

CENTRO DE INVESTIGACIONES EN ÓPTICA, A.C.

Sistema de Reconocimiento Facial Mediante
Técnicas de Visión Tridimensional

Tesis para obtener el grado de:

Maestro en Optomecatrónica

Presenta:

Ing. Miguel Ángel Vázquez López

Dirigido por:

Dr. Francisco Javier Cuevas de la Rosa

León, Guanajuato; marzo de 2014.

One picture is worth more than ten thousand words.

Agradecimientos

Doy gracias al Creador Divido por la fortaleza y haberme permitido llegar hasta este momento importante de mi formación profesional.

A mi familia por su amor, su comprensión y el apoyo incondicional que me ha brindado en todo momento.

A mi asesor de tesis, Dr. Francisco Javier Cuevas de la Rosa, por compartir sus conocimientos, experiencias y tiempo en la realización de éste trabajo.

A mis sinodales, Dr. Manuel de la Torre Ibarra y Dr. Moises Cywiak Garbarcewicz, por haber compartido parte de su tiempo en revisar este trabajo y por brindarme sus comentarios para bien de la tesis.

Al Maestro Guillermo Garnica por compartir su conocimiento y experiencia en el tema de reconstrucción tridimensional de objetos.

A mis amigos y compañeros que coincidimos en espacio y tiempo, de corazón agradezco su amistad y apoyo.

A todas aquellas personas que aceptaron participar en la creación de la base de dato de rostros tridimensional. Ustedes son el elemento fundamental en este trabajo, mis más sinceros agradecimientos.

Miguel Ángel.

Resumen

El presente trabajo de tesis se ha realizado como requisito para obtener el grado de Maestro en Optomecatrónica en el Centro de Investigaciones en Óptica, A.C.

Se describe el desarrollo de un sistema de reconocimiento facial a partir del análisis de rostros digitalizados en tres dimensiones, el cual es una aplicación biométrica que se apoya de diferentes áreas del conocimiento tales como: el reconocimiento de patrones, la óptica, la visión artificial, geometría diferencial y la estadística. Las principales contribuciones del trabajo son: la aplicación de técnicas de proyección de luz estructurada para el cálculo de la nube de puntos relacionada al rostro, el uso de la fase de la señal demodulada para realizar el reconocimiento, la alineación del rostro a partir del perfil de un rostro modelo y la tolerancia de reconocimiento frente a gestos y expresiones faciales.

El cálculo de la información tridimensional del rostro se lleva a cabo mediante un sistema de proyección de luz estructurada el cual despliega un conjunto de patrones de franjas con la técnica de desplazamiento de fase, los descriptores del rostro se obtienen con el método de eigenrostros o Análisis de Componentes Principales (PCA), finalmente la clasificación e identificación se realiza con la técnica de Análisis Discriminante Lineal.

Organización de la tesis

La tesis está integrada por seis capítulos y una sección de apéndices. A continuación se realiza una breve descripción de éstos.

Capítulo 1. Introducción: Describe los objetivos, realiza el planteamiento del problema y la justificación de la tesis. Así mismo se hace una breve revisión de las técnicas existentes en la solución del problema de reconocimiento facial.

Capítulo 2. Introducción a la biometría: Hace una revisión general sobre la biometría en el cual se describe la importancia de los rasgos del cuerpo humano en el desarrollo de los sistemas de reconocimiento de personas.

Capítulo 3. Teoría de reconocimiento de patrones: En este capítulo se hace una introducción acerca del reconocimiento de patrones, se describe el fundamento biológico que lo sustenta y se hace una revisión de los conceptos fundamentales.

Capítulo 4. Técnicas de reconstrucción tridimensional de objetos: El capítulo tiene como objetivo hacer una revisión de los métodos para obtener la información de la profundidad de los objetos, entre las cuales se encuentran: la visión estereoscópica y la proyección de luz estructurada. Se analizan algunos métodos para la obtención y desenvolvimiento de fase.

Capítulo 5. Diseño de un Sistema de Reconocimiento Facial a partir del Análisis Tridimensional del Rostro. Describe la aplicación de los algoritmos y técnicas, descritos en los distintos capítulos, en el desarrollo del sistema propuesto. Se describen los resultados obtenidos y se revisa el desempeño general del sistema.

Capítulo 6. Conclusiones y trabajos a futuro: En este capítulo se analiza la contribución del trabajo, sus posibles aplicaciones, se describen los trabajos pendientes y las posibles mejoras del sistema.

