinomial: $K(x, y) = (\gamma x \cdot y + r)^d, \gamma > 0$
- *Radial Basis Function* (RBF): $K(x, y) = e^{-\gamma \|x-y\|^2}, \gamma > 0$
- *Sigmoid*: $K(x, y) = \tanh(\gamma x \cdot y + r)$

3.4.2 LIB-SVM

LIB-SVM [28] es una librería integrada para Máquinas de Soporte Vectorial. El clasificador utiliza esta librería para encontrar el hiperplano de separación de la

SVM, de cada uno de las clasificaciones deseadas. La librería encuentra el modelo del hiperplano, con la cual posteriormente se determina el porcentaje de clasificación, mediante un conjunto de prueba.

4. RESULTADOS

4.1 Base de Datos

El algoritmo propuesto es evaluado mediante un conjunto de rostros de la base de datos FERET (Fig.23), previamente clasificados manualmente para cada una de las tres clasificaciones.



Figura 23. Imágenes de la base de datos.
Fuente: FERET Database, Autores del Proyecto

En total de la base de datos de FERET [11] se usó un grupo de 323 imágenes, el cual fue dividido en 3 subconjuntos, la distribución de los subconjuntos se observa en la tabla 5. La cantidad total de imágenes por clasificador se estableció de

acuerdo al menor número de imágenes que se encontraron para cada clase en la base de datos, así por ejemplo en la clasificación por color la cantidad total de imágenes (100) fue establecida por la clase de morenos ya que de esta clase solo se encontraron 46 imágenes. Partiendo de esta cantidad se determinó que para un balanceo adecuado de datos la mejor cantidad de imágenes indicada para la clase de no morenos fuese 54. De igual forma se procedió con el total de imágenes para la clasificación por género y geometría, quedando distribuidas así:

- GÉNERO: (97 mujeres, 143 hombres)
- GEOMETRÍA: (141 no cuadrados, 182 cuadrados)
- COLOR: (46 morenos, 54 no morenos)

GRUPO	GEOMETRIA	COLOR	GENERO
ENTRENAMIENTO	232	70	168
PRUEBA	91	30	72
TOTAL	323	100	240

Tabla 5. Distribución de imágenes en subconjuntos de acuerdo al tipo de clasificación.

Una vez definidos los subconjuntos, cada grupo de imágenes es pasado por el algoritmo clasificador, con el fin de extraer las características pertinentes para cada clasificación ya establecidas antes.

4.2. Selección SVM

Una SVM con función *Kernel* RBF (Función de base radial) se seleccionó por tener menos dificultades numéricas ya que esta función *Kernel* tiene sólo un parámetro (γ) por definir [25]. Así los únicos parámetros a seleccionar para el modelo clasificador, son C y γ . Como no son conocidos de antemano qué valores para C y γ son mejores para el problema, se debe hacer un tipo de selección (búsqueda

de parámetros). La meta es identificar correctamente C y γ para que el clasificador pueda predecir exactamente datos desconocidos. La forma recomendada para hallar estos parámetros es la validación cruzada, la cual consiste en dividir el grupo de entrenamiento en v subconjuntos de igual tamaño, posteriormente uno a uno los subconjunto son probados usando el clasificador previamente entrenado con los restantes $v-1$ subconjuntos, así la precisión de hacer validación cruzada está dada por el porcentaje de datos que son correctamente clasificados. Mediante una búsqueda de grilla (*grid-search*) se selecciona aquel par de parámetros que mejor validación cruzada arrojen. El procedimiento que se llevó a cabo para lograr los resultados obtenidos fue el sugerido en [25] en donde plantean un barrido con crecimiento exponencial para el par de parámetros así: $C = 2e-5, 2e-3, \dots, 2e15$, $\gamma = 2e-15, 2e-13, \dots, 2e3$, la variación de los parámetros se hizo inicialmente por parejas hasta encontrar los parámetros que mejor porcentaje de acierto arrojaban, seguidamente con el propósito de obtener mejores resultados se dejó fijo uno variando el otro y viceversa. Los resultados obtenidos se relacionan en las Figuras 24,25 y 26.

