



Instituto Politécnico Nacional

Centro de Investigación en Computación

«Reconocimiento de expresiones faciales
empleando memorias asociativas Alfa-Beta»

TESIS

Que para obtener el grado de Maestro
en Ciencias de la Computación

PRESENTA:

Moisés Vicente Márquez Olivera

Directores de Tesis

Dr. Amadeo José Argüelles Cruz
Dr. David Jaramillo Viguera



MÉXICO, D. F.

DICIEMBRE DE 2012



INSTITUTO POLITÉCNICO NACIONAL

SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

ACTA DE REVISIÓN DE TESIS

En la Ciudad de México, D.F. siendo las 14:00 horas del día 12 del mes de noviembre de 2012 se reunieron los miembros de la Comisión Revisora de la Tesis, designada por el Colegio de Profesores de Estudios de Posgrado e Investigación del:

Centro de Investigación en Computación

para examinar la tesis titulada:

“Reconocimiento de expresiones faciales empleando memorias asociativas Alfa-Beta”

Presentada por el alumno:

MÁRQUEZ

Apellido paterno

OLIVERA

Apellido materno

MOISÉS VICENTE

Nombre(s)

Con registro:

B	1	0	1	6	5	7
---	---	---	---	---	---	---

aspirante de: **MAESTRÍA EN CIENCIAS DE LA COMPUTACIÓN**

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

LA COMISIÓN REVISORA

Directores de Tesis

Dr. Amadeo José Argüelles Cruz

Dr. David Jaramillo Viguera

Dr. Cornelio Yañez Márquez

Dr. Oscar Camacho Nieto

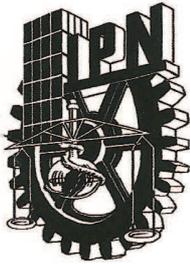
Dr. Luis Octavio López Leyva

PRESIDENTE DEL COLEGIO DE PROFESORES



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA DE INVESTIGACIÓN Y POSGRADO
DIRECCIÓN

Dr. Luis Alfonso Villa Vargas



INSTITUTO POLITÉCNICO NACIONAL
SECRETARIA DE INVESTIGACIÓN Y POSGRADO

CARTA DE CESIÓN DE DERECHOS

En la ciudad de México el día 27 del mes de Noviembre del año 2012, el que suscribe **Moisés Vicente Márquez Olivera** alumno del Programa de Maestría en Ciencias de la Computación con un registro B101657, intelectual del presente trabajo de Tesis bajo la dirección de Dr. Amadeo José Argüelles Cruz y de Dr. David Jaramillo Viguera, cede los derechos del trabajo intitulado **"Reconocimiento de expresiones faciales empleando memorias asociativas Alfa-Beta"**, al Instituto Politécnico Nacional para su difusión, con fines académicos y de investigación.

Los usuarios de la información no deben reproducir el contenido textual, gráficas o datos del trabajo sin el permiso expreso del autor y/o directores del trabajo. Este puede ser obtenido escribiendo a la siguiente dirección yom91@hotmail.com o mmarquezo@ipn.mx . Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.

Moisés Vicente Márquez Olivera

(nombre y firma)

Resumen

El presente trabajo de tesis orienta su atención en la compleja tarea del reconocimiento automático de las expresiones faciales, el propósito principal es probar la asertividad y competitividad de las memorias asociativas de tipo Alfa-Beta en el reconocimiento automático de expresiones faciales.

La primera parte de la tesis brinda una breve introducción, la cual enfatiza la importancia de las expresiones faciales en la comunicación humana, incluyendo los estudios más relevantes dedicados a la búsqueda e interpretación de dichas expresiones, con el objetivo de entender la sinergia entre estos conocimientos y las recientes innovaciones tecnológicas que dan como resultado el diseño de algoritmos que permiten reconocer las expresiones faciales de forma automática.

Posteriormente se presenta el estado del arte de los trabajos que en los diez últimos años han mostrado un significativo avance en ésta área, seguidamente se realiza un análisis de las dos fases más importantes del proceso automático de reconocimiento de las expresiones faciales, incluyendo una breve descripción de los métodos típicamente empleados en la solución de cada fase y los principales problemas que atañen a este proceso.

Finalmente se presentan los resultados obtenidos del diseño e implementación de la memoria asociativa de tipo Alfa-Beta, comparándose con los resultados de otros investigadores que han utilizado algún método distinto de clasificación para reconocer las expresiones faciales. Concluyendo que el alto desempeño de las memorias asociativas Alfa-Beta brindan una solución viable y realmente competitiva en la tarea de reconocimiento automático de expresiones faciales.

Abstract

The present thesis work focuses its attention on the complex task of automatic recognition of facial expressions, the main purpose is to test the assertiveness and competitiveness of the alpha-beta associative memories in the automatic recognition of facial expressions.

The first part of the thesis provides a brief introduction, which emphasizes the importance of facial expressions in the human communication, including relevant studies devoted to the pursuit and interpretation of those expressions, with the aim to understand the synergy between these knowledge and recent technological innovations that result in the design of algorithms that allow us to recognize facial expressions automatically.

Later the state of the art is presented in the work that in the last ten years have shown a significant step forward in this area, then carried out an analysis of the two most important phases of the process of automatic recognition of facial expressions, including a brief description of the methods typically used in the solution of each phase and the main issues of concern to this process.

Finally, we present the results of the design and implementation of the alpha-beta associative memory, comparing with the results of other researchers who have used a different classification method to recognize facial expressions. Concluding that the high performance of the alpha-beta Associative Memories provide a competitive and viable solution in the task of automatic recognition of facial expressions.

Agradecimientos

Este trabajo de investigación no sería posible sin el apoyo de muchas personas e instituciones que en mucho colaboraron con la culminación de este trabajo y es por ello que deseo agradecer a:

El Instituto Politécnico Nacional, que me ha abierto las puertas hacia increíbles experiencias profesionales y personales, por darme la oportunidad de crecer y aprender en sus aulas, con maestros de alta calidad por eso y más soy orgullosamente politécnico.

Al Centro de Investigación en Computación CIC-IPN por guiarme en el aprendizaje de una nueva área de conocimiento, siempre brindando calidad de enseñanza y servicio en los procesos administrativos, por eso doy gracias a mis profesores y a todas las secretarias en donde verdaderamente se preocupan por ti.

Al grupo Alfa-Beta por ser abierto a propuestas y por compartir el conocimiento que se tiene de las Memorias Asociativas Alfa-Beta piedra angular del presente trabajo de investigación.

A el Dr. Amadeo Argüelles por ser un director de tesis integral, brindándome siempre el apoyo y confianza, permitiéndome innovar e impulsando siempre mi creatividad. Gracias por ser un director de tesis excepcional.

A el Dr. David Jaramillo por ser una persona que admiro profundamente y a la que debo lo que tengo y lo que soy, su humildad y bondad infinita me alienta todos los días a seguir adelante e inspirarme en sus conocimientos y enseñanzas para no rendirme y pensar siempre que la vida nos regala una nueva misión para superarnos y ayudar a que el mundo sea un poco mejor cada día.

Agradecimientos especiales

A Dios, por darme la oportunidad de vivir y por estar conmigo en cada paso que doy, por fortalecer mi corazón e iluminar mi mente y por haber puesto en mi camino a aquellas personas que han sido mi soporte y compañía durante todo el periodo de estudio. por haberme permitido llegar hasta este punto y haberme dado salud para lograr mis objetivos, además de su infinita bondad cuidado y amor.

A mi mami Paty que peleó como guerrera por mi vida pese a los diagnósticos médicos de que yo sería un vegetal. Gracias mamita por darme la oportunidad de conocer la felicidad, la tristeza, el amor, la risa, el llanto, el sufrí y las luces de colores. Te amo mamá y quiero que sepas que si hay una persona que se merece el titulo de la “mejor madre del mundo” esa eres tú.

A mi papá Carlos quién siempre tenia un consejo para darme y tiempo para escucharme, gracias por pensar siempre en mi bienestar, por formar parte de mi vida pero sobre todo por ser mi papá.

A mi tío Joel por ayudarme a entender y amar las matemáticas.

A Lili y Jimmy por abandonarlos en mis largas horas de estudio, por todas las fechas pospuestas, gracias por comprender chavos, ustedes saben que se les quiere mucho.

A Carlos y Abril por las horas de desvelo haciendo tarea en la casa y por incoherencias que surgían de las innumerables horas de estudio.

Al gran amor de mi vida por obligarme a hacer la maestría que hoy gracias a Dios llega a su fin. Nena gracias por ser mi compañera, amiga y confidente, por entenderme y quererme como soy, por seguir mis locuras, por vivir y sufrir esta maestría a mi lado, como siempre y por siempre juntos.

Índice general

1. Introducción	11
1.1. Aspectos generales	11
1.2. Justificación	14
1.3. Objetivo general	15
1.4. Objetivos específicos	15
1.5. Estructura de la tesis	16
2. El problema del Reconocimiento de Expresiones Faciales	17
2.1. Estado del arte de los 10 últimos años	18
2.2. Proceso del reconocimiento de expresiones faciales	20
2.3. Análisis e Identificación	20
2.3.1. Identificación facial y la extracción de características	22
2.3.1.1. Detección del rostro	22
2.3.1.2. Extracción de características.	23
2.3.2. Secuencia general de la fase de análisis e identificación	23
2.3.3. Problemas básicos presentes en la fase de análisis e identificación	24
2.3.3.1. Número de caras	24
2.3.3.2. Localización de elementos de interés	24
2.3.3.3. Pose de captura	24
2.3.3.4. Iluminación	25
2.3.3.5. Seguimiento de rostro en video	25
2.3.4. Clasificación de las técnicas empleadas en la fase de análisis e identificación	25
2.3.4.1. Clasificación propuesta por Hjelmás y Low [1].	26
2.3.4.2. Clasificación propuesta por Yang et. al.	28
2.3.5. Base de Datos	28
2.3.6. Descripción de las técnicas empleadas en el Análisis e Identificación	30
2.3.6.1. Seguidor basado en plantillas	30
2.3.6.2. Método de PCA (Análisis de Componentes Principales)	32
2.3.6.3. Método de LDA (Análisis Discriminante Lineal)	36
2.3.6.4. Método de seguimiento basado en el color	37
2.4. Clasificación de las expresiones faciales	40
2.4.1. Métodos utilizados en el reconocimiento	41
2.4.2. Enfoques para la clasificación	41
2.4.2.1. Redes Neuronales Artificiales	43
2.4.2.2. Máquinas de Soporte Vectorial (SVM)	47
2.4.2.3. Redes bayesianas	48
2.4.2.4. Memorias asociativas	49

3. Materiales y Métodos	50
3.1. Memorias asociativas Alfa-Beta	50
3.1.1. Memorias Asociativas y la inteligencia artificial (IA)	50
3.1.1.1. Conceptos básicos	51
3.2. Operaciones binarias α y β : definiciones y propiedades	51
3.2.1. Propiedades algebraicas de los operadores binarios Alfa-Beta	52
3.2.2. Operaciones matriciales	53
3.2.3. Memorias heteroasociativas Alfa-Beta	56
3.2.3.1. Algoritmo de la Memoria Heteroasociativa Alfa-Beta tipo Max (\vee)	56
3.2.3.2. Algoritmo de la memoria heteroasociativa Alfa-Beta tipo Min (\wedge)	57
3.2.4. Memorias autoasociativas Alfa-Beta	58
3.2.4.1. Algoritmo de la memoria autoasociativa Alfa-Beta tipo Max (\vee)	58
3.2.4.2. Algoritmo de la memoria autoasociativa Alfa-Beta tipo Min (\wedge)	60
4. Resultado y Discusión	62
4.1. Fases de experimentación	62
4.2. Paso 1.	64
4.2.1. Banco de datos seleccionado “Cohn- Kanade +”	65
4.2.2. Características de la información que proporciona Cohn-Kanade plus	66
4.2.3. Seleccionar el método de validación	66
4.3. Paso 2.	68
4.3.1. Selección de características	70
4.4. Paso 3.	71
4.4.1. Diagrama de Flujo y algoritmo de memoria Autoasociativa Alfa-Beta tipo Max	74
4.5. Paso 4.	79
4.5.1. Comparar los resultados	80
5. Conclusiones y Trabajo Futuro	82
5.1. Síntesis	82
5.2. Conclusiones	84
5.3. Trabajo a futuro	84

Índice de algoritmos

2.1. Algoritmo básico de PCA (Análisis de Componentes Principales)	33
2.2. Algoritmo básico de <i>SOM's</i> (<i>Self-Organizing Map</i>)	34
2.3. Cálculo de eigenface presentado por Matthew Turk y Alex Pentland[2]	35
2.4. Algoritmo general de una red neuronal aplicada en el reconocimiento de expresiones faciales.	47
2.5. Algoritmo general de las Redes Bayesianas	49
3.1. Algoritmo de memoria heteroasociativa Alfa-Beta tipo Max (\vee)	56
3.2. Algoritmo de la memoria heteroasociativa Alfa-Beta tipo Min (\wedge)	57
3.3. Fase de aprendizaje de las memorias autoasociativas tipo Max (\vee)	58
3.4. Fase de aprendizaje de la memoria autoasociativa Alfa-Beta tipo Min (\wedge)	60
4.1. Fase de aprendizaje de la memoria Autoasociativa Alfa-Beta, utilizando código Johnson-Möbius.	75
4.2. Fase de recuperación de la memoria Autoasociativa Alfa-Beta tipo Max, utilizando código Johnson-Möbius.	76
4.3. Fase de prueba de la memoria Autoasociativa Alfa-Beta, utilizando código Johnson-Möbius.	78

Índice de figuras

2.1. Proceso general del reconocimiento automático de expresiones faciales.	20
2.2. Esquema general de la primera fase del proceso de reconocimiento automático de expresiones faciales: Análisis e Identificación.	21
2.3. Clasificación de métodos de detección de rostros según Hjelmás y Low [1]	26
2.4. Clasificación de los métodos de detección de rostros según Yang et al. [3]	28
2.5. Arquitectura para la detección del rostro por Paul Viola & Michael Jones[4]	32
3.1. El patrón de entrada representado por un vector columna x y el patrón de salida representado por y	51
4.1. Proceso general del reconocimiento automático de expresiones faciales	63
4.2. Proceso general del reconocimiento automático de expresiones faciales	64
4.3. Secuencia de imágenes durante la emoción de sorpresa	65
4.4. Selección de los patrones que integran el conjunto de aprendizaje y prueba empleando K-fold-cross-validation.	67
4.5. Esquema extraído del trabajo de investigación de Lucey et al.[5], en el que se muestra la utilización del Modelo de Apariencia Activa (AAM), en la fase de seguimiento del rostro y en la fase de extracción de características se hace uso de Similitud normalizada de la forma (similarity-normalized shape, SPTS) y (canonical appearance, CAPP) , en el que se marcan 68 landmarks incluidas en el compendio de archivos del banco de datos de Cohn Kanade plus.	68
4.6. Esquema que presenta la fase de analisis e identificación extraído del trabajo de investigación de Lucey et al.[5], que se encuentra incluido en la base de datos de Cohn Kanade Plus.	69
4.7. Esquema que muestra los 59 puntos considerados en la presente tesis para el reconocimiento de expresiones faciales.	70
4.8. Esquema que muestra los 59 puntos considerados en la presente tesis para el reconocimiento de expresiones faciales.	71
4.9. Esquema que muestra la fase de aprendizaje y de prueba de una memoria asociativa Alfa-Beta Autoasociativa	72
4.10. Diagrama de flujo diseñado para la fase de aprendizaje de la memoria Autoasociativa Alfa-Beta, utilizando código Johnson-Möbius.	74
4.11. Diagrama de flujo diseñado para la fase de recuperación de la memoria Autoasociativa Alfa-Beta.	76
4.12. Diagrama de flujo diseñado para la fase de clasificación de la memoria Autoasociativa Alfa-Beta.	77
4.13. Asertividad expresada en porcentaje ordenada por fechas correspondiente a los trabajos de investigación de los 10 últimos años en el reconocimiento automático de expresiones fáciles, en el que se incluye el resultado obtenido en el presente trabajo de tesis con las memorias asociativas Alfa-Beta.	80

4.14. Asertividad expresada en porcentaje ordenada de mayor a menor correspondiente a los trabajos de investigación de los 10 últimos años en el reconocimiento automático de expresiones fáciles, en el que se incluye el resultado obtenido en el presente trabajo de tesis con las memorias asociativas Alfa-Beta. 81

Capítulo 1

Introducción

1.1. Aspectos generales

Las expresiones faciales forman parte de la comunicación no verbal y surgen como consecuencia del proceso de la evolución humana, permitiendo al hombre establecer un sistema robusto de intercambio de conocimientos, ideas, emociones y pensamientos, lo cual permite enriquecer la forma en la que el hombre se comunica con otros seres vivos. Las tareas involucradas en el reconocimiento e interpretación de las expresiones faciales no son triviales, no obstante han sido por siglos una labor exclusiva del humano, el cual es capaz de brindar significado a cada uno de los gestos del rostro.

Es interesante hacer notar que el ser humano es capaz de expresar un mismo mensaje bajo un conjunto de combinaciones gesticulares diferentes y en consecuencia, el receptor puede ser capaz de interpretar el mensaje de maneras distintas, e.g. cuando alguien dice “se nota que eres un gran científico” y lo acompaña una expresión de ironía, se puede intuir que se trata de un sarcasmo, sin embargo si se acompaña de una expresión de felicidad, orgullo, respeto, seriedad o bien asombro, entonces se puede interpretar que el mensaje es un alago por el desempeño del científico.

El psicólogo Albert Mehrabian[6] realizó una investigación de la comunicación no verbal, dando a conocer que sólo un 7% de la comunicación entre dos personas se realiza mediante palabras. Un 38% se comunica mediante la voz y todos sus componentes (volumen, entonación...) y el 55% restante se lleva a cabo a través del lenguaje corporal (gestos, posturas, mirada...). De lo que se puede concluir que las expresiones faciales son un factor de gran importancia en la interpretación de los mensajes que recibimos, por ello cuando se intercambian palabras con otra persona, el mensaje no solo depende de la semántica, ya que al mismo tiempo se reciben mensajes subjetivos que complementan la comprensión del mismo.

Otra función importante de las expresiones faciales es la de transmitir «emociones», las cuales son esenciales en las complicadas relaciones personales (amigo, compañero, pareja, entre otros), es por ello que en la literatura pueden encontrarse importantes esfuerzos de investigación que presentan un conjunto de teorías que intentan identificar cada una de las emociones, además de explicar los factores que influyen directamente en su generación e interpretación.

Ahora bien, las expresiones faciales son producto de las contracciones de los diferentes músculos que constituyen el rostro humano, dichas contracciones pueden ser *voluntarias o involuntarias*, por lo tanto también existen expresiones faciales *simuladas y naturales*.

Las expresiones naturales son aquellas que provienen de emociones auténticas que el ser humano siente durante la emisión de un mensaje, mientras que las simuladas son aquellas que de forma intencional buscan fingir una emoción o bien ilustrar mediante gestos signos sintácticos que refuerzan la comprensión de un mensaje (e.g. ilustrar con gestos los signos de interrogación, admiración, exclamación, entre otros).

Uno de esos estudios es el que relaciona las emociones con las expresiones faciales, el psicólogo Paul Ekman[7] uno de los pioneros en el estudio de las emociones y su directa relación con las expresiones faciales, sus investigaciones son, hoy en día de las más citadas en estudios sobre el tema. Una de sus premisas más relevantes se sustenta en estudios previamente realizados por Charles Darwin[8], en donde la idea fundamental es que “existe un código general para la expresión y percepción de las emociones humanas”, es decir un ser humano es capaz de percibir correctamente las emociones de otro ser humano, sin importar su ubicación geográfica, cultura, idioma, raza, entre otras. Ekman[9] también propone que existen 6 emociones básicas: *felicidad, sorpresa, miedo, disgusto, cólera y tristeza*.

Ekman y Friesen desarrollaron un sistema de clasificación de acciones faciales llamado FACS (Facial Action Coding System)[10, 11], el cual se basa en los movimientos asociados a los músculos del rostro. Durante el desarrollo de su trabajo, enfatizan que una de las tareas más complejas es identificar y establecer en una escala continua de tiempo los rangos transitorios entre una emoción y otra.

De lo mencionado anteriormente, se puede deducir que las expresiones faciales representan un alto impacto e importancia en la interacción humana, cabe señalar que todo el proceso de generación de expresiones faciales y su adecuado reconocimiento es efectuado entre seres humanos de forma natural y universal.

No obstante el ser humano pretende conseguir que estas complejas tareas sean realizadas por las máquinas de forma automática, buscando con ello resultados iguales o mejores a los obtenidos en el reconocimiento humano; es por ello que las máquinas en este sentido juegan un papel muy importante, ya que permiten el desarrollo e implementación de algoritmos que brindan la posibilidad de lograr un reconocimiento automático de las expresiones faciales y aunque falta mucho para llegar a superar la efectividad del ser humano, hoy en día existen importantes esfuerzos de investigación entre ellos el presente trabajo de tesis, los cuales están empeñados en superar el reconocimiento nato del ser humano

El estudio de las expresiones faciales presenta importantes aportaciones de diversas áreas de estudio entre las que se incluyen principalmente la *Psicología, Anatomía, Sociología y en la actualidad la Inteligencia Artificial (IA)*, siendo esta última la de mayor interés en el presente trabajo de tesis ya que aporta propuestas vanguardistas para la automatización del proceso de reconocimiento de las expresiones faciales.

La Inteligencia Artificial (IA) esta integrada por un conjunto de enfoques que son empleados en la solución de problemas relacionados con el aprendizaje, clasificación y en algunos casos la predicción de patrones. Cada enfoque ostenta principios matemáticos propios, que son la base del diseño de algoritmos empleados en la solución de problemas de reconocimiento de patrones, es por ello que dichos algoritmos pueden ser ampliamente utilizados en el problema del reconocimiento de expresiones faciales.

Los enfoques con mayor auge en los últimos años son:

Enfoque estadístico—probabilístico. Los principios fundamentales de esta clase de métodos se encuentran en las teorías de probabilidad y estadística, algunos ejemplos son: Redes Bayesianas (Bayesian Network), Modelos Ocultos de Markov (Hidden Markov Model, HMMs), Máquinas de Soporte Vectorial (Support Vector Machines, SVMs), entre otros. Algunos trabajos relevantes encaminados en el reconocimiento de expresiones faciales que hacen uso de este enfoque pueden encontrarse en [12, 13, 14, 15, 16, 17, 18, 19, 20]

Clasificadores basados en métricas. Este enfoque hace uso de las mediciones espaciales que entre cada patrón, por lo que hace uso de principios métricos y espaciales, algunos ejemplos son: vecino más cercano (K-Nearest Neighbor, KNN), Distancia Euclidiana, Distancia de Mahalanobis, Distancia Dirigida de Hausdorff (Directed Hausdorff Distance, DHD), entre otros. Algunos trabajos relevantes encaminados en el reconocimiento de expresiones faciales que hacen uso de este enfoque pueden encontrarse en [21, 22, 23, 24, 25, 26]

Enfoque neuronal. Estos métodos están inspirados en los modelos biológicos del comportamiento de una red neuronal, algunos ejemplos son: Perceptrón, Backpropagation, Redes Neuronales Convolucionales (convolutional neural network, CNN), entre otros. Algunos trabajos relevantes encaminados en el reconocimiento de expresiones faciales que hacen uso de este enfoque pueden encontrarse en [27, 28, 29].

Enfoque asociativo. Este enfoque hace uso de los modelos de memorias asociativas para dar nacimiento a reconocedores robustos en la clasificación de patrones con alteraciones, dichos modelos son capaces de recuperar los patrones aprendidos de una forma óptima y además de clasificar patrones nuevos con alto grado de asertividad, algunos ejemplos son: Memorias Asociativas Alfa-Beta, Lernmatrix, Linear Associator, Memoria Hopfield, Memoria tipo CHAT, entre otras. Las memorias asociativas ya han probado ser un potente clasificador de patrones cuando es aplicado en áreas de estudio distintas al reconocimiento de expresiones faciales, es por ello que el principal interés de este trabajo de tesis es probar su asertividad en esta línea de investigación, algunos de los trabajos más destacados pueden encontrarse en: [30, 31, 32, 33, 34, 35, 36, 37].

1.2. Justificación

Debido a la importante función que desempeñan las expresiones faciales en la interacción humana, se ha convertido en todo un reto diseñar métodos que permitan interpretar estas expresiones de forma automática, lo cual nos hace pensar en la respuesta a las siguientes preguntas *¿Cuáles son los motivos que llevan al ser humano a crear métodos cada vez más eficientes en el reconocimiento de las expresiones faciales?*, dicho en otras palabras *¿Para qué desearía el ser humano que una máquina interprete sus expresiones de la misma manera que lo hace él?*.

La respuesta en gran medida esta relacionada con la actual necesidad de realizar éste proceso de forma rápida, certera, objetiva y automática, el propósito es el de contribuir en el mejoramiento de procesos de diversas áreas de estudio tales como, la mercadotecnia, seguridad, educación, psicología, psiquiatría, entre muchas otras; en donde sería de gran utilidad contar con un eficiente reconocedor automático de expresiones faciales. Por esta razón se puede decir que la importancia de contar con un óptimo reconocedor automático de expresiones faciales es directamente proporcional al extenso campo de aplicación. A continuación se presentan algunas de las múltiples aplicaciones que un sistema de reconocimiento de expresiones puede tener en ámbitos tales como:

Mercadotecnia. Se pueden realizar estudios basados en la impresión que un producto tiene sobre el consumidor, lo interesante es que las expresiones faciales nos brindan información del efecto que tiene el consumidor con respecto a un producto que por citar algún ejemplo podría ser comida, perfumes, cosméticos, juguetes o bien un producto no tangible como un comercial publicitario. Por lo que representa una herramienta importante para conocer los efectos que tiene el consumidor sobre los productos que este vende.

Seguridad. En materia de seguridad existe una amplia gama de aplicación, e.g. el programa de control de pasajeros mediante técnicas de observación (Screening of Passengers by Observation Techniques, SPOT)[38], el programa detecta expresiones faciales y comportamientos “anormales” de los pasajeros que transitan en los aeropuertos por medio de cámaras colocadas estratégicamente, SPOT busca detectar conductas que pueden ser indicativos de estrés, el miedo o el engaño, la finalidad es identificar a los individuos que representen un peligro potencial.

Educación. Actualmente la creciente demanda de aspirantes para el ingreso a las universidades ha superado la capacidad física de las instituciones públicas, lo que ha llevado a estas instituciones a establecer mecanismos que permitan estudiar vía internet (modo no presencial), un sistema de reconocimiento de expresiones representaría una herramienta para la evaluación de los materiales didácticos con los que el alumno aprende sus diferentes lecciones, ya que se podrían identificar rasgos como: el promedio en que el alumno pone atención al material didáctico, en qué secciones muestra gestos de confusión, si el material es del agrado del alumno, entre otros aspectos.

Psicología y Psiquiatría. El sistema sería un apoyo en las investigaciones que actualmente se realizan en la determinación de conductas patológicas.

Robótica. Actualmente existe un especial interés por la interacción hombre-máquina, lo cual esta ligado a la idea de que el robot debe ser capaz de establecer una eficiente comunicación con el ser humano, para ello el robot debe tener incorporado un reconocedor automático de expresiones humanas, ya que la comunicación humana se ve modificada radicalmente por los gestos que acompañan un mensaje.

1.3. Objetivo general

Implementar un algoritmo que basado en los fundamentos matemáticos de las memorias asociativas de tipo Alfa-Beta, el cual permita ofrecer un porcentaje de asertividad competitivo en la tarea de clasificación automática de expresiones faciales.

1.4. Objetivos específicos

Para alcanzar el objetivo de la tesis, es preciso cumplir los siguientes objetivos específicos:

1. Investigar sobre diferentes bancos de datos que se encuentran enfocados en la tarea de reconocimiento de expresiones faciales y seleccionar uno de ellos para ser utilizado como banco de pruebas durante los experimentos.
2. Implementar un algoritmo haciendo uso de las memorias asociativas empleando los operadores α y β y sus propiedades.
3. Probar el banco de datos seleccionado con el algoritmo de la memoria asociativa alfa Beta y comparar los resultados con otros trabajos existentes en la literatura actual.

1.5. Estructura de la tesis

En la actualidad la línea de investigación de la Inteligencia Artificial presenta un conjunto de métodos que son utilizados en la clasificación de patrones con aplicación en múltiples áreas de investigación y las expresiones faciales no son la excepción, razón por la cual este trabajo de tesis se encuentra encaminado en incursionar en el empleo de las memorias asociativas Alfa-Beta en esta tarea, con la finalidad de brindar una solución competitiva en la tarea de reconocimiento de las expresiones faciales.

Capítulo I

Presentan una breve introducción de los aspectos generales del tema de la tesis, así como la justificación, objetivos y estructura del trabajo de tesis.

Capítulo II

Presenta una recopilación de los métodos mayormente utilizados en el reconocimiento automático de expresiones faciales, así como los principales problemas que forman parte del proceso de reconocimiento. La primera parte enfatiza la importancia de las expresiones faciales en la comunicación humana, así como los esfuerzos de investigación encaminados en la búsqueda e interpretación de dichas expresiones. En la segunda parte se realiza un análisis de las dos fases más importantes del proceso automático del reconocimiento de las expresiones faciales y se describen los métodos típicamente empleados en la solución de cada fase.

Capítulo III

Se enfoca en el diseño y descripción de las memorias asociativas Alfa-Beta las cuales se apoyan en los modelos matemáticos de los operadores α y β . En este capítulo se presentan las memorias Heteroasociativas y Autoasociativas, describiendo sus operadores y algoritmos.

