



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

EVALUACIÓN DE ESTRATEGIAS DE FUSIÓN PARA UN SISTEMA DE IDENTIFICACIÓN
DE PERSONAS MULTIMODAL UTILIZANDO IMÁGENES DE ROSTRO Y ZONA
PERIOCLAR

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL ELÉCTRICO

FELIPE EDUARDO SMITH ALBORNOZ

PROFESOR GUÍA:
CLAUDIO PÉREZ FLORES

MIEMBROS DE LA COMISIÓN:
CARLOS ARAVENA CERECEDA
HÉCTOR AGUSTO ALEGRÍA

SANTIAGO DE CHILE
2015

RESUMEN DE LA MEMORIA
PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL ELÉCTRICO
POR: FELIPE SMITH ALBORNOZ
FECHA: 2015
PROF. GUÍA: CLAUDIO PÉREZ FLORES

EVALUACIÓN DE ESTRATEGIAS DE FUSIÓN PARA UN SISTEMA DE IDENTIFICACIÓN DE PERSONAS MULTIMODAL UTILIZANDO IMÁGENES DE ROSTRO Y ZONA PERIOcular

La biometría corresponde al estudio de medidas en base a diferentes características humanas, tales como huellas digitales, iris, rostro y voz, entre otros. Un gran campo de aplicación de esta corresponde al reconocimiento de rostro para seguridad y control de identidad. Recientemente, se han realizado estudios que indican que la zona peri-ocular, segmento que rodea al ojo, puede ser usado en un sistema de reconocimiento con buenos resultados.

Este trabajo de título propone como objetivo estudiar la fusión de información entre dos sistemas de reconocimiento, basado en imágenes de rostro e imágenes de zona peri-ocular, a nivel de características, puntaje y decisión. Para ello se usan las bases de datos AT&T de rostro y una base propia de imágenes de zona peri-ocular pertenecientes al laboratorio de procesamiento de imágenes del Departamento de Ingeniería Eléctrica de la Universidad de Chile.

Se implementan sistemas de reconocimiento basándose en 3 métodos de extracción de características diferentes, PCA, LDA y LBP, en C++ utilizando la librería FaceRecognizer de OpenCV. Se implementa además un sistema de fusión para cada nivel de información: característica, puntaje y decisión.

Se realizan pruebas de desempeño a los sistemas de reconocimiento de forma individual, fusionados por nivel e integrados totalmente y se comparan mediante el área bajo la curva ROC y la curva de Precision-Recall. Se crea además un sistema fusionado por puntaje válido y moda de decisión llegando a un 98.57% de clasificación correcta.

Finalmente se concluye que el mejor tipo de fusión es en el nivel de decisión, considerando el costo computacional de los algoritmos, y se entregan detalles sobre las direcciones a seguir en una posible investigación futura.

A mi familia y amigos

Agradecimientos

Quisiera agradecer primero a mi familia, por apoyarme en todo y estar siempre cuando los necesité. A mi padre por mostrarme el principio del camino que me llevó a ser ingeniero, a mi madre por enseñarme lo que es el trabajo y el esfuerzo en pos de los demás, y a mi hermana por ayudarme a levantar la cabeza de los libros y mostrarme que hay cosas más importantes en la vida.

Quisiera agradecer también a Pablo, quien me acompañó desde el colegio con su amistad y apoyo, estando siempre para escuchar mis descargas contra el mundo y aconsejarme. También a Fahmi, por acompañarme durante mi paso por Beauchef y por esas interminables conversaciones de la vida que tanto me han servido durante mis momentos más negros. A Ignacio y Estefanía por todos los buenos momentos que pasamos durante la universidad y todas las veces que me acogieron en su hogar.

Quisiera darles las gracias al equipo de Escuela de Verano, que me enseñó lo maravillosa que puede ser la docencia. En especial a Daniela, que siempre está con una sonrisa dispuesta a ayudar a todos, y a Liz, que siempre me acompañó, apoyó y alegró durante mi paso por EdV y la memoria, es simplemente la mejor monito del mundo! (#delBueno).

Quiero agradecer también a la gente con la que trabajé en el laboratorio de robótica de la facultad, en especial a Mauricio, que ha sido siempre un ejemplo a seguir, y su ayuda durante las RoboCup de Turquía y Mexico. Y a los Nerds del laboratorio de imágenes, en especial a Carlos, que estuvo pendiente de este trabajo y me ayudó cada vez que nada funcionaba.

A mis amigos de la U. A la Familia (Clau, Kata, Pancho, Pipo, Jorge, Sven, Fahmi y Nico) por no perder el contacto durante tanto tiempo y brindar siempre esa cuota de alegría y sarcasmo, que son tan necesarios en mi vida. A Sebastian Diaz, por esas largas noches de estudio y fiestas (no necesariamente en ese orden) y por la amistad que quedó después de entrar a especialidad. A Riffo, por aportar con su visión de mundo y sus historias locas, y por no dejar que muriera en la 210, espero que vuelva de Francia algún día. A Pablo Castelletto y Matías Galleguillos por todas esas noches de trabajos, tarreos, jugos y fracasos varios, simplemente gracias porque... OW. Finalmente a Romina por su compañía y cariño durante gran parte de mi paso por eléctrica, me enseñaste muchas cosas y estaré siempre muy agradecido, solamente te deseo lo mejor en la vida y espero que nos topemos de nuevo en algún lugar del mundo.

Agradecer también a todas las personas que alguna vez me recibieron en sus casa por culpa de la universidad. A ellos les estaré siempre inmensamente agradecido.

Terminar diciendo que este trabajo es el reflejo de todas las experiencias que pasé durante mi etapa universitaria y que, como tal, cada persona con la que he compartido algo ha tenido impacto en mi vida y le estoy muy agradecido.

Tabla de Contenido

1. Introducción	8
1.1. Motivación	8
1.2. Alcances	9
1.3. Objetivo general	10
1.4. Objetivos específicos	10
1.5. Estructura del trabajo	10
2. Marco teórico	12
2.1. Reconocimiento de patrones	12
2.1.1. Clasificación	12
2.1.2. Desempeño de un clasificador [14]	13
2.1.3. Errores de adquisición de datos	17
2.2. Sistemas biométricos	17
2.2.1. Operación de un sistema biométrico	18
2.2.2. Verificación e identificación	19
2.2.3. Características de una muestra biométrica	20
2.2.4. Limitantes de los sistemas biométricos	21
2.3. Multibiometría	22
2.3.1. Tipos de multibiometría	22
2.3.2. Modos de operación	24

2.3.3.	Niveles de fusión	24
2.4.	Reconocimiento basado en rostro	27
2.5.	Reconocimiento basado en zona peri-ocular	27
2.6.	Extracción de características	28
2.6.1.	Análisis de componentes principales	28
2.6.2.	Análisis de componentes lineales	29
2.6.3.	Patrones locales binarios	30
2.7.	Técnicas de fusión	31
2.7.1.	Fusión a nivel de características	31
2.7.2.	Fusión a nivel de puntajes	32
2.7.3.	Fusión a nivel de decisión	34
2.8.	Estado del arte en fusión de reconocimiento rostro-periocular	35
3.	Implementación	37
3.1.	Bases de datos	37
3.1.1.	AT&T face database	37
3.1.2.	Base de imágenes de zona peri-ocular	38
3.1.3.	Base de imágenes multimodal	39
3.2.	Reconocimientos unimodales	40
3.2.1.	Reconocimiento de rostro	41
3.2.2.	Reconocimiento de zona peri-ocular	42
3.3.	Reconocimiento multimodal	42
3.3.1.	Fusión por características	42
3.3.2.	Fusión por puntaje	42
3.3.3.	Fusión por decisión	44
3.4.	Sistema fusionado de decisión con puntaje válido	44
4.	Análisis y discusión de resultados	46

4.1.	Resultados del reconocimiento de rostro	46
4.1.1.	PCA	46
4.1.2.	LDA	46
4.1.3.	LBP	47
4.2.	Resultados para el reconocimiento de zona peri-ocular	47
4.2.1.	PCA	47
4.2.2.	LDA	47
4.2.3.	LBP	47
4.3.	Resultados para el reconocimiento multi-modal	48
4.3.1.	Fusión de características	48
4.3.2.	Fusión de puntajes	48
4.3.3.	Fusión de decisión	48
4.4.	Sistema fusionado de decisión con validación de puntajes	48
4.4.1.	Sin esquema de normalización	50
4.4.2.	Esquema de normalización MinMaxNorm	50
4.4.3.	Esquema de normalización zNorm	50
4.4.4.	Visualización de curvas ROC	51
4.5.	Estrategias de fusión	51
5.	Conclusiones y trabajo futuro	64
	Bibliografía	65

Índice de figuras

2.1. Diagrama indicando el porcentaje del universo de muestras dividido según la Precision y el Recall.	16
2.2. Direcciones de mejor desempeño según la curva de evaluación. Verde indica mejora, mientras que rojo indica pérdida de desempeño. El gráfico de la izquierda corresponde a una curva ROC, mientras que a la derecha se observa la curva precision-recall	17
2.3. Diferentes esquemas de fusión de sistemas biométricos multimodales. Figura tomada de [17].	26
2.4. Ejemplo de codificación LBP utilizando los 8 pixeles adyacentes.	30
2.5. Concatenación de histogramas en imagen con 20 secciones.	31
2.6. Fusión simple de características: Características concatenadas	32
2.7. Fusión de puntajes: se crea un puntaje $\bar{S} = f(s_1, \dots, s_n)$ donde n corresponde al número de puntajes obtenido por los clasificadores uni-modales.	33
2.8. Fusión de decisión: se obtiene una decisión final utilizando las predicciones de cada clasificador por separado. En este caso, las características y los puntajes quedan iguales, ninguno de los sub-sistemas es intervenido. Sólo importa su salida.	34
2.9. Muestra de las imágenes utilizadas para reconocimiento periocular versus reconocimiento de rostro para diferentes expresiones. Figura extraída de [18]	35
2.10. Resultados obtenidos en [18] para reconocimiento periocular versus reconocimiento de rostro para diferentes expresiones	36
3.1. Muestra de la base de imágenes de rostro AT&T [15]	38
3.2. Imagen de alta resolución capturada con cámara AVIGILON 16MP ProHD.	39

3.3.	Recortes de zona peri-ocular utilizados para el trabajo, se cuenta con 971 imágenes separadas en 52 clases diferentes.	39
3.4.	Diagrama de bloques de la solución unimodal. Luego de ingresar las imágenes se extraen las características con 3 métodos diferentes, posteriormente se crea un modelo a partir de los individuos enrolados para finalmente medir la distancia de cada nueva predicción a cada elemento de la base de datos en el espacio de características.	41
3.5.	Diagrama de bloques de la fusión de puntajes. Se hacen 3 pruebas en total, una para cada esquema de normalización (sin norma, minMaxNorm y zNorm). Los resultados se muestran en la sección 4.3.	43
3.6.	Diagrama de bloques de la fusión por decisión. Los resultados se muestran en la sección 4.3.	44
3.7.	Diagrama de bloques del sistema fusionado de decisión con puntaje válido.	45
4.1.	Curva ROC y Precision-Recall para el sistema de reconocimiento de rostro sin normalizar puntajes, utilizando PCA como método de extracción de características según los parámetros en 3.2.	53
4.2.	Curva ROC y Precision-Recall para el sistema de reconocimiento de rostro sin normalizar utilizando LDA como método de extracción de características según los parámetros en 3.2.	54
4.3.	Curva ROC y Precision-Recall para el sistema de reconocimiento de rostro sin normalizar utilizando LBP como método de extracción de características según los parámetros en 3.2.	55
4.4.	Curva ROC y Precision-Recall para el sistema de reconocimiento de zona periocular sin normalizar utilizando PCA como método de extracción de características según los parámetros en la sección 3.2.2.	56
4.5.	Curva ROC y Precision-Recall para el sistema de reconocimiento de zona periocular sin normalizar utilizando LDA como método de extracción de características según los parámetros en la sección 3.2.2	57
4.6.	Curva ROC y Precision-Recall para el sistema de reconocimiento de zona periocular sin normalizar utilizando LBP como método de extracción de características según los parámetros en la sección 3.2.2.	58
4.7.	Curvas ROC y Precision-Recall para el sistema de reconocimiento de características concatenadas, sin normalizar.	59

4.8. Curva ROC para el sistema de reconocimiento multimodal basado en moda y validación de puntaje utilizando el esquema de normalización minMaxNorm. . . . 60

4.9. Curva Precision-Recall para el sistema de reconocimiento multimodal basado en moda y validación de puntaje utilizando el esquema de normalización minMaxNorm. 60

4.10. Curva ROC para el sistema de reconocimiento multimodal basado en moda y validación de puntaje utilizando el esquema de normalización zNorm. 61

4.11. Curva Precision-Recall para el sistema de reconocimiento multimodal basado en moda y validación de puntaje utilizando el esquema de normalización zNorm. . . . 61

4.12. Comportamiento de los verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos según el umbral para el sistema fusionado de decisión con validación de puntaje (minMaxNorm). Se observa el comportamiento no monótono de los falsos positivos. 62

4.13. Comportamiento de los verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos según el umbral para el sistema fusionado de decisión con validación de puntaje (zNorm). Se observa el comportamiento no monótono de los falsos positivos. 63

Índice de tablas

1.1. Diferentes softwares que contienen reconocimiento de rostro y sus respectivos desarrolladores.	9
4.1. Resumen de las áreas bajo la curva ROC y la curva Precision-Recall para los sistemas de reconocimiento unimodales.	48
4.2. Tasa de aciertos para cada fusión de puntajes, en cada esquema de normalización según lo descrito en la sección 2.7.2.	49
4.3. Resumen de la tasa de aciertos y tasa de error en los diferentes métodos comparados con la moda de las decisiones	49
4.4. Resumen de las áreas bajo la curva ROC y la curva Precision-Recall para los diferentes esquemas de normalización en el sistema de reconocimiento por moda de decisión con validación por puntaje.	50

Capítulo 1

Introducción

1.1. Motivación

La biometría corresponde al estudio de medidas en base a diferentes características humanas, tales como huellas digitales, iris, rostro y voz, entre otros. El uso de biometría se ha extendido rápidamente en la industria a medida que la necesidad de identificar individuos de forma certera crece: el control de acceso por reconocimiento biométrico, la detección de rostros en cámaras fotográficas y la validación de transacciones bancarias con huellas digitales son sólo un ejemplo de lo importante que esta ciencia puede llegar a ser, y de lo inmersa que estará en la sociedad en un futuro no muy lejano. [3]

El reconocimiento de rostro es uno de los sistemas biométricos más populares. Empresas de todo rubro han comenzado a incluir tecnología basada en biometría para llevar a cabo las tareas más diversas: control de acceso, validación de identidad, auto-enfoque en cámaras inteligentes, filtros fotográficos inteligentes, ahorro de energía en computadores y celulares son algunos ejemplos de aplicaciones del reconocimiento de rostro [3]. Algunos software asociados a grandes empresas mundiales que utilizan el reconocimiento de rostro se describen en la tabla 1.1.

El uso de la zona peri-ocular como patrón biométrico es más reciente, los estudios realizados en [13] muestran que la zona peri-ocular tiene un uso potencial para la identificación de personas sin depender de los ojos o la boca, lo que es de interés en caso de imágenes en ambientes no controlados, donde la postura facial no siempre es fija. Este patrón biométrico es, por tanto, de gran interés en el estudio de tecnologías que permitan prescindir, o soportar, según sea el caso, al reconocimiento de rostro.

Todo sistema de identificación biométrica tiene una incertidumbre asociada. Esta incertidumbre está asociada a múltiples aspectos del problema tales como, por ejemplo, el ruido

Tabla 1.1: Diferentes softwares que contienen reconocimiento de rostro y sus respectivos desarrolladores.

Software	Desarrollador
digiKam	KDE
iPhoto	Apple
OpenCV	Willow Garage
Photoshop Elements	Adobe Systems
Picasa	Google
Picture Motion Browser	Sony
Windows Live Photo Gallery	Microsoft
Lightroom	Adobe

intrínseco asociado a las mediciones de los sensores o la baja disponibilidad de información en la característica biométrica asociada al sistema. La multibiometría, definida como la fusión de uno o más sistemas biométricos, aparece como una alternativa para reducir incertidumbre y aumentar la eficiencia total del sistema [14].

1.2. Alcances

En este trabajo se propone una metodología basada en la fusión de información de un sistema de reconocimiento de personas basado en imágenes de rostro y un segundo sistema de reconocimiento basado en imágenes de zona peri-ocular, con el fin de estudiar el desempeño final de los sistemas por separado y el fusionado.

Se implementarán tres sistemas de reconocimiento de rostro junto a otros tres sistemas de reconocimiento de zona peri-ocular. Se fusionarán luego las bases de datos de cada sistema para crear una nueva base para un sistema de clasificación multimodal.

La implementación del trabajo se realizará en C++, utilizando la librería OpenCV y, más específicamente, la librería FaceRecognizer.

El resultado de cada método de fusión será comparado con el desempeño de los métodos de rostro y peri-ocular por separado, para posteriormente concluir sobre los resultados obtenidos.

1.3. Objetivo general

- Evaluar las estrategias de fusión de información provenientes de características extraídas de imágenes digitales del rostro y de la zona peri-ocular para la identificación de personas a distancia.

1.4. Objetivos específicos

- Estudiar la fusión de información de rostros frontales y de la zona periocular para identificar personas.
- Comparar los resultados de la fusión de información con la identificación obtenida por cada una de las técnicas, rostros y periocular, por separado.
- Evaluar la fusión de información a nivel de decisión y comparar con la fusión a nivel de características y a nivel de puntaje.

1.5. Estructura del trabajo

En el capítulo 1 se presenta un resumen de la problemática general que abarca este trabajo, así como la motivación, alcance y objetivos, generales y específicos.

En el Capítulo 2 se describen los conceptos necesarios para entender el problema a abordar y la implementación realizada, todo esto con la intención de dar al lector una mejor contextualización para comprender y conocer la teoría tras este trabajo. Se explica ampliamente el concepto de multi-biometría y se hace una revisión del estado del arte en reconocimiento de zona peri-ocular y rostro. Dentro de los conceptos a abordar se encuentran los métodos de extracción de características: análisis de componentes principales, análisis de discriminante lineal y patrones locales binarios. Así como también métodos de fusión de características, puntajes y decisiones.

En el capítulo 3 se muestra la implementación de las pruebas realizadas tanto para la fusión a nivel de características, puntaje y decisión. Así como una descripción de los algoritmos realizados.

En el capítulo 4 se muestran los resultados, en forma de gráficos, tablas y curvas de característica operativa del receptor y de Precision-Recall, se discute además sobre el resultado obtenido y se determina la mejor solución según un contexto dado.