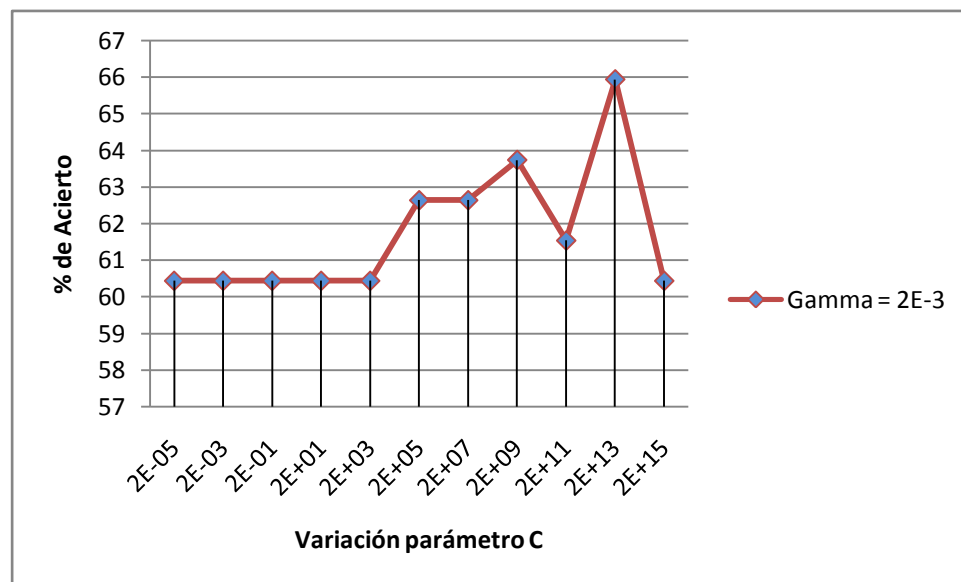


Figura 24. Barrido de los parámetros C y γ con sus respectivos porcentajes de acierto para la clasificación por geometría.

DISEÑO E IMPLEMENTACIÓN DE UN CLASIFICADOR DE IMÁGENES FACIALES PARA SER UTILIZADO EN UN SISTEMA DE RECONOCIMIENTO DE ROSTROS EN TIEMPO REAL

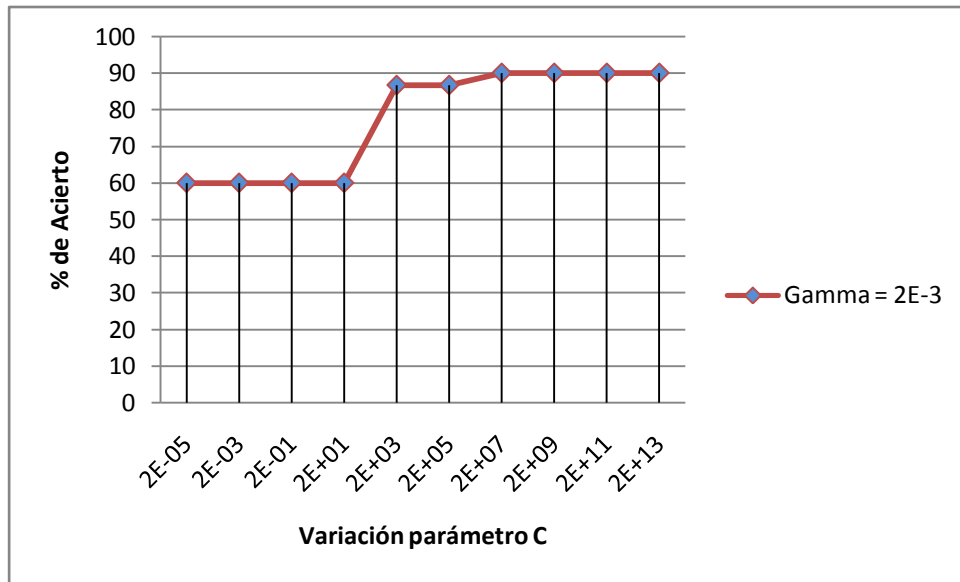


Figura 25. Barrido de los parámetros C y γ con sus respectivos porcentajes de acierto para la clasificación por color de piel.