Capítulo IV

Se describe el desarrollo y diseño del sistema de reconocimiento de expresiones faciales incluyendo las condiciones de experimentación en la implementación de la memoria asociativa Alfa-Beta. También se documentan los resultados obtenidos en la prueba realizada bajo el modelo propuesto, del mismo modo estos resultados son comparados con los obtenidos en otros métodos probados bajo las mismas condiciones de experimentación. El objetivo es obtener un marco comparativo entre la eficiencia de las memorias asociativas Alfa-Beta y la de otros métodos que también pertenecen a la línea de investigación de la Inteligencia Artificial y que son comúnmente empleados en la comunidad científica en la tarea de reconocimiento automático de expresiones faciales.

Capítulo V

Finalmente, en el **capítulo 5** se presentan las conclusiones, así como una recapitulación de los resultados obtenidos durante la fase de experimentación, además se propone una lista de posibles mejoras para consideración de los trabajos a futuro que deriven de este trabajo de investigación.

Capítulo 2

El problema del Reconocimiento de Expresiones Faciales

Hoy en día se pueden encontrar un gran número de investigaciones que presentan estudios relevantes sobre cómo las personas identifican e interpretan la información emocional de otras personas, haciendo destacar a las expresiones faciales de entre todas las formas de comunicación no verbal, las cuales proveen la mayor cantidad de información para determinar el estado emocional de otras personas.

Las expresiones faciales son un elemento esencial en el proceso de la comunicación humana, ya que funcionan como modificadores radicales durante la interpretación de un mensaje. Debido a que las expresiones faciales son producto de las contracciones musculares del rostro, éste representa una valiosa fuente de información para el estudio, identificación e interpretación de las diferentes expresiones faciales que actuando de forma individual o bien en conjunto proporcionan información indirecta sobre lo que un ser humano siente y piensa. Sin embargo, no son tareas triviales el leer y comprender el significado de todas las expresiones faciales que el ser humano es capaz de realizar.

La comunidad científica ha mostrado gran interés en el desarrollo de algoritmos que permitan que las máquinas realicen el reconocimiento de expresiones de manera automática, buscando con ello resultados iguales o mejores a los obtenidos en el reconocimiento humano, lo cual representa un desafiante reto con un extenso campo de aplicación. La Inteligencia Artificial (IA) ha promovido la posibilidad de tener reconocedores automáticos de expresiones faciales, esto se debe a que cuenta con nuevos algoritmos que son capaces de aprender, clasificar e incluso predecir un conjunto de patrones; los cuales pueden ser usados para aprender y reconocer los diferentes tipos expresiones.

Gracias a algoritmos de este tipo no es extraño encontrar múltiples investigaciones encaminadas en el desarrollo de la visión artificial, muchas de ellas interesadas en la identificación de rostros y en el reconocimiento de expresiones faciales. Sin embargo, el desarrollo de algoritmos capaces de procesar imágenes depende en gran parte del desarrollo de hardware y software, los cuales proporcionan las condiciones óptimas para lograr que los algoritmos sean rápidos, precisos y útiles.

Si bien el avance de la tecnología es necesario para el diseño de reconocedores automáticos capaces de identificar los diferentes gestos que el ser humano es capaz de realizar, también se debe considerar que no es sencillo realizar estudios que ayuden en la búsqueda del significado de cada expresión facial, así como su directa relación con las emociones y pensamientos. Por lo tanto se puede concluir que el reconocimiento de expresiones faciales depende de forma paralela del resultado investigaciones de áreas tan diversas como la psicología, anatomía, la tecnología de hardware y software empleados, entre

muchas otras.

2.1. Estado del arte de los 10 últimos años

A continuación se muestra una revisión de los trabajos con mayor relevancia en el reconocimiento automático de las expresiones faciales, los cuales son resultado de los esfuerzos realizados en los últimos 10 años por parte de investigadores interesados en esta área de estudio.

Tabla 2.1: Tabla de análisis de trabajos relevantes en el reconocimiento automático de expresiones faciales concernientes a los 10 últimos años.

Referencia	Análisis e identificación	Clasificación	Base de datos	Precisión	Código	No./Tipo de expresión	Observaciones
Zhu et al. 2002[20]	Filtros de Kalman (momentos invariantes)	HMM	Propia	96.70 %	S/D	4/A/D/H/Su	Menor: A= 80 % Mayor: D/H/Su=100 %
Pardas and Bonafonte 2003[19]	Active Contour Algorithm (snake) con PCA y motion estimation	HMM	Cohn-Kanade	84 %	MPEG-4 FAPs	6/A/D/F/H/Sa/Su	Menor: Sa= 59.60 % Mayor: Su=100 %
Cohen et al. 2003[39]	Motion Units (MUs) usando Piecewise Bézier Volume Deformation (PBVD)	*NB-Gaussian, NB-Cauchy, TAN, Single HMM, Multi-level HMM	Propia y Cohn-Kanade	TAN 83 %	FACS	6/A/D/F/H/Sa/Su	TAN y base Propia: Menor: Sa= 61.26 % Mayor: Su=93.93 %
Abboud et al. 2004[26]	PCA	Distancia Euclidiana, Distancia de Mahalanobis	Propia	84 %	MPEG-4 FAPs	6/A/D/F/H/Sa/Su	Menor: Sa= 66.67 % Mayor: Su=100 %
Ma and Khorasani 2004[29]	2-D discrete cosine transform (DCT)	Feedforward Neural Networks	Propia	93.75 %	S/D	4/A/H/Sa/Su	Menor: Sa= 80 % Mayor: H/Su=100 %
Pantic et al. 2004[40]	Varios filtros de transformación	Basado en reglas de clasificación	Propia	86 %	FACS	4/A/H/Sa/Su	Hace uso de vista frontal y de perfil
Deng et al. 2005[41]	Local Gabor Filter Bank and PCA plus LDA	Distancia Euclidiana y Cityblock	JAFFE	97.33 %	FACS	4/A/H/Sa/Su	Hace énfasis en la extracción de características
Shan et al. 2005[24]	Local Binary Patterns (LBP) and Raw image data (IMG)	K-NN	Cohn-Kanade	90.79 %	FACS	6/A/D/F/H/Sa/Su	LBP muestra mayor asertividad con K=11
Aleksic and Katsaggelos 2005[24]	PCA	Multistream-HMM	Cohn-Kanade	93.66 %	MPEG-4 FAPs	7/A/D/F/H/Sa/Su/N	Menor: A= 70.6 % Mayor: Su=100 %
Yu and Bhanu 2006[42]	Local Gabor Filter Bank and Genetic programming	SVM	JAFFE	80.95 %	FACS	6/A/D/F/H/Sa/Su	Menor: F= 33.33 % Mayor: H/Su=100 %
Zheng et al. 2006[43]	Gabor Wavelet	KCCA	JAFFE	98.36 %	FACS	6/A/D/F/H/Sa/Su	Presenta resultados con diferentes kernels

Referencia	Análisis e identificación	Clasificación	Base de datos	Precisión	Código	No./Tipo de expresión	Observaciones
Sebe et al. 2007[28]	Piecewise Bézier volume deformation (PBVD)	NB, K-NN, MC4 boosting, ID3,TAN, SVM, Perceptron, etc.	Propia y Cohn-Kanade	K-NN en base propia 95 %	MU	4/D/H /N/Su	El UM (unidades de Movimiento) es similar una UA en FACS. Realiza pruebas con varios clasificadores.
Kotsia and Pitas 2007[17]	Candide facial grid, PFEF grid	Support Vector Machine (SVM) and Set of FAU Rules	Cohn-Kanade	PFEG y SVM 99.70 %	FACS	6/A/D/F/H/Sa/Su	Menor: A= 96.7 % Mayor: D/F/H/Sa/Su=100 %
Noh and et al. 2007[44]	Adaptive Facial Feature Region, Detection and Feature-Adaptive, Motion Energy Analysis Genetic	ID3 decision tree	JAFFE	75.40 %	FACS	6/A/D/F/H/Sa/Su	Menor: F= 63.3 % Mayor: Su=83.8 %
Cheng et al. 2007[23]	Analytical Hierarchy Process (AHP)	K-NN	JAFFE	85.20 %	MPEG-4 FAPs	5/A/H/Sa/Su/N	Menor: Sa= 76 % Mayor: H=94 %
Xiang et al. 2008[45]	FCM	fuzzy spatio-temporal (FST)	Cohn-Kanade	88.80 %	FACS	6/A/D/F/H/Sa/Su	Menor: Sa=77.8 % Mayor: H=95.7 %
Shih et al. 2008[16]	2D-LDA, LDA, ICA, PCA, 2D-PCA	SVM	JAFFE	95.71 %	FACS	7/A/D/F/H/Sa/Su/N	Menor: Sa=90.32 % Mayor: A/N/Su =100 %
Cheon and Kim. 2009[22]	AAM y DAF	Directed Hausdorff Distance (DHD) K-NN	POSTECH, FED06 y Cohn-Kanade	96.06 %	S/D	4/A/H/N/Su	Menor: A=90.31 % Mayor: N=100 %
Park and Kim. 2009[15]	Diferential-AAM	SVM	SFED2007	88.12 %	FACS	4/A/H/N/Su	Menor: Su=82.51 % Mayor: N=92.5 %
Shan et al. 2009[46]	LBP, Boosted-LBP y Gabor	SVM (linear, polynomial y RBF)	Cohn-Kanade, JAFFE y MMI	95.10 %	FACS	7/A/D/F/H/Sa/Su/N y 6 quitando N	Menor: Sa=74.7 % Mayor: D/H=97.5 %
Fu et al. 2011[13]	Log-Gabor Feature (IGF), PCA y ICA	SVM	CUN	91.47 % frontal	S/D	7/A/D/F/H/Sa/Su/N	Menor: Sa=80 % Mayor: A/H/Su=100 %
Sha et al. 2011[21]	GLF, PCA, SCF y FS	K-NN	BU-3DFE	83.50 %	S/D	6/A/D/F/H/Sa/Su	Menor: F=69.76 % Mayor: Su=95.41 %
Zavaschi et al. 2012[12]	Gabor filters(GF) y Local Binary Patterns (LBP)	SVM	JAFFE	96.20 %	FACS	7/A/D/F/H/Sa/Su/N	Menor: Sa=93.54 % Mayor: F/N Su=100 %

A= Anger, D=Disgust, H=Happiness, Su=Surprise, F=Fear, S=Sad, N=Neutral, S/D =Sin Dato

2.2. Proceso del reconocimiento de expresiones faciales

El reconocimiento automático de las expresiones faciales es considerado una extensión del problema del “*Reconocimiento Facial*” ya que la forma tradicional de realizar el reconocimiento de expresiones faciales comienza con la detección del rostro para después seguir los pasos 2, 3, 4 y 5 de la figura 2.1. Sin embargo también hay algoritmos que pueden ser capaces de omitir la detección del rostro, tal como se explica en la sección 2.3.2.

La figura 2.1 muestra de forma general la estructura básica para resolver el problema de reconocimiento automático de expresiones faciales, el cual se divide en dos fases generales el *Análisis e Identificación* y la *Clasificación* (véase en 2.1). Ambas fases se integran por un conjunto de técnicas y métodos que son utilizados para resolver problemas particulares de cada fase. Se hablará más descriptivamente de la fase de análisis e identificación, así como de los métodos típicamente empleados en esta fase en la sección 2.3, mientras que en la sección 2.4 serán descritos los métodos comúnmente utilizados para la solución del problema de clasificación de expresiones faciales.

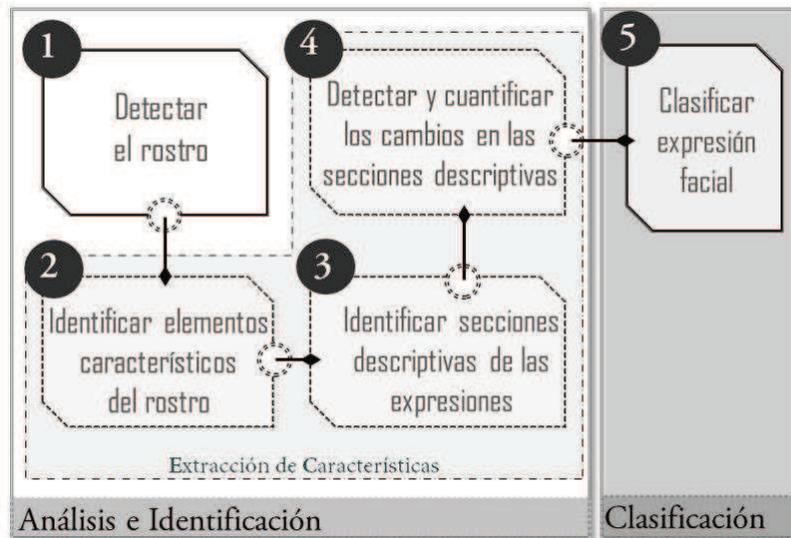


Figura 2.1: Proceso general del reconocimiento automático de expresiones faciales.

2.3. Análisis e Identificación

Esta fase se dedica a la obtención de datos (valores numéricos) a partir de imágenes, por lo que es todo un reto extraer dichos datos de la manera precisa, rápida y fidedigna. La fase de Análisis e Identificación persigue los siguientes objetivos:

- **Seleccionar.** Determinar las características que son realmente representativas y que faciliten la búsqueda y localización de un elemento, algunas tareas asignadas a la selección son:
 - Evaluar y priorizar los elementos que mejor caracterizan a un rostro (información útil en el paso 2 de la figura 2.2).

- Determinar el conjunto de contracciones musculares que mejor describen a cada expresión que se pretende reconocer y que a su vez permitan distinguir entre una expresión y otra.
- Determinar las secciones (puntos, líneas o áreas generalmente) en donde visiblemente se detecta un pronunciada deformación del rostro en el momento de realizar una gesticulación (información útil en los pasos 3 y 4 de la figura 2.2).

➤ **Búsqueda y Localización.** Obtener un conjunto selecto de información que sea útil en el reconocimiento de expresiones faciales, algunas tareas asignadas a la selección son:

- Buscar y localiza el rostro objetivo (paso 1 y 2 de la figura 2.2)
- Buscar y localizar las secciones descriptivas (determinadas durante el análisis) en el rostro objetivo (paso 3 de la figura 2.2).
- Detectar el momento en el que se realiza una expresión facial, esto se logra mediante la localización y medición del desplazamiento de las secciones descriptivas con respecto a una posición neutral de las mismas (paso 4 de la figura 2.2).

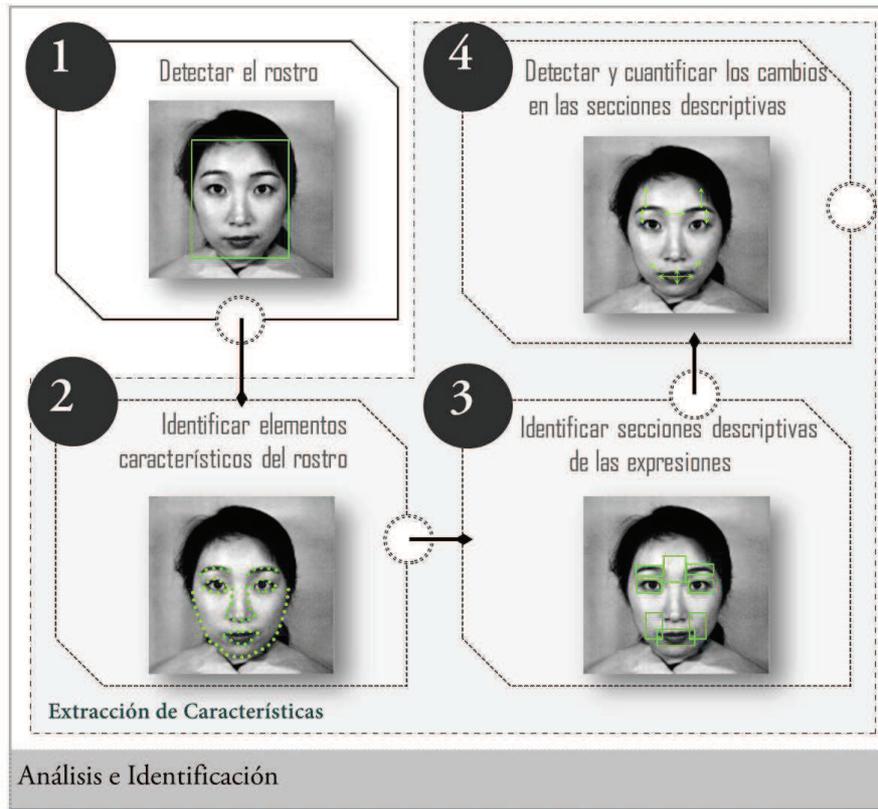


Figura 2.2: Esquema general de la primera fase del proceso de reconocimiento automático de expresiones faciales: **Análisis e Identificación.**

El análisis e identificación consta de cuatro pasos en su forma tradicional las cuales serán descritas a continuación (véase la figura 2.2).

Paso 1.- Detectar el rostro. La detección del rostro es una de las tareas fundamentales puesto que permite aislar el rostro del resto de la imagen, disminuyendo así el costo computacional en las tareas que se realicen posteriormente. El objetivo de este paso es detectar de forma rápida y confiable el rostro de un ser humano dentro de una imagen completa, para ello discrimina el espacio que alberga al rostro objetivo del espacio integrado por elementos que deben ser ignorados, este proceso constituye un verdadero reto.

Después de detectar la presencia de un rostro en la imagen, el siguiente objetivo es extraer la información de este, que incluye a los pasos 2 y 3 descritos a continuación.

Paso 2.- Identificar elementos característicos del rostro. Se realiza una selección, búsqueda y localización de elementos característicos del rostro humano, usualmente los elementos más utilizados son las cejas, los ojos, la nariz y la boca. Este paso nos sirve para realizar el paso 3, ya que considera la localización de los elementos característicos y los toma como punto de referencia para determinar la ubicación de las secciones descriptivas seleccionadas.

Paso 3.- Identificar secciones descriptivas de una expresión. Consiste en realizar la selección, búsqueda y localización de las secciones del rostro que proporcionen información relevante para el reconocimiento de una expresión facial determinada y que a su vez nos permitan distinguir entre una expresión y otra. Ahora bien, considerando que durante una expresión facial se puede notar visiblemente que ciertas secciones del rostro se deforman, así que se debe determinar y localizar las secciones del rostro que mejor describen a cada expresión facial con base en la intensidad de la deformación. Las secciones suelen ser representadas por puntos, líneas o áreas del rostro.

Paso 4.- Detectar y cuantificar los cambios en las secciones descriptivas. Tiene la finalidad de parametrizar las variaciones que sufren las secciones descriptivas durante una expresión facial. Ahora bien las expresiones faciales son generadas por las contracciones de los músculos faciales en intervalos de tiempo que pueden ir desde 250 ms hasta 5 s, lo que genera deformaciones temporales en las secciones descriptivas, lo cual trae como consecuencia cambios o variaciones que pueden ser medidas y cuantificadas. Algunos aspectos importantes a considerar en la parametrización son:

- Duración. Es el intervalo de tiempo en el que se presenta una expresión facial. Tradicionalmente el intervalo que comienza en una posición neutral o sin deformación (músculos faciales relajados), pasa por el momento en el que el gesto está en su máxima expresión y termina cuando el rostro regresa a la posición neutral.
- Intensidad. Siguiendo una escala (bajo, medio, alto, muy alto) se categoriza la intensidad de una expresión facial. La intensidad depende del grado de deformación que sufren los elementos característicos.
- Selección de parámetros. Existen esfuerzos de investigación muy importantes que ayudan en la selección de las secciones descriptivas del rostro que mejor representan una expresión facial, así como en la interpretación de las deformaciones de estos y su relación con su correspondiente expresión facial. Una de las publicaciones más citadas en la literatura son los trabajos de Ekman y Friesen [47, 7, 48], en su trabajo de investigación proponen un sistema de clasificación de acciones faciales llamado FACS (Facial Action Coding System), el cual es considerado uno de los sistemas más utilizados para la descripción de las expresiones faciales. FACS utiliza 44 unidades de acción (UA) las cuales permiten la descripción de una expresión facial.

2.3.1. Identificación facial y la extracción de características

2.3.1.1. Detección del rostro

Otros elementos a considerar en la identificación facial y la extracción de características que tienen que ver con la detección de rostros involucran aspectos biométricos, a saber:

- Sexo.** Ya que hay diferencias significativas entre el aspecto de un hombre y una mujer.
- Edad.** Es de importancia considerar el área promedio del rostro, así como el tamaño y distancias entre las características faciales con respecto a la edad del individuo, por ejemplo: la dimensión del rostro de un bebe, un niño y un adulto varían con respecto a la edad.
- Raza.** La clasificación de las razas humanas consideran las características intrínsecas que comparan un conjunto de personas y algunas de ellas tienen una relación directa con el aspecto del rostro, por ejemplo: las dimensiones de los elementos faciales y el color de piel.
- Vestimenta.** Es común el uso de aretes, sombreros, anteojos, etc.
- Otros.** El uso de bigote, peinados, ser calvo, etc.

2.3.1.2. Extracción de características.

Consiste en realizar la identificación y extracción de los elementos que caracterizan el rostro humano y también las secciones que representen un alto impacto en el proceso de clasificación de las expresiones faciales. Usualmente los elementos más identificados son las cejas, los ojos, la nariz, la boca, entrecejo, comisura de labios, entre otros.

Con la ayuda de algunos conceptos básicos de biometría del rostro humano se puede realizar la ubicación de las secciones específicas del rostro, considerándose algunos parámetros fijos tales como el número de elementos que conforman un rostro (dos ojos, una nariz y una boca); también se puede hacer uso de la relación paramétrica que los elementos del rostro tienen entre sí, es decir cada elemento mantiene posición relativa con respecto a los otros, e.g. las cejas están una arriba de cada ojo, y la nariz se encuentra entre y por debajo de los ojos, así como la boca se encuentra debajo de la nariz, entre otras consideraciones biométricas. Otra posibilidad sería considerar la forma general distintiva de cada uno de los elementos, el color, o incluso la textura.

2.3.2. Secuencia general de la fase de análisis e identificación

Se debe aclarar que la figura 2.1 presenta una secuencia general y tradicional de la fase de “análisis e identificación”; sin embargo, la secuencia así como la omisión de los pasos dependerá de las técnicas que sean empleadas para la extracción de características. A manera de ejemplo, considérese las diferencias entre las secuencias A y B descritas a continuación:

Secuencia A (identificación facial y después extracción de características). Siguiendo los pasos 1, 2, 3 y 4 de la figura 2.1. Es la secuencia tradicional en la que es de importancia realizar la identificación del rostro (paso 1) de manera precisa y confiable, de lo contrario los esfuerzos que realice el sistema en la extracción de características serán por demás absurdos.

Secuencia B (comenzando desde la extracción de características). Siguiendo los pasos 2, 3, 4 y 5 de la figura 2.1. En esta clase de identificación NO es necesario localizar primero el rostro (paso 1), ya que es posible identificar que se trata de un rostro con solo localizar los elementos característicos faciales, para ello se puede hacer uso de la biometría del rostro en donde se consideran parámetros fijos del rostro humano tales como:

- Número de elementos que conforman una cara (dos ojos, una nariz y una boca).
- La posición que cada elemento característico tienen con respecto a los otros, es decir las cejas están una arriba de cada ojo, la nariz se encuentra entre y por debajo de los ojos, la boca se ubica debajo de la nariz, entre otros.

- La forma, textura o color de los elementos. Los ojos son ovalados, la nariz tiende a ser triangular, las cejas son dos líneas, la boca es rosada, entre otros.

De la secuencia B se puede concluir que, gracias a las características invariables del rostro, es posible determinar la localización de los elementos característicos (paso 2) y después seguir con los pasos 3, 4 y 5 de la figura 2.1, sin tener que haber realizado la detección facial (paso 1)

2.3.3. Problemas básicos presentes en la fase de análisis e identificación

Son muchos los aspectos que se deben considerar en el momento de localizar un rostro humano o bien sus elementos característicos, ya que en una imagen o video es necesario diferenciar entre el rostro objetivo y el escenario. Aquí se presentan algunos de los problemas básicos propios de esta fase.

2.3.3.1. Número de caras

No importando si se trata de identificar un rostro en una imagen fija o en un video, se debe definir el alcance de la identificación, es decir delimitar el número de personas a las que se les desea realizar el reconocimiento de expresiones faciales, ya que de ello depende el diseño de un algoritmo que cumpla con el objetivo deseado. Es muy importante que el algoritmo sea capaz de detectar de forma precisa el rostro objetivo, ya que de ello depende el éxito posterior de la búsqueda y extracción de las características específicas (solo en caso de seguir la secuencia B explicada en 2.3.2).

2.3.3.2. Localización de elementos de interés

Para abordar este problema se supondrá que se ha solucionado el problema de la detección del rostro dentro de una imagen, ahora se debe identificar los elementos básicos que son parte de un rostro, y son utilizados para identificar de mejor manera las expresiones faciales.

Tomando en cuenta las recomendaciones del Dr. Ekman[9, 10] los elementos del rostro que tienen un mayor impacto para la clasificación de los seis estados de ánimo generales son la boca, los ojos, las cejas y la nariz, sin embargo en esta etapa se debe considerar que el ruido del entorno suele ser difícil de eliminar.

2.3.3.3. Pose de captura

Este problema refiere principalmente a la variación del ángulo y la distancia a la que se observa una imagen dada dentro de las capturas de la cámara, lo cual interviene directamente en la variación de escalas y de ejes de rotación, además de la posibilidad de extraer datos parciales de un rostro, debido a que el ángulo de rotación del rostro puede llegar a ocasionar que parte del rostro no sea visible, esto es un gran problema sobretodo en modelos 2D. Y para modelos en 3D afecta cuando se determina la dimensión de los elementos del rostro, e.g. la distancia entre la nariz y un ojo no será la misma si se considera una toma de vista directa a la cámara, a una toma en la que la cámara tiene un ángulo de inclinación.

Por lo que puede intuirse que será relativamente más sencillo, si se cuenta con una toma frontal a la cara, a diferencia de la dificultad que representa la identificación en una toma donde el rostro esta de perfil. Es por ello que cuando se decide trabajar en el reconocimiento de expresiones faciales suele ser un factor a considerar que la base de datos que se seleccione para trabajar este integrada por caras

solo en toma frontal, en caso contrario es necesario seleccionar un método que se acople a los ángulos de rotación del rostro, o bien algún procesamiento previo para la normalización.

2.3.3.4. Iluminación

Es un problema que tiene que ver con la variación de la luz en el momento de tomas de video o fotografía. La falta o exceso de iluminación en una imagen puede influir directamente en la localización de elementos, por lo que suele ser necesaria una fase de pre-procesamiento que nos permita normalizar las imágenes utilizando métodos de filtrado, por ejemplo Gabor Wavelets. Otro problema de iluminación es la reflexión de la luz en los ojos.

2.3.3.5. Seguimiento de rostro en video

Este problema representa una consideración extra cuando se debe trabajar con video, ya que en una imagen fija el proceso de detección se realizara una sola vez, mientras que en el video la persona tiene movimiento continuo, por lo que se debe seguir su rostro, no obstante la persona puede realizar una serie de gesticulaciones en cuestión de segundos, de modo que el tiempo de análisis de la expresión facial depende de la correcta detección del rostro, ya que si se pierde esa ubicación el proceso de identificación de elementos no tendrá mayor relevancia.

La estimación de la posición y orientación de la cabeza supone la localización del rostro en diferentes planos, por lo que es un factor que debe considerarse en el diseño del algoritmo durante la fase de detección del rostro.

El número de grados de libertad con los cuales trabajan los algoritmos de seguimiento también debe ser un aspecto a determinarse durante la etapa de alcance y delimitación total del sistema. Una posible clasificación de estos es la siguiente[49]:

- *Seguimiento en 2D y $2\frac{1}{2}D$.* Realiza un seguimiento del rostro en un plano de (x, y) , en donde hay poca flexibilidad en la detección de movimientos de giro, por lo que se restringe al usuario del sistema, la utilización de este en modo rostro-frontal.
- *Seguimiento en 3D.* Realiza un seguimiento del rostro en un plano (x, y, z) , estimándose los seis grados de libertad de la cabeza y orientación, es de notarse que el procesamiento es más robusto pero la flexibilidad que ofrece es considerablemente mejor que el trabajo en $2D$.

2.3.4. Clasificación de las técnicas empleadas en la fase de análisis e identificación

El creciente interés despertado en el área de visión artificial, ha dado como resultado una gran cantidad de algoritmos empleados en la fase de “Análisis e Identificación” la cual incluye el problema de la detección del rostro y la extracción de características (véase la figura 2.2). Y justamente esa extensa variedad de algoritmos dificulta realizar una clasificación flexible que incluya a los algoritmos existentes y también incluya a futuros algoritmos.

En los “surveys” dedicados al problema de detección facial se pueden encontrar 2 clasificaciones relevantes, que son comúnmente referidas en el problema de detección y localización del rostro:

2.3.4.1. Clasificación propuesta por Hjelmás y Low [1].

Publicada en 2001 (véase la figura 2.3) esta integrada por dos grandes grupos, el primero es el “enfoque basado en rasgos faciales” reúne a todos los algoritmos que buscan detectar el rostro mediante características que definen un rostro humano tales como boca, ojos, nariz, color de piel, forma, etc. El segundo es el “enfoque basado en la imagen” esta integrado por los algoritmos que hacen uso de toda la imagen como un conjunto, el objetivo es establecer si hay una cara o no. A continuación se presenta una descripción general de cada uno de los rubros de la clasificación realizada por Hjelmás y Low[1].