Finalmente, el capítulo 5 presenta las conclusiones del trabajo junto a una lista con posibles consideraciones de trabajo a futuro siguiendo esta línea.

Capítulo 2

Marco teórico

2.1. Reconocimiento de patrones

El reconocimiento de patrones es el área de la ciencia que se enfoca en clasificar objetos en un número de categoría o *clases*. Dependiendo de la aplicación, estos objetos pueden ser imágenes, señales de onda o cualquier tipo de medición que requiera ser clasificada. [16]

El reconocimiento de patrones se categoriza generalmente de acuerdo al tipo de aprendizaje utilizado para generar el valor de salida. El aprendizaje supervisado asume que un conjunto de datos de entrenamiento (en adelante conjunto de entrenamiento) se ha entregado como dato, este conjunto consiste en un set de instancias que han sido propiamente catalogadas a mano con el resultado correcto (conocido como objetivo). Posteriormente un procedimiento de aprendizaje genera un modelo que tratará de cumplir los siguientes objetivos: Trabajar lo mejor posible en el conjunto de entrenamiento, y ser lo más general posible ante nuevos datos (generalmente, esto significa ser lo más simple posible). El aprendizaje no supervisado, por otro lado, trabaja con datos que no han sido catalogadas a mano e intenta encontrar patrones inherentes a los datos mismos que pueden ser usados para determinar el resultado para nuevas instancias de datos.

Ejemplos de métodos supervisados son: redes neuronales, máquinas de soporte vectorial (*support vector machines*, en inglés), clasificadores bayesianos y distintos métodos de kernel [12].

Ejemplos de métodos no-supervisados son: clustering, k-means, c-means, fuzzy c-means y self-organizing maps [12].

2.1.1. Clasificación

Cuando se habla de un clasificador, generalmente se hace referencia al conjunto de algoritmos que permiten obtener alguna etiqueta, valor o clasificación desde un conjunto de datos con un orden.

Espacio de características: Una muestra a clasificar, puede ser representada como un punto en un espacio n -dimensional, donde n está determinado por el número de características que describen las muestras, dado esto, muestras similares se encontrarán en zonas similares del espacio, lo que permite clasificar y agrupar muestras usando diferentes medidas de distancia. A este espacio n -dimensional, se le conoce generalmente como espacio de características [8].

Medidas de distancia: Una vez representado cada elemento dentro del espacio de características, podemos utilizar distintas técnicas para obtener el grado de similitud entre ellas, siendo la más común la medida de distancia. La idea de clasificar usando medidas de distancia es bastante simple: a menor distancia en el espacio de características, mayor similitud entre las muestras. La clasificación se puede realizar utilizando cualquier medida de distancia, aunque diferentes distancias conllevarán a distintos resultados [8].

2.1.2. Desempeño de un clasificador [14]

El objetivo de un sistema de clasificación es identificar, o a veces simplemente verificar, de forma correcta una muestra entre un listado de clases. Sin embargo, dicha clasificación puede ser correcta o no, dependiendo de que tan bueno sea el sistema. Para evaluar el desempeño de un clasificador, es crucial definir primero los siguientes conceptos:

- **Ground truth:** Es la etiqueta de pertenencia de cada elemento en el conjunto de entrenamiento a una clase particular. Este etiquetado se hace generalmente a mano y se asume como un dato verdadero y absoluto para el clasificador.
- **Verdadero Positivo (VP):** Es cuando el sistema asigna a la muestra de entrada una clase de forma correcta, coincidiendo con la clase asignada en el *ground truth*.
- **Falso Positivo (FP):** Es cuando el sistema asigna a la muestra de entrada una clase, no coincidiendo con la clase asignada en el *ground truth*.
- **Verdadero Negativo (VN):** Es cuando el sistema indica que la muestra ingresada no es válida y no pertenece a ninguna clase cuando la muestra efectivamente no pertenece a ninguna clase del conjunto de entrenamiento. Suele llamarse clase impostora a este tipo de clases.
- **Falso Negativo (FN):** Es cuando el sistema indica que la muestra ingresada es impostora, siendo que pertenece al conjunto de clases de entrenamiento.

Medidas de desempeño

Para las definiciones que siguen a continuación, se usará la siguiente nomenclatura:

- T_p : Verdadero positivo
- F_p : Falso positivo
- T_n : Verdadero negativo
- F_n : Falso negativo

Se tienen por tanto, las siguientes definiciones:

- **Tasa de verdaderos positivos (True Positive Rate, TPR):** también llamada **sensitividad**, mide la proporción entre la cantidad de positivos detectados y positivos correctos. Se calcula de la forma $\frac{T_p}{T_p+F_n}$.
- **Tasa de falsos positivos (False Positive Rate, FPR):** también llamada *especificidad*, mide la proporción entre la cantidad de falsos detectados y falsos correctos. Se calcula de la forma $\frac{F_p}{F_p+T_n}$.
- **Curva ROC:** La relación entre sensitividad y especificidad permite obtener un punto de operación para un sistema de clasificación. Esto resulta útil en problemas donde el costo de detectar incorrectamente, o de no detectar, son no despreciables. En un control de acceso, por ejemplo, puede ser más importante no dar a acceso a un posible sospechoso, que dar acceso a un posible positivo. Para visualizar estos puntos de operación se utilizan las curvas ROC, llamadas así por su sigla en inglés *Receiver Operating Characteristic*. Un buen sistema de clasificación tenderá a presentar una Sensitividad alta, manteniendo una especificidad lo más baja posible. En la figura 2.2 se muestra la dirección de crecimiento de desempeño en una curva ROC. El punto de operación ideal corresponde a la posición (0, 1) del plano, es decir *Sensitividad* = 1 y *Especificidad* = 0.
- **Precisión:** Fracción de las clasificaciones totales que pertenecen al conjunto de clases de entrenamiento. En términos matemáticos, corresponde a la relación entre los verdaderos positivos T_p y el número de positivos totales en el universo de muestras, $T_p + F_p$ [11].
- **Exhaustividad:** También llamado Recall, término que se usará a lo largo de este trabajo. Fracción de individuos que fueron correctamente clasificados del universo de muestras positivas. En la figura 2.1 se muestra una explicación gráfica de los conceptos de precision y recall, en conjunto en tonos de verde corresponde al universo de los elementos positivos (elementos cuyas clases existen en el conjunto de entrenamiento) mientras que en tonos de rojo se encuentran los elementos perteneciente al universo de elementos irrelevantes (impostores). El recall se calcula siguiendo la fórmula $\frac{T_p}{T_p+F_n}$ [11].

- **Curva Precision-Recall:** Unifica la información obtenida de la precisión y el recall y permite elegir un punto de operación para un sistema de clasificación. El punto de operación ideal para un clasificador en la curva Precision-Recall corresponde al punto $(1, 1)$, es decir $Precision = 1$ y $Recall = 1$, que corresponde al punto en que todas las detecciones son relevantes y además se ha detectado todo el conjunto de muestras. En la figura 2.2 se observa la dirección de mejor desempeño tanto para la curva precision-recall como para la curva ROC descrita previamente.
- **Área bajo las curvas:** Es posible comparar el desempeño de un clasificador con otro mediante el área bajo la curva ROC o el área bajo la curva Precision-Recall. Una mayor área indica una mejor performance general ante diferentes puntos de operación.

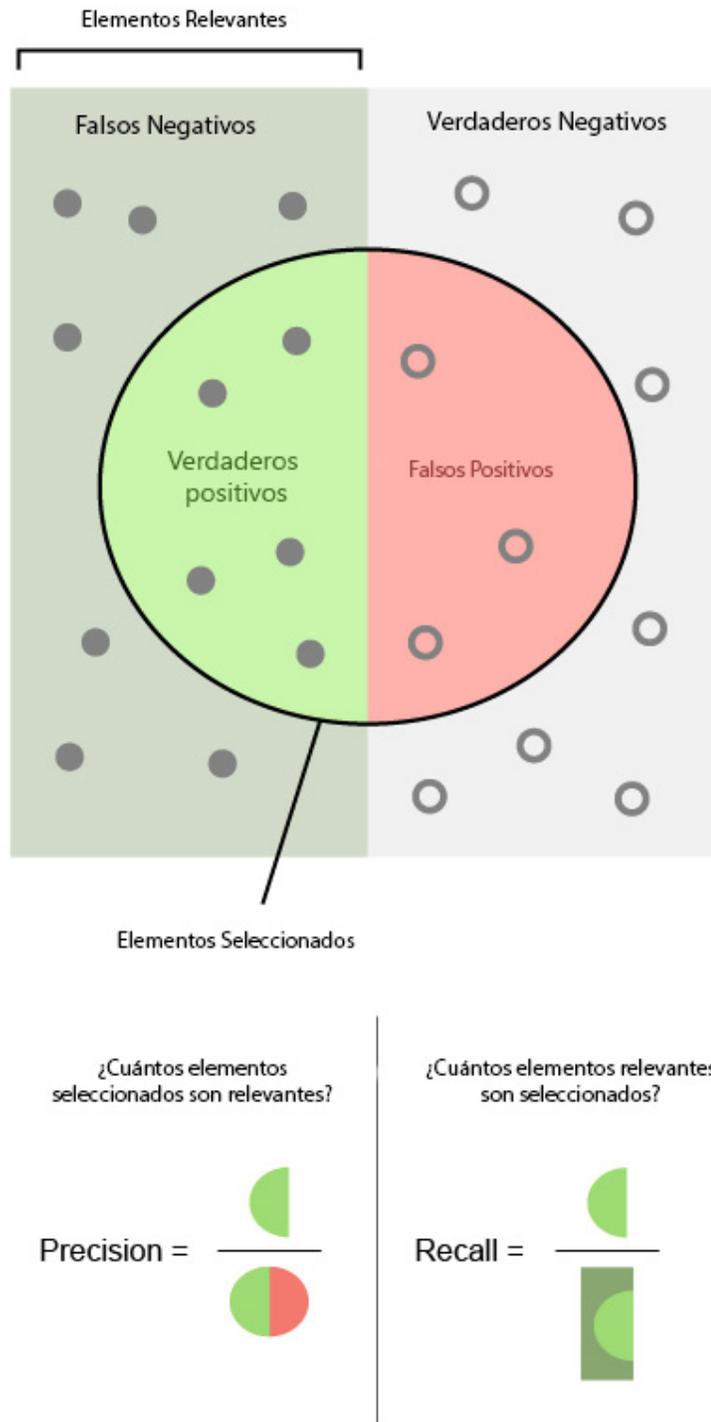


Figura 2.1: Diagrama indicando el porcentaje del universo de muestras dividido según la Precision y el Recall.

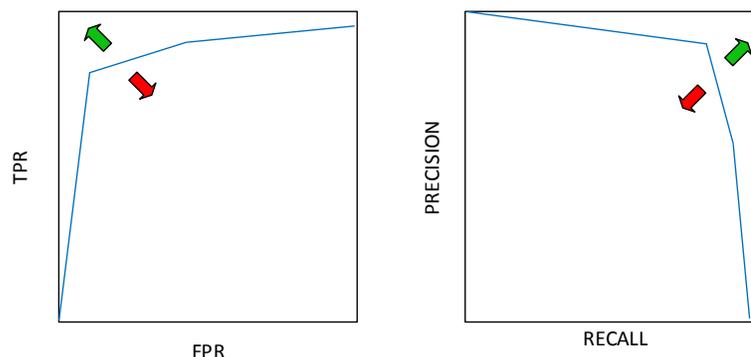


Figura 2.2: Direcciones de mejor desempeño según la curva de evaluación. Verde indica mejora, mientras que rojo indica pérdida de desempeño. El gráfico de la izquierda corresponde a una curva ROC, mientras que a la derecha se observa la curva precision-recall

2.1.3. Errores de adquisición de datos

La eficiencia o la precisión de un sistema biométrico es dependiente de los datos, usualmente influenciado por factores ambientales y de rendimiento. Los factores ambientales incluyen temperatura y humedad (que pueden afectar al sensor) y condiciones de iluminación (que afectan la medición), mientras que los factores de rendimiento incluyen calidad de la captura, composición de la población de usuarios, tiempo entre enrolamiento y verificación y robustez de los algoritmos de reconocimiento. Sin embargo, al ser fundamentalmente un sistema de reconocimiento de patrones, los sistemas biométricos pueden ser medidos en términos de adquisición de muestras y errores de desempeño, tal como se describen a continuación [17].

Condiciones ambientales adversas fomentan la adquisición errónea de datos en el sistema. Por ejemplo, (i) Enrolamiento defectuoso: que corresponde al porcentaje de veces en que el sistema rechaza la muestra debido a muestras ruidosas o de baja calidad y (ii) Captura defectuosa: esto se refiere a la inhabilidad de los sensores de obtener una muestra biométrica válida. Muy a menudo estos errores prevalecen en los sistemas que capturan y enrolan de forma automática. Una causa común de este tipo de error es la mala mantención de los sensores (i.e., suciedad y desperfectos).

2.2. Sistemas biométricos

La tarea más ardua de manejar en un sistema de identificación es la determinación (o verificación) de la identidad (o la supuesta identidad) de un individuo. Esta acción puede ser necesaria por diversas razones, pero la principal de ellas es evitar que impostores tengan acceso

a recursos protegidos. Los métodos tradicionales para establecer la identidad de una persona incluyen los métodos basados en conocimiento (e.g., contraseñas) y los basados en posesiones (e.g., tarjetas de identificación), el problema con estos métodos es que pueden ser fácilmente perdidos, compartidos, manipulados o robados, debilitando la seguridad que se busca [8, 14].

La biometría ofrece una solución confiable a algunos aspectos del manejo de identidades al utilizar esquemas automáticos o semi-automáticos para reconocer individuos basándose en sus características físicas y/o conductuales [14]. Utilizando biometría es posible establecer una identidad basada en la persona misma (o bien, en quien *es* esa persona) en lugar de centrarse entorno a lo que la persona posee o recuerda. En algunas aplicaciones, la biometría puede ser usada como complemento a las contraseñas y tarjetas de identificación, añadiendo un nivel de seguridad adicional, lo que es llamado en ocasiones autenticación con estrategia de factores duales. Aún así, la biometría no tiene por qué reemplazar a las tarjetas ni a las contraseñas en todas las aplicaciones

Los sistemas biométricos usan una variedad de características físicas o conductuales, incluyendo huellas digitales, rostro, geometría de la mano, iris, retina, firma, palmas, voz, oreja e información de ADN, entre otros, para establecer la identidad de un sujeto. En la literatura asociada, se le llama a cada una de estas características una modalidad. Mientras los sistemas biométricos tienen sus propias limitantes, tienen una ventaja sobre los métodos de seguridad tradicionales al no ser fácilmente robados o compartidos, sin contar además la conveniencia de eliminar la necesidad de diseñar y recordar contraseñas, o de construir y movilizar objetos de identificación [14].

2.2.1. Operación de un sistema biométrico

Un sistema biométrico es esencialmente un sistema de reconocimiento de patrones que obtiene datos biométricos desde un individuo, extrae un conjunto de características de dichos datos, los compara contra conjuntos de características almacenados en bases de datos, y finalmente ejecuta una acción basado en el resultado de la comparación. Dicho esto, un sistema biométrico genérico se puede ver como si estuviera conformado por cuatro módulos principales: un módulo de sensor; un módulo de extracción de características; un módulo de comparación (o de pareo); y un módulo de bases de datos [14]. Cada módulo se describirá a continuación:

1. **Módulo de sensor:** Un lector biométrico es requerido para adquirir los datos biométricos sin procesar de un individuo. El sensor define entonces la interfaz humano-máquina y es, por lo tanto, crucial en el desempeño del sistema. Una interfaz pobremente diseñada puede resultar en una baja aceptación por parte del usuario. Dado que la gran mayoría de las modalidades biométricas son adquiridas como imágenes, la calidad de los datos sin procesar también se ve impactada por las características de la tecnología de la cámara que se esté utilizando.
2. **Módulo de extracción de características:** Primero se analiza la calidad de los datos adquiridos por el sensor para determinar si es posible un procesamiento posterior.

Normalmente, los datos obtenidos son pre-procesados para obtener una mejor calidad. Sin embargo, ante datos pobres en calidad, se puede solicitar al usuario presentar los datos biométricos nuevamente. Los datos biométricos son posteriormente procesados y un conjunto de características distintivas son extraídas para representar la información relevante. Durante un enrolamiento en el sistema, estos conjuntos de características son almacenados en bases de datos y son referidos comunmente como *templates*

3. **Módulo de comparación y toma de decisiones:** Las características extraídas son comparadas contra los templates almacenados para generar puntajes de comparación (o *match scores*). El puntaje puede depender de la calidad de los datos adquiridos con el sensor. Este módulo además encapsula el módulo de toma de decisiones, en el cual los match scores son usados para validar una identidad presentada por el sujeto, o bien entregar un ranking de las identidades enroladas para así proceder a identificarlo.
4. **Módulo de base de datos:** La base de datos actúa como un repositorio de información biométrica. Durante el proceso de enrolamiento, los conjuntos de características extraídos desde la muestra biométrica (i.e., los templates) son almacenados en la base de datos junto con (la mayoría de las veces) cierta información biográfica del sujeto (nombre, numero de identificación, rut, dirección, etc.) caracterizándolo. Los datos obtenidos durante el enrolamiento pueden ser supervisados por un humano o no, dependiendo de la aplicación. Por ejemplo, un usuario intentado crear una nueva clave de acceso a su computador procederá a enrolarse y crear la cuenta de forma autónoma; por otro lado, el enrolamiento para el uso de un cajero automático activado por biometría requerirá de un funcionario capacitado para confirmar la identidad del sujeto como un paso de seguridad adicional.

Por otro lado, cabe mencionar que el template de un sujeto particular puede ser extraído desde una sola muestra biométrica (una sola imagen, o una sola unidad de tiempo con información de voz, por ejemplo) o bien utilizando múltiples muestras. Algunos sistemas almacenan múltiples templates para compensar las variaciones intra-clase asociadas a un sujeto dado. Los sistemas de reconocimiento facial, por ejemplo, pueden almacenar diferentes imágenes del rostro de un sujeto, con cada template correspondiendo a una pose diferente respecto a la cámara, de esta forma se incrementa la eficiencia del sistema ante variaciones en la postura del usuario al momento de identificarse [8, 14].

2.2.2. Verificación e identificación

Dependiendo del contexto, un sistema biométrico puede operar entre estados de verificación e identificación. El problema de verificación corresponde a la validación de la identidad de un sujeto comparando los datos biométricos capturados con el template en la base de datos correspondiente a quien el sujeto dice ser. En un sistema de este tipo, un individuo se presenta como alguien al sistema, generalmente mediante tarjetas inteligentes o con un sistema de usuario/contraseña,

y el sistema ejecuta posteriormente una comparación uno a uno con el template específico, determinando si la afirmación de identidad del sujeto es verdadera o no (esto también es conocido como reconocimiento positivo). El principal enfoque de este modo es impedir que múltiples sujetos compartan la misma identidad [8].