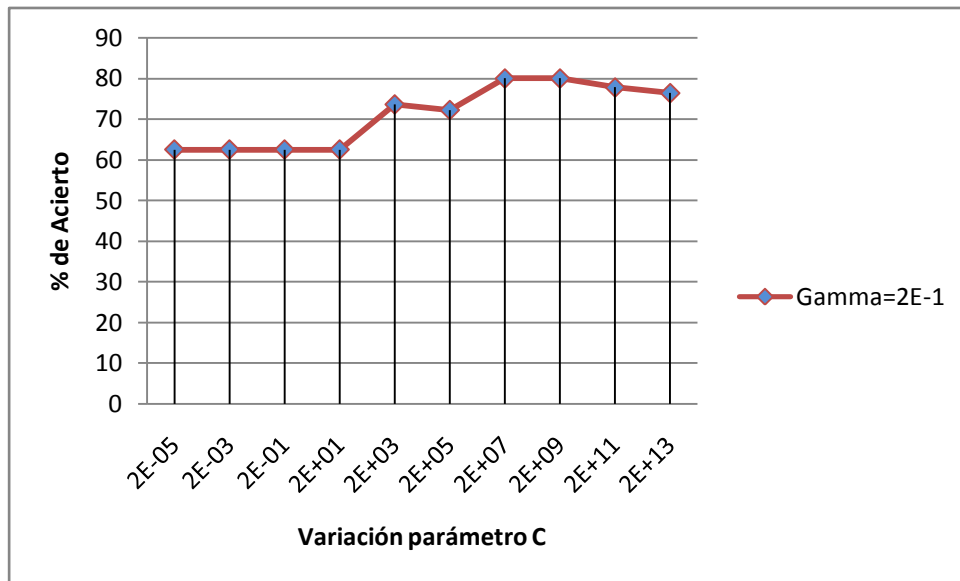


Figura 26. Barrido de los parámetros C y γ con sus respectivos porcentajes de acierto para la clasificación por género

4.3. Modelos Obtenidos

Establecidos los mejores parámetros para la función *Kernel* y el modelo se encontraron los siguientes modelos con sus respectivos porcentajes de clasificación.

	GEOMETRÍA	COLOR	GÉNERO
C	2e13	2e6	2e7
γ	2e-3	2e-3	2e-1
Grupo de Entrenamiento	232 imágenes	70 imágenes	168 imágenes
Grupo de Prueba	91 imágenes	30 imágenes	72 imágenes
Imágenes Totales	323 imágenes(141 no cuadradas, 182 cuadradas)	100 imágenes (46 morenos, 54 blancos)	240 imágenes (97 mujeres, 143 hombres)
% Acierto	65,9361% (60/91)	90% (27/30)	80% (58/72)

Tabla 6. Modelos finales obtenidos para el clasificador

4.4 Comparación con trabajos previos

Pocos trabajos se han dedicado específicamente a la clasificación por color de piel. Más que todo se han concentrado en el reconocimiento a partir de la detección de la piel [30], por lo cual se hace complejo realizar una comparación entre trabajos previos. Trabajos previos sobre clasificación por color de piel muestran efectividades menores al 75% [31] y otros no muestran el porcentaje de acierto de los algoritmos [20], por lo cual los resultados obtenidos superan los aquellos descritos previamente. La clasificación por género ha sido ampliamente implementada obteniendo resultados entre 80% y 90% [22], por lo cual la propuesta de este trabajo se encuentra en los estándares previos. No fue posible

encontrar trabajos previos sobre clasificación por geometría del rostro por lo que este trabajo puede ser considerado un punto de partida (Tabla 6).

Clasificación	Trabajos Previos	Resultado de este trabajo
Color de Piel	- Detección de piel sin concentrarse en la clasificación.[30]	Porcentaje de acierto del 90%
	- No muestran porcentaje de acierto.[31]	
Género	- Eficiencia por debajo del 75%[20]	Porcentaje de acierto del 80%
	Porcentajes de acierto entre 80% y 90% [22]	
Geometría	No se encontraron	Porcentaje de acierto del 66%

Tabla 7. Comparación con trabajos previos

5. CONCLUSIONES

En este trabajo, se presentó la clasificación de imágenes de rostros humanos basada en tres características faciales: Color, Género y Geometría. Se usó el algoritmo ASM para facilitar la extracción de características y posteriormente SVM para la toma de decisiones.

Los resultados experimentales mostraron que la clasificación por color tuvo una eficiencia bastante alta. Más aún, es de resaltar la utilización del algoritmo ASM para eliminar las regiones que no son parte de la piel puesto que no se encuentran referencias con respecto a la aplicación de dicha técnica que mejora considerablemente el resultado de la clasificación. Por su parte la clasificación por género fue buena, dados los antecedentes que muestran eficiencia entre el 80% y el 90% [22].

Con respecto a la clasificación por geometría hay aún muchos aspectos para mejorar ya que la eficiencia no fue la esperada. Este resultado se puede explicar por la dificultad de evaluar claramente la geometría de una persona por parte del ser humano, ya que existe algo de subjetividad en la clasificación. Sin embargo, la clasificación por geometría marca un punto de partida debido a la ausencia de literatura sobre este tema en particular, razón por la cual se destaca la utilización de esta técnica en este trabajo y el punto de referencia para futuras investigaciones.