Anexos: En esta sección se desarrollan algunos fundamentos teóricos y matemáticos que no fueron descritos en los capítulos previos, pero que son de utilidad en el desarrollo del trabajo.

Índice

Capítulo 1. Introducción	1
1.1 Objetivo de la tesis	1
1.2 Objetivos particulares	1
1.3 Planteamiento del problema	1
1.4 Antecedentes.....	3
1.4.1 Enfoques en el reconocimiento facial según el tipo de datos usados	3
1.4.2 Enfoques en el reconocimiento facial según la técnica análisis.	3
Capítulo 2. Introducción a la Biometría.....	7
2.1 ¿Qué es la biometría?	7
2.1.1 Rasgos biométricos	7
2.1.2 Propiedades de los rasgos biométricos	9
2.2 Sistema biométrico	9
2.2.1 Diseño de un sistema biométrico.....	10
2.2.2 Verificación e identificación	11
2.3 Variabilidad de los rasgos biométricos	11
2.4 Tipos de errores y su medición.....	13
Capítulo 3. Reconocimiento de Patrones.....	15
3.1 Conceptos básicos del reconocimiento de patrones	16
3.1.1 Patrones.....	16
3.1.2 Clases	16
3.1.3 Características	16
3.2 Similitud	17
3.2.1 Clasificación.....	18
3.2.2 Regresión	18
3.2.3 Descripción	18
3.3 Paradigmas del reconocimiento de patrones.....	19
3.4 Diseño de un sistema de reconocimiento de patrones.....	20
3.5 Análisis de componentes principales (Eigenrostros)	21
3.6 Análisis discriminante	24
3.6.1 Función discriminante lineal	26
Capítulo 4. Técnicas de Reconstrucción Tridimensional de Objetos	29
4.1 Visión estereoscópica.....	30
4.1.1 La visión estereoscópica artificial	31
4.1.2 Localización de puntos conjugados.	33
4.2 Proyección de luz estructurada	33

4.2.1	Reconstrucción tridimensional mediante proyección de franjas	34
4.3	Métodos de obtención de la fase	37
4.3.1	Método de la Transformada de Fourier.....	37
4.3.2	Détodo de Desplazamiento de Fase.....	39
4.3.3	Demodulación directa de fase	41
4.4	Técnicas de desenvolvimiento de fase	42
4.4.1	Desenvolvimiento de fase unidimensional	43
4.4.2	Desenvolvimiento de fase de funciones bidimensionales mediante mapas de calidad siguiendo rutas discretas	44
Capítulo 5.	Diseño de un Sistema de Reconocimiento Facial a partir del Análisis Tridimensional del Rostro	47
5.1	Introducción	47
5.1.1	Conceptos claves	48
5.1.2	Aspectos técnicos	48
5.2	Módulo de reconstrucción tridimensional del rostro.....	50
5.2.1	Submódulo de Adquisición de Datos	50
5.2.2	Submódulo de Preprocesamiento.....	55
5.2.3	Submódulo del cálculo de la información tridimensional.....	58
5.3	Módulo de alineación y segmentación.....	64
5.3.1	Alineación.....	64
5.3.2	Segmentación	81
5.4	Módulo de Extracción de Características	83
5.5	Módulo de clasificación/reconocimiento	89
5.5.1	Decisión	90
5.6	Resultados.....	93
5.6.1	Obtención del modelo tridimensional de los rostros mediante proyección de franjas.....	93
5.6.2	Alineación del rostro a partir del perfil de un rostro modelo.	97
5.6.3	Reconocimiento a partir de la señal demodulada.	102
5.6.4	Tolerancia de reconocimiento frente a gestos y expresiones faciales.....	105
Capítulo 6.	Conclusión y trabajo a futuro	111
6.1	Conclusión.....	111
6.2	Trabajo a futuro	114
Apendice A.	Manipulación de matrices	116
A.1	Producto interno de dos vectores	117
A.2	Producto de un escalar y un vector	117
A.3	Vector ortogonal	117
A.4	Vector normal	118
A.5	Matriz ortogonal.....	118

A.6	Determinante de una matriz	118
A.7	Producto matricial.....	118
A.8	Eigenvectores y eigenvalores.....	119
Apendice B. Procesamiento digital de imágenes		120
B.1	Resolución de una imagen	121
B.2	Profundidad de color.....	122
B.3	Frecuencia espacial de una imagen digital.....	122
B.4	Filtro pasa bajo ideal.....	123
B.5	Filtro pasa bajo Butterworth	123
B.6	Transformada de Fourier	123
B.7	Transformada Discreta de Fourier	124
Apendice C Análisis Multivariante		125
C.1	Análisis de componentes principales.....	126
C.2	Derivación de componentes principales	127
Bibliografía		129

Capítulo 1. Introducción

1.1 Objetivo de la tesis

Desarrollar un sistema de identificación facial a partir del análisis de la información tridimensional del rostro para determinar la identidad de personas.