En el estado de identificación, el sistema reconoce a un sujeto buscando y comparando contra toda la base de datos, en una búsqueda uno-contra-todos, sin que el sujeto diga ser una persona en específico. Es un modo generalmente menos preciso y computacionalmente más costoso que el modo de verificación, pero resuelve el problema de reconocimiento negativo ya que se puede establecer si el sujeto es quien, implícita o explícitamente, niega ser. Con este sistema se puede impedir que un solo sujeto utilice múltiples identidades. Cabe destacar que, mientras los métodos tradicionales de contraseñas, llaves o tarjetas solucionan el problema de reconocimiento positivo, solamente un sistema biométrico tiene la capacidad de resolver el problema de reconocimiento negativo.

A lo largo de este trabajo, se usarán los términos genéricos *reconocimiento* o *autenticación* en los momentos en que no se desee hacer distinción entre los modos de verificación e identificación.

2.2.3. Características de una muestra biométrica

Actualmente, se utilizan varios modos biométricos diferentes en diversas aplicaciones. Cada modo tiene sus pros y contras y, por lo tanto, la elección de este está vinculado fuertemente a la naturaleza del problema. Según [9] Se han identificado siete factores claves para la idoneidad de una muestra física o conductual para ser usada en una aplicación biométrica. Dichos factores son:

1. **Universalidad:** Cada individuo que acceda a la aplicación deberá poseer la muestra.
2. **Unicidad:** La muestra deberá ser suficientemente diferente entre individuos acorde a la población comprendida por el sistema.
3. **Persistencia:** La muestra deberá ser suficientemente invariante sobre un periodo de tiempo respecto al algoritmo de comparación. Una muestra que cambia constantemente en el tiempo no es útil en biometría.
4. **Medición** Debe ser posible obtener y digitalizar la muestra biométrica usando dispositivos acordes que no causen inconveniente al sujeto. Además, los datos adquiridos deberán ser procesables para obtener conjuntos de datos representativos.
5. **Desempeño:** La precisión del reconocimiento y los recursos requeridos para alcanzar dicha precisión deben ser compatibles con las restricciones impuestas por la aplicación misma.
6. **Aceptación:** Los individuos que son objetivos para la aplicación deberán estar dispuestos a presentar sus muestras biométricas al sistema.

7. **Evasión:** Esto se refiere a la facilidad con que la muestra de un individuo puede ser imitada por otro utilizando artefactos (e.g., dedos falsos, fotografías).

2.2.4. Limitantes de los sistemas biométricos

Mientras los sistemas biométricos entregan varias ventajas sobre los sistemas de autenticación basados en objetos y contraseñas, es necesario considerar las limitaciones de estos sistemas al desarrollar aplicaciones en el mundo real que incluyan una gran cantidad de usuarios. Algunos de los problemas usualmente encontrados en los sistemas biométricos se muestran a continuación:

1. **Ruido en los datos adquiridos:** Huellas digitales con cicatrices, o una muestra de voz de un sujeto resfriado son ejemplos de muestras con ruidosas. Este ruido, además puede venir del sensor mismo, o de condiciones desfavorables del ambiente. Datos biométricos ruidosos pueden no ser satisfactoriamente asociados con sus templates correspondientes en la base de datos, lo que puede resultar en un sujeto siendo rechazado erróneamente.
2. **Variaciones intra-clase:** Estas variaciones son distintas del ruido normal y pueden ser creadas a partir de una mala interacción con el sensor (e.g., mala postura del rostro al adquirir los datos), o debidos a cambios en las características biométricas de una persona pasado un periodo de tiempo. Estas variaciones pueden ser corregidas almacenando múltiples templates de cada usuario. La actualización de los templates en el tiempo es también un componente esencial de cualquier sistema biométrico, sin embargo, es en la biometría conductual donde más variabilidad de este tipo existe, ya que está ligada al estado psicológico del individuo mismo, haciendolo más difícil de predecir [14].
3. **Similitudes inter-clase:** Esto se refiere al solapamiento del espacio de características correspondiente a múltiples clases o individuos. En un sistema de identificación que comprende un gran número de sujetos enrolados, la similitud inter-clase entre individuos subirá la tasa de falsas detecciones en el sistema. Por lo tanto, existe una cota superior para la cantidad de sujetos a enrolar en un sistema biométrico.
4. **No-Universalidad:** Es posible que el sistema no sea capaz de obtener los datos biométricos de un subconjunto de los usuarios. Es imposible, por ejemplo, obtener los datos de ambos iris de una persona tuerta, u obtener las huellas digitales de una persona a la que le ha sido amputado el brazo. En este contexto, es necesario que se considere la universalidad de los datos biométricos con los que se trabajarán en el sistema tratando de ser lo más general posible.
5. **Problemas de inter-operabilidad:** La calidad e idoneidad de los datos biométricos adquiridos dependen usualmente del sujeto mismo más que del sensor. Sin embargo, la elección del sensor, así como sus diferentes reemplazos y reparaciones, pueden tener un

fuerte impacto en el desempeño del sistema. Es necesario revisar que los datos capturados luego de un cambio de sensor sean procesables nuevamente, y además que sean capaces de obtener datos tales que el espacio de las características no varíe demasiado.

2.3. Multibiometría

Si bien la muestra de un sujeto está definida por un conjunto de características distintivas, en muchas instancias este conjunto puede ser descrito por un vector de características de un largo d fijo. El conjunto d -dimensional usualmente reside en un espacio en el que una multitud de vectores similares, representando a una multitud de sujetos independientes. Es por esto, que dado un número muy grande de individuos enrolados, el desempeño del clasificador disminuye en el tiempo; es por esto que la cantidad de individuos para un clasificador dado está restringida para una tasa de error dada. Aunque la sintonización correcta de parámetros, y la modelación correcta del problema y de la extracción de características pueden mejorar el desempeño de un clasificador, no es posible esperar una mejora continua e ilimitada [14].

Esta limitante puede ser aliviada fusionando información presentada por múltiples fuentes. Múltiples imágenes de un rostro; combinaciones de rostro, iris y huella digital; voz y forma de caminar; todas las combinaciones posibles pueden ser usadas para mejorar el desempeño del sistema de autenticación. La fusión de modos permite ampliar el espacio de características, lo cual sirve tanto para ampliar la cantidad de individuos que pueden ser enrolados como para hacer más separables los conjuntos actuales en la base [14].

Un sistema que consolida la evidencia presentada por múltiples fuentes biométricas es lo que se llama un sistema multibiométrico, o sistema biométrico multimodal. Se espera teóricamente que, bien implementado, un sistema multimodal sea más confiable debido a la mayor disponibilidad de información [14].

Sin embargo, no es necesario solamente unir las características, es posible integrar más tipos de información tales como, puntajes, datos sin procesar o decisiones del clasificador desde diferentes fuentes biométricas. El término biometría multimodal denota la fusión de diferentes tipos de información y es, por lo tanto, de un enfoque mucho más amplio que la biometría normal. El problema de consolidar esta información es conocido como *fusión de información*.

2.3.1. Tipos de multibiometría

El término multibiometría sin embargo, es mucho más amplio que lo anteriormente mencionado. Se presenta a continuación una breve explicación de cada tipo de multibiometría según el estudio realizado en [17]

Multi sensor

Los sistemas multisensor adquieren una muestra de una modalidad dada a través de diferentes sensores o unidades de captura. Esto se basa en el supuesto de que la información de diferentes sensores será complementaria, y por lo tanto más detallada, lo que elevará finalmente el desempeño del sistema. La adquisición multisensor además disminuye la posibilidad de obtener muestras distorsionadas o de mala calidad.

Aunque este esquema de sistema maneja eficientemente el problema de datos pobres o distorsionados, también resulta en un mayor costo total. Además, al requerir múltiples lecturas se necesita una mayor cooperación del usuario, lo que hace al sistema más invasivo. Se recomienda además que un sistema de estas características requiera de sensores que no requieran contacto físico, ya que el sentimiento de invasión en el usuario crece; además, las condiciones higiénicas que presentan los sensores por contacto pueden dificultar la tarea de tomar muestras debido a un alto rechazo de los usuarios.

Multi algoritmo

Estos sistemas utilizan diferentes métodos de extracción de características y algoritmos de pareo sobre una sola muestra, de una sola modalidad, adquirida con un solo sensor. Posteriormente, los resultados individuales de cada clasificador se combinan para obtener una decisión final. Datos pobres o distorsionados pueden afectar fuertemente la eficiencia del sistema. Además, se debe tener en cuenta que diferentes algoritmos entregarán además diferentes resultados en diferentes tiempos, lo que hace necesario un módulo de fusión, y además se quedará limitado por el algoritmo más lento.

Multi instancia

Estos sistemas, como su nombre indica, capturan diferentes instancias de una misma modalidad utilizando un solo sensor, desde diferentes ángulos y posiciones. Las múltiples instancias entregan información completa de la muestra a medir. Sin embargo, el hecho de que sea una muestra mono-sensor hace que este método sea susceptible a perturbaciones.

Multi muestra

Como el nombre sugiere, estos sistemas utilizan un solo sensor para adquirir múltiples ejemplos de un mismo modo biométrico. Un ejemplo de sistema utilizando este método sería un módulo para reconocimiento ocular que capture ambos ojos. Uno de los limitantes más importantes de estos sistemas son los tiempos de cómputo y el espacio requerido en la memoria.

Multi modo

Estos sistemas utilizan múltiples muestras biométricas adquiridas a través de diferentes sensores. Por ejemplo, iris y rostro combinados para establecer una identidad. Estos sistemas ofrecen mayor precisión y presentan la robustez necesaria ante perturbaciones en el sensor, mal manejo del sistema por parte del usuario o malas condiciones ambientales. Sin embargo, se requiere de métodos efectivos de reducción de dimensionalidad. Además, múltiples modalidades deben ser seleccionadas de diferentes regiones que contengan información biométrica.

2.3.2. Modos de operación

Un sistema de biometría multimodal puede funcionar en dos modos, según lo descrito en [14]:

- **Modo serial:** también llamado *en cascada*, las muestras biométricas son procesadas una tras otra, en secuencia. La salida de cada sub-sistema es utilizado además como entrada en el sub-sistema siguiente. Generalmente se utiliza un orden de complejidad decendiente, utilizando los primeros sub-sistemas para reducir el espacio de búsqueda, y los finales para distinguir de forma precisa rápidamente entre pocos sujetos de la base de datos.
- **Modo paralelo:** múltiples modalidades son procesadas simultáneamente y los resultados obtenidos son posteriormente combinados para un resultado final más preciso. Este tipo de arquitectura entrega generalmente mejores resultados, pero requiere a su vez de un mayor coste computacional.

2.3.3. Niveles de fusión

Para un sistema biométrico multimodal existen diferentes estrategias en cuanto a obtener el resultado se refiere. A continuación se mencionarán las estrategias principales, según [17], las cuales además pueden observarse en la figura 2.3.

- **Nivel de sensores:** Los datos sin procesar de diferentes sensores son combinados para formar un nuevo conjunto de datos, de los cuales se obtendrán las características.
- **Nivel de características:** Las características de cada modalidad se fusionan en un nuevo conjunto de características. Es importante que estas características sean independientes entre sí, y que además se encuentren en la misma escala de medición. Un ejemplo típico es la concatenación de vectores de características de rostro e iris.
- **Nivel de puntajes:** Los vectores de características son procesados de forma independiente por cada clasificador, las salidas de estos clasificadores son posteriormente fusionadas para determinar la identidad del sujeto en cuestión.

- **Nivel de decisión:** Múltiples clasificadores comparan el vector de características con los templates de la base de datos. Las salidas de cada uno son posteriormente fusionadas para llegar a una decisión final empleando técnicas como mayoría por votación.

Según la aplicación que se requiera, cada estilo de fusión presenta ventajas y desventajas. En general, los métodos de fusión de puntaje son los más eficientes en sistemas en tiempo real, mientras que los métodos de fusión de decisión presentan un mayor rendimiento general.

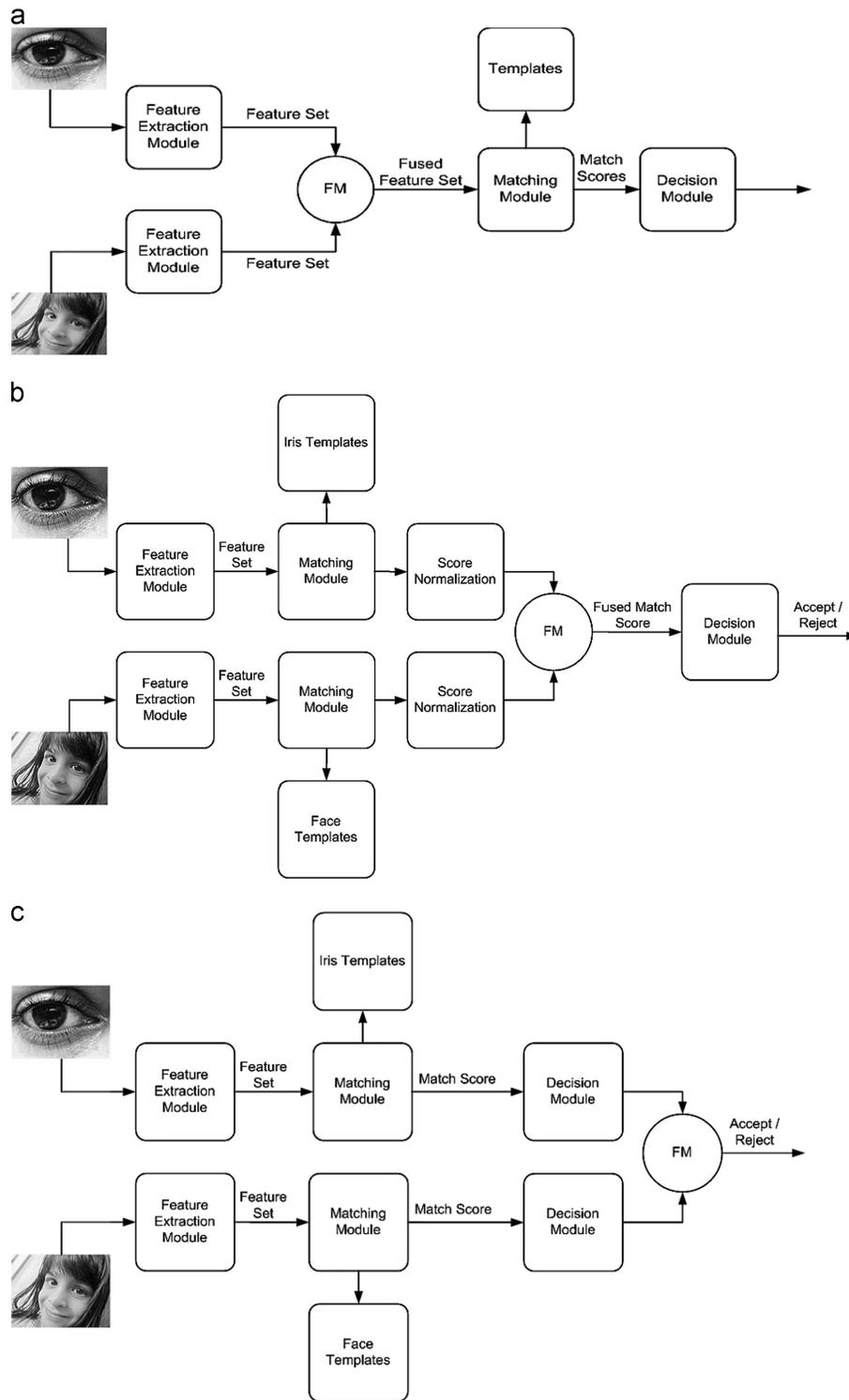


Figura 2.3: Diferentes esquemas de fusión de sistemas biométricos multimodales. Figura tomada de [17].

2.4. Reconocimiento basado en rostro

El reconocimiento de rostro es una tarea que los humanos realizan de forma rutinaria y sin esfuerzo en sus vidas. El reconocimiento de rostro tiene como motivación no solamente los desafíos que posee, sino que también las numerosas aplicaciones que este entrega en cualquier área donde se requiera identificar correctamente a las personas [8].

El reconocimiento de rostro se ha popularizado rápidamente como una de las tecnologías biométricas dominantes debido a sus grandes ventajas sobre el resto: Es intuitiva, fácil de usar por parte del sujeto, no intrusiva y además permite hacer reconocimiento de individuos en masa y a distancia. Se espera que un sistema de identificación de este tipo identifique rostros a partir de imágenes o videos de entrada, y que sea capaz de funcionar en modo de verificación o de identificación. Otros sistemas de este tipo añaden además una etapa previa de chequeo de seguridad, donde el rostro a identificar es comparado con templates específicos correspondientes a personas sospechosas o de acceso limitado, en una comparación uno-contra-pocos [8].

Se ha avanzado bastante desde el primer sistema de reconocimiento de rostro. desarrollado por Kanade [10], permitiendo incluso múltiples reconocimientos simultáneos o reconocimiento en tiempo real.

2.5. Reconocimiento basado en zona peri-ocular

La biometría ocular es ampliamente utilizada en sistemas biométricos debido a los avances en reconocimiento de iris. La invarianza en el tiempo, así como la gran cantidad de información contenida es la razón por la cual el iris es una de las modalidades biométricas más estudiadas, y es una de las más populares [13]. Sin embargo, el reconocimiento de iris puede presentar algunas limitantes:

- El iris es un objeto móvil respecto al rostro, lo cual hace difícil su encuadre.
- El iris es pequeño en comparación al resto del rostro, por lo que su obtención sin métodos controlados se dificulta.
- El iris es muy susceptible a la iluminación.
- El iris es difícil de identificar en luz visible (requiere de luz cercana al infrarojo para su detección óptima).

En [13] se define la zona periocular como una pequeña zona alrededor de los ojos, y propone el reconocimiento de zona periocular como una alternativa al reconocimiento de iris. Las fortalezas de esta modalidad por sobre iris son:

- Existen imágenes donde el iris no puede ser capturado, ya sea por tamaño o poca cooperación del usuario. La zona periocular, en cambio, puede ser adquirida siempre.
- La zona periocular representa un buen trade-off entre velocidad de cómputo y precisión.
- La zona periocular puede ser usada en diferentes resoluciones.
- Puede entregar información sobre la forma del ojo, lo cual puede ser usado como biometría débil (*soft biometrics*).
- Es más robusto a oclusiones.
- No es necesario rediseñar el módulo de captura, ya que rostro y periocular pueden ser obtenidos con la misma cámara.