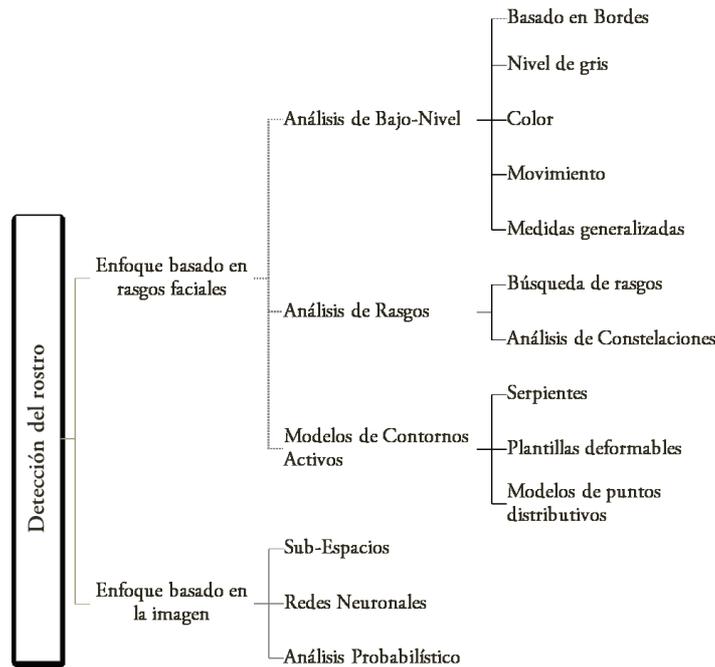


Figura 2.3: Clasificación de métodos de detección de rostros según Hjelmás y Low [1]

Enfoque basado en rasgos faciales Conjunta todas aquellas técnicas que hacen uso de los rasgos que definen una cara humana así como de los parámetros que se pueden extraer a partir de la relación y variación entre ellos, algunos rasgos comunes son ojos, cejas, labios, boca, mentón, líneas de contorno, entre otros.

➤ **ANÁLISIS DE BAJO NIVEL.** Son algoritmos que trabajan a nivel de pixel.

- **Bordes.** Buscan detectar las líneas que componen los bordes de una cara y las utiliza para detectar los rasgos faciales.
- **Niveles de gris.** Hace uso de la variación de tonalidades grises que se presentan de forma predefinida en el rostro, por ejemplo: es considerar que las cejas, pupilas y labios son más oscuras que las regiones a su alrededor, entonces algunos algoritmos localizan los mínimos locales dentro de una imagen segmentada (la imagen puede ser tratada previamente con variaciones de contraste).

- **Color.** Utiliza el conocimiento que se tiene sobre los rangos de color que ciertas secciones del rostro suelen tener. Algunos rasgos que mantienen cierto rango de manera constante son la piel, los labios, ojos, entre otros.
- **Movimiento.** Se concentra en el tratamiento de secuencias de video, las técnicas buscan analizar los cambios entre las secuencias y parametrizar las diferencias.
- **Medidas Generalizadas.** Son algoritmos que hacen uso de la simetría del rostro humano y de las dimensiones generalizadas entre los componentes de la cara humana.

∧ ANÁLISIS DE RASGOS. Hacen uso de la geometría de la cara.

- **Búsqueda de rasgos.** El objetivo de estos algoritmos es buscar rasgos que destacan del resto, lo que hace que su localización sea relativamente más sencilla, después se hace uso de conocimientos antropométricos de la geometría de la cara, para determinar la localización de los rasgos menos notorios.
- **Análisis de Constelaciones.** Se basa en el uso de modelos probabilísticos aplicados en la búsqueda de la posición espacial de los rasgos faciales, los cuales intentan buscar patrones que se asemejen a una cara humana.

∧ MODELOS DE CONTORNOS ACTIVOS. Los algoritmos intentan representar la imagen en rasgos de alto nivel y después interactuar con rasgos locales de la imagen (ojos, brillo), para finalmente deformarla poco a poco hasta adaptarla a la forma de los rasgos.

- **Serpientes.** Usualmente se utilizan para localizar el contorno de la cara. Primero se inicializa una serpiente en las proximidades de la cara y después se deforma hasta obtener una figura lo más similar al contorno de la cara.
- **Plantillas Deformables.** Es una variante del uso de serpientes, ya que una vez localizado el contorno del rostro, el siguiente paso es localizar rasgos del rostro utilizando la técnica de serpientes.
- **Modelos de puntos distributivos.** Se trata de una discretización estadística parametrizada de las formas de la cara representada por un conjunto de puntos.

Enfoque basado en la Imagen A diferencia del enfoque basado en rasgos faciales, este enfoque estudia a la imagen como un conjunto y no solo partes de la misma. Para ello se realiza una representación de la imagen, la cual es parte de la fase de entrenamiento de algoritmos que son capaces de aprender y/o analizar, dicha información.

- ∧ SUB-ESPACIOS. Se fundamentan en los trabajos realizados por Sirovich y Kirby [50] que son retomados tiempo después por Matthew Turk y Alex Pentland[2]. En donde el rostro esta representado en un sub-espacio del espacio que representa a la imagen original, reduciendo de esta forma el consumo computacional de procesamiento.
- ∧ REDES NEURONALES. Las redes neuronales forman parte de los métodos de reconocimiento de patrones (clasificadores), son utilizadas para reconocer si en una imagen hay un rostro o bien rasgos particulares del mismo. En las redes neuronales existen dos pasos generales, el primero es la fase de entrenamiento (aprende de una base de datos) y después esta la fase de Clasificación (determina si hay o no un rostro o en su caso los rasgos faciales).
- ∧ ANÁLISIS PROBABILÍSTICO. Como su nombre lo indica hace uso de cálculos de probabilidad y estadísticos. Algunos de los métodos más empleados son las Máquinas de Soporte Vectorial (SVM) y reglas de decisión de Bayes.

2.3.4.2. Clasificación propuesta por Yang et. al.

La segunda clasificación fue propuesta en 2002 por Yang et al. [3] y puede observarse en la figura 2.4.

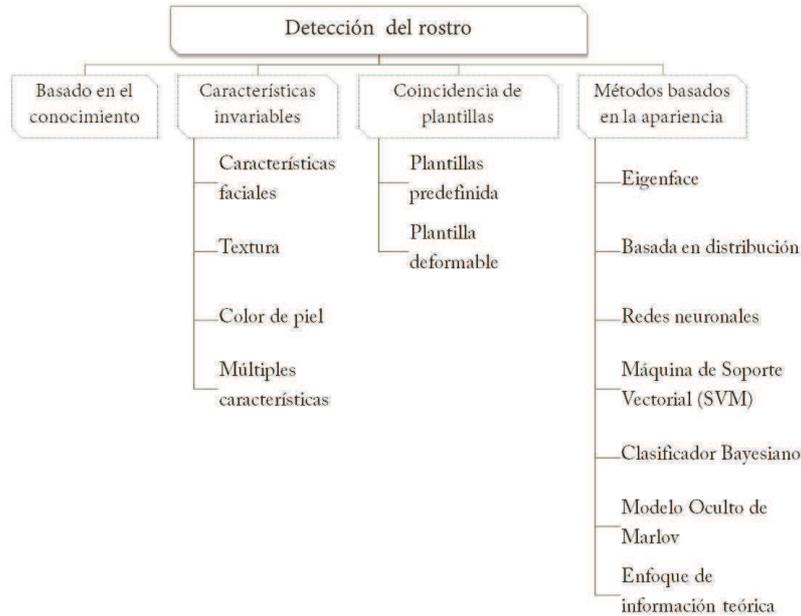


Figura 2.4: Clasificación de los métodos de detección de rostros según Yang et al. [3]

1. **Métodos basados en conocimiento.** Hacen uso de los parámetros que previamente se conocen sobre la forma de un cuerpo humano. Este conocimiento se usa para formar reglas que permitan distinguir entre rostros y no-rostros, considerando características descriptivas como la forma, el tamaño, la textura y otros rasgos faciales como los ojos, la nariz, la barbilla, las cejas; así como las relaciones entre ellos (posiciones relativas y distancias)
2. **Métodos basados en características invariantes.** Buscan las características del rostro que son persistentes bajo diferentes condiciones de iluminación y pose como lo es el color de la piel y la textura.
3. **Métodos de correspondencia de plantillas.** Utilizan varios patrones del rostro que describen completa o parcialmente sus características. Para detectar un rostro se calcula la correlación entre una imagen de entrada y los patrones con los que se cuenta.
4. **Métodos basados en la apariencia.** Se crean patrones o modelos a partir de un conjunto de imágenes de entrenamiento tomando los valores de sus píxeles.

2.3.5. Base de Datos

Una base de datos es en sí misma la materia prima con la que aprende determinado método dedicado al reconocimiento de expresiones faciales, cualquiera que este sea. Las bases de datos, dedicadas a la tarea de la clasificación automática de expresiones faciales, están integradas por imágenes fijas o bien en algún formato de video (secuencias de imágenes).

A continuación se presentan algunas de las bases de datos típicamente referidas en los artículos concernientes a las investigaciones del presente tema.

Tabla 2.2: Tabla Comparativa de los bancos de datos más citados en la literatura de reconocimiento de expresiones faciales [51, 52] .

Banco de Datos	N. Imágenes	N. Sujetos	Expresión/ Pose/ Iluminación	Color	Resolución	Año	Observaciones
FERET [53, 54, 55]	14,126*	1,010	2/9-20/2	Gris	256x384	1997	Son 1,564 conjuntos. (que contienen un total de 14,126 imágenes), de los cuales 365 sets son repeticiones de tomas en diferentes días.
JAFFE [56, 57]	213*	10	7/1/1	Gris	256x256	1998	Solo mujeres japonesas.
AR DataBase [58]	3,288*	116	4/1/4	Color	768x567	1998	70 hombres y 56 mujeres. Presenta tomas con parte del rostro parcialmente oculto (oclusión).
Cohn-Kanade [59, 60]	486**	97	6/1/1	Gris	640x490	2000	Versión 1, cada secuencia se inicia desde posición neutral, basada en FACs[11, 10].
CMU-PIE [61, 62]	41,368**	68	4/13/43	Color	640x480	2000	La expresión es captada por 13 cámaras.
CAS-PEAL DataBase[63, 64]	30,900*	1,040	6/21/9-15	Color	360x480	2003	595 hombres y 445 mujeres, se usan anteojos y gorras.
Korean Face DataBase (KFDB)[65]	52,000*	1000	5/7/16	Color	640x480	2003	Las tomas tienen 3 estilos: natural (Sin accesorios), con banda en el cabello y anteojos. Son 52 imágenes por individuo. Anexa localización de un máximo de 26 puntos faciales por imagen.
BU-3D FE[66]	2,500†	100	7/2/4	Color	1040x1329	2006	56 % mujeres y 44 hombres, entre los 18 y 70 años, con variedad étnica y racial. Son modelos animados en 3D.
BU-4D FE[67]	606††	101	6/1/1	Color	1040x1329	2008	Se le agrega la dimensión del tiempo para crear secuencias, además están en alta definición.
CMU Multi-PIE [68, 69]	755,370**	337	6/15/19	Color	3072x2048	2008	Imágenes de alta resolución. El sujeto es captado por 15 cámaras, 13 de ellas están al frente a la altura de la cabeza con 15° entre ellas y las 2 restantes a una altura superior.
Radboud Face DataBase[70]	8,040*	67	8/5/1	Color	1024x681	2010	Incluye niños, basada en FACs. Los ángulos de pose son: 180°,135°,90°, 45° y 0°.
Cohn- Kanade+ [71, 60]	593**	123	6/1/1	Gris	640x490	2010	En la versión 2 se incrementa el número de secuencias en un 22 % y de sujetos en un 27 %.

* Imágenes, **Secuencias de Imágenes, †Imágenes de modelos 3D, ††Secuencias de modelos 3D

Existen bases de datos gratuitas y comerciales, además estas bases de datos son de apoyo en la evaluación de los algoritmos diseñados por investigadores de todo el mundo para el reconocimiento de expresiones faciales.

Gracias a las bases de datos un investigador puede ahorrarse el arduo trabajo que implica la elaboración de una base de datos propia, en la que debe contarse con un estudio fotográfico o de filmación, también se necesitan actores que sean capaces de imitar las expresiones faciales, entre muchas otras consideraciones. Razón por la cual el hecho de contar con un conjunto de imágenes o secuencia de imágenes es de gran apoyo ya que permiten probar los algoritmos diseñados en la fase de *Análisis e Identificación* (etapa en la que se obtienen los datos numéricos) y también en la fase de *Clasificación* (Véase la figura 2.1). Se debe recordar que la eficacia y eficiencia de los métodos de clasificación depende de la cantidad, el tipo y la calidad de los datos de entrada, por lo que se debe seleccionar de forma cuidadosa la información de entrada empleada para el aprendizaje de la fase de clasificación. Otro aspecto importante a considerar es que las bases de datos nos sirven como referencia comparativa, ya que permite medir y comparar la asertividad de un determinado método con respecto a otros métodos, los cuales hacen uso de la misma base de datos.

2.3.6. Descripción de las técnicas empleadas en el Análisis e Identificación

Es de importancia considerar que las técnicas a continuación descritas no son las únicas que existen para solucionar el problema de la identificación y localización de elementos característicos de expresiones faciales, ya que cada día surgen nuevos algoritmos, además de que existen variantes híbridos o modificaciones de las técnicas originales. No obstante en la siguiente sección se presenta una descripción general de las técnicas que típicamente son utilizadas en la solución de la fase de Análisis e Identificación integrada por los problemas de detección del rostro y extracción de características.

Ahora bien como ya se ha expuesto la detección de rostros suele ser el primer paso a superar, por ello es fundamental conocer los principales métodos enfocados en la detección de rostros, siendo de vital importancia la selección de un algoritmo eficiente, ya que el éxito de la extracción de características depende de la correcta detección del rostro, así que es muy importante utilizar un algoritmo que cumpla de mejor manera con las necesidades del reconocedor de expresiones faciales.

2.3.6.1. Seguidor basado en plantillas

Fisher introdujo un sistema automático que hacía uso de plantillas (máscaras) de rasgos de partes de la cara de modo que se tuviera un mapa global.

Los métodos basados en plantillas utilizan imágenes predefinidas de las regiones faciales para realizar su localización. Los trabajos más relevantes referentes a estos métodos son los propuestos de Hallinan[72] y Yuille[73].

El método de uso de plantillas para el reconocimiento de rostros es muy sencillo, el primer paso es extraer regiones de las imágenes que abarcan ciertas secciones del rostro: ojos, nariz, boca, etc. Y las cuales son almacenadas en una base de datos, es decir el rostro es dividido en secciones que se consideran necesarias en el seguimiento de la cara de una persona, después se separan y se guardan en un vector como parte de una imagen completa.

La idea principal del método de plantillas consiste en encontrar patrones a partir de una imagen, es evidente que dentro de la búsqueda se debe dejar un margen de error para permitir al sistema la adaptación, esta técnica también es llamada en la literatura “*Template Matching*”.

Este método tiene un sinnúmero de aplicaciones en el reconocimiento general de objetos en la visión artificial, existen diferentes algoritmos para efectuar el template matching, estos varían dependiendo de su rapidez de filtrado y otros se encuentran basados en el óptimo reconocimiento del patrón.

El método de identificación básico consiste en un barrido de todos los píxeles de la imagen original; de modo que se mueve una plantilla previamente seleccionada, realizando un recorrido en todas las posibles posiciones a lo largo y ancho de la imagen a analizar, a su vez se realiza el procesamiento para determinar el grado de similitud que existe en determinada porción de la imagen en ese instante, hasta llegar a un grado de semejanza aceptable el cual indica que se ha localizado el template dentro de la imagen analizada.

Una de las formas más simples de realizar la búsqueda es la siguiente:

Para empezar es necesario tener una imagen original T de donde se extrae una plantilla t con un tamaño $maxn$. La selección de la plantilla representa un segmento de la imagen total, e.g. la boca, un ojo, una ceja, entre otros elementos característicos del rostro.

Ahora bien se tiene también la imagen de búsqueda representada por $S(x, y)$, donde (x, y) son las coordenadas de cada pixel de la imagen de búsqueda, y la plantilla $t(x_t, y_t)$, en donde (x_t, y_t) representan las coordenadas de cada pixel en la plantilla.

El siguiente paso es realizar el barrido de la plantilla, simplemente moviendo el punto de origen de la plantilla $t(x_t, y_t)$, encima de cada (x, y) de la imagen de búsqueda S , haciendo una búsqueda de similitudes a su vez se realiza el cálculo de la suma de productos entre los coeficientes en $S(x, y)$ y $t(x_t, y_t)$, que son extraídos durante todo el recorrido necesario para abarcar todo lo largo y ancho de la imagen de búsqueda S .

Ahora queda una pregunta por contestar ¿Cómo se realiza la similitud entre los píxeles?, pues bien actualmente hay muchas formas una de ellas es por medio de la comparación de intensidades de cada pixel, utilizando *SAD* (suma de diferencias absolutas), esto es que un pixel de la imagen de búsqueda S tiene las coordenadas (x_s, y_s) y tiene la intensidad $I_s(x_s, y_s)$, entonces hay un pixel del template t con coordenadas (x_t, y_t) y una intensidad $I_t(x_t, y_t)$. Teniendo estos datos se puede calcular la diferencia absoluta de la intensidad de los píxeles, la cual es definida por $Diff(x_s, y_s, x_t, y_t) = |I_s(x_s, y_s) - I_t(x_t, y_t)|$.

$$SAD(x, y) = \sum_{i=0}^{T_{Fil}} \sum_{j=0}^{T_{Col}} Diff(x + i, y + j, i, j) \tag{2.1}$$

Mientras se recorre la imagen S , se toma la medida *SAD* de la siguiente forma:

$$\sum_{x=0}^{S_{Fil}} \sum_{y=0}^{S_{Col}} SAD(x, y) \tag{2.2}$$

Donde S_{Fil} y S_{Col} se puede observar que son la búsqueda entre filas y columnas de la imagen de búsqueda S , mientras que t_{Fil} y t_{Col} son las filas y columnas de la plantilla.

Éste es un método relativamente fácil de implementar, no obstante su ejecución es lenta aun procesando imágenes en escala de grises y el tiempo se incrementa en imágenes que tiene buena resolución.

Existen formas de disminuir la complejidad computacional del método de plantillas original, esto es mediante el uso de filtrado de dominio de los espacios, generalmente se basa en el método de convolución[74], un ejemplo lo presentan Viola y Jones en los denominados “Haar-like features”[4], el cual es ampliamente utilizado, este método se trata de un operador definido en una subregión rectangular de la plantilla, donde una parte de los píxeles se suman y los otros se restan. Si el valor del filtro está por encima de cierto umbral establecido previamente, este es aceptado o es rechazado. De esta forma se puede distinguir una forma de otra, son capaces de hacer eso gracias a que realizan una codificación de diferencia de intensidades en la imagen, generando características significativas para la identificación de contornos, mediante contrastes entre regiones.

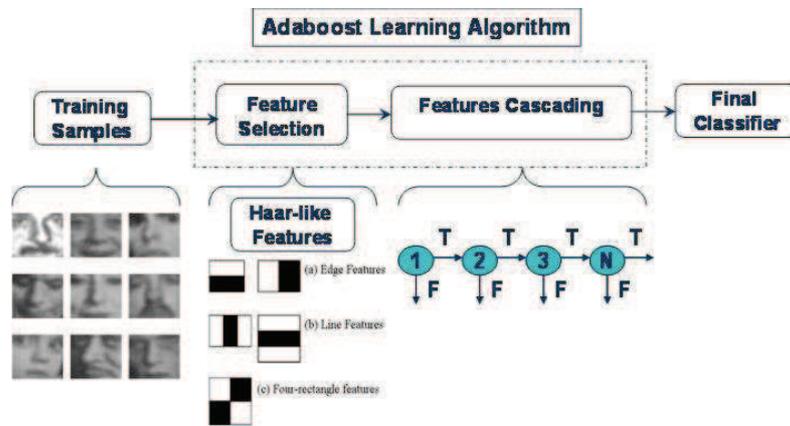


Figura 2.5: Arquitectura para la detección del rostro por Paul Viola & Michael Jones[4]

2.3.6.2. Método de PCA (Análisis de Componentes Principales)

Es una técnica cuyo objetivo principal es la reducción de la dimensionalidad de los datos de entrada. El PCA es un método de reducción dimensional en el que se busca extraer a partir de un conjunto de imágenes de entrenamiento un sub-espacio cuya base maximice la varianza del espacio original. El PCA una de las técnicas más utilizadas para la detección de rostros, ya que realiza una identificación rápida, sencilla y con resultados aceptables, sin embargo el método muestra serias deficiencias para cambios de iluminación y contrastantes.

Por ello se dice que el PCA pertenece a los métodos de proyección sobre sub-espacios, los espacios vectoriales definidos en los sub-espacios son vectores ortogonales propios de una matriz de covarianza calculada a partir de los vectores originales. En el PCA se suelen utilizar los vectores denominados «Eigenface» los cuales sirven para reducir la cantidad de procesamiento necesario y mejora de precisión en los resultados, lo cual ahorra costo computacional durante la identificación.

El término de «Eigenface» en el reconocimiento facial fue acuñado con Lawrence Sirovich y Robert M. Kirby[50, 75], en su trabajo de investigación realizaron un algoritmo en donde se hacía uso de los Eigenfaces como base del Análisis de Componentes Principales PCA. El algoritmo está integrado por un conjunto de imágenes de rostros almacenados en una base de datos, después se realizan los cálculos para determinar el sistema de coordenadas pertinente, el propósito es comprimir las imágenes a partir

de la obtención de un sub-conjunto de coordenadas que son ponderadas asignándole un peso, por lo tanto cada cara o rasgo facial (nariz, ojos, boca, etc.) tiene asignado un peso. De tal manera los rostros pueden ser reconstruidos mediante cálculos ponderados de una pequeña pero sustancial colección de características singulares de cada rostro. El principal problema de este método es encontrar los vectores ortogonales (eigenface) que mejor representen a una imagen. El cálculo de eigenface fue presentado por Matthew Turk y Alex Pentland[2].

Este método se encuentra íntimamente relacionado con los Eigenface¹ los cuales son descritos más detalladamente en la sección 2.3.6.2.

Algoritmo de PCA (Análisis de Componentes Principales)

Los trabajos de L. Sirovich y M. Kirby[50] en 1987, así como de Matthew Turk y Alex Pentland[2] en 1991, presentan un algoritmo del Análisis de Componentes Principales.

Algoritmo 2.1 Algoritmo básico de PCA (Análisis de Componentes Principales)

1. *Preparar el set de entrenamiento.* Se debe contar con un conjunto de imágenes de rostros, las cuales deben cumplir con dimensiones en píxeles de $w \times h$, donde w es el ancho y h el alto. Las imágenes no deben ser estrictamente cuadradas, pero si deben tener las mismas dimensiones.
 2. *Calcular los Eigenface.* Se calculan los *eigenfaces* del set de entrenamiento M , la elección se hace basados en los datos que describan de mejor manera el conjunto de datos de entrenamiento. Recordando que este eigenface puede ser recalculado, sin embargo debe tenerse en cuenta que eso implicaría un costo computacional alto.
 3. *Calcular los pesos.* Los cuales serán calculados a partir de la imagen original y de los *Eigenface* calculados en el paso anterior, proyectando la imagen de entrada sobre cada *Eigenface*.
 4. *Umbral.* Determinar si la imagen es un rostro, así como saber si hay similitud con el conjunto de entrenamientos o no. Se hace determinando la cercanía dentro del espectro del espacio-face.
 5. *Determinación de conocimiento.* Si se trata de un rostro, se debe clasificar el peso del patrón para saber si la persona es conocida o no. De tal manera que los valores de las distancias o bien diferencias entre los vectores son lo que determinan la similitud o disimilitud entre los rostros, mientras que la matriz de covarianza y los vectores característicos son calculados considerando las imágenes entrada previamente almacenadas, finalmente el umbral debe ser adaptativo con respecto a la distancia entre las caras.
-

Mapas autoorganizados de Kohonen

Dentro de las primeras aplicaciones útiles de los vectores en el reconocimiento del rostro destacan los dos trabajos realizados por Teuvo Kohonen[76] “Learning Vector Quantization (LVQ)” y “self-organizing feature maps (SOMs)”, su investigación esta motivada por la idea de que un conjunto de información asignada a un conocimiento en particular, es asimilado por el cerebro como un mapa de características organizado geoméricamente dentro del mismo.

¹El término eigen se introdujo por primera vez en este contexto por David Hilbert en 1904, significa característico o individual, en alemán. Los eigen vectors son los vectores propios

Los *SOM's* (*Self-Organizing Map*) son mapas autoorganizados que hacen uso de los modelos de redes neuronales. Los mapas autoorganizados de Kohonen tienen dos capas: la capa de entrada y capa de competición; en donde la capa de entrada (formada por N neuronas, una por cada variable de entrada) tiene dos funciones la primera es recibir y la segunda es la de transmitir la información procedente del exterior a la capa de competición. La capa de competición (formada por M neuronas) es la encargada de procesar la información y formar el mapa de rasgos, esto se hace gracias que cada célula produce una salida cuando se compara la entrada y sus pesos, de tal suerte que la célula que produzca una salida más pequeña, será la célula ganadora; de forma general las neuronas de esta última capa están organizadas en un espacio bidimensional.

Algoritmo 2.2 Algoritmo básico de *SOM's* (*Self-Organizing Map*)

1. **Inicializar pesos.** Asignar a los pesos valores pequeños aleatorios.
 2. **Presentar una nueva entrada.** El conjunto de imágenes son empleadas para el aprendizaje de la red neuronal presentándose cíclicamente hasta llegar a la convergencia de la red, actualizar α .
 3. **Propagar el patrón de entrada hasta la capa de competición.** Obtener los valores de salida de las células de dicha capa.
 4. **Seleccionar la célula ganadora C cuya salida sea menor.**
 5. **Actualizar.** Se actualizan las conexiones entre la capa de entrada y la célula C , así como las de su vecindad, según su grado de vecindad.
 6. Si α esta por encima de cierto umbral volver al paso 2, en caso contrario el *SOM* ha aprendido.
-

Método de Eigenface

Los Eigenface son un conjunto de datos representados por vectores, los cuales son empleados típicamente en problemas de visión artificial y con mayor frecuencia en el problema de reconocimiento facial. Los Eigenface sirven para reducir el tiempo de procesamiento y mejorar la precisión en los resultados, lo cual disminuye el costo computacional durante la identificación del rostro. Los vectores propios son denominados eigenvalores o autovectores, en donde cada uno posee valores escalares conocidos como eigenvalores, ya que el sistema interpreta cada imagen como un conjunto bidimensional de patrones brillantes y oscuros.

Algunos autores se refieren al *PCA* y a los *Eigenfaces* de forma indistinta, lo cual no es del todo correcto ya que el *PCA* es un método general de análisis de datos, por otra parte *Eigenfaces* es un método de reconocimiento que utiliza *PCA* con alguna variación.

Matthew Turk y Alex Pentland[2] realizaron un sistema capaz de reconocer una cara en tiempo casi real, basándose en la idea de que los rostros humanos se encuentran en posición vertical, lo que brinda una serie de características dimensionales constantes.

Calcular los Eigenface

Supóngase una imagen de un rostro I el cual está representado por un sistema de coordenadas bidimensional (x, y) , también se tiene un arreglo de bits de $N \times N$ que son la cantidad de bits que integran la imagen, donde cada bit está caracterizado por un valor de intensidad, de esa forma se tiene un vector bidimensional con dimensión N^2 . Sin embargo procesar una imagen completa representaría un costo

computacional alto. Por lo que el objetivo es lograr que un rostro se represente en un *sub-espacio*, pero con una dimensión más pequeña.

Algoritmo 2.3 Cálculo de eigenface presentado por Matthew Turk y Alex Pentland[2]

1. **Tener un conjunto de entrenamiento.** El conjunto de entrenamiento es representado por los vectores $\Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_m$.
2. **Calcular la media.** Se debe calcular el promedio de todas las imágenes que pertenecen al set de entrenamiento.

$$\psi = \frac{1}{M} \sum_{i=0}^M \Gamma_i \tag{2.3}$$

3. **Substraer la media.** Este paso tiene el propósito de centrar la imagen y se hace restando a cada imagen (Γ_i) la imagen media o vector promedio (ψ).

$$\phi_n = \Gamma_n - \psi \tag{2.4}$$

4. **Valores y vectores propios.** Los vectores son muy grandes, por lo que hay que determinar las componentes principales, se hace buscando un conjunto de vectores ortonormales v_k , que mejor describan la distribución de los datos en la imagen original, el k_{th} vector, v_k , se eligió de tal manera que:

$$\lambda_k = \frac{1}{M} \sum_{n=1}^M (v_k^\tau \phi_n)^2 \tag{2.5}$$

Es un máximo sujeto a:

$$v_l^\tau v_k = \delta_{lk} = \begin{cases} 1, & \text{if } l = k \\ 0, & \text{de otra forma} \end{cases}$$

v_k es vector y λ_k es el valor propio de la matriz de M .

5. **Determinar la matriz de covarianza.** Ahora bien cada imagen puede verse como un vector N -dimensional, donde $N = wxh$, después se hace una matriz M , en donde cada columna está formada por el vector N , los cuales son vectores que describen la distribución de características de cada una de las imágenes de entrenamiento, la N es determinada de la siguiente forma:

$$C = \frac{1}{M} \sum_{n=1}^M \phi_n \phi_n^\tau \tag{2.6}$$

$$C = AA^\tau \tag{2.7}$$

Donde la matriz $A = \phi_1 \phi_2 \phi_3 \dots \phi_n$, es una matriz normalizada tiene una dimensión de $dXd(N^2x N^2)$, lo que hace de extraer sus vectores y valores propios un proceso computacionalmente muy costoso, incluso para imágenes de tamaño normal.

Esto puede ser solucionado considerando que si el número de puntos en el espacio de imágenes es menor que la dimensión del espacio $m < N^2$, habría como mucho $m - 1$ vectores propios significativos. Puede resolverse el problema tomando apropiadamente combinaciones lineales de las imágenes. Si en lugar de calcular AA^τ se consideran v_i como los vectores propios de $A^\tau A$ (y α_i sus valores propios) resultaría que:

$$A^\tau A v_i = \alpha_i v_i \tag{2.8}$$

Premultiplicado a ambos lados por A , se tiene:

$$AA^T Av_i = \alpha_i Av_i \quad (2.9)$$

Donde Av_i son los vectores propios de $S = AA^T$

Entonces a partir de éste análisis, se construye la matriz $A^T A$ de dimensión $m \times m$ y se encuentran los m vectores propios.