1.2 Objetivos particulares

- Diseñar un prototipo de adquisición de rostros, basado en técnicas de proyección de luz estructurada.
- Implementar algoritmos de proyección de franjas con perfil cosenoidal y binario para obtener la información tridimensional del rostro.
- Construir una base de datos de imágenes faciales 3D.
- Implementar algoritmos basados en análisis multivariante para obtener las características principales del rostro.
- Implementar algoritmos de análisis discriminante en el proceso de reconocimiento.

1.3 Planteamiento del problema

El sentido de la vista, uno de los sistemas más complejos de nuestro organismo, es capaz de percibir el color, la forma y la distancia de los objetos. Los ojos pueden distinguir unos 10 millones de colores, permitiéndole distinguir detalles en diferentes circunstancias; tiene la posibilidad de cambiar el foco desde el infinito a centímetros en milisegundos y alcanza una visión panorámica de casi 180º. La interpretación de la información visual se realiza en la corteza visual del cerebro mediante un proceso cognoscitivo que requiere de tres cuartas partes de su capacidad de procesamiento. El sentido de la visión ha sido tema de interés en diversas áreas como la psicología, la psicosociología, la biología y la neuropsicología [1, 2].

El rostro es uno de los principales focos de atención en nuestras interacciones sociales cotidianas, centramos nuestra atención visual en las características y expresiones faciales. Somos capaces de reconocer cientos de rostros, incluyendo el de aquellas personas que no hemos visto durante algún tiempo, llevando a cabo esta tarea en una fracción de segundo [3].

Por otro lado, la emulación del proceso de reconocimiento facial mediante sistemas artificiales ha sido un desafío para algunas disciplinas como la informática, la visión artificial y la biometría. En condiciones controladas, estos sistemas tienen la ventaja de ser rápidos, precisos, económicos y utilizan métodos no invasivos; sin embargo en condiciones no controladas se enfrentan con diferentes problemas relacionados con los datos a analizar como las variaciones de escala, de orientación, expresión facial, condiciones de iluminación, occlusiones, entre otros [4-8].

Por otra parte, tradicionalmente se ha utilizado una gran variedad de métodos para llevar a cabo la identificación de personas como son: usuario y contraseña, número de identificación personal (NIP), tarjeta de identificación por radio frecuencia (RFID), llave, pasaporte, licencia de manejo, por mencionar algunos; estos métodos utilizan recursos que se pueden perder, olvidar, compartir, manipular o robar; ocasionando problemas desde económicos hasta asuntos de seguridad. Debido a estos inconvenientes, es de gran importancia desarrollar sistemas de identificación que no presenten tales desventajas [9].

Las técnicas de identificación basadas en la biometría ofrecen una solución de identificación robusta, dado que utiliza los rasgos físicos o de conducta del cuerpo humano que son únicos, permanentes e intransferibles. Los rasgos físicos pueden ser extraídos del ojo (iris, retina), de la manos (huellas dactilares, geometría, venas) o de las características faciales; así mismo pueden ser utilizados los rasgos de conducta como la manera de caminar, el timbre de la voz, la escritura, la firma, la dinámica de presionar el teclado, entre otros [10].

De acuerdo a lo expuesto, en este trabajo de tesis se pretende desarrollar un sistema de reconocimiento de personas basado en la biometría el cual use como patrón biométrico la información de profundidad del rostro.

1.4 Antecedentes

1.4.1 Enfoques en el reconocimiento facial según el tipo de datos usados

Según la naturaleza de los datos usados en el proceso de reconocimiento, existen dos enfoques: a partir de mapas de intensidad y mapas de profundidad [11].

Los basados en mapas de intensidad realizan el análisis de imágenes digitales del rostro (imágenes 2D), sin embargo éstos presentan algunas desventajas como la sensibilidad a los cambios de perspectiva, de escala, iluminación y la posibilidad de suplantar a un usuario a partir de otra imagen 2D.