La clasificación de esta modalidad se puede hacer en base a dos tipos de características diferentes:

- **Características de nivel uno:** Correspondiente a características de nivel geométricos y de forma. Estas características comprenden desde la forma de los ojos, a los lagrimales, cejas, pestañas, pliegues del párpado, etc.
- **Características de nivel dos:** Correspondiente a información general sobre la zona, son características de nivel más estadístico. Estas características comprenden desde histogramas de color, a textura de la piel, mediciones de iluminación, etc.

2.6. Extracción de características

2.6.1. Análisis de componentes principales

En adelante PCA, por sus siglas en inglés (*Principal Component Analysis*), busca resolver un problema de dimensionalidad en la clasificación de imágenes. Una imagen de rostro de, por ejemplo, 100×100 píxeles será representado como un vector de 10,000 elementos. Esto puede ser un problema a la hora de computar algoritmos de reconocimiento, por lo que surge la necesidad de reducir la dimensionalidad del problema disminuyendo lo menos posible su información. PCA transforma un conjunto de variables posiblemente correlacionadas en un conjunto más pequeño de variables no correlacionadas entre sí. La idea principal consiste en eliminar las dimensiones que aportan poca información al problema, dejando sólo las componentes realmente relevantes, de mayor varianza, llamadas componentes principales [2, 4].

Descripción algorítmica

Sea $X = \{x_1, x_2, \dots, x_n\}$ un vector aleatorio con muestras $x_i \in \mathbb{R}^d$ los pasos a seguir son:

1. Calcular la media $\mu = \frac{1}{n} \sum_{i=1}^n x_i$.
2. Calcular la matriz de covarianza $S = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)(x_i - \mu)^T$.
3. Calcular los valores propios λ_i y vectores propios v_i de S .
4. Ordenar los vectores propios según su valor propio asociado de forma descendente. Las K componentes principales corresponden a los vectores asociados a los K mayores valores propios.
5. Con esto, las K componentes principales para cualquier muestra x observada es $y = W^T(x - \mu)$, donde $W = (v_1, v_2, \dots, v_k)$.
6. Finalmente, la reconstrucción de la imagen desde la base PCA está dada por $x = Wy + \mu$.

El método de componentes principales efectúa el reconocimiento de la siguiente forma:

1. Proyecta todas las muestras de entrenamiento al sub-espacio de componentes principales.
2. Proyecta la imagen de entrada al sistema (prueba) al sub-espacio de componentes principales
3. Encuentra el vecino más cercano entre las imágenes de entrenamiento proyectadas y la proyección de la imagen de entrada.

2.6.2. Análisis de componentes lineales

El análisis de componentes lineales, también conocido como análisis de Fisher o análisis LDA, encuentra una combinación lineal de características que maximizan la varianza de los datos. Esta es una herramienta poderosa para minimizar la dimensionalidad de los datos pero, sin embargo, no necesariamente se incluye información respecto de las clases [5].

El análisis de discriminante lineal entrega una reducción de dimensionalidad específica por clases. Para encontrar la combinación de características que separa mejor las clases, el método de discriminante lineal maximiza la relación entre la varianza inter-clase y la varianza intra-clase. La idea resulta bastante directa, las imágenes de la misma clase deberían verse muy juntas en el espacio de características, mientras que las que pertenecen a distintas clases deberían estar alejadas entre si [2].

Descripción algorítmica

Sea $X = \{X_1, X_2, \dots, X_c\}$ un vector aleatorio con muestras obtenidas de c clases, con $X_i = \{x_1, x_2, \dots, x_n\}$ los pasos a seguir son:

1. Calcular las matrices de autocorrelación y correlación cruzada $S_B = \sum_{i=1}^c N_i (\mu_i - \mu)(\mu_i - \mu)^T$ y $S_W = \sum_{i=1}^c \sum_{x_j \in X_i} (x_j - \mu_i)(x_j - \mu_i)^T$ donde μ es la media total $\mu = \frac{1}{N} \sum_{i=1}^N x_i$ y μ_i es la media de cada clase $i \in \{1, \dots, c\}$, $\mu_i = \frac{1}{|X_i|} \sum_{x_j \in X_i} x_j$.
2. Se busca ahora una proyección W que maximice el criterio de separabilidad $W_{opt} = \operatorname{argmax}_W \frac{|W^T S_B W|}{|W^T S_W W|}$.

2.6.3. Patrones locales binarios

Los patrones locales binarios (en adelante **LBP**, por su sigla en inglés) cuantifican los patrones de intensidad en vecindades locales de píxeles y resultan útiles para identificar puntos, fines de líneas, esquinas, bordes y otros patrones de textura distintivos. Su uso en biometría es utilizado como método de extracción de características y ha sido usado en métodos como reconocimiento facial, reconocimiento de expresión y reconocimiento de iris [1].

Como el nombre sugiere, calcular el puntaje de LBP de un píxel involucra contar binariamente los cambios de intensidad relativos entre una vecindad de radio p y el píxel en cuestión. Restando el valor de cada píxel con el valor central, y binarizando la función signo del resultado se obtiene un set de valores que permiten codificar cada punto de la imagen en una tira de bits. El vector de características puede hacerse obteniendo cada valor individual del LBP en la imagen, como en la figura 2.4, o bien concatenando los histogramas de diferentes áreas, como indica la figura 2.5.

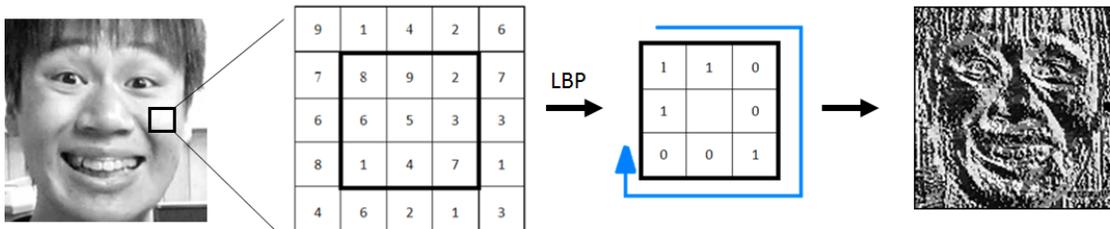


Figura 2.4: Ejemplo de codificación LBP utilizando los 8 píxeles adyacentes.

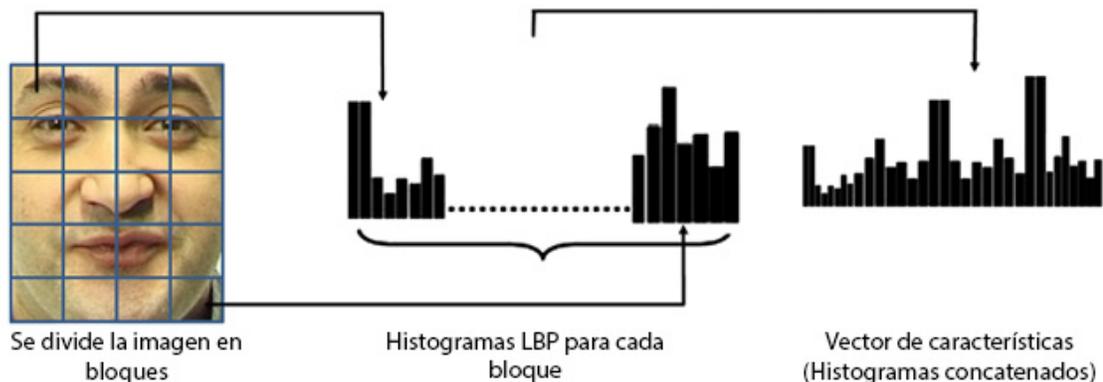


Figura 2.5: Concatenación de histogramas en imagen con 20 secciones.

2.7. Técnicas de fusión

Para fusionar información en un sistema biométrico, se puede ver el problema desde 3 posiciones claves: características, puntaje y decisión.

2.7.1. Fusión a nivel de características

El primer nivel de fusión que se puede hacer en un sistema de clasificación está en las características. La fusión de características corresponde al pre y post-procesamiento de las características de cada sub sistema por separado para llegar a un nuevo conjunto de características que pasarán a un clasificador distinto [17].

El diagrama de la figura 2.6 muestra el orden de pasos antes de llegar al clasificador. En este trabajo, se considerará solo el caso de las características concatenadas.

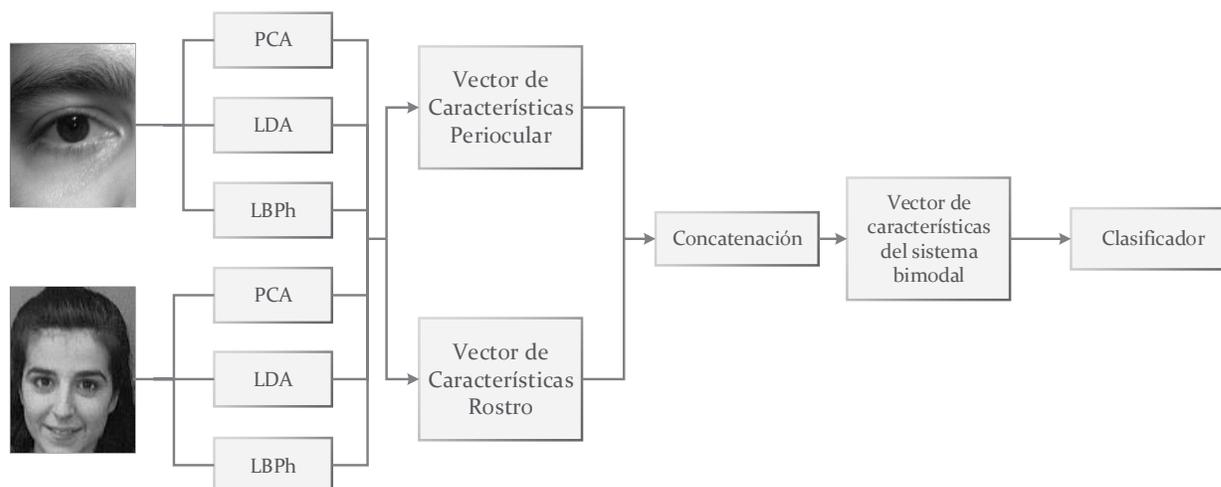


Figura 2.6: Fusión simple de características: Características concatenadas

2.7.2. Fusión a nivel de puntajes

La fusión de puntajes es, en su expresión más simple, determinar el grado de certeza de una predicción basado en diferentes grados de certezas asociados a sub-predicciones [17].

Un ejemplo simple sería determinar si una persona es quien dice ser basándose en 2 rasgos, imagen de rostro y huella digital. Un sistema de acceso por rostro indica positivo con 90% de certeza, mientras que la huella digital indica positivo con 98% de certeza. Usando una regla de suma simple, se podría estimar la certeza del sistema completo como el promedio de las certezas de sus sub-sistemas por separado. En este caso la certeza del sistema completo sería de 94%. En un segundo ejemplo, donde rostro y huella digital indiquen positivo con 90% y 30% de certeza respectivamente, la certeza del sistema completo sería de 60%, lo cual podría no ser aceptable, entregando finalmente una respuesta negativa. Otro tipo de fusión posible corresponde a seleccionar la respuesta del sistema que entregue mayor certeza, con lo cual se puede obtener una respuesta más cercana a lo deseable. La implementación de reglas de fusión de este estilo busca crear un sistema más robusto ante cambios externos donde, por ejemplo, una de las características podría no presentarse o ser obtenida con errores.

Si bien la fusión por puntaje se utiliza principalmente para la verificación en sistemas biométricos, puede ser adaptada para sistemas de identificación usando el nuevo puntaje para validar una predicción hecha por los sub-sistemas [17].

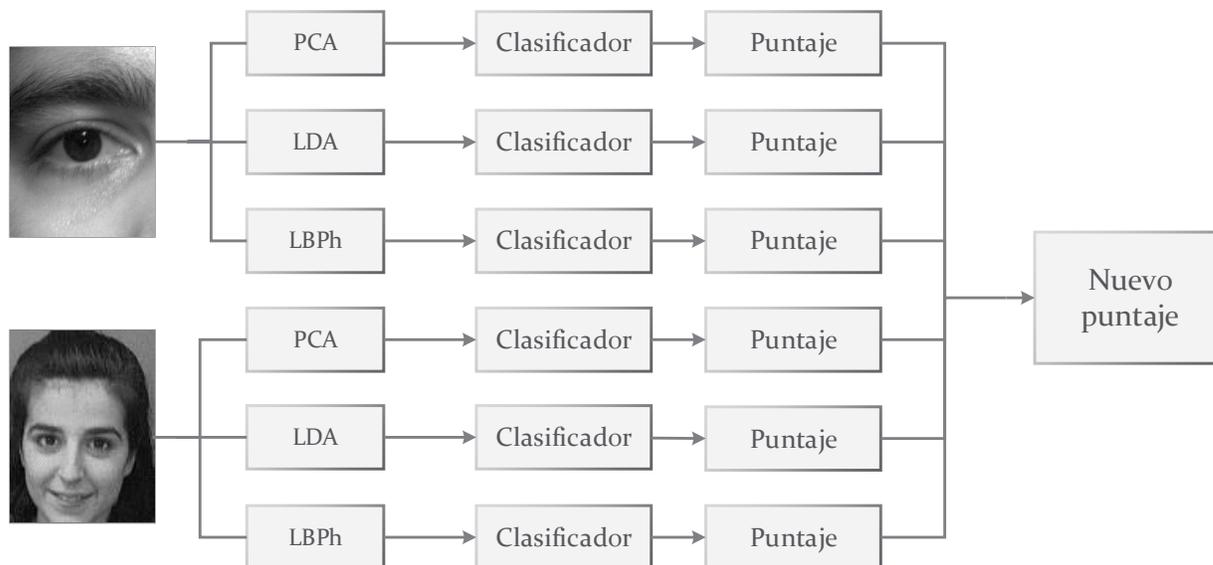


Figura 2.7: Fusión de puntuajes: se crea un puntaje $\bar{S} = f(s_1, \dots, s_n)$ donde n corresponde al número de puntuajes obtenido por los clasificadores uni-modales.

Esquemas de normalización

Un tema relevante durante la fusión de puntuajes es el esquema de normalización a usar. La diversa naturaleza de los puntuajes (que pueden ser distancias, probabilidades, frecuencias u otras medidas), hace que la normalización de los datos sea crucial para el desempeño correcto del sistema [6]. En este trabajo, se utilizarán dos tipos de normalización diferentes, descritos por las ecuaciones 2.1 y 2.2.

Normalización minMax: corresponde al estiramiento lineal de los datos, normalizándolos en el intervalo $[0, 1]$, la ecuación de la normalización se entrega en 2.1, donde \hat{S}_i corresponde al i -ésimo puntaje normalizado, S_i al i -ésimo puntaje sin normalizar, y donde $Min(S)$ y $Max(S)$ corresponden a los máximos y mínimos puntuajes del conjunto, respectivamente [6].

$$(2.1) \quad \hat{S}_i = \frac{S_i - Min(S)}{Max(S) - Min(S)}$$

Normalización z-score: Corresponde a ajustar los puntuajes como si pertenecieran a una distribución normal según la ecuación 2.2. En este caso, se tiene que \hat{S}_i corresponde al i -ésimo puntaje normalizado, S_i al i -ésimo puntaje sin normalizar, \bar{S} corresponde a la media del conjunto de puntuajes y $\sigma(S)$ a su desviación estándar [6].

$$(2.2) \quad \hat{S}_i = \frac{S_i - \bar{S}}{\sigma(S)}$$

2.7.3. Fusión a nivel de decisión

En este caso, las características y los puntajes de cada sub-sistema son irrelevantes. Ninguno de los sub-sistemas es intervenido a nivel de características o puntaje, lo que importa en este esquema de fusión es la salida de cada clasificador. En la figura 2.8 se entrega el diagrama de la solución fusionada en decisión. Es el esquema más simple de los 3 métodos de fusión estudiados en este trabajo [6].

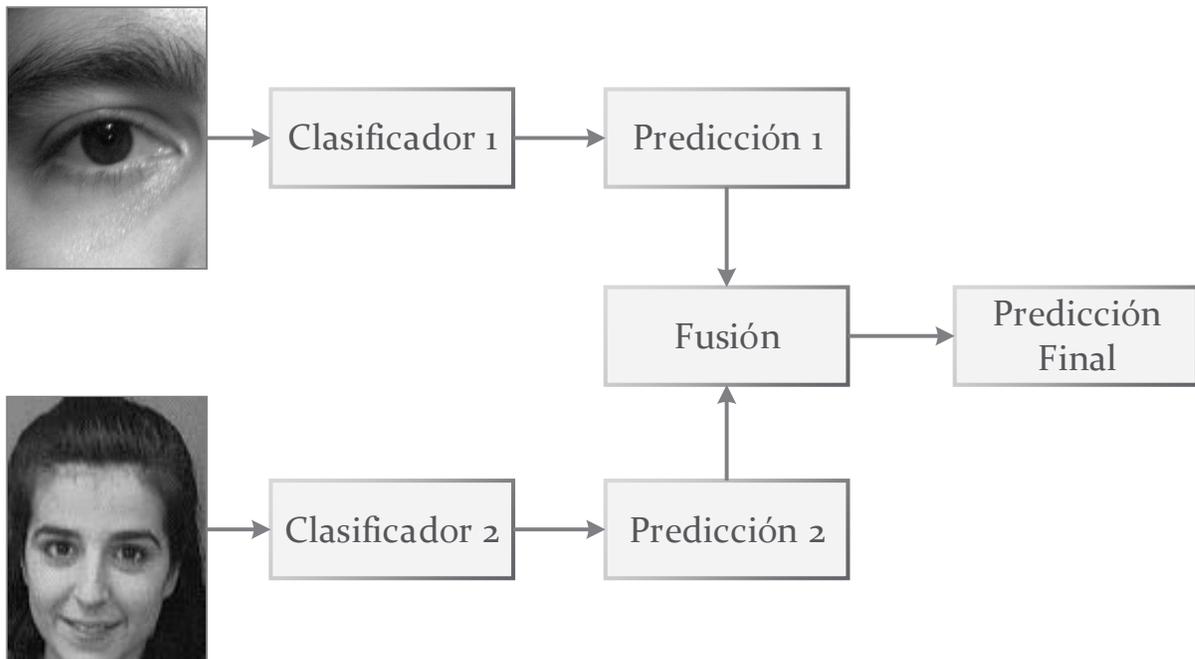


Figura 2.8: Fusión de decisión: se obtiene una decisión final utilizando las predicciones de cada clasificador por separado. En este caso, las características y los puntajes quedan iguales, ninguno de los sub-sistemas es intervenido. Sólo importa su salida.