2.3.6.3. Método de LDA (Análisis Discriminante Lineal)

La idea básica de LDA [77, 78], es obtener una proyección de los datos en un espacio de menor o en ocasiones de igual dimensión que los datos originales de entrada, por lo que se dice que el método de LDA [77] [78] es otro método de proyección sobre un sub-espacio al igual que PCA.

La diferencia entre ambos métodos radica, en que mientras el PCA es un método no supervisado que busca los vectores propios o Eigenfaces que mejor representan a la imagen, el LDA es un método supervisado, que está diseñado para buscar los vectores que proporcionen mayor diferencia entre cada una de las clases después de la proyección [79], en otras palabras el método pretende una vez que se tienen los vectores extraídos de los rostros llevarlos a un sub-espacio de baja dimensionalidad que aumente la separabilidad de las clases presentes (cada expresión facial representa una clase).

Una de las aplicaciones más importantes de este método, se hizo en *Fisher-LDA* [80], ya que es una modificación del método original del *LDA* que proporciona ventajas significativas con respecto al original. Para comenzar a comprender como funciona el *LDA*, se analizará el método original.

Lo primero que realiza el método es encontrar un vector w de proyección, que caracterice los datos de la imagen en un espacio unidimensional, con la finalidad de que se encuentre una diferencia significativa de la clase a la que pertenece.

Entonces se tiene un conjunto de patrones $p_1, p_2, p_3, \dots, p_n$, con dimensión N etiquetados con su clase correspondiente $C_1, C_2, C_3, \dots, C_n$. De modo que cada clase tiene un número de patrones n_c . Ahora es necesario encontrar el vector w , y después se debe obtener Z_i que mejor representa las proyecciones vectoriales de los patrones.

$$Z_i = w^T p_i \quad (2.10)$$

Es tiempo de aplicar las modificaciones de Fisher, el objetivo de la modificación es el maximizar la siguiente función objetivo.

$$F(w) = \frac{w^T S_B w}{w^T S_w w} \quad (2.11)$$

Donde S_B es la matriz de dispersión inter-clase y S_w es la matriz de dispersión intra-clase, las cuales son calculadas de la siguiente forma:

$$S_B = \sum_c n p (\mu_c - \mu) (\mu_c - \mu)^T \quad (2.12)$$

$$S_w = \sum_c \sum_{i \in c} (x_i - \mu_c) (x_i - \mu_c)^T \quad (2.13)$$

Donde μ_c es la media de la clase, y μ la media de todos los datos, nP el número de patrones de la clase C . Fisher tiene como objetivo encontrar el vector w de proyección que asegure la maximización del cociente de las matrices de dispersión inter-clase e intra-clase.

$$S_B w = \lambda S_w w \tag{2.14}$$

$$S_w^{-1} S_B w = \lambda w \tag{2.15}$$

En el caso en el que la matriz S_B es NO singular, es decir tiene inversa (problema generalizado de eigen) se tiene un problema de valores propios para la matriz $S_w^{-1} S_B$, el cual se puede resolver sustituyendo la ecuación 2.15 en la función 2.11 como se observa a continuación:

$$S_w^{-1} S_B w = \lambda w \longrightarrow F(w) = \frac{w^T S_B w}{w^T S_w w} = \lambda_k \frac{w_k^T S_B w_k}{w_k^T S_w w_k} \text{ con } k = 1..d \tag{2.16}$$

Donde d es la dimensión de las imágenes.

En la expresión anterior se puede observar que el vector propio que maximiza la función es aquel que tenga un mayor valor propio (mayor λ_k).

La asertividad del PCA y LDA han sido ampliamente comparadas en la literatura, se recomienda el trabajo de Delacque y Grgic[81] que realiza pruebas con los diferentes modelos de proyección sobre sub-espacios utilizando la base de datos FERET [55]

2.3.6.4. Método de seguimiento basado en el color

Conceptos generales del color

El estudio del color es muy amplio por lo que la presente sección se enfocará en conceptos básicos que son necesarios para comprender de mejor manera los métodos de seguimiento de rostros por medio del color.

Lo primero es obtener un concepto general de lo que significa el color.

«El color es la calidad de los fenómenos visuales que depende de la impresión distinta que producen en el ojo, las luces de distinta longitud de onda, la ausencia total de luz (negro), o la suma de todos los colores (blanco)»[82].

Algunos otros conceptos fundamentales son los modelos y espacios de color, descritos a continuación.

- **Los modelos de color.** Son aquellos que describen las características de los colores de las imágenes digitales, e.g. RGB (rojo, verde, azul), CMYK (cian, magenta, amarillo, negro), CIELab (basado en CIE L* a* b*) y GRAY (tonalidades de grises), todos ellos son métodos que representan un método diferente (por lo general, numérico) de descripción de los colores.

- **Los espacios de color.** Son una variante de un modelo de color que tiene una gama (rango) específica de colores, e.g. en el modelo de color RGB, hay un número de espacios de color: Adobe RGB, sRGB, ProPhoto RGB, etc. Cada dispositivo (ya sea un monitor o una impresora) dispone

de su propio espacio de color, por lo que sólo puede reproducir los colores de su gama. Cuando una imagen pasa de un dispositivo a otro, los colores de la imagen pueden cambiar porque cada dispositivo interpreta los valores RGB o CMYK según su propio espacio de color. La gestión del color se emplea al mover imágenes con el fin de garantizar que la mayoría de los colores sean iguales o lo bastante similares como para mostrar cierta consistencia.

Modos del color

Modo de color RGB

El modo de Color RGB de manera general utiliza el modelo RGB y asigna un valor de intensidad a cada píxel. En imágenes de 8 bits por canal, los valores de intensidad varían de 0 (negro) a 255 (blanco) para cada uno de los componentes RGB (rojo, verde, azul) de una imagen en color.

Por ejemplo, un color rojo fuerte podría tener un valor R de 246, un valor G de 20 y un valor B de 50. Si los valores de los tres componentes son idénticos, se obtiene un tono de gris neutro. Si los valores de todos los componentes son 255, el resultado es blanco puro, y negro puro si el valor es de 0.

Las imágenes RGB utilizan tres colores o canales para reproducir los colores en la pantalla. En imágenes de 8 bits por canal, los tres canales se convierten en 24 (8 bits x 3 canales) de información del color por píxel. En imágenes de 24 bits, los tres canales pueden reproducir hasta 16,7 millones de colores por píxel. En imágenes de 48 bits (16 bits por canal) y 96 bits (32 bits por canal), pueden reproducirse más colores por píxel.

Modo de color CMYK

En el modo CMYK, a cada píxel se le asigna un valor de porcentaje para las tintas de cuatricromía. Los colores más claros (iluminaciones) tienen un porcentaje pequeño de tinta, mientras que los más oscuros (sombras) tienen porcentajes mayores. Por ejemplo, un rojo brillante podría tener 2% de cian, 93% de magenta, 90% de amarillo y 0% de negro. En las imágenes CMYK, el blanco puro se genera si los cuatro componentes tienen valores del 0%.

Modo de color Lab

El modelo de color $L^*a^*b^*$ (Lab), a veces denominado CIE² lab, es el modelo de color más completo y convencionalmente utilizado, se basa en la percepción humana del color. Los valores numéricos de Lab describen todos los colores que ve una persona con una capacidad de visión normal.

Como Lab describe la apariencia del color en lugar de la cantidad de colorante necesaria para que un dispositivo (como un monitor, una impresora de escritorio o una cámara digital) produzca el color, Lab se considera un modelo de color independiente de dispositivo.

Los tres componentes del modelo son: componente de luminosidad (L) que varía entre 0 – 100, el componente a (eje verde-rojo) y el componente b (eje azul-amarillo) pueden estar comprendidos entre +127 y – 128.

La forma de conversión de espacio de color RGB al espacio de color Lab es:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0.433910 & 0.0376220 & 0.189860 \\ 0.212649 & 0.715169 & 0.072182 \\ 0.017756 & 0.109478 & 0.872915 \end{bmatrix} x \begin{bmatrix} R/255 \\ G/255 \\ B/255 \end{bmatrix}$$

²CIE: Establecida en 1913 y con sede en Viena, Austria, la Comisión Internacional I de Iluminación es la autoridad en luz, iluminación, color y espacios de colores.

$$\begin{aligned}
L &= 116xY^{\frac{1}{3}} \text{ for } Y > 0.008856 \\
L &= 903.3xY \text{ for } Y > 0.008856 \\
a &= 500x(f(x) - f(y)) \\
b &= 200x(f(y) - f(z)) \\
\text{donde} \\
f(t) &= t^{\frac{1}{3}} \text{ for } t > 0.008856 \\
f(t) &= 7.787xt + \frac{16}{116} \text{ for } t \leq 0.008856
\end{aligned}$$

Modo de escala de grises

El modo Gray (en inglés) utiliza distintos tonos de gris en una imagen. La razón para la conversión de imágenes a color a una escala de grises, se debe a que requiere menos información para cada píxel. En imágenes de 8 bits, puede haber hasta 256 tonos de gris. Cada píxel de una imagen en escala de grises tiene un valor de brillo comprendido entre 0 (negro) y 255 (blanco). En imágenes de 16 y 32 bits, el número de tonos de una imagen es mucho mayor que en las imágenes de 8 bits.

De hecho, una imagen de color gris es donde rojo, verde y azul tienen la misma intensidad en el espacio RGB, por lo que solo es necesario especificar un valor único de intensidad para cada píxel, mientras que tres valores son necesarios para cada píxel en una imagen a color. Una fórmula común de conversión de RGB a Gray es:

$$[Gray] = \begin{bmatrix} 0.212 & 0.715 & 0.072 \end{bmatrix} x \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

La utilización de imágenes en tonalidades grises es muy utilizada actualmente en aplicaciones tecnológicas y algoritmos de procesamiento de imágenes. Por que como ya ha mencionado este tipo de imágenes requiere menos información para cada píxel y hay muchas áreas en las que no es necesaria la imagen en color, que es más complicada y difícil de procesar.

Detección de rostros por medio del color

Las investigaciones realizadas en la segmentación de imagen para el seguimiento de objetos[49] son muy utilizadas sobre todo porque el método representa una forma fácil y fiable, sin embargo no se puede decir lo mismo de la precisión.

La utilización de la segmentación de una imagen basada en el color[83] es utilizada con el objetivo de obtener una primera estimación de la ubicación del rostro en una imagen, o bien como un método auxiliar cuando otros seguidores más precisos son muy pesados y con un procesamiento computacional que se vuelven lento en el seguimiento de rostros.

Como todo método no todo son ventajas, ya que la información visual basada en los colores puede ser confundida con objetos que tengan un color similar al rostro [84, 49]. Sin embargo es una opción viable cuando lo que se requiere es en una primera aproximación de la posición del rostro. Por lo que hay que tener en cuenta que un método de este tipo debe ser acompañado de otros métodos que aseguren una mejor precisión en el seguimiento del rostro.

El diseño de algoritmos de seguimiento basados en el color con un rango limitado de iluminación son sencillos de realizar [85, 86]. El problema radica en el diseño del algoritmo con funciones que permitan interpretar diferentes tipos de iluminación, ya que la luz afecta directamente los valores de RGB registrado por la cámara [87].

Existen algunos algoritmos que son capaces de lograr una mayor precisión [86]. Como ya se mencionó antes se tienen dos opciones básicamente: la primera es detectar la cara y después los elementos de esta, y la otra opción es detectar algún elemento de la cara primero para que con estos datos se pueda lograr una estimación de la ubicación del rostro.

De cualquier forma la identificación del elemento que se haya elegido debe ser en tiempo real y no debe absorber demasiado tiempo en el cálculo.

En caso de elegir la identificación del rostro primero se debe realizar un seguimiento por medio del color de la piel. Los trabajos [88, 89] están basados en seguimiento por color, pero consumen demasiados recursos ya que usa correlación de colores, filtro de Kalman, predicción del contorno y otras herramientas.

Uno de los algoritmos que presenta buenos resultados en el seguimiento de una cara humana en una secuencia de imágenes es el CAMSHIFT [86, 90] (Continuously Adaptive Mean Shift). CAMSHIFT forma parte de una adaptación del algoritmo de Mean Shift [91] esta modificación se realizó para poder procesar las distribuciones de probabilidad dinámicas (con cambios en su tamaño y suposición) las cuales son características de cualquier objeto en movimiento. El modelo que se utiliza por conveniencia es el RGB por lo que si las imágenes se encuentran en otro formato será necesario transformarlas en este tipo de método.

Es importante considerar que existen otros métodos en el seguimiento de detección de rostros y también técnicas para el tratamiento de las imágenes. Por ejemplo [92]:

Rasmussen [93], propone que en un espacio de color RGB el color de la piel fuera encontrada por medio de un elipsoide 3D, con lo que implícitamente utiliza una distribución cercana a la normal.

Kentaro Toyama [49], utiliza directamente el espacio de color RGB y establece como color de la piel a aquel que se encuentra dentro de la región piramidal definida por el mismo investigador.

Yang [85], plantea que la distribución de probabilidad del color de la piel en un espacio cromático RGB es una normal bi-variente para una persona en concreto y una iluminación particular.

Bergasa [94], expone que la distribución de probabilidad de los colores de una imagen completa, incluyendo el fondo, es una mezcla de normales de la que el color de la piel es una componente más.

2.4. Clasificación de las expresiones faciales

La *Clasificación de expresiones faciales* tiene como objetivo lograr la asimilación de los datos previamente almacenados durante la etapa de análisis e identificación (véase la figura 2.1), la tarea generalmente es llevada a cabo por medio de la comparación de los datos extraídos con respecto a una base de datos, con el propósito de identificar las similitudes y diferencias entre cada característica, como ya se ha referido existen trabajos importantes que ayudan en la selección de características que mejor representan a cada expresión facial, uno de las publicaciones más referidas en la literatura con respecto a las características que definen a las seis expresiones faciales básicas es propuesta por Ekman y Friesen [7, 48, 47].

La fase de clasificación indistintamente del método que se emplee consiste en tomar decisiones lo más asertivas posibles, que permitan determinar a que clase pertenece un patrón dado, la idea básica y fundamental es determinar el grado de similitud y/o diferencia entre el patrón a clasificar y los patrones que dan representación a cada clase o en este caso a cada expresión facial. Entonces los métodos involucrados en esta fase tienen como principal objetivo decidir a qué clase pertenece el patrón de prueba (fotografía o video de la que se desea saber la expresión facial que presenta una persona en un instante dado).

Es importante destacar que el tiempo de procesamiento computacional depende del algoritmo utilizado en la fase de clasificación y este a su vez se verá severamente influido por el tamaño de la base de datos con la que se esté trabajando y del número de características que se consideren para realizar la comparación (fase de análisis e identificación). Es por ello que es muy importante la selección de características que de manera real impacten en los resultados, ya que si se eligen aspectos con poca relevancia se corre el riesgo de hacer más lento el proceso de clasificación, lo cual se verá reflejado directamente en el tiempo de procesamiento.

En este punto existen algunas vertientes que se pueden seguir independientemente de la implementación del algoritmo que se seleccione, una de ellas es medir el desplazamiento de determinados músculos faciales durante una gesticulación, otra podría ser realizar comparaciones con una base de datos integrada por una galería de imágenes con diferentes expresiones faciales, y en donde el algoritmo tendrá que determinar con cual existe mayor parecido.

En cualquiera de las formas es imprescindible determinar el momento en el que los músculos faciales comienzan a moverse, en otras palabras el sistema debe ser capaz de determinar, cuando los ojos están cerrados o abiertos, si la boca está fruncida, relajada o abierta, cuando los dientes pueden verse, la posición de las cejas, etc.

La interpretación de movimientos musculares faciales puede ser muy compleja, no obstante la literatura presenta extensas investigaciones al respecto, en especial en la identificación de expresiones faciales. El presente trabajo de tesis ha considerado apegarse a las investigaciones realizadas por el psicólogo Paul Ekman[10, 9] el cual plantea que existen 6 emociones básicas: *felicidad*, *sorpresa*, *miedo*, *disgusto*, *cólera* y *tristeza*.

Ekman y Friesen desarrollaron un sistema de clasificación de acciones faciales llamado FACS (Facial Action Coding System)[10, 11], el cual se basa en los movimientos asociados a los músculos del rostro. Durante el desarrollo de su trabajo, enfatizan que una de las tareas más complejas es identificar y establecer en una escala continua de tiempo los rangos transitorios entre una emoción y otra.

2.4.1. Métodos utilizados en el reconocimiento

En la siguiente sección se realizará una descripción general de los métodos más utilizados en la literatura de reconocimiento de expresiones faciales, no obstante se debe mencionar que no son los únicos métodos existentes para la solución de este problema, además de que hoy en día los investigadores han centrado su atención en la fusión de estos métodos, a los que se les llama comúnmente métodos híbridos, esto con el fin de obtener lo mejor de cada método y de esta manera aumentar la asertividad de el reconocimiento.

2.4.2. Enfoques para la clasificación

En esta sección se abordarán algunos de los métodos que son mayormente utilizados en la solución del problema de clasificación de expresiones faciales, aunque cabe aclarar que los métodos citados en

el presente trabajo de tesis no son los únicos que existen para el tratamiento de dicho problema, sin embargo proporcionan un panorama general de cuáles son los métodos que son mayormente citados en trabajos de vanguardia relacionados con el tema.

Actualmente el tema general de reconocimiento de patrones presenta un conjunto de métodos que son utilizados en la clasificación de patrones de múltiples áreas de investigación, por lo que actualmente el trabajo científico se ve influido en utilizar estos mismos métodos pero aplicados en distintas áreas de investigación, con el propósito de saber cuál es su nivel de asertividad en dichas áreas.

Ahora bien, una de las premisas que son producto del trabajo de la comunidad científica relacionada en el estudio de reconocimiento de patrones, es que los resultados de un clasificador dependen de la base de datos utilizada. Por lo cual los resultados obtenidos de la implementación de determinado método en un problema dado puede verse severamente influido por dos aspectos básicamente: la selección de la base de datos y la selección del método de clasificación (fase de análisis e identificación véase la figura 2.2).

La selección de la base de datos. Lo que nos lleva a pensar en ¿Cuáles son los requisitos que debe cumplir una buena base de datos? Es realmente complejo contestar a dicha pregunta, ya que actualmente se realizan investigaciones activas acerca del tema, por el momento se puede decir que existen diversos puntos de vista al respecto, no obstante la mayoría coinciden en que la base de datos debe contener los datos que mejor representen a los patrones a reconocer y que al mismo tiempo brinde una sustancial diferencia entre ellos lo cual permite determinar la clase a la que pertenece determinado patrón.

La selección del método de clasificación. Los métodos utilizados para la resolución de problemas de clasificación son diversos, sin embargo una característica común, es el hecho de que todos tienen el objetivo final de clasificar, claro cada método lo hace bajo diferente metodología, pero no hay duda de que todos buscan saber cuál es la clase a la que pertenece cierto patrón y ya que el nivel de asertividad se ve afectado por la base de datos que se utilice, algunos investigadores se dan a la tarea de probar los métodos existentes en diferentes bancos de datos, con la finalidad de determinar la asertividad promedio del clasificador, para ejemplificar la idea supóngase que un método puede ser muy buen clasificador de figuras geométricas pero no presentar buena asertividad al implementar el mismo método en el reconocimiento de expresiones faciales.

Sabiendo esto es necesario aclarar que cada uno de los métodos tiene ventajas y desventajas, éstas varían dependiendo del área de investigación a la que pertenece el problema, así como de la base de datos con la que se decida trabajar.

Existen diversos métodos para tratar los problemas de clasificación de expresiones faciales y los principios fundamentales de cada método comúnmente suele llamarse enfoques, a continuación se presenta un compendio de los enfoques mayormente utilizados, sin embargo existe controversia con respecto al tema ya que muchos de los métodos son complementados de dos o más enfoques. Por lo que su clasificación no se encuentra formalmente sustentada, no obstante el cuadro conceptual de los enfoques nos brinda un panorama general de los principios básicos de los métodos referidos durante el desarrollo de la sección 2.1.

Enfoques de IA Los enfoques con mayor auge en los últimos años son:

- **Enfoque estadístico—probabilístico.** Es de los más empleados ya que son pioneros en el área, los principios fundamentales de esta clase de métodos se encuentran en las teorías de probabilidad y estadística, y de forma puntual en el teorema de Bayes, Markov, entre otros.

Algunos trabajos encaminados en el reconocimiento de expresiones faciales que hace uso de este enfoque son:[12, 13, 14, 15, 16, 17, 18, 19, 20].

- **Clasificadores basados en métricas.** Este enfoque hace uso de las mediciones espaciales que existen entre cada patrón, por lo que hace uso de principios métricos y espaciales, este tipo de métodos se caracterizan por su sencillez y por su alto porcentaje de asertividad, quizás uno de los métodos más conocidos de este enfoque es el vecino más cercano o K-Nearest Neighbor (KNN) en inglés. Algunos trabajos encaminados en el reconocimiento de expresiones faciales que hace uso de este enfoque son:[21, 22, 23, 24, 25, 26]
- **Enfoque neuronal.** Son métodos inspirados en los modelos biológicos del comportamiento de una red neuronal, Los modelos matemáticos empleados hacen posible que este tipo de métodos sean capaces de clasificar patrones con gran asertividad. Algunos trabajos encaminados en el reconocimiento de expresiones faciales que hace uso de este enfoque son: [27, 28, 29].
- **Enfoque asociativo.** Este enfoque fue creado en México en el año 2002 dentro del Instituto Politécnico Nacional IPN, en uno de sus centros de investigación llamado “Centro de Investigación en Computación CIC”. En esencia hace uso de los modelos de memorias asociativas para dar nacimiento a reconocedores robustos en la clasificación de patrones con ruido, dichos modelos son capaces de recuperar los patrones aprendidos de una forma optima y además de clasificar patrones nuevos con alto grado de asertividad. Las memorias asociativas ya han probado ser un potente clasificador de patrones cuando es aplicado en áreas de estudio distintas al reconocimiento de expresiones faciales, es por ello que el principal interés de este trabajo de tesis es probar su asertividad en esta línea de investigación , algunos de las investigaciones más destacadas pueden encontrarse en: [30, 31, 32, 33, 34, 35, 36, 37].

2.4.2.1. Redes Neuronales Artificiales

Contexto general de las redes neuronales artificiales

Las redes neuronales tienen su inspiración en el funcionamiento del cerebro humano, por lo que el modelo se desarrolla en base a la emulación de los sistemas nerviosos biológicos.

Las redes neuronales surgen bajo el contexto de computación cognitiva y evolutiva, lógica difusa, razonamiento aproximado, teoría del caos, teoría del aprendizaje, y su funcionamiento se basa en la emulación de los sistemas nerviosos biológicos.

Pero ¿Cuales son los motivos que llevaron al surgimiento de esta clase de métodos?, bueno el cerebro humano es la parte central de nuestro sistema, se comporta como un sistema robusto en donde se procesan y toman todas las decisiones con respecto a el cuerpo. Una de las grandes ventajas de nuestro cerebro es que la gran parte de los procesos son realizados en paralelo, por lo que los fallos que se puedan ocasionar durante el proceso, en la mayoría de los casos no representan un efecto sustancial en el resultado final, otra de las grandes ventajas es su capacidad para adaptarse a el entorno, así como interpretar información ambigua o incompleta, lo mejor de todo es que es pequeño y hace todo los proceso con un nivel bajo de consumo de potencia.

Es por ello que la idea de simular de forma artificial la estructura y funcionamiento del cerebro, es una idea que durante muchos años se ha tratado de investigar. Lo primero que hay que saber es la estructura básica del cerebro y después tratar de hacer sistemas que puedan emular el mismo comportamiento. El sistema que el cerebro utiliza para realizar sus funciones es el sistema nervioso el cual desarrolla una

gran cantidad de funciones a través de diversas subdivisiones. Estas están constituidas de elementos primitivos conocidos como células nerviosas o neuronas.

La neurona es una unidad celular funcional básica de este sistema, la cual está encargada de procesar y almacenar información. Existen una variedad de tipos de neuronas las cuales tienen diferentes formas y tamaños sin embargo la mayoría de ellas comparten características en común.

Comportamiento de una neurona biológica

Ya que una red neuronal artificial está inspirada en red neuronal biológica, es necesario adentrarnos un poco en el funcionamiento de las neuronas cerebrales.

Las redes neuronales son capaces de aprender, gracias a su razonamiento inductivo: esto es que dado una serie de patrones de entrenamiento son capaces de aprender y clasificar cada uno de los patrones aprendidos, y en el caso de presentarle un patrón que no se encuentra dentro del conjunto de patrones de entrenamiento, la red neuronal es capaz de emitir un resultado basado en el parecido con algún patrón de entrenamiento.

La neurona biológica está constituida por tres regiones morfológicamente especializadas: las dendritas, el cuerpo de la célula o soma y el axón.

La información se almacena en las interconexiones entre las neuronas, ya que dependiendo de una serie de factores la información es distribuida y redundante y se direcciona por contenido. Estas operaciones son realizadas en paralelo y de manera asíncrona.

El proceso de transmisión de la información en una neurona comienza cuando las dendritas reciben información de otras neuronas y la dendrita es la encargada de enviarla al cuerpo celular de la neurona; aquí las señales son sumadas y después transmitidas a lo largo del axón para ser transferidas a otras neuronas.

Esta transmisión se realiza por medio de reacciones químicas y a una transmisión exitosa se le denomina sinapsis, durante esta etapa los axones están muy cerca de las dendritas y de las hendiduras sinápticas de otras neuronas; es aquí donde se genera cualquier tipo de representación interna de la información.

Ahora bien las neuronas se encuentran agrupadas en algo que se le da el nombre de Redes Neuronales las cuales internamente a su vez se encuentran agrupadas en capas.

Otra diferencia con respecto a otros métodos, es que las redes neuronales deben ser sometidas a un proceso de aprendizaje o entrenamiento de patrones, una vez que se asegure el aprendizaje de la red neuronal artificial se le puede someter a prueba, presentando patrones aprendidos o bien un patrón nuevo, con el propósito de determinar su nivel de asertividad.

Antecedentes de las redes neuronales artificiales

McCulloch y Pitts[95] son pioneros en la línea de investigación del cerebro como un organismo computacional, de donde surgen los primeros resultados sobre la forma de simular las tareas que lleva a cabo el cerebro. Poco tiempo después en enero 1957 Rosenblatt [96] presenta un reporte, en el que se plantea un modelo de neurona al que denomina “El perceptrón”. Ya en 1960 Donald Hebb[97], en su primer libro proporciona mayor comprensión sobre los mecanismos de interpretación así como las diferentes reacciones químicas del cerebro.

Sin embargo Minsky y Papert[98] en 1969 ponen en duda la eficacia de las redes neuronales para resolver problemas computacionales complejos, ya que en su trabajo de investigación se deja ver la incapacidad del perceptrón para resolver el problema de la función or-exclusiva, lo cual abre un abismo en el tiempo ya que la mayoría de los investigadores alejan su interés de las redes neuronales.

Debido al cuestionamiento de Minsky y Papert, el estudio de mecanismos neuronales se vio desfavorecido, menguando así la motivación de muchos investigadores de la época, por lo que en un periodo de casi veinte años, las RNA cayeron en el olvido.

Cabe destacar que durante la década de los 50's y 60's se desarrollaron grandes avances de las redes neuronales, no obstante la mayoría de ellas solo eran teóricas ya que el avance de la tecnología les impedía ser implementarlas y probarlas en las computadoras de la época, razón por la que fue en décadas posteriores cuando se realiza una implementación de forma física y funcional de las redes neuronales.

No obstante en la década de los 90 se reaviva el interés por los mecanismos neuronales, uno de los trabajos que dieron lugar a su resurgimiento es el realizado por Hopfield[99, 100] donde propone una red neuronal distinta que lleva su nombre, su arquitectura y funcionamiento hacen que sea el sistema se asemeje más al modelo biológico.

En el marco del renacimiento de las redes neuronales también se destaca el trabajo de Rumelhart y McClelland [96] los cuales popularizan el algoritmo de aprendizaje de retropropagación para el perceptrón multicapa, trabajo basado en una investigación hecha 12 años atrás por Werbos[101]. Estas nuevas arquitecturas se merecieron el nombre de “conexionistas”, las cuales esencialmente se basan en un gran número de elementos simples que trabajan en conjunto y son reguladas por “pesos”, además son capaces de procesar información emulando el comportamiento real de una neurona biológica.

Otra razón es debido al creciente desarrollo de máquinas cada vez más rápidas e inclusive de procesamiento paralelo.

Topología de las redes neuronales artificiales

Desde el punto de vista computacional, una RNA puede ser descrita como un conjunto de autómatas celulares (neuronas), por el cual se establece un flujo de información mediante una topología de interconexiones (sinapsis).

Una clasificación topología puede ser:

Redes monocapa. Establecen conexiones laterales entre las neuronas que pertenecen a la única capa que constituye la red. Las redes monocapa se utilizan típicamente en tareas relacionadas con lo que se conoce como auto-asociación; por ejemplo, para regenerar informaciones de entrada que se presenta como incompleta o distorsionada.

Redes multicapa. Estas son las más utilizadas, disponen de neuronas agrupadas en varios niveles. Dado que este tipo de redes se integran por varias capas las conexiones entre neuronas pueden ser del tipo feedforward (conexión hacia adelante) o del tipo feedback (conexión hacia atrás).

La arquitectura de las redes multicapa están formadas por tres componentes principales:

Capa de Entrada. Es la encargada de recibir directamente la información de entrada, la información que será procesada por la red.

Capas Ocultas. Son capas que están entre las capas de entrada y salida y no tienen contacto directo con la información de entrada. El número de capas ocultas se determina manualmente y depende del problema a resolver.

Capa de Salida. Es la encargada de transferir el resultado del procesamiento de la información hacia el exterior de la red.