Los sistemas basados en mapas de profundidad (imágenes 3D), realizan el análisis de la información de profundidad y/o la geometría del rostro. Algunas de las ventajas de estos sistemas son la tolerancia a los cambios de perspectiva, a cambios de iluminación y a las expresiones faciales. Por otro lado la obtención de la información tridimensional del rostro no es tan trivial, de manera que estos sistemas son robustos ante el intento de suplantación o falsificación.

1.4.2 Enfoques en el reconocimiento facial según la técnica análisis.

Según la manera de realizar el análisis de los datos, los algoritmos de reconocimiento facial se clasifican en dos grupos: métodos basados en características y métodos basados en la apariencia.

Los métodos basados en características, realizan el análisis de las propiedades y la geometría del rostro, tales como áreas, distancias y ángulos entre los elementos del rostro.

Los métodos basados en apariencia, conocidos también como métodos holísticos, consideran al rostro como un todo y realizan un análisis global utilizando herramientas estadísticas. Estos métodos buscan un nuevo subespacio de menor dimensión para proyectar los rostros. Entre estos métodos se encuentran: el análisis de componentes principales (PCA), el análisis discriminante lineal (LDA), el análisis de componente independiente (ICA), el análisis de características locales (LFA), Tensorfaces, Redes neuronales (NN) y las máquinas de soporte (SVM). A continuación se realiza una breve descripción éstas.

- a) **Análisis de componentes principales** (PCA, por sus siglas en inglés), también se le conoce como eigenrostros. Este método es de particular interés en este trabajo de tesis ya que se implementa en el proceso de definición de los vectores de características de los rostros. El PCA es una técnica estadística de síntesis de datos el cual determina los componentes principales de la matriz de covarianza de un conjunto de rostros de entrenamiento. Los componentes principales definen un nuevo subespacio de menor dimensión donde los rostros de entrenamiento son proyectados. Una imagen de prueba es proyectada en este subespacio, y mediante un proceso de comparación se realiza la clasificación.
- b) **Análisis discriminante lineal** (LDA, por sus siglas en inglés), establece proyecciones que mejor discriminan a las clases, es decir LDA busca vectores de proyección que maximicen la dispersión extra-clase y minimicen la dispersión intra-clase.
- c) **Análisis de componente independiente** (ICA, por sus siglas en inglés), establece una base no ortogonal para representar los rostros, de tal manera que los rostros proyectados en dicha base son estadísticamente independientes entre sí, a diferencia del PCA que busca una base ortogonal para representar los rostros.
- d) **Análisis de características locales** (LFA, por sus siglas en inglés), constituye grupos de características correlacionadas a partir de la descomposición de eigen-

espacios. Mediante un proceso de selección se obtiene un conjunto de características con una correlación mínima que define un nuevo subespacio. Una representación local es robusta frente a variabilidades debido a cambios de iluminación así como a la variación de poses.

e) **Elastic Bunch Graph Matching** (EBGM), representa los rostros mediante el cálculo de *imágenes gráficas*. Una imagen gráfica se define a partir de la definición de nodos ubicados en puntos característicos del rostro como son las pupilas, las comisuras de los labios y la posición de la nariz. A partir de las imágenes gráficas se realiza el cálculo de los Coeficientes de Gabor para definir las características locales y así determinar múltiples orientaciones y frecuencias de cada nodo.

f) **Redes Neuronales** (NN, por sus siglas en inglés) y **Máquinas de Vectores de Soporte** (SVM, por sus siglas en inglés), comúnmente son usados para clasificar características de baja dimensión debido a la compleja carga computacional que demandan. Las NN han sido ampliamente usadas en sistemas de reconocimiento facial, sin embargo debido a la gran cantidad de usuarios de un sistema de reconocimiento, estas requieren de una mayor cantidad de recursos computacionales. Las SVM establecen hiperplanos que maximizan dispersión extra-clase con la intención de minimizar los errores de clasificación.

g) **Tensorfaces**, una imagen facial está sujeta a variaciones debido a cambios de pose, condiciones de iluminación o expresiones faciales. El método de tensorfaces intenta modelar estas variaciones mediante una estructura multilineal, los tensores de alto orden permiten diseñar modelos para analizar los factores que producen las variaciones en las imágenes.