2.8. Estado del arte en fusión de reconocimiento rostro-periocular

Esta sección se enfoca en el trabajo realizado por [18]. En este, se realizaron diversas pruebas para sistemas biométricos basados en reconocimiento periocular, así como también sistemas multimodales que incluían combinaciones de rostro, iris y zona periocular. Se trabajó con la base de datos FRGC (*Face Recognition Grand Challenge database*), de la cual se utilizaron imágenes obtenidas con 2 expresiones y 2 condiciones de iluminación diferentes, y se analizó el desempeño del reconocimiento periocular por sobre rostro, demostrando que se obtuvieron resultados similares para condiciones normales y que presenta un mejor desempeño ante cambios de expresión. Una muestra de la base de datos puede verse en la figura 2.9. La extracción de características se hizo con LBP y la clasificación con SVM. La tasa de reconocimiento de cada prueba se puede apreciar en la figura 2.10, el reconocimiento periocular presenta resultados similares ante un rostro neutro, y supera con creces al reconocimiento de rostro al cambiar la expresión del sujeto. En total se comparó zona periocular derecha, zona periocular izquierda, zona periocular derecha más zona periocular izquierda, y rostro completo.

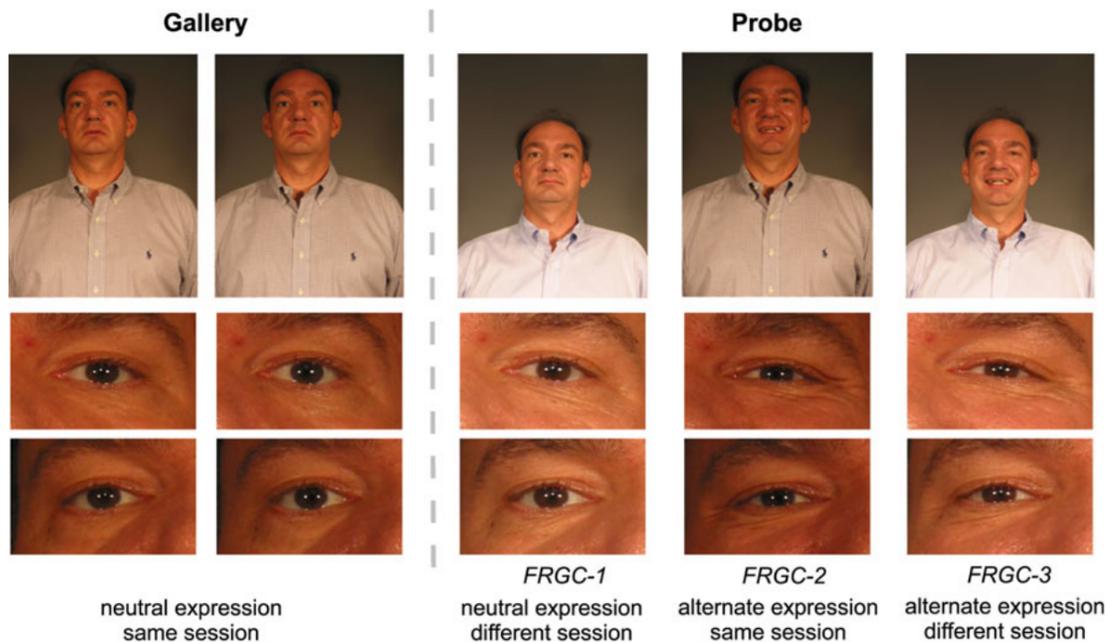


Figura 2.9: Muestra de las imágenes utilizadas para reconocimiento periocular versus reconocimiento de rostro para diferentes expresiones. Figura extraída de [18]

2.8. ESTADO DEL ARTE EN FUSIÓN DE RECONOCIMIENTO ROSTRO-PERIOCULAR

Table 1 Rank-1 recognition rate for the three FRGC experiments described in this paper. © [2010] IEEE

	<i>FRGC-1</i> : neutral expression (different session)			<i>FRGC-2</i> : alternate expression (same session)			<i>FRGC-3</i> : alternate expression (different session)		
	Texture	Color	Texture + Color	Texture	Color	Texture + Color	Texture	Color	Texture + Color
Left periocular	84.6	72.9	87.1	95.6	95.4	96.8	81.2	72.4	87.1
Right periocular	87.1	76.3	88.3	95.6	96.1	96.8	83.4	74.2	87.1
Left + right	91.5	82.9	91.0	97.6	97.6	98.3	88.8	82.0	91.2
Face	94.6	63.2	91.7	98.8	92.2	98.8	90.7	52.4	85.1

Figura 2.10: Resultados obtenidos en [18] para reconocimiento periocular versus reconocimiento de rostro para diferentes expresiones

Capítulo 3

Implementación

3.1. Bases de datos

Para este problema, se utiliza una base de datos de imágenes de zona peri-ocular obtenidas de imágenes de rostro de alta resolución, y una base de rostros en diferentes poses y expresiones.

3.1.1. AT&T face database

La base de datos AT&T contiene un set de imágenes de rostro tomadas entre Abril de 1992 y Abril de 1994. Es generalmente usada en el contexto de reconocimiento de rostro, y fue desarrollada en conjunto por los laboratorios de Voz, Visión y Robótica del departamento de ingeniería de la universidad de Cambridge.

La base consiste en imágenes de 40 sujetos distintos. Para algunos de ellos, las imágenes fueron tomadas a diferentes tiempos, diferentes luminosidades, diferentes expresiones (ojos abiertos, cerrados, sonriendo, no sonriendo) y diferentes detalles faciales (lentes). Todas las imágenes fueron tomadas contra un fondo oscuro homogéneo con los sujetos en posición recta y de frente (con tolerancia de movimiento hacia los lados). Los archivos se encuentran en formato PGM, de 92x112 pixeles cada una. Organizados en 40 directorios.



Figura 3.1: Muestra de la base de imágenes de rostro AT&T [15]

3.1.2. Base de imágenes de zona peri-ocular

La base de datos de imágenes de zona peri-ocular es propiedad del Laboratorio de Procesamiento de imágenes del Departamento de Ingeniería Eléctrica de la Universidad de Chile. Fue obtenida a partir de imágenes de rostro de alta resolución capturadas con una cámara Avigilon 16 MP Pro HD.

Para cada ojo en la base se realiza un recorte de 1000x1200 píxeles alrededor del centro del ojo (zona periocular). Ese recorte se redimensiona a 500x600 píxeles con fin de reducir el costo computacional de los algoritmos de reconocimiento. Finalmente, se aplica un filtro gaussiano de 11x11 píxeles para eliminar el ruido asociado a la medición.

Se considera para este estudio que ambas zonas peri-oculares, izquierda y derecha, son independientes y pueden ser tratados como sujetos diferentes. Con esto, se obtiene un total de 52 sujetos diferentes, o clases, distribuidas en 971 imágenes.



Figura 3.2: Imagen de alta resolución capturada con cámara AVIGILON 16MP ProHD.



Figura 3.3: Recortes de zona peri-ocular utilizados para el trabajo, se cuenta con 971 imágenes separadas en 52 clases diferentes.

3.1.3. Base de imágenes multimodal

Para el sistema multimodal se crea una base de datos compuesta, generalmente llamada base de individuos quiméricos, utilizando las imágenes de rostro de la base AT&T y las imágenes peri-oculares de la base del Laboratorio de Procesamiento de Imágenes. Una base de imágenes

quimérica corresponde, en el caso bimodal, a pares de imágenes asociados a una nueva clase $\bar{x} = (x_1, x_2)$ donde x_1 pertenece a la base AT&T y x_2 a la base de imágenes peri-oculares.

La base de datos AT&T tiene el menor número de individuos, por lo que la base de datos quimérica se conforma con 40 clases. El número de muestras por clase se calcula de la forma $|\bar{x}| = \min(|x_1|, |x_2|)$, a modo de obtener la mayor cantidad posible de muestras.

3.2. Reconocimientos unimodales

Los sistemas de reconocimiento unimodal se diseñan específicamente para este trabajo, tanto para el caso de rostro como para zona peri-ocular. En ambos casos se utiliza la librería FaceRecognizer de openCV. Es preciso recalcar que la medida de distancia utilizada para la clasificación es distancia euclidiana, esta distancia se utiliza en todos los clasificadores para ver con mayor claridad el efecto de las fusiones y su efecto en el reconocimiento.

Para cada sistema, se extraen características de cada imagen utilizando los métodos explicados en la sección 2.6, posteriormente se entrena el modelo correspondiente a cada método de extracción. Finalmente, para cada sistema, se calcula la curva ROC y la curva Precision-Recall. El diagrama de flujo de la solución unimodal se puede apreciar en la figura 3.4.

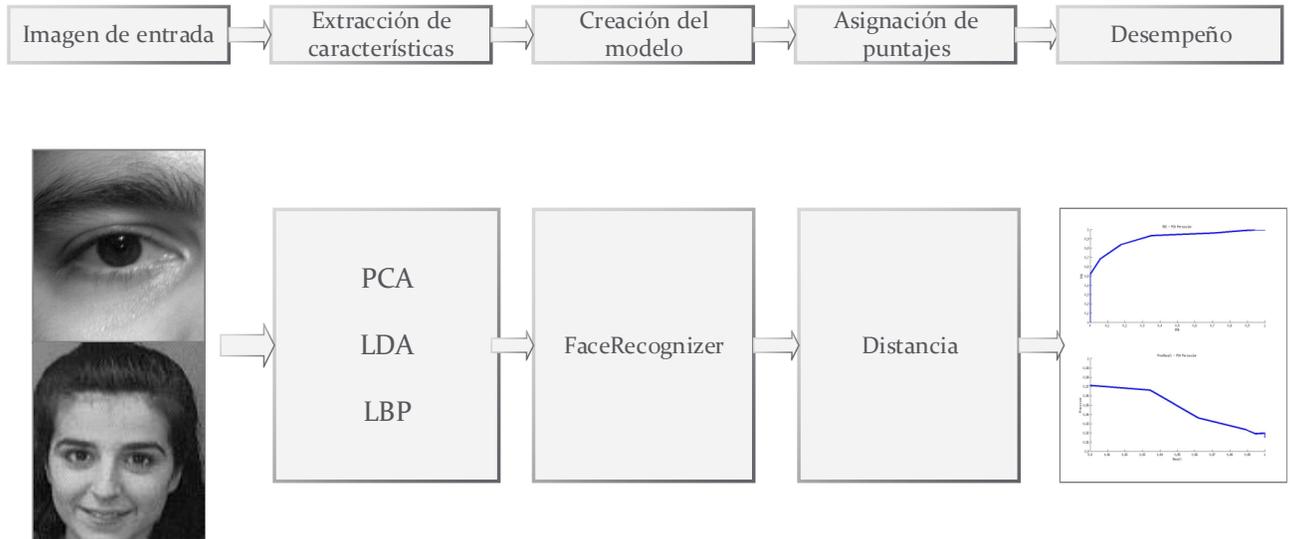


Figura 3.4: Diagrama de bloques de la solución unimodal. Luego de ingresar las imágenes se extraen las características con 3 métodos diferentes, posteriormente se crea un modelo a partir de los individuos enrolados para finalmente medir la distancia de cada nueva predicción a cada elemento de la base de datos en el espacio de características.

3.2.1. Reconocimiento de rostro

Para el sistema de reconocimiento de rostro se utilizan 3 métodos de extracción de características: PCA, LDA y LBP. Se utiliza la base de datos AT&T para el sistema, dejando un 50% de la base para el entrenamiento, y un 50% para las pruebas. Los métodos, que ya fueron descritos en la sección 2.6, utilizan los parámetros descritos a continuación.

PCA

Las características tomadas corresponden a los coeficientes asociados a cada vector propio en la base. Para esto, se realizan pruebas con el modelo utilizando entre 2 y 30 vectores propios. Los resultados deseados se obtienen con 10 vectores propios en el modelo.

Fisher

Las características tomadas corresponden a los coeficientes asociados a cada vector propio en la base. Para esto, se realizan pruebas con el modelo utilizando entre 2 y 15 vectores propios. Los

resultados deseados se obtienen con 10 vectores propios en el modelo.

LBPh

Las características obtenidas corresponden a los histogramas concatenados de distintos cuadrantes de la imagen, con un número fijo de vecinos por pixel y un radio también fijo. Para este trabajo se fijan los siguientes parámetros: 8 vecinos, radio de 1 pixel, 8 cuadrantes en la dirección x , 8 cuadrantes en la dirección y .

3.2.2. Reconocimiento de zona peri-ocular

Los parámetros usados para el sistema de reconocimiento de imágenes de zona peri-ocular son los mismos que para el sistema de reconocimiento de rostro. A fin de estandarizar las medidas de desempeño, para estos sistemas también se calculó la curva ROC, y la curva precision-recall.

3.3. Reconocimiento multimodal

El sistema de reconocimiento multimodal se implementa usando la base de datos quimérica descrita previamente en 3.1.3. Se separó el 50 % de la base para el entrenamiento, y el 50 % restante para pruebas.

3.3.1. Fusión por características

Para la fusión de características, tal como indica la figura 2.6, se realiza una concatenación de características para evaluar su impacto en el sistema de clasificación. Utilizando la misma medida que en los clasificadores unimodales, se implementa un sistema de reconocimiento para las características concatenadas de las imágenes quiméricas formadas a partir de las bases de datos descritas en las secciones 3.1.1 y 3.1.2 con un total de 32.809 características. Los resultados se muestran en la sección 4.3.

3.3.2. Fusión por puntaje

Para cada método de extracción de características, y para cada modo, se entrena el modelo correspondiente. Posteriormente, se compara cada imagen en el conjunto de prueba y se obtiene

su distancia a cada elemento del conjunto de entrenamiento en el espacio de las características. Se utiliza distancia euclidiana como medida de discriminación. Para cada método, se considera la distancia más corta como predicción y se utiliza el módulo de la distancia como puntaje.

Una vez calculada la predicción de cada método por separado, se normalizan los puntajes, para cada método, utilizando norma MinMax y zNorm, las cuales están explicadas en la sección 2.7.2. Una vez calculado el puntaje normalizado para cada método y cada elemento del conjunto de pruebas, se selecciona una predicción según el puntaje, utilizando los siguientes criterios:

- MinScore: Mínimo puntaje de entre los candidatos
- MaxScore: Máximo puntaje de entre los candidatos
- AbsMinScore: Mínimo puntaje absoluto de entre los candidatos
- AbsMaxScore: Máximo puntaje absoluto de entre los candidatos

El diagrama de bloques de la solución se muestra en la figura 3.5 . Mientras que el resultado de estos métodos se muestra en la sección 4.3.

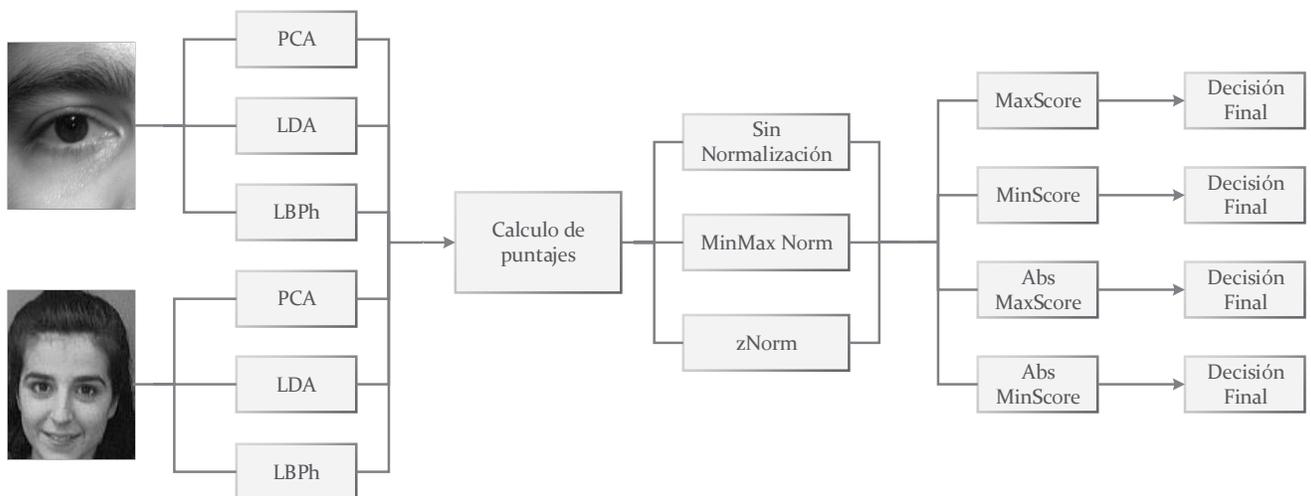


Figura 3.5: Diagrama de bloques de la fusión de puntajes. Se hacen 3 pruebas en total, una para cada esquema de normalización (sin norma, minMaxNorm y zNorm). Los resultados se muestran en la sección 4.3.

3.3.3. Fusión por decisión

Una vez obtenidas las predicciones de cada clasificador, se utiliza la moda de las predicciones, sin considerar el puntaje, como predicción final. El diagrama de bloques de la solución se muestra en la figura 3.6. Mientras que el resultado de estos métodos se muestra en la sección 4.3.

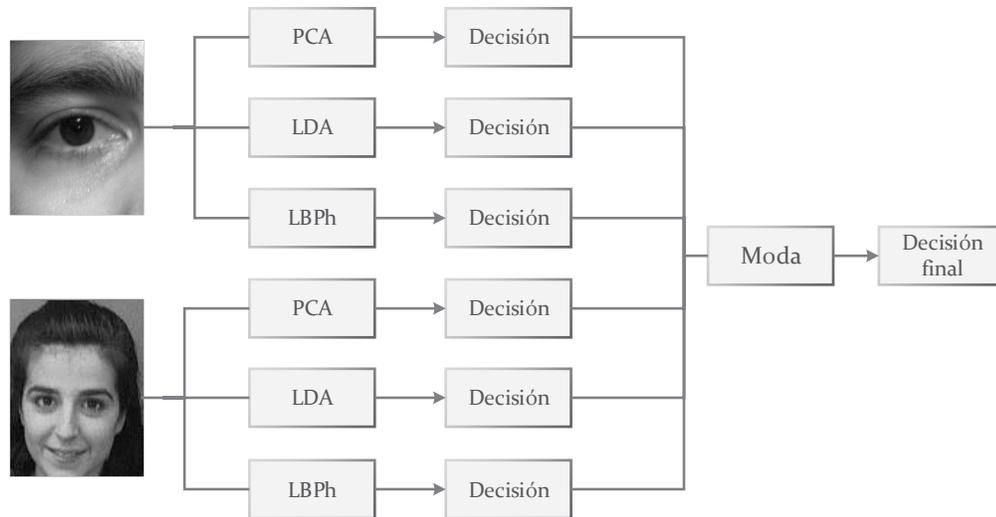


Figura 3.6: Diagrama de bloques de la fusión por decisión. Los resultados se muestran en la sección 4.3.