Funcionamiento de las redes neuronales artificiales

El mecanismo de aprendizaje de una red neuronal es el proceso por el cual una red neuronal modifica sus pesos en respuesta a una información de entrada, mientras aprende con cada uno de los patrones (parámetros que caracterizan a las emociones). Una vez aprendiendo le volvemos a presentar los patrones para ver cuál es su porcentaje de error.

Las redes neuronales tienen la capacidad de generalizar, lo que implica que si se hace una modificación a un patrón original y se le presenta sin que previamente se le haya enseñado, las redes neuronales tienen la capacidad de determinar la clase de expresión a la que se parece más.

El aprendizaje está dividido en dos tipos:

Supervisado. Los ejemplos más utilizados son: el aprendizaje por corrección de error, aprendizaje por refuerzo, aprendizaje estocástico.

No Supervisado. Los ejemplos más representativos son: aprendizaje hebbiano, aprendizaje competitivo y cooperativo.

El proceso de aprendizaje de una RNA, está constituido por elementos de procesamiento o neuronas, su estado de activación, sus funciones de salida y la conexión con otras neuronas.

Los elementos de procesamiento o neuronas. Tienen la función de aceptar las señales de entrada de neuronas previas, y son multiplicadas individualmente por un peso y después sumadas para obtener una salida.

El estado de activación. Existen dos formas de actualizar los estados de una neurona. Uno es el modo síncrono en el cual la información que llega a las neuronas es en forma continua y donde los cambios se efectúan simultáneamente como comandado por un reloj interno que indica cuando debe cambiar su estado. La segunda forma es en modo asíncrono en donde las neuronas verifican su estado continuamente, conforme reciben la información.

Función de Activación. Estas son las encargadas de establecer el nuevo estado de activación la cual depende a su vez del modo de activación seleccionado. Las funciones de activación más utilizadas son:

- *Función Escalón.*
- *Función Lineal y Mixta.*
- *Función Sigmoidal (Continua).*
- *Función Gaussiana.*

Conexión entre Neuronas o Sinapsis. La sinapsis surge cuando a una neurona le llega la señal de otras neuronas, las señales de entrada son combinadas el proceso de combinación se hace sumando el total de entradas y multiplicando por un peso de ponderación.

Un algoritmo resumido se presenta a continuación:

Algoritmo general de una red neuronal

Algoritmo 2.4 Algoritmo general de una red neuronal aplicada en el reconocimiento de expresiones faciales.

1. Adquirir una base de datos con imágenes de caras.
 2. Establecer el tamaño mínimo para reconocer una cara en una imagen, medido en píxeles.
 3. Crear una ventana cuadrada de la imagen del tamaño mínimo establecido en el paso anterior, partiendo de la esquina superior izquierda.
 4. Introducir la porción obtenida de la imagen en la red neuronal, y determinar si hay una cara o no.
 5. Repetir el paso 3 para todas las posiciones posibles de la ventana de imagen, desplazándola horizontal y verticalmente.
 6. Repetir desde el paso 2, aumentando en cada iteración el tamaño de la ventana, hasta que la ventana alcance el tamaño de la imagen original.
 7. Este algoritmo se caracteriza por ser más lento que los demás vistos hasta ahora, pero compensa por su alto porcentaje de aciertos.
-

2.4.2.2. Máquinas de Soporte Vectorial (SVM)

Las Máquinas de Soporte Vectorial (SVM- Support Vector Machines) es un método estadístico utilizado para resolver problemas de regresión o de clasificación. Los aspectos fundamentales del SVM es la transformación del espacio de entrada en un espacio de alta dimensión y la localización en dicho espacio de un hiperplano separador óptimo. El algoritmo de las SVMs tienen sus inicios en la teoría de aprendizaje estadístico, donde son pioneros Vapnik y Chervonenkis, aunque hay diversas controversias con sus orígenes la mayoría coinciden en que su real implementación tiene lugar en Bell Laboratories de AT&T por la década de los 90's [102, 103].

Es necesario contar con un banco de datos de entrenamiento etiquetado por clases (las clases constituyen cada expresión facial a reconocer), con este banco de datos se entrena la máquina, los parámetros de los patrones pertenecientes a cada clase (Conjunto de características paramétricas que representan

a una expresión facial) son representadas por puntos en un espacio y se intenta separar las clases por el mayor espacio posible. Una vez que se ha llevado a cabo la fase de entrenamiento, se le presenta a la SVM un patrón perteneciente a cierta expresión, el SVM debe ser capaz de determinar la clase a la que pertenece dicho patrón, esto se hace al calcular y seleccionar la proximidad más pequeña entre el patrón y conjunto de clases representadas en un hiperplano.

El reto del método es encontrar una forma de separar las clases en un hiperplano, ya que si se cuenta con una separación significativa entre las clases, esto permite tener una clasificación con menor margen de error.

La SVM originalmente pertenecen a la familia de clasificadores lineales, sin embargo también son capaces de buscar un hiperplano que logre transformar un problema de clasificación no-lineal en un problema linealmente separable, esto se realiza mediante la elección de una función kernel adecuada. El problema cuando se hace uso de las funciones de kernel es seleccionar la adecuada, por lo que para la eficiente aplicación del método SVM es necesario construir la función de transformación del espacio original a un espacio de alta dimensión, ya que este punto es crucial para el buen funcionamiento del clasificador.

La aplicación de las SVMs presenta muy buenos resultados en el reconocimiento de expresiones faciales[104, 105, 106, 17, 107].

2.4.2.3. Redes bayesianas

Las redes bayesianas son un método estadísticos basado en los teoremas de Bayes, son los métodos más utilizados en problemas de clasificación o también llamados reconocimientos de patrones [77].

Los algoritmos bayesianos en muchas ocasiones se implementan de forma mixta o híbrida, esto es en complemento con algún otro método de clasificación [108, 109, 110]. Los algoritmos basados en la probabilidad estadística son uno de los pioneros en el área de reconocimiento de patrones por lo que el reconocimiento facial así como el de expresiones faciales no es la excepción por lo que los algoritmos bayesianos presentan un alto nivel de performance al ser utilizados en esta área[108].

El método se basa en calcular la probabilidad de que un suceso ocurra, en la aplicación del reconocimiento de expresiones faciales se realiza una medida de similitud probabilística fundamentada, generalmente haciendo uso de las diferencias entre los niveles de gris de los píxeles de una imagen, de manera que ciertos rangos típicos de la intensidad del color gris representan secciones específicas del rostro, se pueden considerar dos tipos de diferencias.

Variaciones Intrapersonales. Son las variaciones que corresponden a los cambios que existen cuando un mismo individuo realiza diferentes expresiones faciales.

Variaciones Interpersonales. Son variaciones que corresponden a los cambios registrados por diferentes individuos cuando posan una misma expresión facial.

El método requiere el cálculo de la probabilidad a Priori que depende de las fuentes de información. Después para la toma de decisión se hace uso de la regla de Maximum a Posteriori, en donde se pretende considerar válido la identificación de una determinada expresión si cumple con la condición establecida.

Otra medida de similitud ocupada es Maximum Likelihood, esta similitud considera las variaciones intrapersonales. La idea radica en utilizar las relaciones de dependencia y de independencia por complemento, las cuales existen entre variables de un problema de clasificación como el reconocimiento de expresiones faciales, después se calculan los valores numéricos de las probabilidades para determinar la clasificación, es decir asignarle la clase a la que pertenece[111].

Un algoritmo general de este método se presenta a continuación:

Algoritmo 2.5 Algoritmo general de las Redes Bayesianas

1. Se obtienen las muestras que sean representativas M , esto es las características, rasgos o parámetros en un rostro humano que permiten identificar diferentes expresiones faciales (por ejemplo mediciones entre las cejas y el ojo, o bien la distancia de algunos puntos de la boca, etc.).
2. Determinar las clases $C_1, C_2, C_3 \dots C_n$. Es decir cuántas expresiones faciales se van a reconocer.
3. Se determina la probabilidad de cada clase $P(C_n)$, esto se hace tomando en cuenta la muestra y la cardinalidad de las clases.
4. Se determinan los rasgos que son realmente útiles y que representan diferencias entre las clases y se calcula la distribución probabilística $p(X | C_n)$, la cual dependerá del número de cada rasgo de la variable aleatoria vectorial X .
5. La siguiente regla de clasificación se utiliza para determinar la expresión facial a la que pertenece un patrón desconocido X

$$X \in C_n \text{ Si } d_i > d_j \forall i \neq j, \text{ con } d = \ln(P(C_n)) + \ln(p(X | C_n)) \quad (2.17)$$

2.4.2.4. Memorias asociativas

La concepción básica de las memorias surge bajo la idea de que la mente humana es asociativa, por ejemplo un ser humano es capaz de reconocer el rostro de una persona con solo ver un segmento considerable de la cara. O bien se puede relacionar la identificación de un rostro, aun que este haya sufrido cambios parciales como lentes, cambio de peinado, etc.

La ventaja competitiva de las memorias asociativa es su capacidad de aprendizaje y almacenamiento, existen actualmente memorias asociativas con un 100% de eficiencia en la recuperación de patrones [112, 113], otra ventaja es la generalización, es decir su capacidad para reconocer patrones con alteraciones y no solo con los patrones originales con los que se llevo a cabo la fase de aprendizaje[114].

Memorias heteroasociativas: establecen una correspondencia entre el vector de entrada x con el vector de salida y , por ejemplo el vector de entrada puede ser la dimensión paramétrica de una sonrisa, y al vector de salida se le puede asignar un vector propuesto para la felicidad que es la emoción que corresponde a dicha expresión facial.

Memorias autoasociativas: establece la misma correspondencia que la memoria heteroasociativa pero los vectores de entrada x y los vectores de salida y son los mismos, es decir $x = y$.

Las memorias asociativas serán descritas con mayor detalle en el capítulo 3.

Capítulo 3

Materiales y Métodos

El propósito de este capítulo es presentar una descripción del método de clasificación empleado en el proceso de experimentación, permitiendo dar un contexto general del trabajo de investigación realizado en la presente tesis. El capítulo enfoca su atención en la descripción a detalle de las memorias asociativas basadas en modelos Alfa-Beta, las cuales son empleadas como método de reconocimiento de expresiones faciales, ya que el uso de este tipo de memorias en la tarea de reconocimiento de expresiones faciales es la motivación principal del presente trabajo de investigación.

3.1. Memorias asociativas Alfa-Beta

Con el objetivo de comprender de mejor manera el funcionamiento y las ventajas de una memoria asociativa basada en modelos Alfa-Beta, se hace necesario enfatizar el papel de las memorias asociativas dentro de la línea de investigación de la Inteligencia Artificial (IA), así mismo es importante destacar su desarrollo en los últimos años. Razón por la que el presente capítulo se encuentra dividido en 3 partes; en una primera parte se pretende establecer el papel de las memorias asociativas dentro del contexto general de la Inteligencia Artificial (IA), en la segunda parte se describen los conceptos básicos de las memorias asociativas Alfa-Beta, las cuales son la base medular del presente trabajo de tesis ya que el propósito es utilizar esta clase de memorias en la tarea de reconocimiento de expresiones faciales. Finalmente en una tercera parte se describe el modelo matemático de las memorias asociativas Alfa-Beta, incluyendo las fases de aprendizaje y recuperación de las memorias asociativas Heteroasociativas y Autoasociativas.

3.1.1. Memorias Asociativas y la inteligencia artificial (IA)

En la literatura actual se puede encontrar un conjunto de enfoques empleados en la solución de problemas relacionados con la clasificación de patrones, también llamada por muchos autores Inteligencia Artificial (IA). Cabe destacar que cada enfoque presenta un conjunto de métodos que permiten resolver el problema de reconocimiento de expresiones faciales, sin embargo la aplicación de la inteligencia artificial es muy extensa y permite la solución de problemas diversos. Las memorias asociativas son uno de los métodos que pertenecen a la línea de investigación de la IA, es por ello que es conveniente establecer algunos aspectos generales las memorias asociativas tipo Alfa-Beta, las cuales son el principal interés del presente trabajo de investigación (el contexto general de otros enfoques pueden verse en la sección 2.4.2).

3.1.1.1. Conceptos básicos

El interés del presente trabajo de tesis estará enfocado en el último enfoque presentado en la sección 3.1.1, *el enfoque asociativo*.

Una memoria asociativa puede ser vista como un sistema de almacenamiento y recuperación de patrones por medio de la asociación con otros patrones.

En esencia una memoria asociativa tiene el objetivo de recuperar de forma correcta patrones completos, a partir del previo aprendizaje con patrones de entrada, los cuales pueden estar alterados por un ruido de tipo aditivo, sustractivo o combinado, lo cual brinda un mundo de posibilidades para la utilización de dichas memorias en la solución de problemas que presenten dichas características. Por lo que un considerable grupo de investigadores se han dado a la tarea de generar modelos de memorias asociativas enfocándose en mejorar la asertividad en los resultados, así como también en la mejora de la eficiencia de respuesta y/o recuperación de patrones, rapidez y el grado de inmunidad al ruido.

Las memorias asociativas pueden ser dividida en dos fases:

Fase de aprendizaje. Esta fase esta enfocada en la construcción de la memoria asociativa, por lo que el modelo aprende a partir de los patrones de entrada. La idea fundamental se encuentra enfocada en encontrar el o los operadores necesarios que sean capaces de codificar la relación que existe entre los patrones de entrada y de salida, y con dicha codificación se genere la matriz de aprendizaje M .

Fase de recuperación. En la fase de recuperación la memoria pone a prueba lo que ha aprendido previamente. Entonces una vez que se ha conformado la matriz M (en la fase de aprendizaje), se le presenta un patrón de entrada x que previamente se aprendió, M se opera con el operador o los operadores necesarios bajo ciertas circunstancias con el patrón x , y se genera un patrón de salida y .



Figura 3.1: El patrón de entrada representado por un vector columna x y el patrón de salida representado por y .

3.2. Operaciones binarias α y β : definiciones y propiedades

Ya que el interés del presente trabajo de tesis se encuentra enfocado en las memorias asociativas Alfa-Beta, es necesario describir los conceptos básicos y los modelos matemáticos de los operadores α y β , para ello se hará uso de la bibliografía correspondiente a las investigaciones realizadas por el grupo Alfa-Beta [114, 115, 112, 113].

Las memorias asociativas Alfa-Beta hacen uso de dos operadores binarios α y β , además hacen uso de máximos \vee y mínimos \wedge ; este tipo de memorias son capaces de trabajar bajo dos diferentes arquitecturas (autoasociativo y heretoasociativo).

Como ya se ha mencionado en la sección 3.1.1.1 el funcionamiento de las memorias asociativas se divide en dos fases: *fase de aprendizaje* y *fase de recuperación*, de modo que en el caso específico de las memorias asociativas Alfa-Beta el operador α es utilizado para la fase de aprendizaje y el operador β en la fase de recuperación. La utilización de ambos operadores se rigen bajo las directrices que pueden verse en las tablas 3.1 y 3.2 y su correspondiente demostración matemática se puede encontrar en [114].

Para la definición de los operadores binarios α y β se especifican los conjuntos denominados A y B los cuales contienen los elementos:

$$A = 0, 1 \quad y \quad B = 0, 1, 2$$

Entonces la operación binaria $\alpha = A \times A \rightarrow B$ está establecida por la tabla 3.1:

Tabla 3.1: Operación binaria $\alpha = A \times A \rightarrow B$

x	y	$\alpha(x, y)$
0	0	1
0	1	0
1	0	2
1	1	1

Mientras que la operación binaria $\beta = B \times A \rightarrow A$ se define como se muestra en la tabla 3.2:

Tabla 3.2: Operación binaria $\beta = B \times A \rightarrow A$

x	y	$\beta(x, y)$
0	0	0
0	1	0
1	0	0
1	1	1
2	0	1
2	1	1

3.2.1. Propiedades algebraicas de los operadores binarios Alfa-Beta

El operador binario Alfa (α) se rige bajo las propiedades algebraicas mostradas en la tabla 3.3, debe recordarse que \vee es el operador máximo y \wedge es el operador de mínimo.

El operador binario α se rige bajo las propiedades algebraicas mostradas en la tabla 3.4, debe recordarse que \vee es el operador máximo y \wedge es el operador de mínimo.

Tabla 3.3: Propiedades algebraicas de operador binario α

$\alpha 1$ isoargumentos en α	$\alpha(x, x) = 1$
$\alpha 2$.- intercambio de argumentos en α	$(x \leq y) \leftrightarrow \alpha(x, y) \leq \alpha(y, x)$
$\alpha 3$.-es creciente por la izquierda	$(x \leq y) \leftrightarrow [\alpha(x, z) \leq \alpha(y, z)]$
$\alpha 4$.-es decreciente por la derecha	$(x \leq y) \leftrightarrow [\alpha(z, x) \leq \alpha(z, y)]$
$\alpha 5$.-es distributiva por la derecha respecto al \vee	$[\alpha(x \vee y), z] = \alpha(x, z) \vee \alpha(y, z)$
$\alpha 6$.-es distributiva por la derecha respecto al \wedge	$[\alpha(x \wedge y), z] = \alpha(x, z) \wedge \alpha(y, z)$

Tabla 3.4: Propiedades algebraicas de operador binario β

$\beta 1$ isoargumentos en β	$\beta(1, x) = x$
$\beta 2$.- intercambio de argumentos en β	$\beta(x, x) = x \forall x \in A$
$\beta 3$.-es creciente por la izquierda	$(x \leq y) \rightarrow [\beta(x, z) \leq \beta(y, z)]$
$\beta 4$.-es decreciente por la derecha	$(x \leq y) \leftrightarrow [\beta(z, x) \leq \beta(z, y)]$
$\beta 5$.-es distributiva por la derecha respecto al \vee	$[\beta(x \vee y), z] = \beta(x, z) \vee \beta(y, z)$
$\beta 6$.-es distributiva por la derecha respecto al \wedge	$[\beta(x \wedge y), z] = \beta(x, z) \wedge \beta(y, z)$

3.2.2. Operaciones matriciales

Los conjuntos A y B, las operaciones binarias α y β junto con los operadores \wedge (mínimo) y \vee (máximo) conforman el sistema algebraico $(A, B, \alpha, \beta, \wedge, \vee)$ en el que están inmersas las memorias asociativas Alfa-Beta.

Para comenzar se requiere la definición de cuatro operaciones matriciales que se muestran en la tabla 3.5.

Tabla 3.5: Operaciones Matriciales

Operación αmax : $P_{m \times r} \Delta_a Q_{r \times n} = f_{ij}^\alpha $, donde $f_{ij}^\alpha = \bigvee_{k=1}^r \alpha(p_{ik}, q_{kj})$
Operación βmax : $P_{m \times r} \Delta_\beta Q_{r \times n} = f_{ij}^\beta $, donde $f_{ij}^\beta = \bigvee_{k=1}^r \beta(p_{ik}, q_{kj})$
Operación αmin : $P_{m \times r} \nabla_a Q_{r \times n} = h_{ij}^\alpha $, donde $h_{ij}^\alpha = \bigwedge_{k=1}^r \alpha(p_{ik}, q_{kj})$
Operación βmin : $P_{m \times r} \nabla_\beta Q_{r \times n} = h_{ij}^\beta $, donde $h_{ij}^\beta = \bigwedge_{k=1}^r \beta(p_{ik}, q_{kj})$

Los lemas presentados a continuación permiten comprender de mejor manera a las memorias asociativas, ya que presentan la forma de utilizar los operadores para la construcción de la matriz de aprendizaje a partir de patrones de entrada. La literatura original se puede encontrar en [114].

El siguiente lema muestra los resultados obtenidos al utilizar las operaciones correspondientes a la utilización del operador α , las aplicadas al vector columna y vector fila que corresponden a los dos primeros patrones del banco de aprendizaje.

Lema 1. Sean $x \in A^n$ y $y \in A^m$; entonces $y \nabla_{\alpha} x^t$ es una matriz de dimensiones $m \times n$ y que además se cumple que: $y \nabla_{\alpha} x^t = y \Delta_{\alpha} x^t$.

Demostración

$$y \nabla_{\alpha} x^t = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \nabla_{\alpha}(x_1, x_2, \dots, x_n) \quad (3.1)$$

$$y \nabla_{\alpha} x^t = \begin{pmatrix} \nabla_{k=1}^1 \alpha(y_1, x_1) & \nabla_{k=1}^1 \alpha(y_1, x_2) & \dots & \nabla_{k=1}^1 \alpha(y_1, x_n) \\ \nabla_{k=1}^1 \alpha(y_2, x_1) & \nabla_{k=1}^1 \alpha(y_2, x_2) & \dots & \nabla_{k=1}^1 \alpha(y_2, x_n) \\ \vdots & \vdots & \dots & \vdots \\ \nabla_{k=1}^1 \alpha(y_m, x_1) & \nabla_{k=1}^1 \alpha(y_m, x_2) & \dots & \nabla_{k=1}^1 \alpha(y_m, x_n) \end{pmatrix}_{m \times n} \quad (3.2)$$

$$y \nabla_{\alpha} x^t = \begin{pmatrix} \alpha(y_1, x_1) & \alpha(y_1, x_2) & \dots & \alpha(y_1, x_n) \\ \alpha(y_2, x_1) & \alpha(y_2, x_2) & \dots & \alpha(y_2, x_n) \\ \vdots & \vdots & \dots & \vdots \\ \alpha(y_m, x_1) & \alpha(y_m, x_2) & \dots & \alpha(y_m, x_n) \end{pmatrix}_{m \times n} \quad (3.3)$$

$$y \nabla_{\alpha} x^t = \begin{pmatrix} \bigwedge_{k=1}^1 \alpha(y_1, x_1) & \bigwedge_{k=1}^1 \alpha(y_1, x_2) & \dots & \bigwedge_{k=1}^1 \alpha(y_1, x_n) \\ \bigwedge_{k=1}^1 \alpha(y_2, x_1) & \bigwedge_{k=1}^1 \alpha(y_2, x_2) & \dots & \bigwedge_{k=1}^1 \alpha(y_2, x_n) \\ \vdots & \vdots & \dots & \vdots \\ \bigwedge_{k=1}^1 \alpha(y_m, x_1) & \bigwedge_{k=1}^1 \alpha(y_m, x_2) & \dots & \bigwedge_{k=1}^1 \alpha(y_m, x_n) \end{pmatrix}_{m \times n} \quad (3.4)$$

Entonces se puede decir que la matriz $y \nabla_{\alpha} x^t$ es una matriz de dimensiones $m \times n$, y que $y \nabla_{\alpha} x^t = y \Delta_{\alpha} x^t$.

Para representar las operaciones ∇_{α} y Δ_{α} , se ha designado el siguiente simbolo \otimes , el cual se opera un vector columna de dimensión m con un vector fila de dimensión n :

$$y \nabla_{\alpha} x^t = y \otimes x^t = y \Delta_{\alpha} x^t \quad (3.5)$$

Donde ij -ésima componente de la matriz $y \otimes x^t$ está dada por:

$$[y \otimes x^t]_{ij} = \alpha(y_i, x_j) \quad (3.6)$$

Dado un índice de asociación μ , la expresión anterior indica ij-ésima componente de la matriz $y^\mu \otimes (x^\mu)^t$, es posible expresarla de la siguiente manera:

$$\left[y^\mu \otimes (x^\mu)^t \right]_{ij} = \alpha(y_i^\mu, x_j^\mu) \quad (3.7)$$

Hasta este punto se ha analizado al operador α , ahora toca el turno a el análisis del operador β , el cual opera con una matriz de dimensiones $m \times n$ con un vector columna de dimensión n usando las operaciones ∇_β y Δ_β . A continuación se presentan los lemas 2 y 3, los cuales estan relacionados con las i-ésimas componentes de los vectores columna resultantes de dimensión m , a partir de ambas operaciones ∇_β y Δ_β .

Lema 2. Sean $x \in A^n$ y P una matriz de dimensiones $m \times n$. La operación $P_{m \times n} \nabla_\beta x$ da como resultado un vector columna de dimensión m , cuya i-ésima componente tiene la siguiente forma y que además se cumple que: $(P_{m \times n} \nabla_\beta x)_i = \bigvee_{j=1}^n \beta(p_{ij}, x_j)$.

Demostración.

$$P_{m \times n} \nabla_\beta x = \begin{pmatrix} p_{11} & p_{12} & \cdot & \cdot & \cdot & p_{1n} \\ p_{21} & p_{22} & \cdot & \cdot & \cdot & p_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ p_{m1} & p_{m2} & \cdot & \cdot & \cdot & p_{mn} \end{pmatrix} \nabla_\beta \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} \quad (3.8)$$

$$P_{m \times n} \nabla_\beta x = \begin{pmatrix} \beta(p_{11}, x_1) \vee \beta(p_{12}, x_2) \vee \cdot \vee \cdot \vee \beta(p_{1n}, x_n) \\ \beta(p_{21}, x_1) \vee \beta(p_{22}, x_2) \vee \cdot \vee \cdot \vee \beta(p_{2n}, x_n) \\ \cdot \\ \cdot \\ \beta(p_{m1}, x_1) \vee \beta(p_{m2}, x_2) \vee \cdot \vee \cdot \vee \beta(p_{mn}, x_n) \end{pmatrix} = \begin{pmatrix} \bigvee_{j=1}^n \beta(p_{1j}, x_j) \\ \bigvee_{j=1}^n \beta(p_{2j}, x_j) \\ \cdot \\ \cdot \\ \bigvee_{j=1}^n \beta(p_{mj}, x_j) \end{pmatrix} \quad (3.9)$$

Se obtiene un vector columna de dimensión m cuya i-ésima componente es

Lema 3. Sean $x \in A^n$ y P una matriz de dimensiones $m \times n$. La operación $P_{m \times n} \Delta_\beta x$ da como resultado un vector columna de dimensión m , cuya i-ésima componente tiene la siguiente forma y que además se cumple que: $(P_{m \times n} \Delta_\beta x)_i = \bigwedge_{j=1}^n \beta(p_{ij}, x_j)$.

Demostración.

$$P_{m \times n} \Delta_\beta x = \begin{pmatrix} p_{11} & p_{12} & \cdot & \cdot & \cdot & p_{1n} \\ p_{21} & p_{22} & \cdot & \cdot & \cdot & p_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ p_{m1} & p_{m2} & \cdot & \cdot & \cdot & p_{mn} \end{pmatrix} \Delta_\beta \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} \quad (3.10)$$

$$P_{m \times n} \Delta_{\beta} x = \begin{pmatrix} \beta(p_{11}, x_1) \wedge \beta(p_{12}, x_2) \wedge \dots \wedge \beta(p_{1n}, x_n) \\ \beta(p_{21}, x_1) \wedge \beta(p_{22}, x_2) \wedge \dots \wedge \beta(p_{2n}, x_n) \\ \vdots \\ \beta(p_{m1}, x_1) \wedge \beta(p_{m2}, x_2) \wedge \dots \wedge \beta(p_{mn}, x_n) \end{pmatrix} = \begin{pmatrix} \bigwedge_{j=1}^n \beta(p_{1j}, x_j) \\ \bigwedge_{j=1}^n \beta(p_{2j}, x_j) \\ \vdots \\ \bigwedge_{j=1}^n \beta(p_{mj}, x_j) \end{pmatrix} \quad (3.11)$$

3.2.3. Memorias heteroasociativas Alfa-Beta

Se tienen dos tipos de memorias heteroasociativas Alfa-Beta ($\alpha\beta$) tipo Max (\vee) denotada por M y su ij -ésima componente como m_{ij} y las tipo Min, denotadas por W y su ij -ésima componente como w_{ij} . En la generación de ambos tipos de memorias se usará el operador \otimes , el cual como se explicó (véase 3.2.2) tiene la siguiente forma:

$$[y^{\mu} \otimes (x^{\mu})^t]_{ij} = \alpha(y_i^{\mu}, x_j^{\mu}) \mid \mu \in \{1, 2, \dots, p\}, i \in \{1, 2, \dots, m\}, j \in \{1, 2, \dots, n\} \quad (3.12)$$

3.2.3.1. Algoritmo de la Memoria Heteroasociativa Alfa-Beta tipo Max (\vee)

Fase de Aprendizaje de la memoria heteroasociativa $\alpha\beta$ tipo (\vee)

Para la fase de aprendizaje se siguen dos pasos, en el primer paso se utiliza el operador \otimes , mientras que en el segundo paso se hace uso del operador máximo \vee .

Algoritmo 3.1 Algoritmo de memoria heteroasociativa Afa-Beta tipo Max (\vee)

Paso 1. Para cada $\mu = 1, 2, \dots, p$, a partir de la pareja (x^{μ}, y^{μ}) se construye la matriz

$$[y^{\mu} \otimes (x^{\mu})^t]_{m \times n} \quad (3.13)$$

Paso 2. Aplicar el operador binario máximo \vee a las matrices obtenidas en el paso 1:

$$M = \bigvee_{\mu=1}^p [y^{\mu} \otimes (x^{\mu})^t] \quad (3.14)$$

La entrada ij -ésima está dada por la siguiente expresión:

$$m_{ij} = \bigvee_{\mu=1}^p \alpha(y_i^{\mu}, x_j^{\mu}) \quad (3.15)$$

y de acuerdo con que $\alpha = Ax A \rightarrow B$, se tiene que $m_{ij} \in B, \forall i \in \{1, 2, \dots, m\}, \forall j \in \{1, 2, \dots, n\}$.

Fase de recuperación de la memoria heteroasociativa $\alpha\beta$ tipo (\vee)

Para la fase de recuperación de las memorias heteroasociativas presentan dos casos posibles. La diferencia fundamental es que el primer caso el patrón de entrada es un patrón fundamental x^{ω} ; es decir,

la entrada es un patrón $\omega \in \{1, 2, \dots, p\}$; mientras que en el segundo caso, el patrón de entrada NO es un patrón fundamental, sino la versión distorsionada de por lo menos uno de los patrones fundamentales; lo anterior significa que si el patrón de entrada es \tilde{x} , debe existir al menos un valor de índice $\omega \in \{1, 2, \dots, p\}$, que corresponde al patrón fundamental respecto del cual \tilde{x} es una versión alterada con alguno de los tres tipos de ruido: aditivo, sustractivo o mezclado.