BIBLIOGRAFÍA

- [1] J. Flitter. "Reconocimiento de Patrones". http://ie.fing.edu.uy/ense/asign/recpat/material/sistemas_rec_patrones.pdf. 2002
- [2] K. Fukunaga. "Introduction to Statistical Pattern Recognition", Second Edition (Computer Science and Scientific Computing Series), 1990.
- [3] DAME Data Mining and Exploration. "Classification". <http://voneural.na.infn.it/classification.html>
- [4] J. Cáceres. "Reconocimiento de Patrones y el aprendizaje No Supervisado". Escuela técnica Superior de Informática. Universidad de Alcalá
- [5] S. Gunn. "Support Vector Machines for Classification and Regression", Faculty of Engineering, Science and Mathematics. University of Southampton. 1998
- [6] H. Horn, J. Kosche. "Segmentierung von Bilddaten". Seminar: Effiziente Darstellung von Daten. 2004.
- [7] R. Bayona. "Implementación Software de algoritmos para la detección de rasgos faciales utilizados como indicadores de somnolencia en conductores usando secuencias de imágenes Biométricas". Universidad Industrial de Santander. 2007.
- [8] T.F. Cootes C.J. Taylor, D.H. Cooper, J. Graham. "Active shape models - their training and application". Computer Vision and Image Understanding (61): 38--59. 1995.
- [9] M. Kowalewsky. "Lecture 5 – Morphometric Data/Methods". <http://www.paleo.geos.vt.edu/geos5384/Lec5.htm>
- [10] M. Nordstrom, M. Larsen, J. Sieralowski, M. Stegmann. "The IMM Face Database – An Annotated Dataset of 240 Images. Informatics and Mathematical Modelling, Technical University of Denmark. 2004.
- [11] The Color FERET Database <http://face.nist.gov/colorferet>. 2003.
- [12] S. Milborrow: Stasm software library (2007). www.milbo.users.sonic.net/stasm
- [13] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. CVPR, Volume 1, pages 511-518, 2001.
- [14] Open Computer Vision Library. <http://opencvlibrary.sourceforge.net>
- [15] R. Baldock, J. Graham. Chapter 7: "Model-Based Methods in Analysis of Biomedical Images in Image Processing and Analysis". Oxford University Press, 2000, pp223-248.
- [16] O. Jesorsky, K. Kirchberg, and R. Frischholz. Robust Face Detection using the Hausdorff Distance. AVBPA, 2001. <http://www.bioid.com/downloads/facedb>.
- [17] Hausdorff Distance. AVBPA, 2001. <http://www.bioid.com/downloads/facedb>.
- [18] A.M. Martinez and R. Benavente. The AR Face Database. CVC Tech. Report 24, 1998. http://rvl1.ecn.purdue.edu/~aleix/aleix_face_DB.html.
- [19] XM2VTS database <http://www.ee.surrey.ac.uk/CVSSP/xm2vtsdb>

- [20] M. Harville, H. Baker, N. Bhatti. "Consistent Image-Based Measurement And Classification Of Skin Color". Hewlett-Packard Laboratories.
- [21] Obuyuki Otsu. "A threshold selection method from gray-level histograms". *IEEE Trans. Sys., Man., Cyber.* 9: 62–66. 1979
- [22] R. Brunelli and T. Poggio. "Hyperbf networks for gender classification", DARPA Image Understanding Workshop, 1992, pp: 311–314.
- [23] E. Mäkinen, R. Raisamo. "Evaluation of Gender Classification Methods with Automatically Detected and Aligned Faces". *IEEE Transactions on pattern Analysis and Machine Intelligence*, vol 30, No. 3. 2008.
- [24] B. Moghaddam, M.H Yang. "Gender Classification with Support Vector Machines". Mitsubishi Electric Research Laboratories. 2000.
- [25] V. Vapnik and A. Chervonenkis. "On the uniform convergence of relative frequencies of events to their probabilities." *Theory of Probability and its Applications*, 16(2):264--280, 1971.
- [26] H. Chih-Wei, C. Chih-Chung, L. Chih-Jen. "A Practical Guide to Support Vector Classification". Department of Computer Science, National Taiwan University. 2008.
- [27] Corinna Cortes and V. Vapnik, "Support-Vector Networks", *Machine Learning*, 20, 1995. <http://www.springerlink.com/content/k238jx04hm87j80g/>
- [28] M. Aizerman, E. Braverman, and L. Rozonoer (1964). "Theoretical foundations of the potential function method in pattern recognition learning". *Automation and Remote Control* 25: 821–837.
- [29] C.-C. Chang and C.-J. Lin. "LIBSVM: a library for support vector machines". Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>. 2001.
- [30] Jinfeng Yang, Yihua Shi, and Mingliang Xiao. Geometric Feature-Based Skin Image. Tianjin Key Lab for Advanced Signal Processing.
- [31] I. A. G. Boaventura, V. M. Volpe, I. N. da Silva, A. Gonzaga Fuzzy Classification of Human Skin Color in Color Images.