3.4. Sistema fusionado de decisión con puntaje válido

Una vez obtenida la predicción de cada sistema por separado, y sus respectivos puntajes, es posible hacer un nuevo sistema basado en las salidas de los sistemas anteriores. Para este caso, se adopta el siguiente esquema:

- La entrada del sistema corresponde a un par de imágenes rostro-periocular.
- Se extraen características de cada imagen utilizando PCA, LDA y LBP, obteniendo así 6 vectores de características.
- Se entrena un modelo basado en cada característica extraída, lo que genera un total de 6 modelos diferentes.
- Durante la prueba, para cada imagen entrante se obtienen los puntajes asociados a cada modelo, lo que entrega 6 puntajes diferentes. Los cuales son normalizados utilizando zNorm y minMaxNorm.

- Se aplica un umbral a cada puntaje, descartando cualquier puntaje bajo el umbral.
- Finalmente, la decisión final corresponde a la moda de los puntajes válidos. En caso de empate, se elige la decisión cuyo puntaje total sumado sea menor.

Con el método descrito, se hacen pruebas variando el umbral entre $\sigma = 0.01$ y $\sigma = 0.99$ con pasos $\Delta\sigma = 0.01$ para el caso minMax y sin normalización. El umbral para el caso zNorm se mueve entre $\sigma = -1$ y $\sigma = 1$ con pasos $\Delta\sigma = 0.01$.

Con la predicción fusionada, y su grado de certeza, es posible ahora calcular la curva ROC y la curva precision-Recall. considerando lo siguiente:

- **Verdadero Positivo:** Se validó el puntaje y la clasificación fue correcta.
- **Falso Positivo:** Se validó el puntaje y la clasificación fue incorrecta.
- **Verdadero Negativo:** La clasificación es incorrecta, pero no se validó el puntaje. La salida del clasificador es **no detectado**.
- **Falso Negativo:** La clasificación es correcta, pero el puntaje no se valida.

El diagrama de bloques de la solución se muestra en la figura 3.7 . Mientras que el resultado de estos métodos se muestra en la sección 4.4.

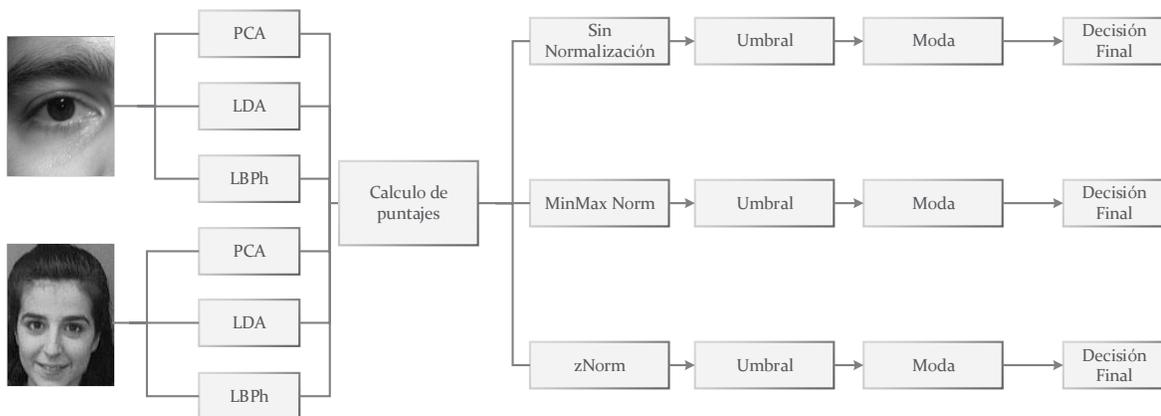


Figura 3.7: Diagrama de bloques del sistema fusionado de decisión con puntaje válido.

Capítulo 4

Análisis y discusión de resultados

En esta sección se discuten los resultados de cada etapa del trabajo, desde el desempeño del reconocimiento unimodal, hasta el desempeño de los sistemas multimodales. El desempeño de los clasificadores unimodales se mide usando curvas ROC y curvas precision-recall, así como las áreas bajo estas, según lo explicado en la sección 3.

4.1. Resultados del reconocimiento de rostro

Se entregan a continuación los resultados obtenidos para el sistema de reconocimiento de rostro utilizando la base de datos AT&T según los parámetros descritos en 3.2.

4.1.1. PCA

En la figura 4.1 se muestra la curva ROC y la curva Precision-Recall para el sistema de reconocimiento facial utilizando PCA como método de extracción de características. El área bajo la curva ROC es de $A_{ROC} = 0.9047$ mientras que el área bajo la curva Precision-Recall es $A_{PrecRecall} = 0.9805$.

4.1.2. LDA

En la figura 4.2 se muestra la curva ROC y la curva Precision-Recall para el sistema de reconocimiento facial utilizando LDA como método de extracción de características. El área bajo la curva ROC es de $A_{ROC} = 0.8338$ mientras que el área bajo la curva Precision-Recall es $A_{PrecRecall} = 0.9588$.

4.1.3. LBP

En la figura 4.3 se muestra la curva ROC y la curva Precision-Recall para el sistema de reconocimiento facial utilizando LBP como método de extracción de características. El área bajo la curva ROC es de $A_{ROC} = 0.9346$ mientras que el área bajo la curva Precision-Recall es $A_{PrecRecall} = 0.9831$.

4.2. Resultados para el reconocimiento de zona peri-ocular

4.2.1. PCA

En la figura 4.4 se muestra la curva ROC y la curva Precision-Recall para el sistema de reconocimiento de zona peri-ocular utilizando PCA como método de extracción de características. El área bajo la curva ROC es de $A_{ROC} = 0.8687$ mientras que el área bajo la curva Precision-Recall es $A_{PrecRecall} = 0.9656$.

4.2.2. LDA

En la figura 4.5 se muestra la curva ROC y la curva Precision-Recall para el sistema de reconocimiento de zona peri-ocular utilizando LDA como método de extracción de características. El área bajo la curva ROC es de $A_{ROC} = 0.9139$ mientras que el área bajo la curva Precision-Recall es $A_{PrecRecall} = 0.9944$.

4.2.3. LBP

En la figura 4.6 se muestra la curva ROC y la curva Precision-Recall para el sistema de reconocimiento de zona peri-ocular utilizando LBP como método de extracción de características. El área bajo la curva ROC es de $A_{ROC} = 0.8789$ mientras que el área bajo la curva Precision-Recall es $A_{PrecRecall} = 0.9884$.

El resumen de las áreas bajo las curvas se puede ver en la tabla 4.1.

Tabla 4.1: Resumen de las áreas bajo la curva ROC y la curva Precision-Recall para los sistemas de reconocimiento unimodales.

Método	A_{ROC}	$A_{Prec-Recall}$
PCA Rostro	0.9047	0.9805
LDA Rostro	0.8338	0.9588
LBP Rostro	0.9346	0.9831
PCA Periocular	0.8687	0.9656
LDA Periocular	0.9139	0.9944
LBP Periocular	0.8789	0.9884

4.3. Resultados para el reconocimiento multi-modal

4.3.1. Fusión de características

Para esta prueba, se obtiene la curva ROC y Precision-Recall para el esquema de características concatenadas, las cuales se muestran en la figura 4.7. El área bajo la curva ROC es de $A_{ROC} = 0.9440$ mientras que el área bajo la curva Precision-Recall es de $A_{PrecRecall} = 0.9887$, mientras que la tasa de aciertos de del 90.05 %.

4.3.2. Fusión de puntajes

Para esta prueba, se obtienen las tasas de acierto de cada clasificador unimodal por separado, para luego obtener la tasa de aciertos de los clasificadores considerando la fusión a nivel de puntaje utilizando los criterios descritos en la sección 3.3.2. Las tasas de acierto se entregan en la tabla 4.2.

4.3.3. Fusión de decisión

Para esta prueba, se obtienen las tasas de acierto de cada clasificador unimodal por separado, para luego obtener la tasa de aciertos de los clasificadores considerando la fusión a nivel de decisión utilizando los criterios descritos en la sección 3.3.3. Las tasas de acierto se entregan en la tabla 4.3.

4.4. Sistema fusionado de decisión con validación de puntajes

En esta sección se entregan los resultados del sistema de identificación basado en moda de decisión con validación de puntajes utilizando los esquemas de normalización descritos en la

Tabla 4.2: Tasa de aciertos para cada fusión de puntajes, en cada esquema de normalización según lo descrito en la sección 2.7.2.

Método	Tasa de aciertos	Tasa de error
PCA Rostro	0.9285	0.071429
LDA Rostro	0.814286	0.185714
LBP Rostro	0.921429	0.078571
PCA Periocular	0.95	0.05
LDA Periocular	0.957143	0.042857
LBP Periocular	0.971429	0.028571
MaxScore (sinNorm)	0.828571	0.171429
absMaxScore (sinNorm)	0.828571	0.171429
minScore (sinNorm)	0.978571	0.021429
absMinScore (sinNorm)	0.978571	0.021429
MaxScore (minMaxNorm)	0.757143	0.242857
absMaxScore (minMaxNorm)	0.757143	0.242857
minScore (minMaxNorm)	0.971429	0.028571
absMinScore (minMaxNorm)	0.971429	0.028571
MaxScore (zNorm)	0.814286	0.185714
absMaxScore (zMaxNorm)	0.350000	0.650000
minScore (zMaxNorm)	0.985714	0.014286
absMinScore (zNorm)	0.928571	0.071429

Tabla 4.3: Resumen de la tasa de aciertos y tasa de error en los diferentes métodos comparados con la moda de las decisiones

Método	Tasa de aciertos	Tasa de error
PCA Rostro	0.9285	0.071429
LDA Rostro	0.814286	0.185714
LBP Rostro	0.921429	0.078571
PCA Periocular	0.95	0.05
LDA Periocular	0.957143	0.042857
LBP Periocular	0.971429	0.028571
Moda	0.985714	0.014286

Tabla 4.4: Resumen de las áreas bajo la curva ROC y la curva Precision-Recall para los diferentes esquemas de normalización en el sistema de reconocimiento por moda de decisión con validación por puntaje.

Método	A_{ROC}	$A_{Prec-Recall}$
Sin normalizar	0.8782	0.9676
minMaxNorm	0.8782	0.9676
zNorm	0.9499	0.9989

sección 3.4. El diagrama de bloques de la solución se encuentra en la figura 3.7.

4.4.1. Sin esquema de normalización

Se hacen pruebas según los parámetros descritos en 3.4 variando el umbral entre $\sigma = 0$ y $\sigma = 500$ utilizando un paso $\Delta\sigma = 0.1$. El área bajo la curva ROC es de $A_{ROC} = 0.8782$ mientras que el área bajo la curva Precision-Recall es de $A_{PrecRecall} = 0.9676$. Las curvas se muestran en las figuras 4.8 y 4.9 respectivamente.

4.4.2. Esquema de normalización MinMaxNorm

Se hacen pruebas según los parámetros descritos en 3.4 variando el umbral entre $\sigma = 0$ y $\sigma = 1$ utilizando un paso $\Delta\sigma = 0.01$. El área bajo la curva ROC es de $A_{ROC} = 0.8782$ mientras que el área bajo la curva Precision-Recall es de $A_{PrecRecall} = 0.967$. Las curvas coinciden punto a punto con la curva sin esquema de normalización, lo que puede ser un artefacto producido por el tamaño reducido de la base de datos y porque la normalización minMax no es capaz de separar las muestras en el espacio de características.

4.4.3. Esquema de normalización zNorm

Se hacen pruebas según los parámetros descritos en 3.4 variando el umbral entre $\sigma = -1$ y $\sigma = 1$ utilizando un paso $\Delta\sigma = 0.01$. El área bajo la curva ROC es de $A_{ROC} =$ mientras que el área bajo la curva Precision-Recall es de $A_{PrecRecall} = 0.9989$. Las curvas se muestran en las figuras 4.10 y 4.11 respectivamente.

El resumen de las áreas bajo las curvas está en la tabla 4.4. Se puede ver que hay una mejora respecto a los sistemas unimodales por separado, siendo el sistema con esquema de normalización gaussiano (zNorm) el con mejor desempeño.

4.4.4. Visualización de curvas ROC

El sistema fusionado de decisión se basa en validar puntajes de entre 6 posibles opciones, para luego obtener la decisión final utilizando la moda de éstas. Esta solución presenta un comportamiento no monótono en el valor de los falsos positivos entregados tal y como se muestra en las figuras 4.12 y 4.13, ya que un umbral de validación más bajo no necesariamente indica un número mayor de clasificaciones, sino que un número mayor de decisiones entre las cuales tomar la moda. Esto dificulta un poco la obtención de una curva ROC que describa todo el sistema en función del umbral de validación, por lo que es necesario limitar el número de puntos. Este número puede ser de interés estudiarlo con una base de datos de mayor complejidad, como, por ejemplo, la base *Labeled Faces in the wild* [7].

El hecho de que la curva Precision-Recall no considere la cantidad de falsos positivos, sino que la eficiencia de la clasificación respecto a lo clasificado, nos permite evaluar mejor un caso como este, en que la dependencia de la curva y los datos es no trivial.

4.5. Estrategias de fusión

Desde los resultados presentados en esta sección se deduce que la fusión de información a cualquiera de los tres niveles (características, puntaje y decisión) es beneficiosa para el sistema de reconocimiento. Incluso sin considerar las ventajas de la multibiometría en general mencionadas en el capítulo 2 es deseable implementar los esquemas de fusión para mejorar la performance del sistema.

Dentro de los sistemas unimodales, el mejor resultado se observa en el sistema de clasificación basado en LBP, siendo mejor en su versión de rostro si se considera el criterio de área bajo la curva ROC, o periocular si se considera el criterio de área bajo la curva Precision-Recall.

Por otro lado, según los resultados obtenidos en el trabajo, se puede afirmar que tanto la fusión por características, la fusión por puntajes y la fusión por decisión entregan una mejor tasa de acierto respecto a los sistemas unimodales por separado, como indican las tablas 4.2 y 4.3.

Bajo el criterio de mayor tasa de aciertos, los mejores métodos corresponden a fusión por puntaje mínimo (zNorm) y moda de decisión, sin embargo, considerando el costo computacional de ambas fusiones, la fusión por decisión es mejor, ya que considera menos comparaciones, y por lo tanto un menor costo computacional, para un mismo caso.

El criterio de área bajo las curvas ROC y Prec-Recall indica que el mejor sistema de identificación es el sistema fusionado de decisión con validación de puntajes. En particular, el sistema con esquema de normalización gaussiana (zNorm) entrega un área bajo la curva Prec-Recall igual a $A_{Prec-Rec} = 0.9989$, lo que la posiciona como el mejor desempeño de todas las pruebas realizadas en este trabajo y un área bajo la curva ROC igual a $A_{ROC} = 0.9499$, posicionándolo solamente bajo la clasificación peri-ocular basado en LDA y LBP.

El caso sin normalizar y el caso con norma minMax presentan los mismos resultados. Esto podría deberse al acotado número de individuos en la base de datos, lo que hace difícil evaluar la diferencia entre ambos, se propone en la sección 5 posibles experimentos para determinarlo.

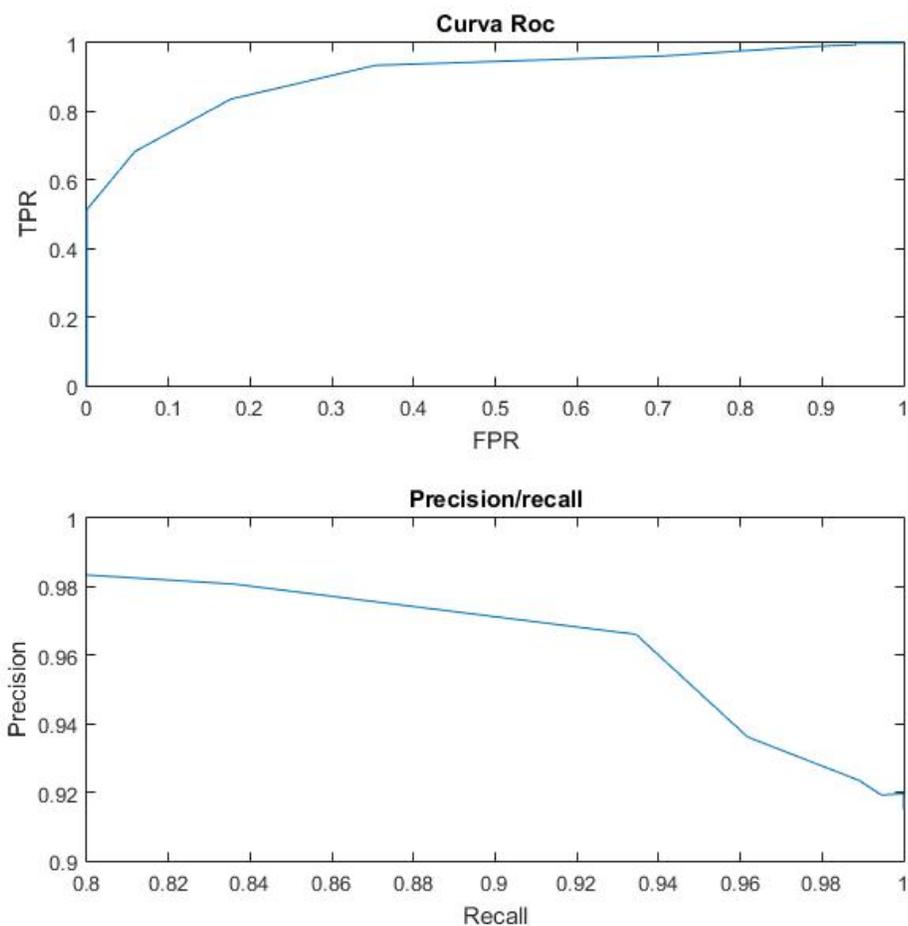


Figura 4.1: Curva ROC y Precision-Recall para el sistema de reconocimiento de rostro sin normalizar puntajes, utilizando PCA como método de extracción de características según los parámetros en 3.2.