Lo primero que se hace es presentarle un patrón x^ω , donde $\omega \in \{1, 2, \dots, p\}$, a la memoria heteroasociativa $\alpha\beta$ la cual es de tipo ∇ , posteriormente se realiza la operación $\Delta_\beta : M \Delta_\beta x^\omega$.

Ahora bien, considerando las dimensiones de la matriz M son de $m \times n$ y el patrón que se desea recuperar x^ω es un vector columna de dimensión n , por lo que el resultado de la operación anterior debe ser un vector columna de dimensión m , cuya i -ésima componente es:

$$(M \Delta_\beta x^\omega)_i = \bigwedge_{j=1}^n \beta(m_{ij}, x_j^\omega) \quad (3.16)$$

3.2.3.2. Algoritmo de la memoria heteroasociativa Alfa-Beta tipo Min (\wedge)

Fase de aprendizaje

Algoritmo 3.2 Algoritmo de la memoria heteroasociativa Alfa-Beta tipo Min (\wedge)

Paso 1. Para constituir la matriz se considera para cada $\mu = 1, 2, \dots, p$, a partir de la pareja (x^μ, y^μ)

$$[y^\mu \otimes (x^\mu)^t]_{m \times n} \quad (3.17)$$

Paso 2. Después se aplica el operador binario mínimo \wedge a las matrices obtenidas en el paso 1:

$$W = \bigwedge_{\mu=1}^p [y^\mu \otimes (x^\mu)^t] \quad (3.18)$$

Por lo que la entrada ij -ésima está dada por la siguiente expresión:

$$w_{ij} = \bigwedge_{\mu=1}^p \alpha(y_i^\mu, x_j^\mu) \quad (3.19)$$

Fase de recuperación

En la fase de recuperación se le presenta un patrón x^ω , donde $\omega \in \{1, 2, \dots, p\}$, a la memoria heteroasociativa $\alpha\beta$ tipo \wedge y se realiza la operación $\Delta_\beta : W \nabla_\beta x^\omega$.

De forma que si se consideran las dimensiones de la matriz Min W (Construida en la fase de aprendizaje) son de $m \times n$, y el patrón que se desea recuperar es x^ω el cual es un vector columna de dimensión n , el resultado de la operación anterior debe ser un vector columna de dimensión m , cuya i -ésima componente es:

$$(W \nabla_\beta x^\omega)_i = \bigvee_{j=1}^n \beta(w_{ij}, x_j^\omega) \quad (3.20)$$

3.2.4. Memorias autoasociativas Alfa-Beta

Si a una memoria heteroasociativa se le impone la condición de que $y^\mu = x^\mu \forall \mu \in \{1, 2, \dots, p\}$, entonces deja de ser heteroasociativa y ahora se le denomina autoasociativa.

A continuación se enlistan algunas de las características de las memorias autoasociativas Alfa-Beta:

1. El conjunto fundamental toma la forma $\{(x^\mu, x^\mu) \mid \mu = 1, 2, \dots, p\}$
2. Los patrones fundamentales de entrada y salida son de la misma dimensión; denotada por n .
3. La memoria es una matriz cuadrada, para ambos tipos, \vee y \wedge . Si $x^\mu \in \mathcal{A}^n$ entonces:

$$\vee = [v_{ij}]_{m \times n} \text{ y } \wedge = [\lambda_{ij}]_{m \times n} \quad (3.21)$$

3.2.4.1. Algoritmo de la memoria autoasociativa Alfa-Beta tipo Max (\vee)

Las fases de aprendizaje y recuperación son similares a las memorias heteroasociativas Alfa-Beta, las cuales se presentan a continuación.

Fase de aprendizaje de la memoria autoasociativa tipo Max (\vee)

Algoritmo 3.3 Fase de aprendizaje de las memorias autoasociativas tipo Max (\vee)

Paso 1. Para cada $\mu = 1, 2, \dots, p$, a partir de la pareja (x^μ, x^μ) se construye la matriz

$$\left[x^\mu \otimes (x^\mu)^t \right]_{m \times n} \quad (3.22)$$

Paso 2. Se aplica el operador binario máximo \vee a las matrices obtenidas en el paso 1:

$$M = \bigvee_{\mu=1}^p \left[x^\mu \otimes (x^\mu)^t \right]_{m \times n} \quad (3.23)$$

La entrada ij -ésima de la memoria está dada así:

$$m_{ij} = \bigvee_{\mu=1}^p \alpha(x_i^\mu, x_j^\mu) \quad (3.24)$$

y de acuerdo con que $\alpha = AxA \rightarrow B$, se tiene que $m_{ij} \in B$, $\forall j \in \{1, 2, \dots, n\}$, $\forall i \in \{1, 2, \dots, m\}$.

Fase de recuperación de la memorias autoasociativa tipo Max (\vee)

Para la fase de recuperación de las memorias autoasociativas Alfa-Beta tipo \vee presentan dos casos posibles. La diferencia fundamental es que en el primer caso el patrón de entrada es un patrón fundamental; es decir, la entrada es un patrón $\omega \in \{1, 2, \dots, p\}$; mientras que en el segundo caso, el patrón de entrada NO es un patrón fundamental, sino la versión distorsionada de por lo menos uno de los patrones fundamentales; lo anterior significa que si el patrón de entrada es \tilde{x} , debe existir al menos un valor de índice $\omega \in \{1, 2, \dots, p\}$, que corresponde al patrón fundamental respecto del cual \tilde{x} es una versión alterada con alguno de los tres tipos de ruido: aditivo, sustractivo o mezclado. Analizando más a detalle ambos casos se tiene que:

Caso 1. En donde se considera un *patrón fundamental*. Se presenta un patrón x^ω , con $\omega \in \{1, 2, \dots, p\}$, a la memoria autoasociativa tipo $\alpha\beta$ y se realiza la operación Δ_β :

$$\bigvee \Delta_\beta x^\omega \quad (3.25)$$

El resultado de la operación anterior será el vector columna de dimensión n .

$$\left(\bigvee \Delta_\beta x^\omega \right) = \bigwedge_{j=1}^n \beta (m_{ij}, x_j^\mu) \quad (3.26)$$

$$\left(\bigvee \Delta_\beta x^\omega \right) = \bigwedge_{j=1}^n \beta \left\{ \left(\bigvee_{\mu=1}^p \alpha (x_i^\mu, x_j^\mu), x_j^\mu \right) \right\} \quad (3.27)$$

Lema 6. Memoria asociativa $\alpha\beta$ «Una memoria autoasociativa Alfa-Beta $\alpha\beta$ tipo Max \bigvee tiene únicamente unos en su diagonal principal»

Demostración. La ij -ésima de la memoria autoasociativa Alfa-Beta $\alpha\beta$ tipo Max \bigvee está dada por $m_{ij} = \bigvee_{\mu=1}^p \alpha (x_i^\mu, x_j^\mu)$, de acuerdo con la fase de aprendizaje (Véase 3.2.4.1). Las entradas de la diagonal principal se obtiene de la expresión anterior haciendo que $i = j$. de tal forma que:

$$m_{ij} = \bigvee_{\mu=1}^p \alpha (x_i^\mu, x_i^\mu), \quad \forall i \in \{1, 2, \dots, n\} \quad (3.28)$$

Esto es posible gracias a la propiedad de α presentada en la tabla 3.1, se tiene que $\alpha(x_i^\mu, x_i^\mu) = 1$, entonces la expresión anterior se convierte en:

$$m_{ij} = \bigvee_{\mu=1}^p (1) = 1, \quad \forall i \in \{1, 2, \dots, n\} \quad (3.29)$$

Teorema 5. Memoria asociativa $\alpha\beta$ y conjunto fundamental . «Una memoria asociativa $\alpha\beta$ tipo \bigvee , recupera de manera correcta el conjunto fundamental completo: además tienen máxima capacidad de aprendizaje ... »

$$\bigvee \Delta_\beta x^\omega = x^\omega, \quad \forall \omega \in \{1, 2, \dots, p\} \quad (3.30)$$

Demostración. Sea $\omega \in \{1, 2, \dots, p\}$ arbitrario y considerando el **Lema 6** descrito anteriormente, para cada $i \in \{1, 2, \dots, n\}$ escogida arbitrariamente:

$$m_{ii} = 1 = \alpha(x_i^\omega, x_i^\omega) \quad (3.31)$$

Lo cual nos indica que la memoria autoasociativa $\alpha\beta$ tipo \bigvee recuperará de forma correcta todo el conjunto fundamental aprendido. Además es la demostración del teorema 5.

Caso 2. En donde se considera un *patrón alterado*. Se presenta el patrón binario \tilde{x} (un patrón alterado de algún patrón fundamental x^ω) el cual es un vector columna de dimensión n , a la memoria autoasociativa Alfa-Beta tipo \bigvee y después se realiza la operación mostrada a continuación.

$$\bigvee \Delta_\beta \tilde{x} \quad (3.32)$$

Al igual que en el caso 1, el resultado de la operación anterior es un vector columna de dimensión n , cuya i -ésima componente se expresa de la siguiente manera:

$$(\bigvee \Delta_{\beta} \tilde{x}) = \bigwedge_{j=1}^n \beta(v_{ij}, \tilde{x}) \quad (3.33)$$

$$(\bigvee \Delta_{\beta} \tilde{x}) = \bigwedge_{j=1}^n \beta \left\{ \left[\bigvee_{\mu=1}^p \alpha(x_i^{\mu}, x_j^{\mu}) \right], \tilde{x}_j \right\} \quad (3.34)$$

De acuerdo con el Teorema 5, que puede consultarse en [114]:

$$\bigvee \Delta_{\beta} x^{\omega} = x^{\omega}, \forall \omega \in \{1, 2, \dots, p\} \quad (3.35)$$

Esto significa que la memoria autoasociativa Alfa-Beta tipo \bigvee recupera de manera correcta el conjunto fundamental completo. Además, en la demostración de este Teorema, en ningún momento aparece restricción alguna sobre p , que es la cardinalidad del conjunto fundamental; y esto quiere decir que el conjunto fundamental puede crecer tanto como se quiera. La consecuencia directa es que el número de patrones que puede aprender una memoria autoasociativa Alfa-Beta tipo \bigvee , con recuperación correcta, es máximo.

3.2.4.2. Algoritmo de la memoria autoasociativa Alfa-Beta tipo Min (\bigwedge)

Fase de Aprendizaje de la memoria autoasociativa tipo Min(\bigwedge)

Algoritmo 3.4 Fase de aprendizaje de la memoria autoasociativa Alfa-Beta tipo Min (\bigwedge)

Fase de Aprendizaje

Paso1. Para cada $\mu = 1, 2, \dots, p$, a partir de la pareja (x^{μ}, x^{μ}) se construye la matriz

$$[x^{\mu} \otimes (x^{\mu})^t]_{n \times n} \quad (3.36)$$

Paso2. Se aplica el operador binario máximo \bigvee a las matrices obtenidas en el paso 1:

$$\bigwedge = \bigvee_{\mu=1}^p [x^{\mu} \otimes (x^{\mu})^t] \quad (3.37)$$

La entrada ij -ésima está dada por la siguiente expresión:

$$\lambda_{ij} = \bigvee_{\mu=1}^p \alpha(x_i^{\mu}, x_j^{\mu}) \quad (3.38)$$

y de acuerdo con que $\alpha: A \times A \rightarrow B$, se tiene que $\lambda_{ij} \in B$, $i \in \{1, 2, \dots, n\}$. $\forall j \in \{1, 2, \dots, n\}$.

Fase de Recuperación de la memoria autoasociativa tipo Min(\bigwedge)

La fase de recuperación de las memorias autoasociativas Alfa- Beta tipo \bigwedge tiene dos casos posibles. En el primer caso el patrón de entrada es un patrón fundamental; es decir, la entrada es un patrón x^{ω} , con $\omega \in \{1, 2, \dots, p\}$. En el segundo caso, el patrón de entrada NO es un patrón fundamental, sino la versión distorsionada de por lo menos uno de los patrones fundamentales; lo anterior significa que si el

patrón de entrada es \tilde{x} , debe existir al menos un valor de índice $\omega \in \{1, 2, \dots, p\}$, que corresponde al patrón fundamental respecto del cual \tilde{x} es una versión alterada con alguno de los tres tipos de ruido: aditivo, sustractivo o mezclado.

Caso 1. Patrón fundamental. Se presenta un patrón x^ω , con $w \in \{1, 2, \dots, p\}$, a la memoria autoasociativa $\alpha\beta$ tipo \vee y se realiza la operación ∇_β :

$$\bigwedge \nabla_\beta x^\omega. \quad (3.39)$$

El resultado de la operación anterior será el vector columna de dimensión n .

$$\left(\bigwedge \nabla_\beta x^\omega \right) = \bigvee_{j=1}^n \beta(\lambda_{ij}, x_j^\omega) \quad (3.40)$$

$$\left(\bigwedge \nabla_\beta x^\omega \right) = \bigvee_{j=1}^n \beta \left\{ \left[\bigvee_{\mu=1}^p \alpha(x_i^\mu, x_j^\mu) \right], x_j^\omega \right\} \quad (3.41)$$

Caso 2. Patrón alterado. Se presenta el patrón binario \tilde{x} (patrón alterado de algún patrón fundamental x^ω) que es un vector columna de dimensión n , a la memoria autoasociativa Alfa-Beta tipo \bigwedge y se realiza la operación

$$\bigwedge \nabla_\beta \tilde{x} \quad (3.42)$$

Al igual que en el caso 1, el resultado de la operación anterior es un vector columna de dimensión n , cuya i -ésima componente se expresa de la siguiente manera:

$$\left(\bigwedge \nabla_\beta \tilde{x} \right) = \bigvee_{j=1}^n \beta(\lambda_{ij}, \tilde{x}_j) \quad (3.43)$$

$$\left(\bigwedge \nabla_\beta \tilde{x} \right) = \bigwedge_{j=1}^n \beta \left\{ \left[\bigvee_{\mu=1}^p \alpha(x_i^\mu, x_j^\mu) \right], \tilde{x}_j \right\} \quad (3.44)$$

De acuerdo con el Teorema 4.7, que puede consultarse en [114]:

$$\bigvee \nabla_\beta x^\omega = x^\omega, \forall \omega \in \{1, 2, \dots, p\} \quad (3.45)$$

Esto significa que la memoria autoasociativa Alfa-Beta tipo \vee recupera de manera correcta el conjunto fundamental completo. Además, en la demostración de este Teorema, en ningún momento aparece restricción alguna sobre p , que es la cardinalidad del conjunto fundamental; y esto quiere decir que el conjunto fundamental puede crecer tanto como se quiera. La consecuencia directa es que el número de patrones que puede aprender una memoria autoasociativa Alfa-Beta tipo \bigwedge , con recuperación correcta, es máximo.

Capítulo 4

Resultado y Discusión

Este capítulo está enfocado en describir los experimentos realizados empleando los modelos de memorias asociativas Alfa-Beta en el problema de reconocimiento de expresiones faciales, así como presentar los resultados experimentales realizados además de la discusión de los mismos.

El proceso sigue la guía de una serie de pasos que han sido propuestos para el desarrollo de la implementación de las memorias asociativas Alfa-Beta en el reconocimiento de expresiones, además resulta ser como un camino viable cuando lo que se desea es incursionar en el diseño e implementación de un algoritmo de reconocimiento de expresiones faciales.

Para la etapa de experimentación se hace uso de la base de datos Cohn Kanade plus, la cual se basa en la codificación de emociones de FACS empleando Unidades de Acción (UA). La memoria asociativa Alfa-Beta fue programada en Visual Basic .NET (VB.NET) utilizando el entorno de desarrollo Microsoft Visual Studio, instalado en una computadora personal (PC) armada, con un procesador Intel Core 2 Quad a 3.1 GHz, 4 GBytes de RAM y 200 GBytes en disco duro, con el sistema operativo Microsoft Windows 7.

4.1. Fases de experimentación

Como se explicó en la sección 2.2 el proceso del reconocimiento de expresiones faciales se encuentra dividido en dos fases generales, el *Análisis e Identificación* y la *Clasificación* tal como se ve en la figura 4.1, y es por ello que se han considerado la base fundamental del proceso de experimentación.

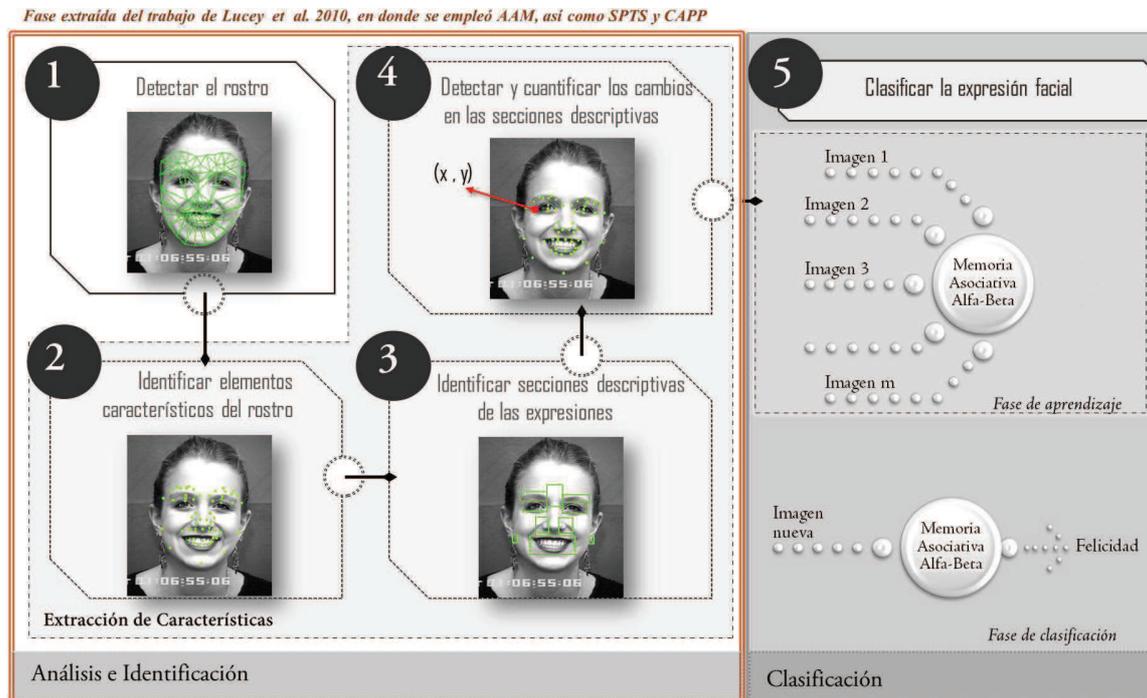


Figura 4.1: Proceso general del reconocimiento automático de expresiones faciales

Ahora bien, se han propuesto cuatro pasos que brindaran un contexto general del desarrollo de la etapa de experimentación.

- **Primer paso.** Tiene como propósito realizar una investigación de las bases de datos disponibles que puedan cumplir con los requerimientos del sistema, no obstante siempre existe la opción de realizar una base de datos propia como lo muestran algunos trabajos presentados en la tabla 2.1. En el caso de este trabajo de investigación se ha optado por utilizar el banco de datos Cohn Kanade plus.
- **Segundo paso.** Éste paso está íntimamente relacionado con la fase de Análisis e Identificación del proceso general de reconocimiento de expresiones faciales, por lo tanto la calidad de los datos que son extraídos influye severamente en la asertividad del clasificador. Debido a que es de nuestro interés mostrar la asertividad de las memorias asociativas Alfa-Beta en el problema del reconocimiento de expresiones, la fase de extracción de características forma parte de un proceso secundario, por lo tanto se hará uso de una extracción de datos que contiene el banco de datos Cohn Kanade plus, el cual fue realizado utilizando el modelo de apariencia activa (AAM) que se explica a detalle en el trabajo de investigación realizado por Lucey et al.[5].
- **Tercer paso.** Éste paso forma parte del objetivo principal del presente trabajo de investigación, ya que en él se diseña e implementa la memoria asociativa Alfa-Beta, en este paso se determina el tipo de memoria asociativa y su algoritmo.
- **Cuarto paso.** Finalmente el cuarto paso tiene la finalidad de determinar el porcentaje de asertividad de la memoria asociativa Alfa-Beta en el problema de reconocimiento de expresiones faciales, además se compara contra los resultados que presentan los trabajos del estado del arte de los 10 últimos años en el tema.

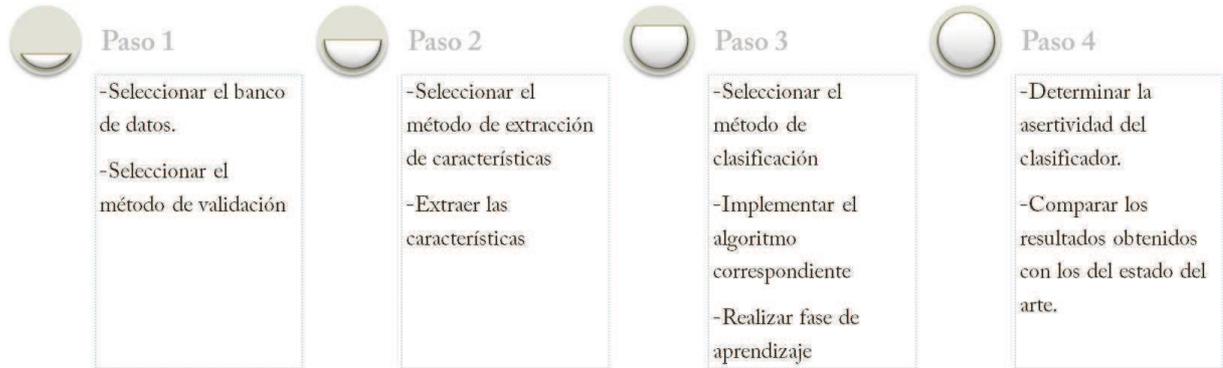


Figura 4.2: Proceso general del reconocimiento automático de expresiones faciales

4.2. Paso 1.

Durante el desarrollo de este primer paso se realiza una revisión de las bases de datos dedicadas al problema de expresiones faciales. Para este paso fue necesario realizar un estudio del arte de las bases de datos que son típicamente empleadas en el tema, con el propósito de conocer y analizar las características de cada una, entre una de las características de importancia se encuentra su disponibilidad, ya que existen algunas de ellas que son comerciales y no es nuestro interés comprar alguna base. En la sección 2.3.5 se muestra el resultado de la búsqueda y estudio de los bancos de datos mayormente referenciados en el problema del reconocimiento de expresiones faciales.

A continuación en la tabla 4.1 se muestra la preselección de los bancos de datos disponibles en el repositorio mostrado en la sección 2.3.5, esta preselección se realizó considerando algunas características de entre las cuales destacan: su disponibilidad y su relevancia en la referencia de trabajos relacionados con el tema.

Tabla 4.1: Preselección de bancos de datos del estado del arte de la sección 2.3.5.

Banco de Datos	N. Imágenes	N. Sujetos	Expresión/ Pose/ Iluminación	Color	Resolución	Año	Observaciones
JAFFE [56, 57]	213*	10	7/1/1	Gris	256x256	1998	Solo mujeres japonesas.
Cohn- Kanade+ [71, 60]	593**	123	6/1/1	Gris	640x490	2010	En la versión 2 se incrementa el número de secuencias en un 22 % y de sujetos en un 27 %.

* Imágenes, **Secuencias de Imágenes.

4.2.1. Banco de datos seleccionado “Cohn- Kanade +”

Hasta este punto se ha realizado un análisis general, considerando aspectos como: el tipo de datos proporcionado por cada uno de los dos bancos de datos, revisar si se incluía información sobre la variación de iluminación, pose, color, resolución, tipo de código de parametrización de las emociones, formato de las imágenes, entre algunos otros aspectos. También se ha realizado una revisión general de los artículos relacionados con cada una de las bases de datos, finalmente se ha concluido que la base de datos que mejor cubre con las características que se desean probar en la memoria asociativa es la segunda base de datos “Cohn-Kanade +”.



Figura 4.3: Secuencia de imágenes durante la emoción de sorpresa

A continuación se presenta un estudio mas detallado del banco de datos de “Cohn-Kanade +”.

El antecedente de «Cohn-Kanade plus (CK+)» es la base de datos «Cohn-Kanade (CK)», esta última creada en el año 2000, con el propósito de promover la investigación sobre la detección automática de las expresiones faciales individuales, por lo que la base de datos CK se convirtió en una de las más utilizadas para probar la asertividad de los algoritmos encaminados en el tema. No obstante, existían tres limitaciones evidentes que podían mejorarse:

1. Si bien los códigos de las UA están bien validados, las etiquetas de la emoción que correspondía a cada imagen no tenían la misma validación.
2. La falta de una representación métrica común con la cual evaluar nuevos algoritmos.
3. No existía un protocolo estándar para la realización de base de datos como esta.

Para hacer frente a estas y otras preocupaciones, se presenta el extendido de la base de datos Cohn-Kanade (CK +) . El número de secuencias se incrementa en un 22% y el número de sujetos en un 27%. La expresión a la que corresponden los gestos en cada secuencia es determinada por FACS, por lo tanto las etiquetas y emociones correspondientes han sido plenamente revisados y validados.

Los sujetos originalmente eran 97 estudiantes universitarios matriculados en las clases de introducción a la psicología, y para CK+ se ha incrementado hasta 123 estudiantes. Las edades comprenden entre 18 a 30 años. El 65% por ciento son mujeres, el 15% afroamericanos y 3% eran asiáticos o latinos.

El estudio en donde se realizaron las tomas estaba equipado con una silla para el sujeto y dos cámaras Panasonic WV3230, cada una conectada a un grabador de vídeo Panasonic S-VHS. Una de las cámaras se encuentra directamente frente al sujeto y la otra a 30° a la derecha. Los sujetos fueron instruidos por un experto para llevar a cabo una serie de expresiones basadas en Unidades de Acción (e.g. Neutral

→AU0) y combinaciones de Unidades de Acción (e.g. Felicidad → AU6+12+25). Los sujetos empezaron cada expresión en una pose neutral y terminan con la expresión en su máximo esplendor. Seis de las tomas se basan en descripciones prototipo de las seis emociones básicas (alegría, sorpresa, ira, miedo, asco y tristeza).

Las secuencias de imágenes fueron digitalizadas en 640 x 480 o 490 matrices de píxeles con una precisión de 8 bits en escala de grises. Las imágenes están disponibles en formatos PNG y JPG. El cuadro final de cada secuencia de imágenes se codifica utilizando FACS (Facial Action Coding System), que describe la expresión del sujeto en términos de unidades de actuación (UA). Los Códigos de FACS se llevaron a cabo por un codificador certificado de FACS. La media de kappa para la concordancia entre observadores fue de 0,86, una hoja de cálculo en Excel contiene estos códigos FACS, la cual está disponible para su descarga en [60].

4.2.2. Características de la información que proporciona Cohn-Kanade plus

El paquete contiene 4 archivos comprimidos. La descripción se encuentra relacionada con el tipo y nombre del archivo:

1. **Las imágenes (Cohn-Kanade-images.zip)** --> hay 593 secuencias de 123 sujetos, todas las expresiones son codificadas por FACS. Todas las secuencias comienzan y terminan con una pose neutral.
2. **Las Landmarks (Landmarks.zip)** --> Se realiza un seguimiento de 68 puntos en cada imagen utilizando el método de Active Appearance Model (AAM).
3. **Los archivos de codificación FACS (FACS_labels.zip)** --> para cada secuencia (593) sólo hay un archivo FACS, el cual corresponde a la última imagen de la secuencia. Cada línea del archivo corresponde a una determinada UA e intensidad. Un ejemplo se da a continuación.
4. **Los archivos de codificación de las emociones (Emotion_labels.zip)** --> Sólo 327 de las 593 secuencias están etiquetadas con su respectiva emoción. Esto se debe a que son las únicas tomas que se ajustan de mejor manera a el prototipo de la emoción. Como en el caso de los archivos de FACS, sólo hay 1 archivo de emoción para cada secuencia, la cual corresponde a la última trama. Los números determinan la emoción, el rango va desde 0 a 7 (neutral = 0, ira = 1, desprecio = 2, asco = 3, miedo = 4, felicidad = 5, tristeza = 6, sorpresa = 7).

4.2.3. Seleccionar el método de validación

Existen diferentes métodos de validación como son el método H (*Holdout*), validación cruzada (*K-fold cross-validation*), LOO (*Leave One Out*), entre otros. A continuación se presenta una breve descripción de los tres métodos de validación mencionados.

Holdout. El método holdout (método H) divide aleatoriamente el conjunto de datos en dos subconjuntos. El primer grupo se integra por 2/3 del total del conjunto de patrones y corresponde al denominado conjunto de entrenamiento, el cual es utilizado para que el clasificador aprenda, y el segundo grupo contiene el restante de los patrones es decir 1/3 y forma parte del grupo de prueba y con estos últimos se determina la asertividad del clasificador.

K-fold-cross-validation. En el método de k-dobles validación-cruzada (k-fold cross-validation) [116], funciona de la siguiente manera: la base de datos es particionada aleatoriamente en k subconjuntos, aproximadamente del mismo tamaño, en donde k-1 subconjuntos constituyen el conjunto de entrenamiento y el restante el conjunto de prueba.

Leave-One-Out. Es un caso especial del método de K-fold-cross-validation. En este método de validación la base de datos es particionada k veces, siendo k el número de casos originales n. Se forma un conjunto de entrenamiento con los n-1 casos dejando un caso fuera, el cual será utilizado en la clasificación. Cabe señalar que este método es el menos sesgado pero presenta una varianza alta, la cual en algunos casos presenta una estimación de la precisión del clasificador no deseada. Este tipo de validación, generalmente es aplicado a bases de datos con relativamente pocos casos.

Para el presente trabajo de tesis se ha utilizado el método de validación **K-fold-cross-validation**, tal como se muestra en la figura 4.4.