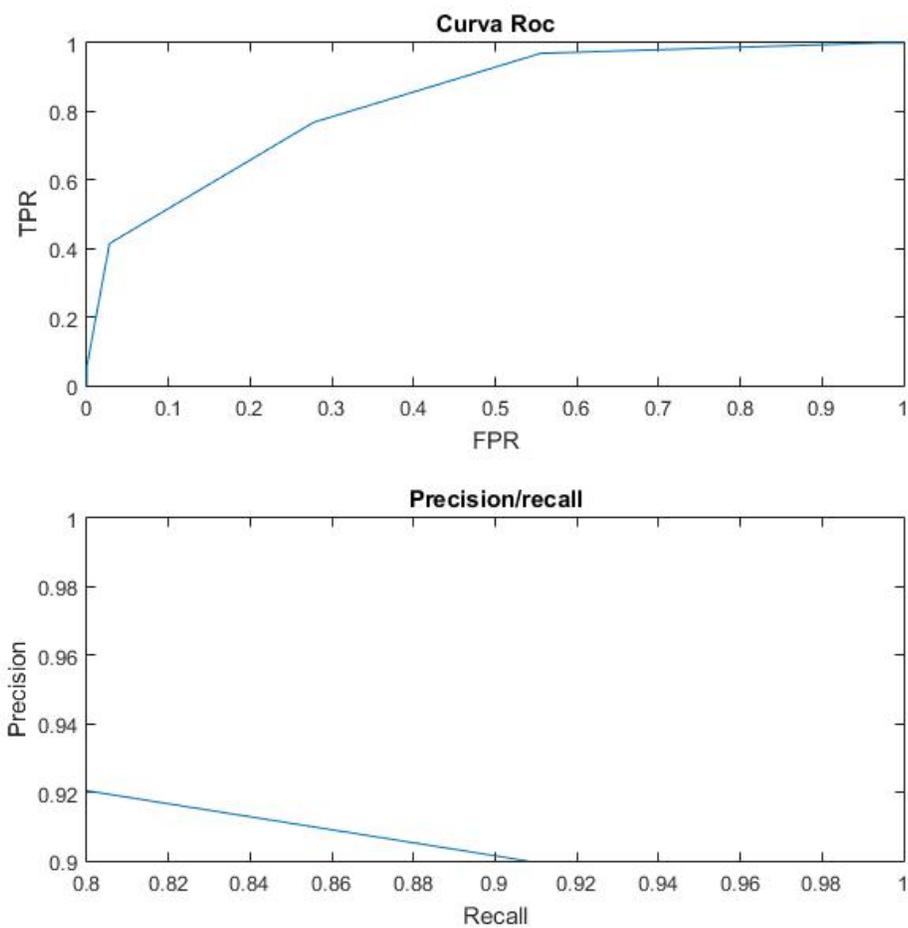


Figura 4.2: Curva ROC y Precisión-Recall para el sistema de reconocimiento de rostro sin normalizar utilizando LDA como método de extracción de características según los parámetros en 3.2.

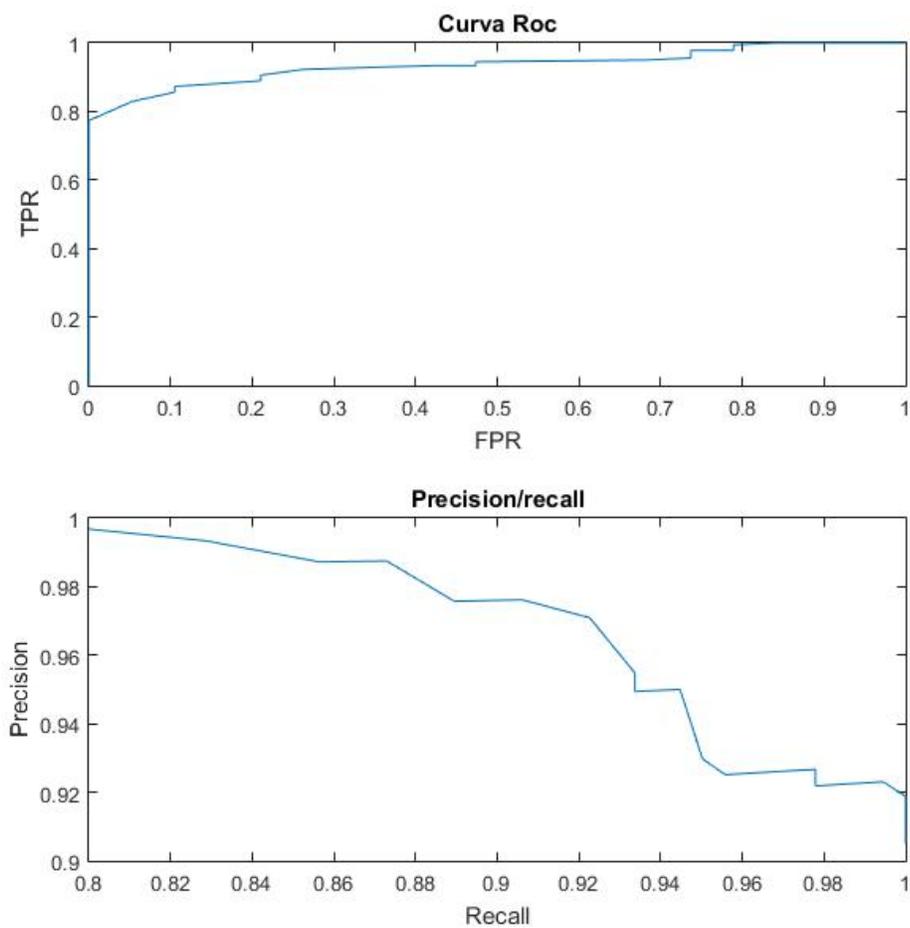


Figura 4.3: Curva ROC y Precisión-Recall para el sistema de reconocimiento de rostro sin normalizar utilizando LBP como método de extracción de características según los parámetros en 3.2.

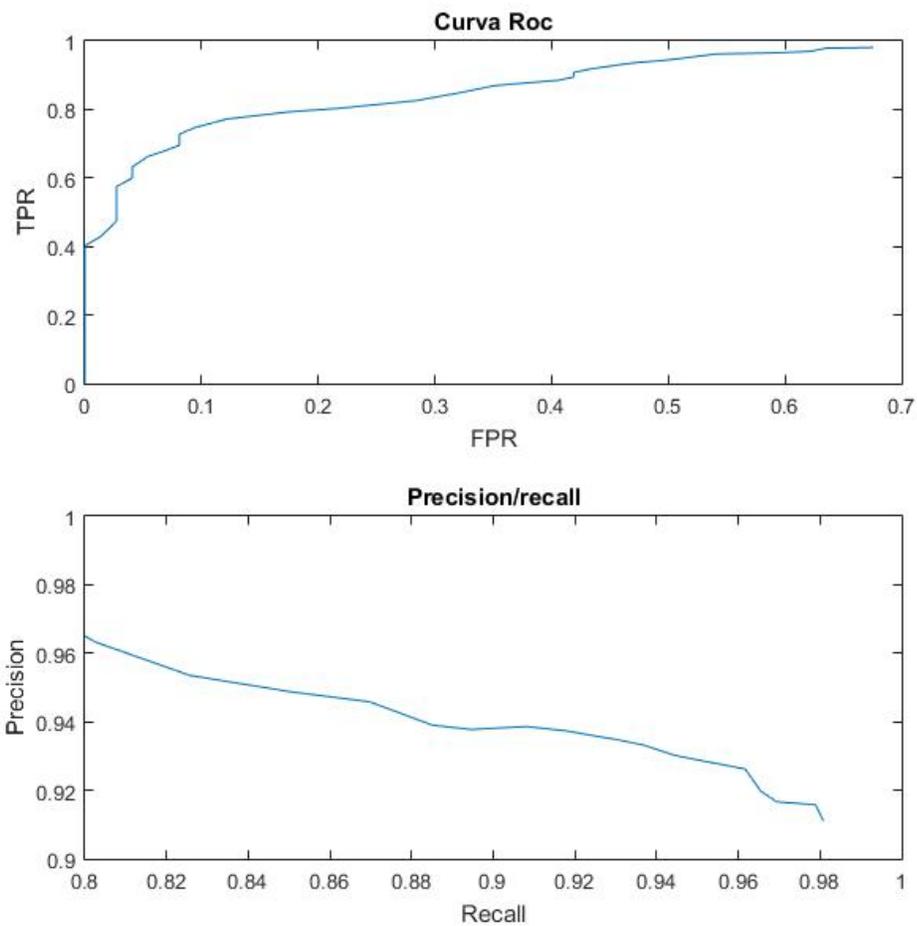


Figura 4.4: Curva ROC y Precision-Recall para el sistema de reconocimiento de zona periocular sin normalizar utilizando PCA como método de extracción de características según los parámetros en la sección 3.2.2.

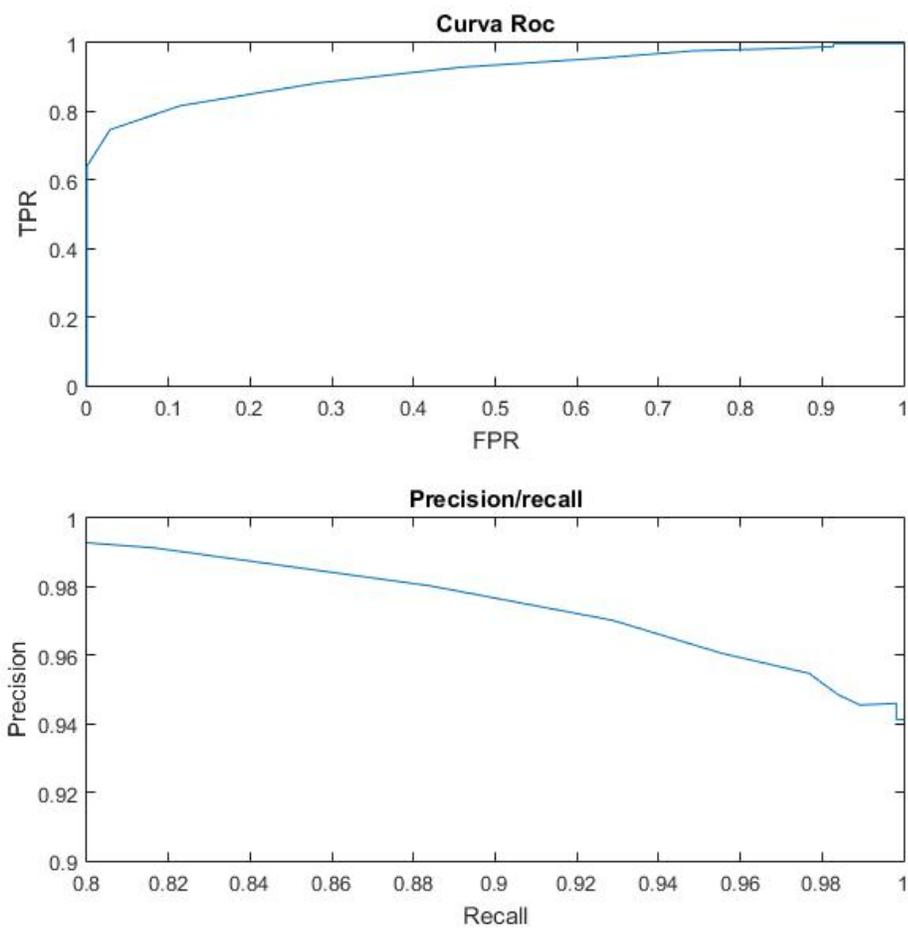


Figura 4.5: Curva ROC y Precision-Recall para el sistema de reconocimiento de zona periocular sin normalizar utilizando LDA como método de extracción de características según los parámetros en la sección 3.2.2 .

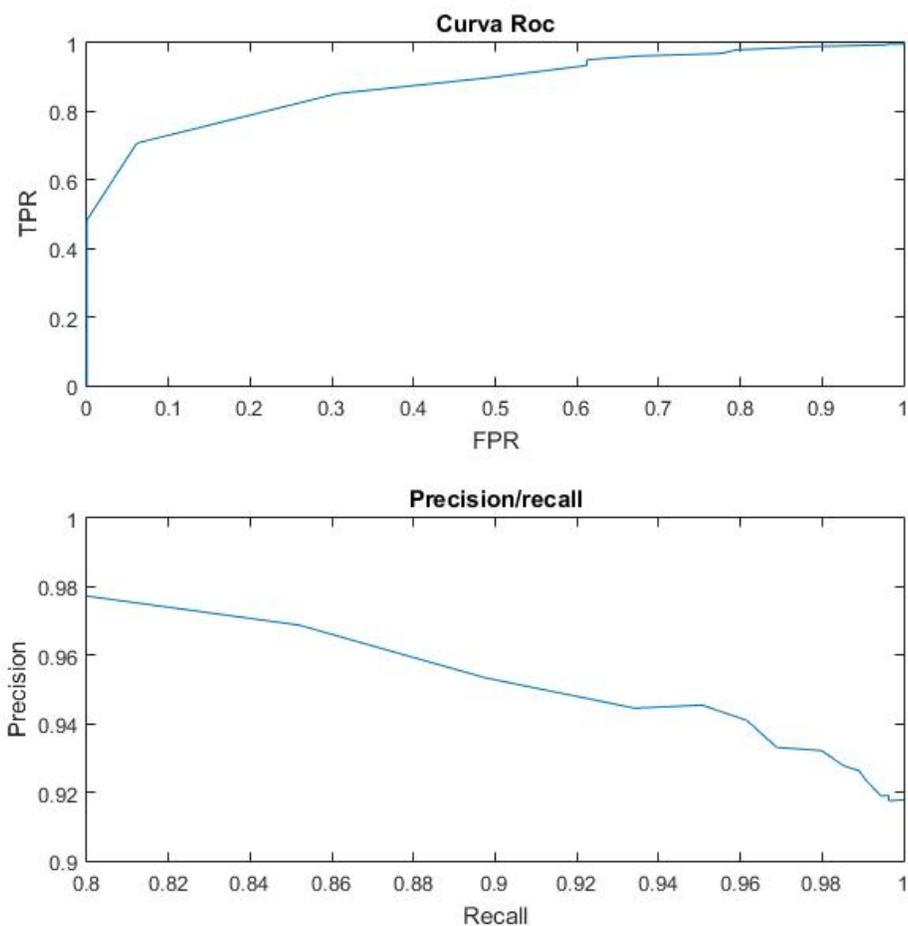


Figura 4.6: Curva ROC y Precision-Recall para el sistema de reconocimiento de zona periocular sin normalizar utilizando LBP como método de extracción de características según los parámetros en la sección 3.2.2.

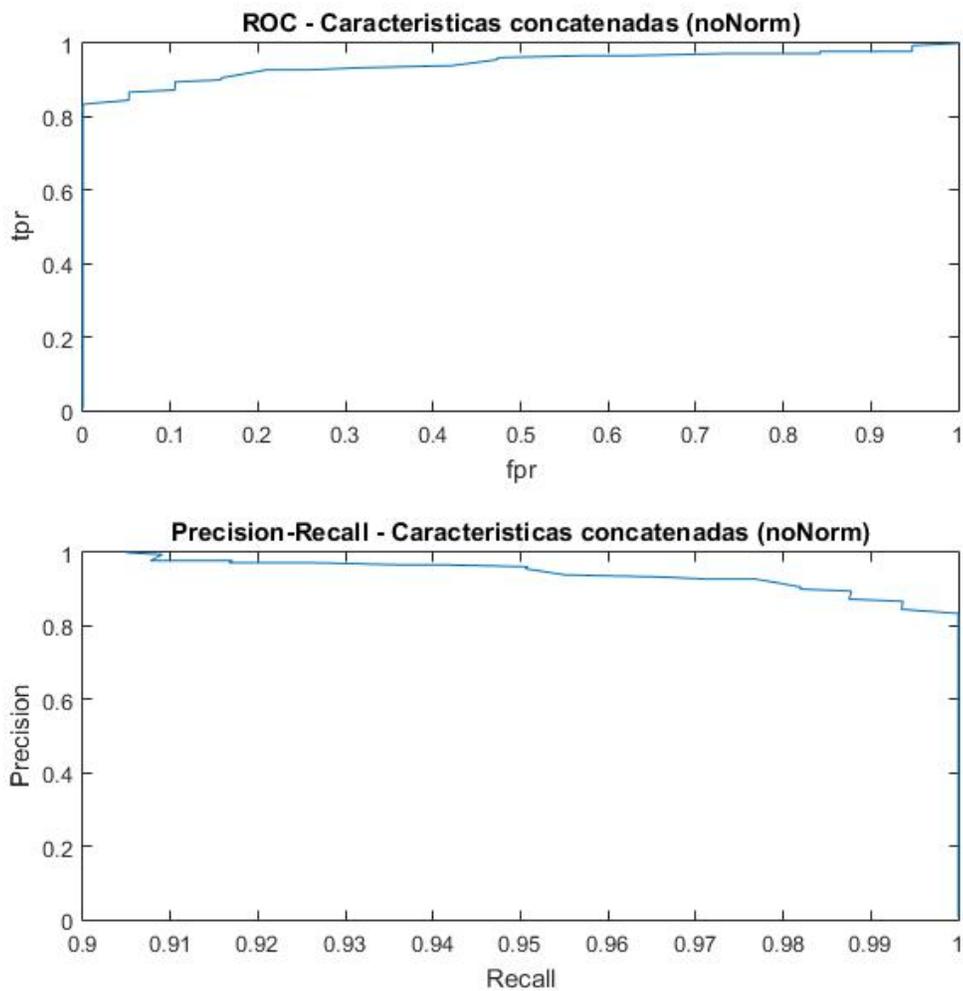


Figura 4.7: Curvas ROC y Precision-Recall para el sistema de reconocimiento de características concatenadas, sin normalizar.

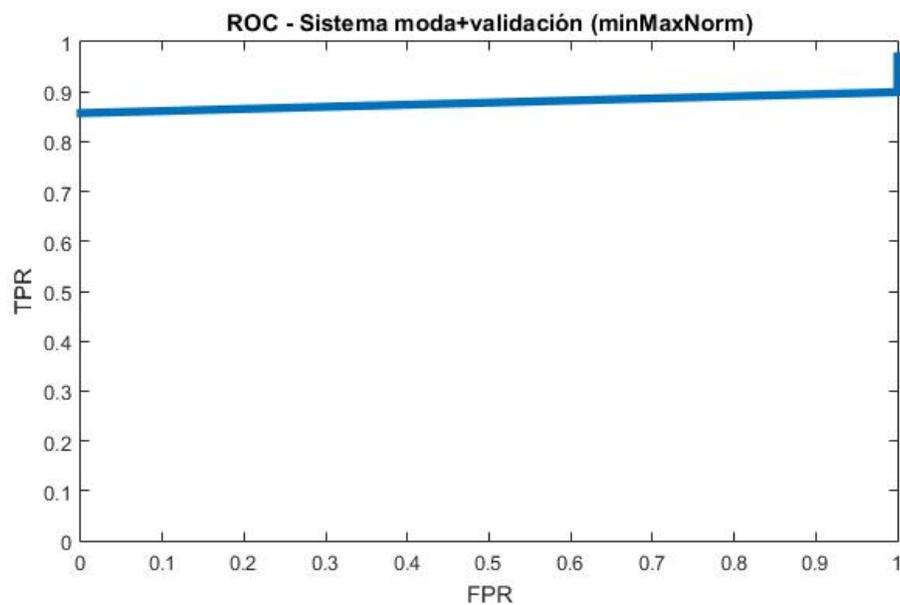


Figura 4.8: Curva ROC para el sistema de reconocimiento multimodal basado en moda y validación de puntaje utilizando el esquema de normalización minMaxNorm.

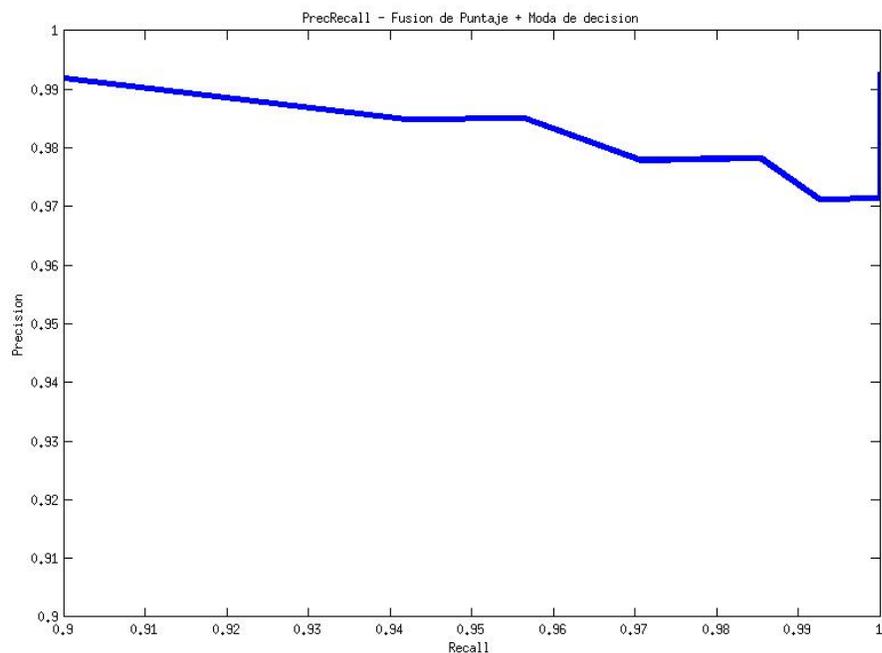


Figura 4.9: Curva Precision-Recall para el sistema de reconocimiento multimodal basado en moda y validación de puntaje utilizando el esquema de normalización minMaxNorm.

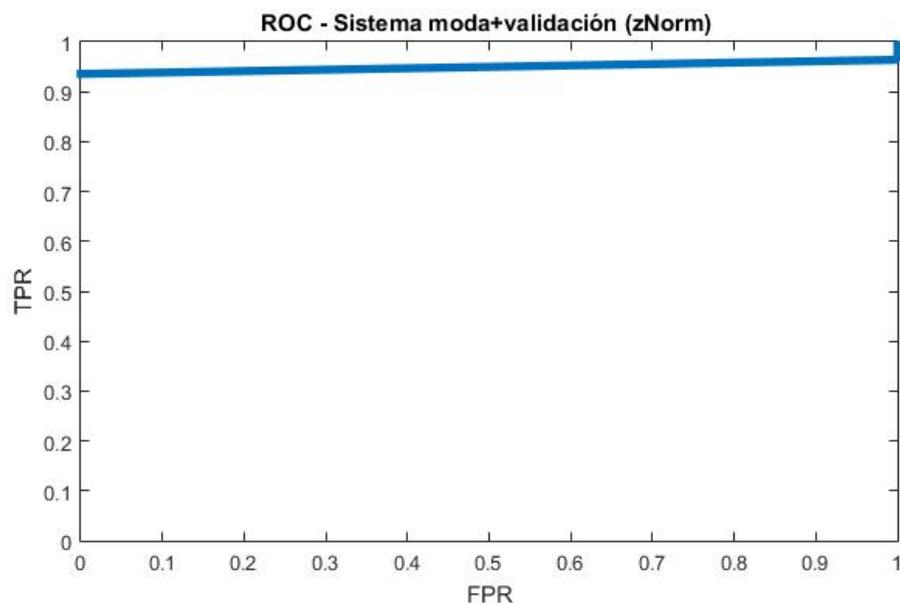


Figura 4.10: Curva ROC para el sistema de reconocimiento multimodal basado en moda y validación de puntaje utilizando el esquema de normalización zNorm.

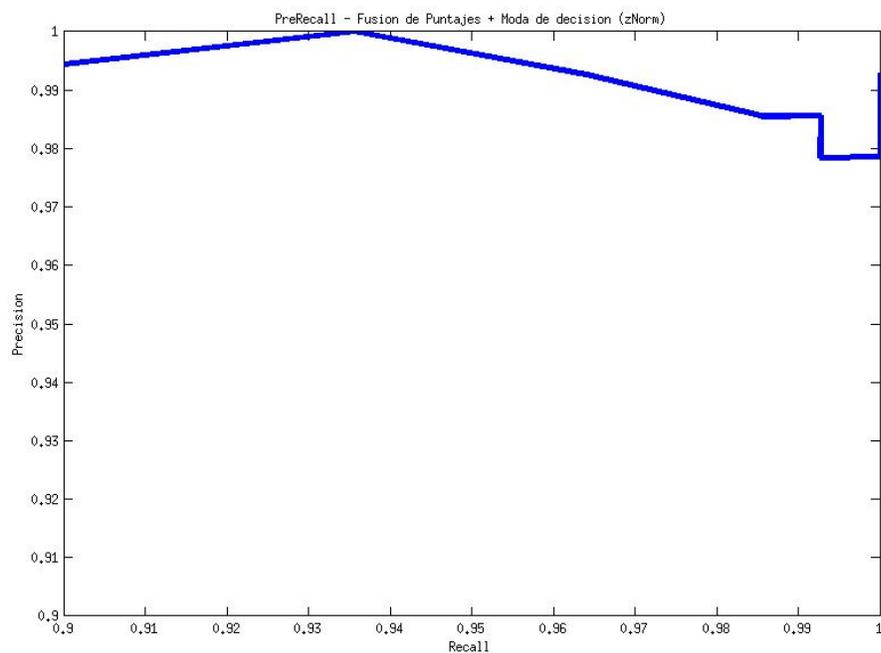


Figura 4.11: Curva Precision-Recall para el sistema de reconocimiento multimodal basado en moda y validación de puntaje utilizando el esquema de normalización zNorm.

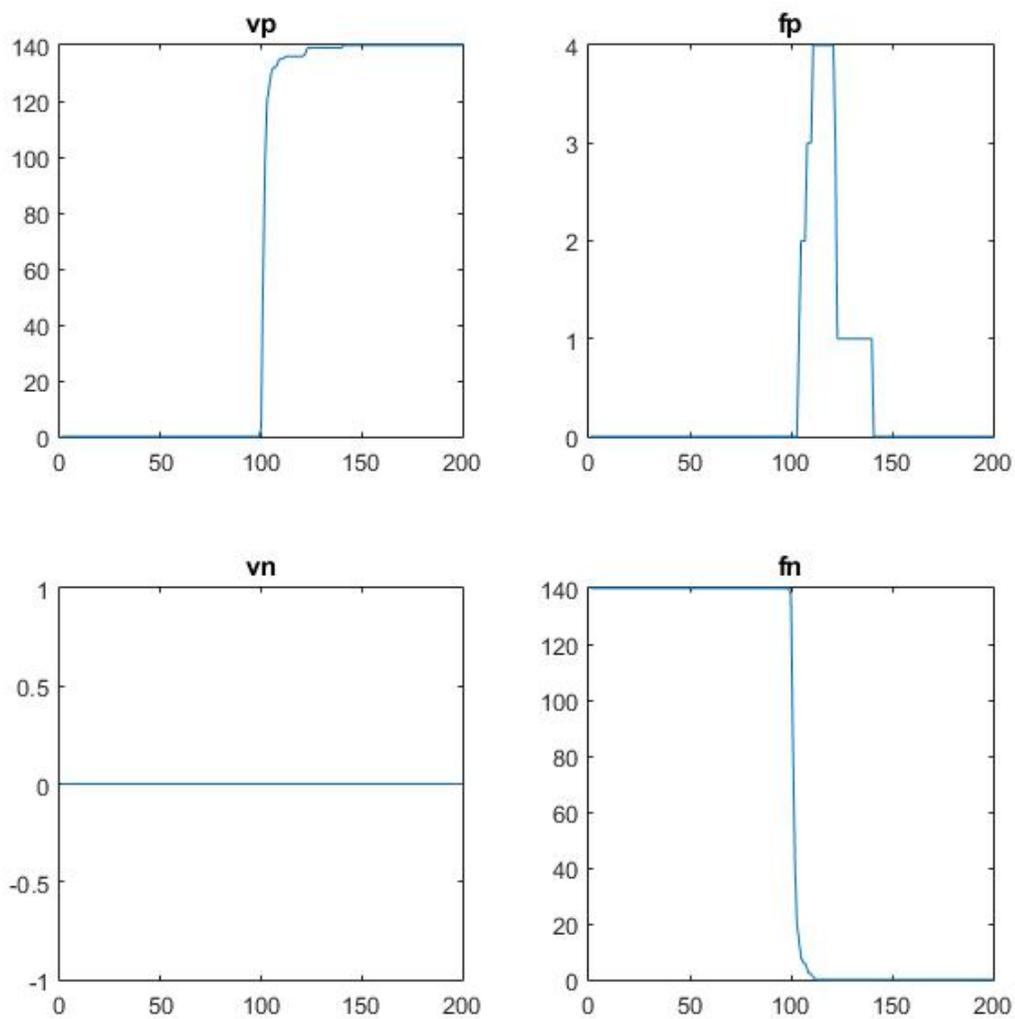


Figura 4.12: Comportamiento de los verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos según el umbral para el sistema fusionado de decisión con validación de puntaje (minMaxNorm). Se observa el comportamiento no monótono de los falsos positivos.

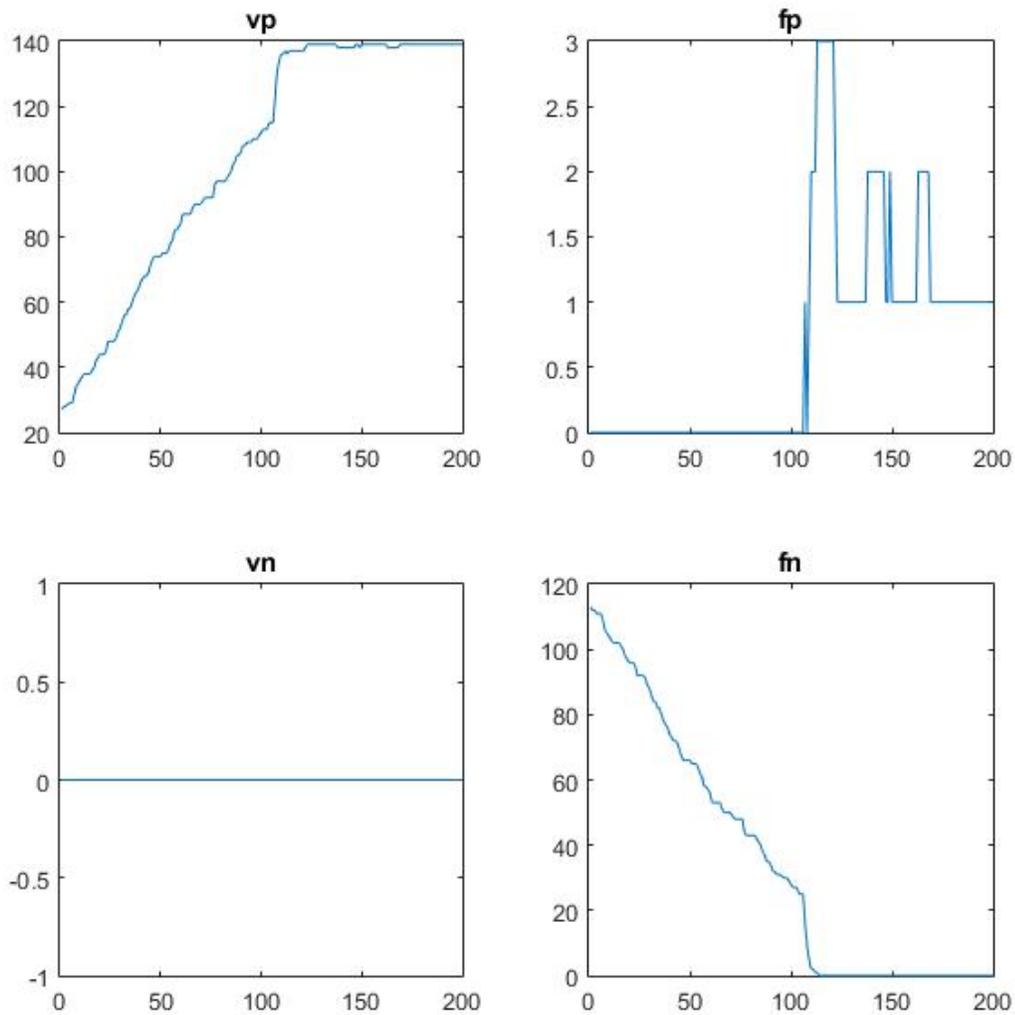


Figura 4.13: Comportamiento de los verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos según el umbral para el sistema fusionado de decisión con validación de puntaje (zNorm). Se observa el comportamiento no monótono de los falsos positivos.

Capítulo 5

Conclusiones y trabajo futuro

Utilizando la base de datos AT&T y una base de imágenes de zona peri-ocular propia, se implementa una base de datos artificial. Considerando pares de rostro-periocular se logra diseñar, además, un sistema de reconocimiento frontal de rostros utilizando tres métodos diferentes de extracción de características: PCA, LDA y LBP. Además, se diseña un sistema de reconocimiento de imágenes de zona periocular utilizando los mismos métodos de extracción de características.

Tomando como base los dos sistemas de reconocimiento, rostro y zona peri-ocular, se implementan varios sistemas de reconocimiento multimodal, utilizando fusión de información a nivel de características, puntajes y decisiones. Con esto se llega a diferentes sistemas, llegando en el mejor caso, correspondiente a la fusión por moda de decisión y mínimo puntaje válido, a un 98.57% de clasificaciones correctas.

Se concluye que, en función de los resultados mostrados en el capítulo 4, la mejor estrategia de fusión está a nivel de decisión, dada la complejidad computacional implicada.

Finalmente, para dar término al presente trabajo, se presentan algunas indicaciones de trabajo futuro:

- Probar distintas bases de datos, más complejas, para la identificación de rostro. Se recomienda la FRGC, citada en el capítulo 2 para comparar los resultados obtenidos con publicaciones como [13] y [18].
- Utilizar una base de datos multi sensor en lugar de imágenes quiméricas. Analizar la posible información redundante entre las imágenes.
- Estudiar cómo influye la distancia utilizada para clasificar en el problema y cómo influye en el desempeño general del sistema.

- Estudiar efecto en cambios de iluminación.
- Estudiar cómo diferentes métodos de extracción de características influyen en la clasificación general, y cómo influyen en la fusión de características.
- Estudiar cómo el desempeño del clasificador varía ante la presencia de impostores.
- Implementar el sistema en tiempo real.
- Estudiar cómo la reducción de dimensionalidad de los histogramas LBP afectan en la fusión concatenada de características.

Bibliografía

- [1] Timo Ahonen, Abdenour Hadid, and Matti Pietikäinen. Face recognition with local binary patterns. In *Computer vision-eccv 2004*, pages 469–481. Springer, 2004.
- [2] Peter N Belhumeur, João P Hespanha, and David J Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(7):711–720, 1997.
- [3] Kevin W Bowyer, Kyong I Chang, Ping Yan, Patrick J Flynn, Earnie Hansley, and Sudeep Sarkar. Multi-modal biometrics: an overview. In *Second Workshop on Multi-Modal User Authentication*, volume 105, 2006.
- [4] Richard O Duda, Peter E Hart, and David G Stork. Pattern classification. 2nd. *Edition. New York*, 2001.
- [5] Ronald A Fisher. The use of multiple measurements in taxonomic problems. *Annals of eugenics*, 7(2):179–188, 1936.
- [6] Mingxing He, Shi-Jinn Horng, Pingzhi Fan, Ray-Shine Run, Rong-Jian Chen, Jui-Lin Lai, Muhammad Khurram Khan, and Kevin Octavius Sentosa. Performance evaluation of score level fusion in multimodal biometric systems. *Pattern Recognition*, 43(5):1789–1800, 2010.
- [7] Gary B. Huang, Manu Ramesh, Tamara Berg, and Erik Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst, October 2007.
- [8] Anil K Jain and Stan Z Li. *Handbook of face recognition*. Springer, 2005.
- [9] Anil K Jain, Salil Prabhakar, and Lin Hong. A multichannel approach to fingerprint classification. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 21(4):348–359, 1999.
- [10] Takeo Kanade. Picture processing system by computer complex and recognition of human faces. 1974.

-
- [11] Allen Kent, Madeline M Berry, Fred U Luehrs, and James W Perry. Machine literature searching viii. operational criteria for designing information retrieval systems. *American documentation*, 6(2):93–101, 1955.
- [12] Sotiris B Kotsiantis, I Zaharakis, and P Pintelas. Supervised machine learning: A review of classification techniques, 2007.
- [13] Unsang Park, Ross Jillela, Arun Ross, and Anil K Jain. Periocular biometrics in the visible spectrum. *Information Forensics and Security, IEEE Transactions on*, 6(1):96–106, 2011.
- [14] Arun A Ross, Karthik Nandakumar, and Anil K Jain. *Handbook of multibiometrics*, volume 6. Springer, 2006.
- [15] F Samaria and A Harter. The orl database of faces. *AT&T Lab. Cambridge*, <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>, 1994.
- [16] Sergios Theodoridis and Konstantinos Koutroumbas. *Pattern Recognition*. 2003.
- [17] JA Unar, Woo Chaw Seng, and Almas Abbasi. A review of biometric technology along with trends and prospects. *Pattern Recognition*, 47(8):2673–2688, 2014.
- [18] Damon L Woodard, Shrinivas J Pundlik, Philip E Miller, and Jamie R Lyle. Appearance-based periocular features in the context of face and non-ideal iris recognition. *Signal, Image and Video Processing*, 5(4):443–455, 2011.