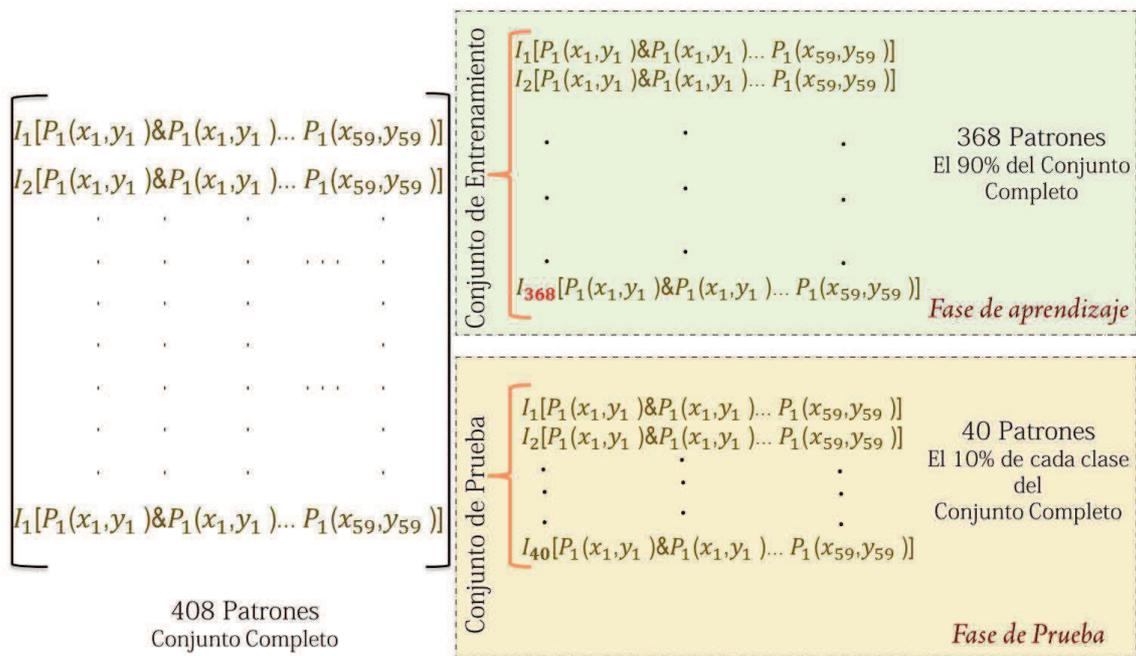


Figura 4.4: Selección de los patrones que integran el conjunto de aprendizaje y prueba empleando K-fold-cross-validation.

4.3. Paso 2.

En este paso se ha decidido hacer uso de la extracción de características que se incluye en el archivo integrado en la base de datos Cohn Kanade plus, esto se debe a que no es el interés principal de este trabajo de tesis enfocar su atención en la extracción de las características, ya que el verdadero reto se encuentra en la fase de clasificación, como se puede ver en la sección 4.4.

Ahora bien, el archivo Landmarks.zip surge del trabajo de investigación realizado por Lucey et al. [5] en el que se emplea un modelo de apariencia activa (AAM), para realizar un seguimiento de la cara y extraer características visuales. Tal como se muestra en la imagen 4.5.

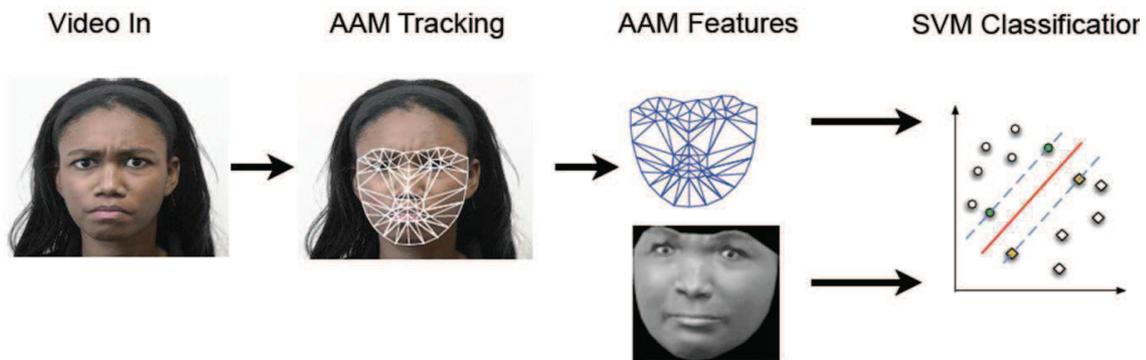


Figura 4.5: Esquema extraído del trabajo de investigación de Lucey et al.[5], en el que se muestra la utilización del Modelo de Apariencia Activa (AAM), en la fase de seguimiento del rostro y en la fase de extracción de características se hace uso de Similitud normalizada de la forma (similarity-normalized shape, SPTS) y (canonical appearance, CAPP) , en el que se marcan 68 landmarks incluidas en el compendio de archivos del banco de datos de Cohn Kanade plus.

En el artículo se justifica el uso del método de extracción de Modelo de Apariencia Activa (AAM), ya que fundamentan que este método ha demostrado dar buenos resultados en el seguimiento del rostro y también en la fase de extracción fidedigna de características faciales.

En la presente tesis se ha prestado especial interés en conocer el formato de los datos extraídos con AAM presentado en el trabajo de Lucey et al.[5], los cuales son incluidos en el archivo número dos que se describe en la sección 4.2.2.

Lucey et al.[5] refieren que el algoritmo es capaz de realizar el seguimiento de un rostro objetivo mediante la estimación de la forma y los parámetros de apariencia (AAM), y que gracias a la información almacenada durante el proceso de seguimiento del rostro se pueden extraer las siguientes características particulares.

SPTS (similarity-normalized shape). La similitud normalizada de la forma S_n , la cual esta relacionada con los 68 vértices que forman puntos dentro de la imagen y debido a que cada imagen puede verse representada como una matriz de píxeles, los 68 puntos o vértices tienen una correspondencia con los píxeles, de modo que cada punto tiene asignados un par de coordenadas ordenadas (x, y) .

Fase extraída del trabajo de Lucey et al. 2010, en donde se empleó AAM, así como SPTS y CAPP

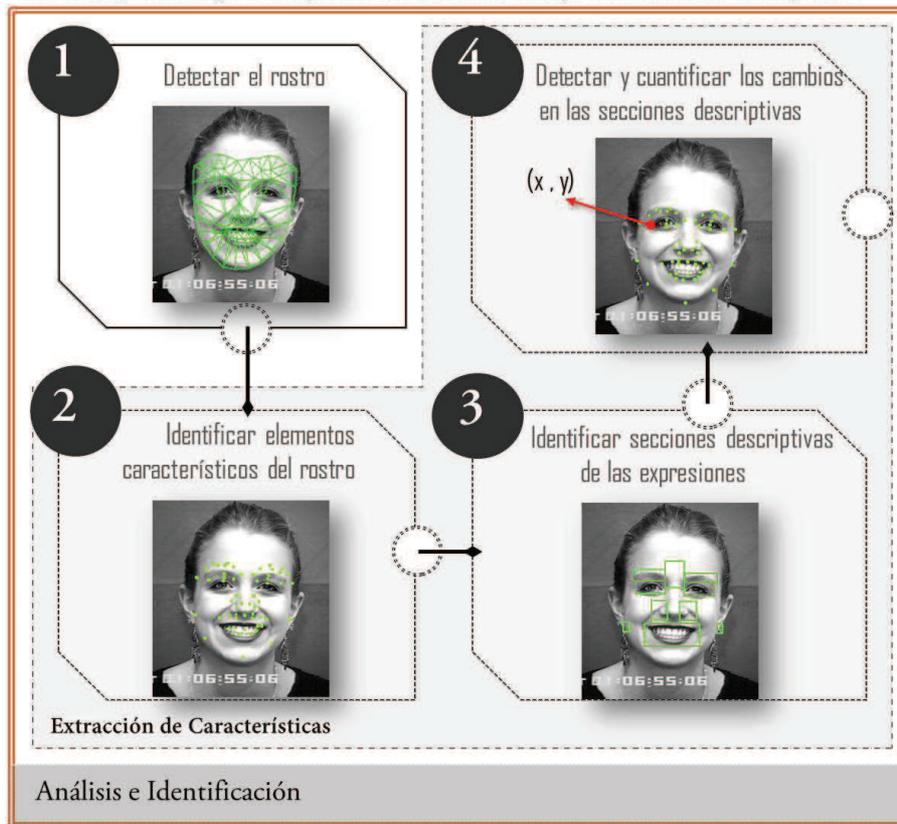


Figura 4.6: Esquema que presenta la fase de análisis e identificación extraído del trabajo de investigación de Lucey et al.[5], que se encuentra incluido en la base de datos de Cohn Kanade Plus.

4.3.1. Selección de características

Para el presente trabajo de investigación se hace uso únicamente de los primeros 59 puntos que se presentan en el archivo de *Landmarks.zip* descrito en la sección 4.2.2, ya que se han considerados solo los puntos con mayor relevancia para el reconocimiento de expresiones faciales. Los 59 considerados se muestran en la figura 4.7.

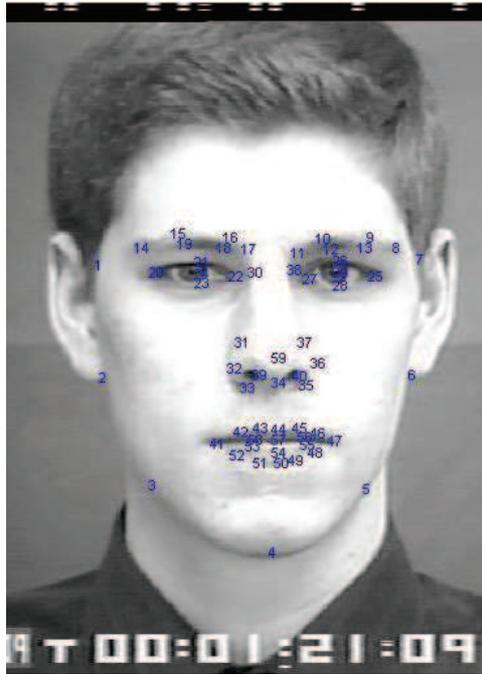


Figura 4.7: Esquema que muestra los 59 puntos considerados en la presente tesis para el reconocimiento de expresiones faciales.

De modo que para la formación de la matriz de características faciales se realiza una concatenación por cada una de las imágenes, para la presente investigación se ha hecho una selección de 408 imágenes, por lo tanto también representan el número de patrones de la matriz, tal como se muestra en la figura 4.8.

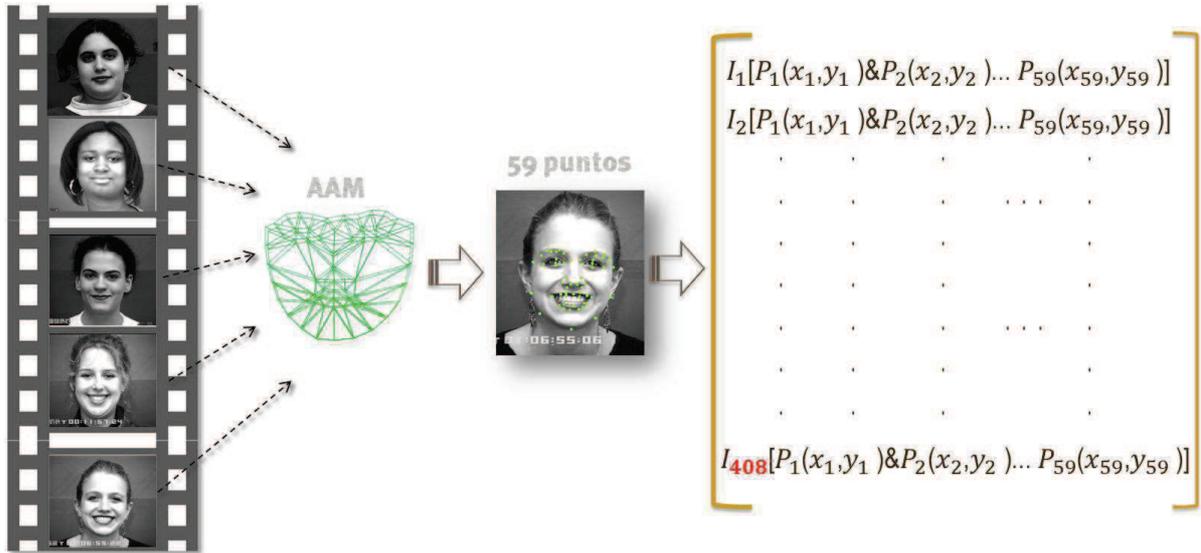


Figura 4.8: Esquema que muestra los 59 puntos considerados en la presente tesis para el reconocimiento de expresiones faciales.

4.4. Paso 3.

En la presente tesis se hace uso de las memorias asociativas Alfa-Beta, las cuales fueron descritas en la sección 3.1.1, de modo que el paso tres tiene como finalidad implementar un algoritmo basado en los modelos matemáticos de las memorias previamente descritos. En la figura 4.9 puede verse que la memoria asociativa presenta dos fases generales. La primera fase esta encargada del aprendizaje de la memoria asociativa, por lo que hace uso del conjunto fundamental o de entrenamiento, si se recuerda durante los experimentos se utilizó un método de validación llamado K-fold Cross Validation el cual esta descrito en la sección 4.2.3, en donde se separa al conjunto de entrenamiento o fundamental del conjunto de prueba.

De modo que el conjunto de entrenamiento será empleado en la primera fase de la memoria asociativa Alfa-Beta (fase de aprendizaje) y el conjunto de prueba será utilizado en la segunda fase de la memoria (fase de prueba). Ambas fases pueden verse en la figura 4.9.

Fase de Aprendizaje. El propósito de esta fase es la de aprender a partir de un conjunto de patrones verdaderos, una analogía que ayuda a comprender esta fase es la siguiente «supóngase un niño al que se le desea enseñar cuatro colores (Rojo, Verde, Amarillo y Negro), de modo que durante la fase de aprendizaje se le debe mostrar cada uno de los colores (incluso con diferentes tonalidades) permitiendo así relacionar lo que el niño observa con el nombre del color», de tal modo que cada tonalidad de rojo que se le muestre al niño será un *patrón* que pertenece a la *clase* de Rojo, y lo mismo con los demás colores . En la memoria asociativa pasa algo parecido solo que ella aprende bajo los modelos matemáticos explicados en la sección 3.1.1. En esta fase se utiliza un conjunto de patrones al cual se denomina conjunto fundamental o conjunto de entrenamiento, la determinación del tamaño de este conjunto suele ser producto de la utilización de algún método de validación (véase la sección 4.2.3).

Fase de Prueba. El objetivo es determinar la asertividad del clasificador (en este caso la memoria asociativa Alfa-Beta), de modo que se debe presentar al clasificador un patrón nuevo y la tarea del clasificador es la de determinar a que clase pertenece. Siguiendo con la analogía de los colores, durante esta fase se le debe presentar al niño un color diferente supóngase el «Rosa Intenso» de tal manera que la tarea del niño es determinar a que clase pertenece este nuevo color (también se puede presentar una variación de tonalidad de algún color aprendido, con la condición de que esta tonalidad no haya sido previamente aprendida por el niño) dependiendo de su similitud con los colores previamente aprendidos, probablemente la respuesta que se espera es que el color Rosa Intenso pertenece a la clase de Rojo.

Durante esta fase se suele hacer uso de un conjunto que no se le ha enseñado al clasificador el cual es llamado «conjunto prueba», el tamaño de este conjunto suele determinarse por medio del método de validación (véase la sección 4.2.3).

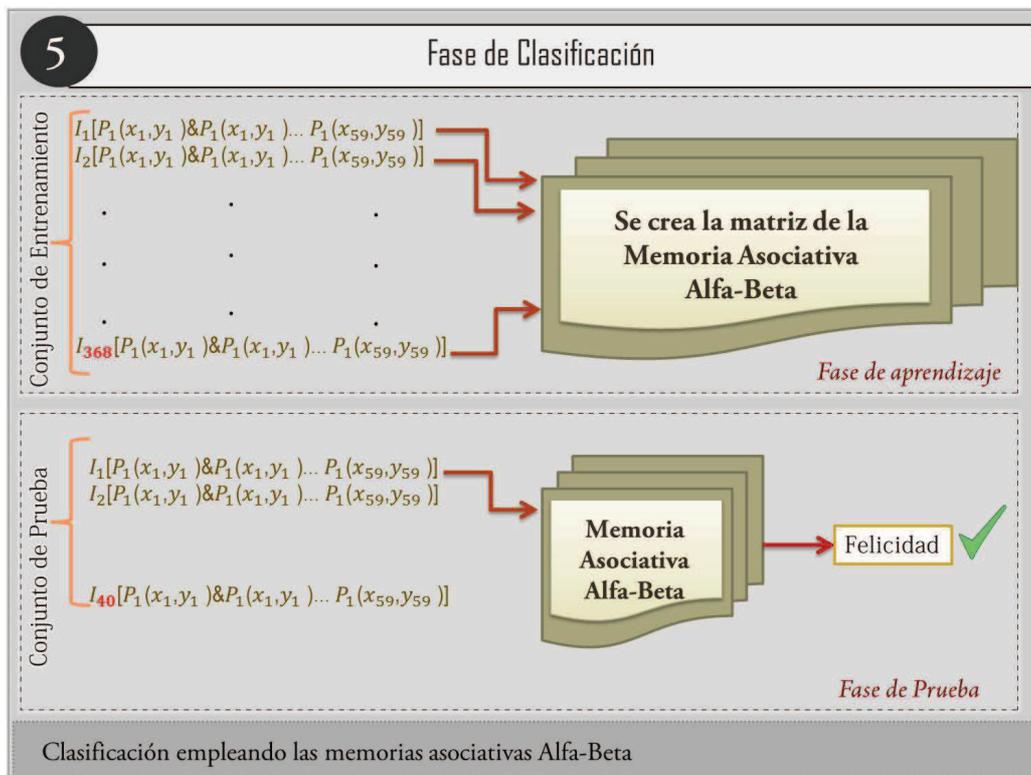


Figura 4.9: Esquema que muestra la fase de aprendizaje y de prueba de una memoria asociativa Alfa-Beta Autoasociativa

Estas dos fases son parte del algoritmo de la memoria asociativa, no obstante existe otra fase que pocas veces es referenciada en los clasificadores, se le llama Fase de Recuperación.

Fase de Recuperación. La fase de recuperación tiene como propósito determinar si todos los patrones que se le enseñaron al clasificador durante la «Fase de Aprendizaje» fueron aprendidos correctamente. De modo que una vez que la memoria asociativa ha pasado la fase de aprendizaje se le muestran nuevamente todos y cada uno de los patrones que previamente ha aprendido, para determinar si

aprendió correctamente. Parece lógico que un clasificador debe ser capaz de aprender correctamente antes de intentar clasificar un patrón nuevo, sin embargo muchos clasificadores no son capaces de recordar todos los patrones que aprendieron, y es justamente esa una de las ventajas de las memorias asociativas Alfa-Beta, ya que es capaz de recordar todo lo aprendido (100 % en la fase de recuperación) y también es capaz de clasificar eficientemente patrones nuevos que no han sido aprendidos. Siguiendo con la analogía del niño que aprende colores durante la fase de aprendizaje, ahora en la fase de recuperación se debe volver a preguntar al niño cada uno de las tonalidades de colores que previamente se le han enseñado, para determinar si aprendió correctamente todo lo que se le enseñó.

A continuación se muestra el diagrama de flujo y el algoritmo empleado en la programación de la memoria asociativa Alfa-Beta. Debe considerarse que se ha utilizado el código Johnson-Möbius modificado, durante la binarización de los datos, este algoritmo se encuentra plenamente detallado en el trabajo de investigación de Rolando Flores [117].

4.4.1. Diagrama de Flujo y algoritmo de memoria Autoasociativa Alfa-Beta tipo Max

El siguiente diagrama de flujo 3.2.4.1 muestra a grandes rasgos la base del programa elaborado para la fase de aprendizaje de la memoria autoasociativa Alfa-Beta tipo Max.

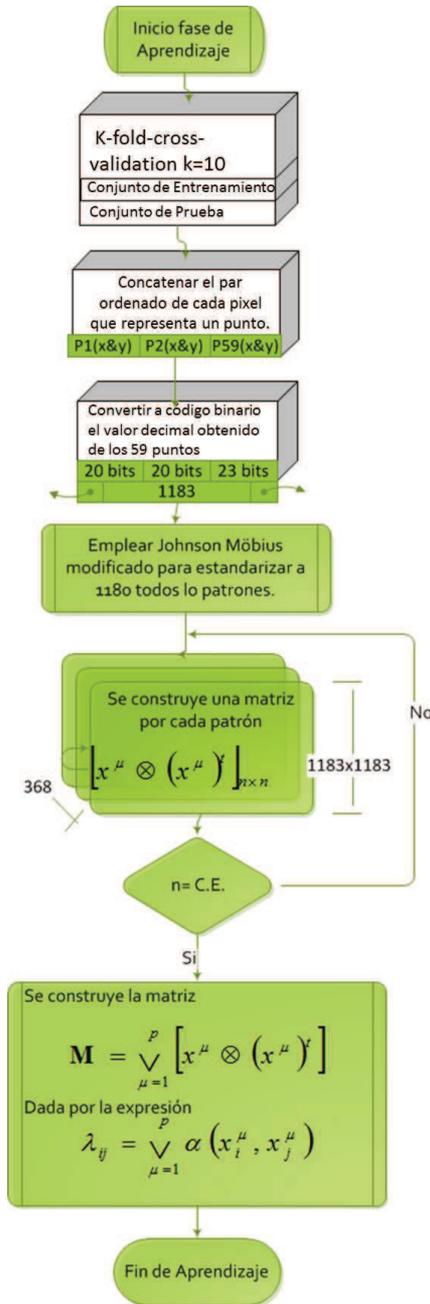


Figura 4.10: Diagrama de flujo diseñado para la fase de aprendizaje de la memoria Autoasociativa Alfa-Beta, utilizando código Johnson-Möbius.

De modo que el algoritmo de la memoria autoasociativa queda como se muestra a continuación.

Algoritmo 4.1 Fase de aprendizaje de la memoria Autoasociativa Alfa-Beta, utilizando código Johnson-Möbius.

Memoria Asociativa Alfa – Beta **Fase de Aprendizaje**

1. Obtener el conjunto de entrenamiento (*C.E.*) como se muestra en la figura 4.9.
2. Convertir a código binario cada una de las componentes de las coordenadas (x, y) que pertenecen a cada punto.
3. Concatenar todas las cadenas binarias de las coordenadas que pertenecen a una misma imagen. De modo que cada patrón esta representado por la concatenación de las cadenas binarias de todos los puntos que forman parte de una sola imagen. Entonces el número de patrones es igual al número de imágenes que pertenecen al conjunto de entrenamiento.
4. Aplicar código Johnson-Möbius modificado a cada cadena binaria concatenada o patrones.
5. Para cada patrón $\mu = 1, 2, \dots, n$. imágenes del *C.E.*, se construye la matriz.

$$[x^\mu \otimes (x^\mu)^t]_{n \times n} \quad (4.1)$$

De modo que se tiene un número de matrices= al número de imágenes del *C.E.*

6. Se aplica el operador binario máximo \vee a las matrices obtenidas en el paso 5 y se obtiene la matriz M , de tal forma que se crea la memoria autoasociativa $\alpha\beta$ tipo \vee

$$M = \bigvee_{\mu=1}^p [x^\mu \otimes (x^\mu)^t] \quad (4.2)$$

La entrada ij -ésima está dada por la siguiente expresión:

$$\lambda_{ij} = \bigvee_{\mu=1}^p \alpha(x_i^\mu, x_j^\mu) \quad (4.3)$$

y de acuerdo con que $\alpha: A \times A \rightarrow B$, se tiene que $\lambda_{ij} \in B$, $i \in \{1, 2, \dots, n\} \cdot \forall j \in \{1, 2, \dots, n\}$

Ahora bien, para la fase de recuperación de la memoria en donde se le presentan los 368 patrones que la memoria aprendió durante la «fase de aprendizaje», se diseñó el algoritmo mostrado en la figura 4.11.



Figura 4.11: Diagrama de flujo diseñado para la fase de recuperación de la memoria Autoasociativa Alfa-Beta.

Así que el algoritmo de la memoria autoasociativa implementado para la fase de recuperación se muestra a continuación.

Algoritmo 4.2 Fase de recuperación de la memoria Autoasociativa Alfa-Beta tipo Max, utilizando código Johnson-Möbius.

Memoria Asociativa Alfa – Beta **Fase de Recuperación**

1. Se le presenta cada uno de los patrones (x^ω , donde $\omega \in \{1, 2, \dots, p\}$) aprendidos a la memoria autoasociativa $\alpha\beta$ tipo \vee y se realiza la operación Δ_β :

$$\bigwedge \Delta_\beta x^\omega. \tag{4.4}$$

$$\left(\bigwedge \Delta_\beta x^\omega \right) = \bigvee_{j=1}^n \beta(\lambda_{ij}, x_j^\mu) \tag{4.5}$$

$$\left(\bigwedge \Delta_\beta x^\omega \right) = \bigvee_{j=1}^n \beta \left\{ \left[\bigvee_{\mu=1}^p \alpha(x_i^\mu, x_j^\mu) \right], x_j^\omega \right\} \tag{4.6}$$

- **Primera Evaluación.** De este modo se determina la asertividad de recuperación de patrones de la memoria autoasociativa $\alpha\beta$ tipo \vee .
-

Finalmente el diagrama de flujo elaborado para la fase de clasificación puede verse en la figura 4.12, en donde se prueba la habilidad de la memoria para clasificar patrones que no fueron previamente aprendidos, en este caso 40 patrones.

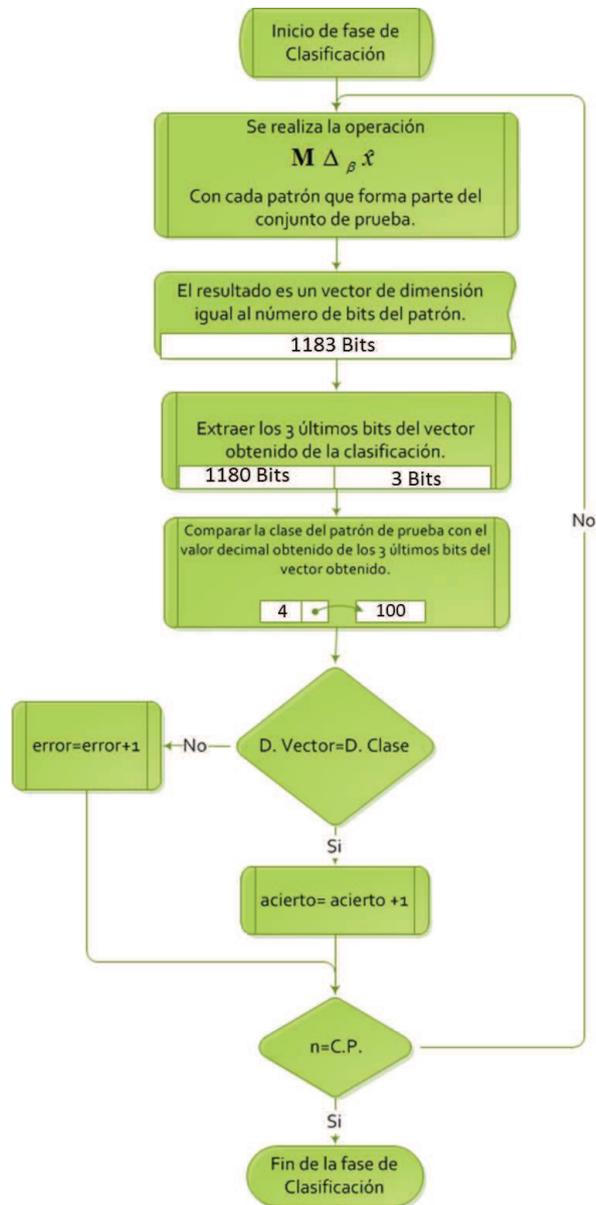


Figura 4.12: Diagrama de flujo diseñado para la fase de clasificación de la memoria Autoasociativa Alfa-Beta.

De modo que el algoritmo de la fase de clasificación queda como se muestra a continuación.

Algoritmo 4.3 Fase de prueba de la memoria Autoasociativa Alfa-Beta, utilizando código Johnson-Möbius.

Memoria Asociativa Alfa – Beta **Fase de Clasificación**

1. Se le presenta cada uno de los patrones \hat{x} que forman parte del conjunto de prueba (C.P.) a la memoria Autoasociativa $\alpha\beta$ tipo \vee y se realiza la operación Δ_β :

$$\bigwedge \nabla_\beta \hat{x} \quad (4.7)$$

Al igual que en la fase de recuperación el resultado es un vector columna de dimensión n expresada:

$$\left(\bigvee \Delta_\beta \hat{x} \right) = \bigwedge_{j=1}^n \beta(v_{ij}, \hat{x}) \quad (4.8)$$

$$\left(\bigvee \Delta_\beta \hat{x} \right) = \bigwedge_{j=1}^n \beta \left\{ \left[\bigvee_{\mu=1}^p \alpha(x_i^\mu, x_j^\mu) \right], \hat{x}_j \right\} \quad (4.9)$$

- ⤵ **Segunda Evaluación.** De este modo se determina la asertividad de clasificación de patrones distintos a los aprendidos por la memoria Autoasociativa $\alpha\beta$ tipo \vee .
-

4.5. Paso 4.

Ahora bien, el algoritmo ha sido programado en Visual Basic .NET (VB.NET) utilizando el entorno de desarrollo Microsoft Visual Studio, instalado en una computadora personal (PC) armada, con un procesador Intel Core 2 Quad a 3.1 GHz, 4 GBytes de RAM y 200 GBytes en disco duro, con el sistema operativo Microsoft Windows 7.

Utilizando el banco de datos Cohn Kanade plus y la extracción de características incluidas en el compendio del archivo de la base, se implementó el algoritmo 4.1 con un método de validación K-fold-cross-validation con k=10 obteniendo:

- **Fase de Aprendizaje:** El conjunto de entrenamiento está integrado por 368 del conjunto fundamental, con 6 clases que representan seis expresiones faciales (*felicidad, sorpresa, miedo, disgusto, cólera y tristeza*).
- **Fase de Recuperación:** 100 % de recuperación de expresiones faciales aprendidas (conjunto de entrenamiento).
- **Fase de Clasificación:** 98.56 % de asertividad considerando 40 patrones en el conjunto de prueba. La expresión facial que mejor porcentaje de asertividad obtuvo fue la felicidad con un 100 % y la expresión facial con menor asertividad fue el disgusto con un 93 %.

4.5.1. Comparar los resultados

Se ha realizado un gráfico a partir del estudio del estado del arte de los 10 últimos años (véase la tabla 2.1) , en el que se ha agregado el resultado obtenido por la memorias asociativas Alfa-Beta, para realizar un marco comparativo y evaluar la competitividad de este tipo de memorias en el problema de reconocimiento de expresiones faciales. En el gráfico 4.13 se puede ver la evolución de los resultados con el paso del tiempo.



Figura 4.13: Asertividad expresada en porcentaje ordenada por fechas correspondiente a los trabajos de investigación de los 10 últimos años en el reconocimiento automático de expresiones fáciles, en el que se incluye el resultado obtenido en el presente trabajo de tesis con las memorias asociativas Alfa-Beta.

Si se ordenan los trabajos de investigación bajo el parámetro de asertividad como se muestra en la figura 4.14, en la que claramente se puede deducir que las memorias asociativas presentan resultados aceptables y competitivos en el problema de reconocimiento de expresiones faciales, claro que también se debe evaluar que según se muestra en la tabla 2.1, los resultados pueden verse afectados por el tipo de base de datos, el método de extracción de características y por supuesto del método de clasificación empleado.



Figura 4.14: Asertividad expresada en porcentaje ordenada de mayor a menor correspondiente a los trabajos de investigaci3n de los 10 3ltimos a os en el reconocimiento autom tico de expresiones f ciles, en el que se incluye el resultado obtenido en el presente trabajo de tesis con las memorias asociativas Alfa-Beta.

Capítulo 5

Conclusiones y Trabajo Futuro

Este capítulo se ha seccionado en tres partes, la primera tiene la finalidad de presentar una breve recapitulación del desarrollo del trabajo de investigación motivo de la elaboración de la presente tesis, después en la segunda parte se exponen las conclusiones a las que se ha llegado a partir de un análisis minucioso de los resultados obtenidos durante la fase de experimentación. Finalmente se propone una serie de ideas encaminadas a robustecer esta línea de investigación, brindando así, un posible camino a seguir para trabajos desarrollados a futuro.

5.1. Síntesis

Para la realización del presente trabajo de tesis se decidió emplear una metodología interrogativa de investigación, la cual permitió delimitar el tema y establecer los esfuerzos necesarios para lograr realizar de forma satisfactoria los objetivos expuestos en la secciones 1.3 y 1.4. Por lo tanto la relación entre preguntas y respuestas es vital durante la elaboración de la tesis, de modo que se propusieron un conjunto de preguntas primarias y secundarias.

La primera pregunta primaria es *¿Qué conocimientos son necesarios para brindar una solución viable al problema de reconocimiento automático de expresiones faciales?*.

La respuesta nos lleva directamente al Capítulo II, en el que se presenta un compendio del trabajo de búsqueda y estudio sobre los problemas generales y particulares que atañen al problema de reconocimiento automático de expresiones faciales, de donde surgen las siguientes preguntas secundarias.

- ¿Qué es una expresión facial?
- ¿Qué relación existe entre las expresiones faciales y las computadoras?
- ¿Cuáles son las fases del proceso de reconocimiento automático de expresiones faciales?
- ¿Cuáles son los métodos típicamente empleados en cada una de las fases del proceso de reconocimiento automático de expresiones?
- ¿Cuál es la tendencia y resultados de los trabajos de investigación que en los últimos años se encuentran encaminados en buscar una solución al problema de reconocimiento de expresiones faciales de forma automática?

- ¿Como emplear las bondades que ofrecen las memorias asociativas Alfa-Beta para brindar una solución viable en el problema de reconocimiento automático de expresiones faciales?.

De modo que en el Capítulo II se busca brindar respuestas a cada una de estas preguntas, algunos aspectos de importancia que se muestran en el desarrollo de este capítulo son la definición de conceptos básicos referidos al tema de las expresiones faciales, el proceso general del reconocimiento de expresiones faciales de forma automática, la descripción de sus fases, los métodos típicamente empleados en cada fase, los bancos de datos que son mayormente empleados como materia prima para la prueba de eficiencia de algoritmos, entre algunas otras consideraciones de importancia.

La siguiente pregunta primaria surge de la premisa de «entre más se conocen las fortalezas y debilidades de un método será más fácil aprovechar las ventajas competitivas que ofrece el mismo», por lo tanto es necesario saber *¿Cómo funcionan las memorias asociativas de tipo Alfa-Beta?*.

La respuesta a esta pregunta se encuentra contenida en el desarrollo del Capítulo III, en el que se presenta la descripción de las memorias asociativas Alfa-Beta, sus operadores α y β , así como los modelos matemáticos y algoritmos de las memorias Alfa-Beta Heteroasociativas y Autoasociativas.

Otra pregunta primaria que debe ser resuelta es *¿Cómo probar la eficiencia de las memorias Alfa-Beta en la tarea de reconocimiento de expresiones faciales?*.

Esta pregunta conlleva a las siguientes preguntas secundarias.

- ¿Qué banco de datos emplear para la fase de experimentación?.
- ¿Cuál es el método de validación que se empleará para la selección del Conjunto de Entrenamiento (C.E.) y el Conjunto de Prueba (C.P.)?.
- ¿Que método de extracción de características emplear?.
- ¿Cómo implementar las memorias asociativas Alfa-Beta en el presente problema?.
- ¿Cómo establecer un marco comparativo de la efectividad de las memorias asociativas Alfa-Beta con respecto a otros trabajos de investigación enfocados en la presente línea de investigación?.

Las respuestas a estas preguntas se encuentran en la fase de experimentación descrita en el Capítulo IV; el desarrollo de este capítulo se apoya en la guía de cuatro pasos, los cuales facilitan la comprensión de cada etapa durante la fase de experimentación.

5.2. Conclusiones

El presente trabajo de investigación trajo como consecuencia las siguientes conclusiones particulares:

1. Se realizó un arduo estudio del contexto general de los aspectos particulares y generales del problema de reconocimiento automático de expresiones faciales.
2. Se propuso un innovador diseño de las fases principales que conforman el proceso de reconocimiento de expresiones faciales de forma automática.
3. Durante la fase de recuperación se pudo comprobar el teorema 5 (véase la sección 3.2.4.1), en el que se concluye que las memorias asociativas Alfa-Beta recuperan de manera óptima el conjunto de entrenamiento aprendido por la memoria.
4. Se comprobó que el uso de las memorias asociativas Alfa-Beta en el problema de clasificación de expresiones faciales representa una solución viable y realmente competitiva.

5.3. Trabajo a futuro

1. Aplicar los modelos asociativos Alfa-Beta en otros bancos de datos, con el propósito de determinar su efectividad bajo un conjunto distinto de condiciones tales como secuencias de imágenes o imágenes fijas.
2. Emplear un método de extracción de características distinto al empleado en este trabajo de tesis, con la finalidad de establecer el comportamiento de la memoria asociativa bajo un conjunto de datos con diferentes formatos y características.
3. Diseñar e implementar un sistema de reconocimiento automático de expresiones faciales en tiempo real empleando las memorias asociativas Alfa-Beta.
4. Emplear las diferentes variantes de los modelos de memorias asociativas Alfa-Beta, para lograr resultados mejores a los obtenidos en este trabajo de tesis.
5. Elevar el grado de dificultad de la línea de investigación de reconocimiento de expresiones faciales y emprender el reconocimiento de microexpresiones faciales, lo cual permitiría detectar cuando una persona esta mintiendo.
6. Propiciar trabajos de investigación encaminados en el empleo de tecnologías modernas, como el sensor Kinect de Microsoft®, con la finalidad de ampliar y motivar la gama de aplicaciones del reconocimiento de expresiones faciales, aprovechando la eficiencia de las memorias asociativas Alfa-Beta en esta tarea.

Referencias

- [1] E. Hjeltnäs and B.K. Low. Face detection: A survey. *Computer vision and image understanding. Elsevier*, 83(3):236–274, 2001.
- [2] Matthew Turk and Alex Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience. MIT Press*, 3(1):71–89, 1991.
- [3] M.H. Yang, D.J. Kriegman, and N. Ahuja. Detecting faces in images: A survey. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(1):34–58, 2002.
- [4] P. Viola and M. Jones. Robust real-time face detection. *International Journal of Computer Vision. Springer*, 2(57):137–154, 2004.
- [5] P. Lucey, J.F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, pages 94–101. IEEE, 2010.
- [6] Albert Mehrabian. A semantic space for nonverbal behaviors. *Journal of Consulting and Clinical Psychology. American Psychological Association*, 35(2):248–57, 1970.
- [7] Paul Ekman and W.V. Friesen. Facial action coding system: Investigator's guide. *Consulting Psychologists Press*, 1978.
- [8] Charles Darwin. *The expression of the emotions in man and animals*. The classics of psychiatry & behavioral sciences library, 1872.
- [9] Paul Ekman. *Emotion in the Human Face*. NY: Cambridge University Press, New York, 2nd edition, 1982.
- [10] P. Ekman and W. Friesen. Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology. American Psychological Association*, 17:124, 1971.
- [11] Paul Ekman and W. V. Friesen. *Unmasking the Face. A guide to recognizing emotions from facial clues*. Palo Alto: Consulting Psychologists Press. Prentice-Hall, 1975.
- [12] T.H.H. Zavaschi, A.S. Britto, L.E.S. Oliveira, and A.L. Koerich. Fusion of feature sets and classifiers for facial expression recognition. *Expert Systems with Applications, Available online 7 August 2012, ISSN 0957-4174, 10.1016/j.eswa.2012.07.074.*, 2012.
- [13] S. Fu, X. Kuai, and G. Yang. Facial expression recognition by independent log-gabor component analysis. *Advances in Neural Networks-ISBN 2011*, pages 305–312, 2011.
- [14] C. Shan, S. Gong, and P.W. McOwan. Facial expression recognition based on local binary patterns: A comprehensive study. *Image and Vision Computing. Elsevier*, 27(6):803–816, 2009.

- [15] S. Park and D. Kim. Subtle facial expression recognition using motion magnification. *Pattern Recognition Letters*, 30(7):708–716, 2009.
- [16] F.Y. Shih, C.F. Chuang, and P.S.P. Wang. Performance comparisons of facial expression recognition in jaffe database. *International Journal of Pattern Recognition and Artificial Intelligence*, 22(3):445–460, 2008.
- [17] I. Kotsia and I. Pitas. Facial expression recognition in image sequences using geometric deformation features and support vector machines. *IEEE Transactions on Image Processing*, 16(1):172–187, 2007.
- [18] P.S. Aleksic and A.K. Katsaggelos. Automatic facial expression recognition using facial animation parameters and multistream hmms. *Information Forensics and Security, IEEE Transactions on*, 1(1):3–11, 2006.
- [19] M. Pardàs and A. Bonafonte. Facial animation parameters extraction and expression recognition using hidden markov models. *Signal Processing: Image Communication*, 17(9):675–688, 2002.
- [20] Y. Zhu, LC De Silva, and C.C. Ko. Using moment invariants and hmm in facial expression recognition. *Pattern Recognition Letters*, 23(1):83–91, 2002.
- [21] T. Sha, M. Song, J. Bu, C. Chen, and D. Tao. Feature level analysis for 3d facial expression recognition. *Neurocomputing*, 2011.
- [22] Y. Cheon and D. Kim. Natural facial expression recognition using differential-aam and manifold learning. *Pattern Recognition*, 42(7):1340–1350, 2009.
- [23] S.C. Cheng, M.Y. Chen, H.Y. Chang, and T.C. Chou. Semantic-based facial expression recognition using analytical hierarchy process. *Expert Systems with Applications*, 33(1):86–95, 2007.
- [24] C. Shan, S. Gong, and P. McOwan. Appearance manifold of facial expression. *Computer Vision in Human-Computer Interaction*, pages 221–230, 2005.
- [25] H.B. Deng, L.W. Jin, L.X. Zhen, and J.C. Huang. A new facial expression recognition method based on local gabor filter bank and pca plus lda. *International Journal of Information Technology*, 11(11):86–96, 2005.
- [26] B. Abboud, F. Davoine, and M. Dang. Facial expression recognition and synthesis based on an appearance model. *Signal Processing: Image Communication*, 19(8):723–740, 2004.
- [27] C.R. De Silva, S. Ranganath, and L.C. De Silva. Cloud basis function neural network: a modified rbf network architecture for holistic facial expression recognition. *Pattern recognition*, 41(4):1241–1253, 2008.
- [28] N. Sebe, M.S. Lew, Y. Sun, I. Cohen, T. Gevers, and T.S. Huang. Authentic facial expression analysis. *Image and Vision Computing*, 25(12):1856–1863, 2007.
- [29] L. Ma and K. Khorasani. Facial expression recognition using constructive feedforward neural networks. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 34(3):1588–1595, 2004.
- [30] A. Argüelles, C. Yáñez, I. López, and O. Camacho. Prediction of co and no x levels in mexico city using associative models. *Artificial Intelligence Applications and Innovations*, pages 313–322, 2011.
- [31] I. López Yáñez, C. Yáñez Márquez, O. Camacho Nieto, and A.J. Argüelles Cruz. Prediction of air contaminant concentration based on an associative pattern classifier. *Revista Facultad de Ingeniería Universidad de Antioquia*, (60):20–30, 2011.

- [32] A. Arguelles-Cruz, I. López-Yáñez, M. Aldape-Pérez, and N. Conde-Gaxiola. Alpha-beta weightless neural networks. *Progress in Pattern Recognition, Image Analysis and Applications*, pages 496–503, 2008.
- [33] I. López Yáñez. *Clasificador automático de alto desempeño*. PhD thesis, Instituto Politécnico Nacional-Centro de Investigación en Computación, 2007.
- [34] ME Acevedo. *Memorias asociativas bidireccionales alfa-beta*. PhD thesis, Tesis de Doctorado. Centro de Investigación en Computación, México, 2006.
- [35] R. Flores' Carapia. *Memorias asociativas Alfa'Beta basadas en el código Johnson'Möbius modificado*. PhD thesis, Tesis de Maestría en: Ciencias de la Computación. CIC IPN, México, 2006.
- [36] M.E.A. Mosqueda. *Memorias Asociativas Bidireccionales Alfa-Beta*. PhD thesis, Tesis de Doctorado. Instituto Politécnico Nacional. Centro de Investigación en Computación. México, 2006.
- [37] J.C. Salgado Ramírez. *Estudio estadístico comparativo entre memorias asociativas clásicas, memorias morfológicas y memorias alfa-beta para el caso binario*. PhD thesis, Instituto Politécnico Nacional. Centro de Investigación en Computación, 2005.
- [38] S.M. Lord. *Aviation Security: Validat. of TSA's Passenger Screening Behavior Detection Program Underway, But Opportun. Exist to Strengthen Validation and Address Operational Challenges*. DIANE Publishing, 2010.
- [39] I. Cohen, N. Sebe, A. Garg, L.S. Chen, and T.S. Huang. Facial expression recognition from video sequences: temporal and static modeling. *Computer Vision and Image Understanding*, 91(1):160–187, 2003.
- [40] M. Pantic and L.J.M. Rothkrantz. Facial action recognition for facial expression analysis from static face images. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 34(3):1449–1461, 2004.
- [41] H.B. Deng, L.W. Jin, L.X. Zhen, and J.C. Huang. A new facial expression recognition method based on local gabor filter bank and pca plus lda. *International Journal of Information Technology*, 11(11):86–96, 2005.
- [42] J. Yu and B. Bhanu. Evolutionary feature synthesis for facial expression recognition. *Pattern Recognition Letters*, 27(11):1289–1298, 2006.
- [43] W. Zheng, X. Zhou, C. Zou, and L. Zhao. Facial expression recognition using kernel canonical correlation analysis (kcca). *Neural Networks, IEEE Transactions on*, 17(1):233–238, 2006.
- [44] S. Noh, H. Park, Y. Jin, and J.I. Park. Feature-adaptive motion energy analysis for facial expression recognition. *Advances in Visual Computing*, pages 452–463, 2007.
- [45] T. Xiang, MKH Leung, and SY Cho. Expression recognition using fuzzy spatio-temporal modeling. *Pattern Recognition*, 41(1):204–216, 2008.
- [46] C. Shan, S. Gong, and P.W. McOwan. Facial expression recognition based on local binary patterns: A comprehensive study. *Image and Vision Computing*, 27(6):803–816, 2009.
- [47] Paul Ekman, W.V. Friesen, and J.C. Hager. *Facial Action Coding System: The Manual*. Network Information Research Corp., Salt Lake City: Research Nexus Division., 2002.
- [48] Paul Ekman. *Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique*. *Psychological Bulletin*, 1994.

- [49] Kentaro Toyama. Prolegomena for robust face tracking. *ECCV Workshop on Advances in Facial Image Analysis and Recognition Technology*, 1998.
- [50] Lawrence Sirovich and M. Kirby. Low-dimensional procedure for the characterization of human face. *Journal of the Optical Society of America A - Optics, Image Science and Vision*, 4(3):519–524, Marzo 1987.
- [51] C. Anitha, B.A. Venkatesha, and B.S. Adiga. A survey on facial expression databases. *International Journal of Engineering Science and Technology*, 2:10, 2010.
- [52] R. Gross. Face databases, chapter in handbook of face recognition. *Springer-Verlag*, pages 301–327, 2005.
- [53] P.J. Phillips, H. Wechsler, J.S. Huang, and P.J. Rauss. The feret database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing*, 16(5):295–306, 1998.
- [54] P. J. Phillips, H. Moon, P.J. Rauss, and S. Rizvi. The feret evaluation methodology for face recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1090–1104, 2000.
- [55] FERET. The facial recognition technology (feret) database. http://www.itl.nist.gov/iad/humanid/feret/feret_master.html, 2010. Última revisión 02 de November de 2011.
- [56] M. J. Lyons, Shigeru-Akamatsu, Miyuki-Kamachi, and J. Gyoba. Coding facial expressions with gabor wavelets. In *Third IEEE International Conference on Automatic Face and Gesture Recognition*, pages 200–205. Nara Japan: IEEE Computer Society, 1998.
- [57] JAFFE. The japanese female facial expression (jaffe) database. <http://www.kasrl.org/jaffe.html>, 2010. Última revisión 02 de November de 2011.
- [58] A. R.Martinez and R. Benavente. The ar face database. Technical report 24, ComputerVision Center(CVC) Technical Report. Barcelona, 1998.
- [59] T. Kanade, J. Cohn, and Y. Tian. Comprehensive database for facial expression analysis. In *4th IEEE International Conference on Automatic Face and Gesture Recognition*, pages 46–53, 2000.
- [60] Cohn-Kanade. Cohn-kanade au-coded facial expression database. http://vasc.ri.cmu.edu/idb/html/face/facial_expression/. Última revisión 02 de November de 2011.
- [61] T. Sim and S. Baker M. Bsat. The cmu pose, illumination, and expression database. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 25(12):1615–1618, 2003.
- [62] CMU-PIE. Cmu-pie database. http://www.ri.cmu.edu/research_project_detail.html?project_id=418&menu_id=261. Última revisión 02 de November de 2011.
- [63] W. Gao, B. Cao, S. Shan, D. Zhou, X. Zhang, and D. Zhao. Cas-peal large-scale chinese face database and evaluation protocols. Technical report jdl-tr-04-fr-001, Joint Research & Development Laboratory, 2004.
- [64] W. Gao, B. Cao, S. Shan, X. Chen, D. Zhou, X. Zhao, and D. Zhang. The cas-peal large-scale chinese face database and baseline evaluations. *IEEE*, 38(1):149–161w, 2008.
- [65] B. W. Hwang, H. Byun, M.C. Roh, and S.W. Lee. Performance evaluation of face recognition algorithms on the asian face database, kfdb. In *Proceedings of the 4th international conference on Audio-and video-based biometric person authentication*, pages 557–565, Springer-Verlag.

- [66] L. Yin, X. Wei, Y. Sun, T. Wang, and M.J. Rosato. A 3d facial expression database for facial behavior research. In *Automatic face and gesture recognition, 2006. FGR 2006. 7th international conference on*, pages 211–216. IEEE, 2006.
- [67] L. Yin, X. Chen, Y. Sun, T. Worm, and M. Reale. A high-resolution 3d dynamic facial expression database. In *Automatic Face & Gesture Recognition, 2008. FG'08. 8th IEEE International Conference on*, pages 1–6. IEEE, 2008.
- [68] R. Gross, I. Matthews, J. Cohn, T T. Kanade, and S. Baker. Multi-pie. In *IEEE International Conference on Automatic Face and Gesture Recognition*, pages 1–8, 2008.
- [69] CMU Multi-PIE. The cmu multi-pie face database. Ultima revisión 02 de November de 2011.
- [70] O. Langner, R. Dotsch, G. Bijlstra, D.H.J. Wigboldus, S.T. Hawk, and A. Van Knippenberg. Presentation and validation of the radboud faces database. *Cognition and Emotion. Taylor & Francis*, 24(8):1377–1388, 2010.
- [71] P. Lucey, J.F. Cohn, T. Kanade, J. Saragih Z. Ambadar, and I. Matthews. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 13(18):94–101, 2010.
- [72] P.W. Hallinan. Recognizing human eyes. *Proceedings of SPIE*, 1570:214–226, 1991.
- [73] A. Yuille, P. Hallinan, and D. Cohen. Feature extraction from faces using deformable templates. *International Journal of Computer Vision. Springer*, 8(2):99–111, 1992.
- [74] Rafael C. Gonzalez, Richard E. Woods, and Steven L. Eddins. *Digital Image Processing using Matlab*. Pearson Education India, 2004.
- [75] Lawrence Sirovich and M. Kirby. Application of the karhunen-loeve procedure for the characterization of human faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence. Published by the IEEE Computer Society*, 12(1):103–108, Enero 1990.
- [76] Teuvo Kohonen. *Self-organization and associative memory*, volume 1. Berlin: Springer-Verlag, 1988.
- [77] R. Duda, P. Hart, and D. Stork. *Pattern Classification*, volume Capítulo 4. New York, NY, USA: John Wiley and Sons, second edition edition, 2001.
- [78] Jieping Ye, Ravi Janardan, and Qi Li. Two-dimensional linear discriminant analysis. In *The Eighteenth Annual Conference on Neural Information Processing Systems (NIPS)*, volume 17, pages 1569–1576, 2004.
- [79] W. Zhao, R. Chellappa, and A. Krishnaswamy. Discriminant analysis of principal components for face recognition. In *In Proceedings of the 3rd International Conference on Automatic Face and Gesture Recognition*, pages 336–341. Japan, April 1998.
- [80] M. Welling. Fisher linear discriminant analysis. *Department of Computer Science, University of Toronto*, (Noviembre), 2006.
- [81] K. Delac, M. Grgic, and S. Grgic. Independent comparative study of pca, ica, and lda on the feret data set. *International Journal of Imaging Systems and Technology*, pages 252–260, 2005.
- [82] Aurora M. Poo Rubio. *El color*. Ed. CYAD-UAM Azcapotzalco, México, 1992.
- [83] Wladyslaw Skarbek, Andreas Koschan, Z. Veroffentlichung, et al. Colour image segmentation - asurvey. *Technical report, Technical University of Berlin, Germany. Citeseer*, 1994.

- [84] Hans J. Andersen, Moritz Storrang, and Erik Granum. Physics-based modeling of human skin colour under mixed illuminants. *Robotics and Autonomous Systems*, 35:131–142, June 2001.
- [85] Jie Yang, Weier Lu, and Alex Waibel. Skin-color modeling and adaptation. *Computer Vision-ACCV98. Springer*, 2:687–694, 1997.
- [86] Gary R. Bradsk. Computer vision face tracking for use in a perceptual user interface. *Intel Technology Journal. Citeseer*, 2:214–219, 1998.
- [87] Bernt Shiele, Graham D. Finlayson, and James L. Crowley. Comprehensive colour image normalization. *Computer Vision-ECCV98. Springer*, 1406:475–490, 1998. Computer Science.
- [88] P. Fieguth and D. Terzopoulos. Color-based tracking of heads and other mobile objects at video frame rates. *In Proc. of IEEE CVPR*, pages 21–27, 1997.
- [89] C. Wren, A. Azarbayejani, T. Darrell, and A.P. Pentland. Pfinder: Real-time tracking of the human body. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(7):780–785, July 1997.
- [90] J.G. Allen, R.Y.D. Xu, and J.S. Jin. Object tracking using camshift algorithm and multiple quantized feature spaces. *En: Pan-Sydney Area Workshop on Visual information processing. Australian Computer Society, Inc.*, pages 3–7, 2004.
- [91] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence. IEEE Computer Society*, 24(5):603–619, 2002.
- [92] José M. Buenaposada. *Análisis de expresiones faciales mediante visión por computador*. PhD thesis, Universidad Politécnica de Madrid, Madrid, Spain, 2005.
- [93] C. Rasmussen and G. Hager. An adaptative model for tracking objects by color alone. *Technical Report DCS-TR-1200, Yale University*, 1998.
- [94] L. M. Bergasa-Pascual. *Seguimiento facial mediante visión artificial, orientado a la ayuda a la movilidad*. PhD thesis, Universidad de Alcalá de Henares, Madrid, Spain, 1999.
- [95] W. S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in neurons activity. *Bulletin of Mathematical Biology. Springer*, 5(4):115–133, 1943.
- [96] F. Rosenblatt. The perceptron, a perceiving and recognizing automaton project para. *Cornell Aeronautical Laboratory Report*, 85-460-1, 1957.
- [97] D. O. Hebb. The organization of behavior: A neuropsychological theory. *New York: Wiley.*, 1949.
- [98] M. Minsky and S. Papert. Perceptrons. *An Introduction to Computational Geometry. The MIT Press, Cambridge, MA.*, 79:2554–2558, 1969.
- [99] J.J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences. National Acad Sciences*, 79(8):2554–2558, 1982.
- [100] J.J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences. National Acad Sciences*, 81(10):3088–3092, 1984.
- [101] P.J. Werbos. *Beyond Regression: New Tools for Prediction and Analysis in the Behavioural Sciences*. PhD thesis, Dept. of Applied Mathematics, Harvard University, Cambridge, Mass., 1974.

- [102] B. E. Boser, I.M. Guyon, and V.N. Vapnik. A training algorithm for optimal margin classifiers. In D. Haussler, editor, *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, pages 144–152, 1992.
- [103] C. Cortes and V. N. Vapnik. Support vector network. *Machine Learning. Springer*, 20(3):273–297, 1995.
- [104] W. Lu and J. Yang A. Waibel. Skin-color modeling and adaptation. In *In Proc. of Third Asian Conference on Computer Vision*, volume 2, pages 687–694, 1997.
- [105] K. Anderson and P.W. McOwan. A real-time automated system for the recognition of human facial expressions. *IEEE Transactions on Systems, Man and Cybernetics, Part B: Cybernetics*, 36(1):96–105, 2006.
- [106] H. C. Lu, Y. J. Huang, and Y. W. Chen. Real-time facial expression recognition based on pixel-pattern-based texture feature. *Electronics Letters. IET*, 43(17):916–918, 2007.
- [107] N. Zhiguo and Q. Xuehong. Facial expression recognition based on weighted principal component analysis and support vector machines. In *3rd Conference on Advanced Computer Theory and Engineering (ICACTE)*, volume 3, pages 174–178, 2010.
- [108] I. Cohen, N. Sebe, F. Cozman, M. Cirelo, and T.S. Huang. Learning bayesian network classifiers for facial expression recognition using both labeled and unlabeled data. In *IEEE computer society conference on Computer vision and pattern recognition*, pages 595–601, 2003.
- [109] L. Shang and K.P. Chan. Temporal exemplar-based bayesian networks for facial expression recognition. In *Seventh International Conference on Machine Learning and Applications*, pages 16–22, 2008.
- [110] C. Fei, Y. Jiangsheng, and X. Huilin. Facial expression recognition in jaffe dataset based on gaussian process classification. *IEEE Transactions on Neural Networks*, 21:1685–1690, 2010.
- [111] J. Pearl. Fusion, propagation and structuring in belief networks. *Artificial Intelligence*, 29:241–288, 1986.
- [112] C. Yáñez-Márquez and J.L. Díaz de León Santiago. Memorias asociativas basadas en relaciones de orden y operaciones binarias, computación y sistemas. *Revista Iberoamericana de Computación incluida en el Índice de CONACyT.*, 6(4):300–311, 2003.
- [113] C. Yáñez-Márquez and J.L. Díaz de León Santiago. Memorias autoasociativas morfológicas mínimas: condiciones suficientes para convergencia, aprendizaje y recuperación de patrones. *México: IT-177, Serie Azul, CIC-IPN*, 2003.
- [114] C. Yáñez-Márquez. *Memorias Asociativas basadas en Relaciones de Orden y Operadores Binarios. Tesis doctoral, CIC-IPN*. PhD thesis, CIC-IPN, México, 2002.
- [115] C. Yáñez-Márquez and J.L. Díaz de León Santiago. Introducción a las memorias asociativas. *Research on Computing Science, Instituto Politécnico Nacional*, 6, 2003.
- [116] M. Stone. Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 111–147, 1974.
- [117] R. Flores' Carapia. *Memorias asociativas Alfa'Beta basadas en el código Johnson'Möbius modificado*. PhD thesis, Tesis de Maestría en: Ciencias de la Computación. CIC IPN, México, 2006.