Capítulo 2. Introducción a la Biometría

En las civilizaciones antiguas las personas vivían en comunidades pequeñas donde se reconocían sin dificultad, sin embargo con la movilidad y la rápida expansión de la población, la identificación se convirtió en un proceso complejo de manera que en las sociedades modernas ha sido necesaria la implementación de sofisticados sistemas de gestión de identidad. La identidad se refiere al conjunto de la información asociada a una persona, como son: el nombre, el apellido, la fecha de nacimiento, el domicilio, entre otras.

Los sistemas de gestión de identidad son utilizados en diferentes aplicaciones tal como: las aduanas, la restricción de acceso a instalaciones, el control de acceso a recursos informáticos, en las transacciones financieras y en el abordaje de vuelos comerciales, por mencionar algunos. Dentro de los sistemas de gestión de identidad se encuentran los que basan su funcionamiento en la biometría que realizan el análisis de los rasgos del cuerpo humano.

2.1 ¿Qué es la biometría?

La palabra biometría proviene del vocablo griego: *bios* que significa vida y *metrón* cuyo significado es medición, por lo tanto la biometría es la ciencia que establece la identidad de un individuo a partir del análisis de las características fisiológica o del comportamiento del cuerpo humano [9].

2.1.1 Rasgos biométricos

Los rasgos biométricos, son las características fisiológicas o de conducta del cuerpo humano. Éstos deben ser elementos que puedan ser registrados, cuantificados y procesados mediante ordenadores. La Figura 2.1 describe algunos de los rasgos del cuerpo humano.

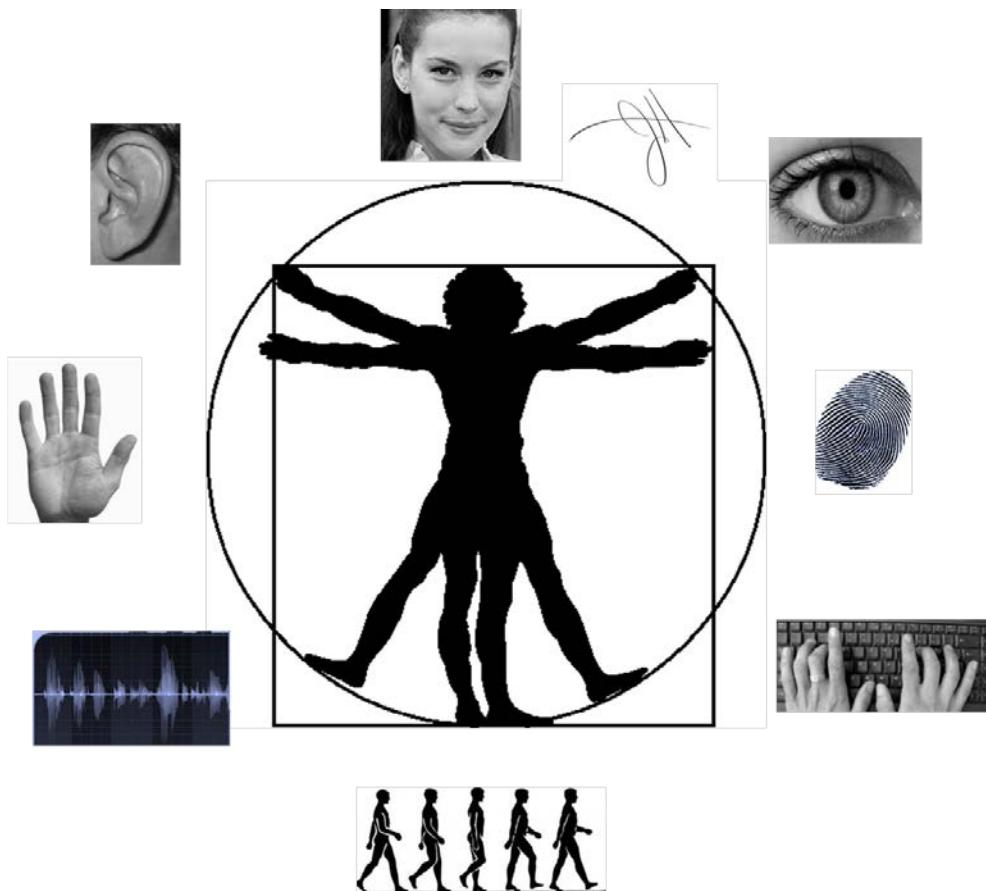


Figura 2.1. Ejemplos de los rasgos del cuerpo humano.

Los rasgos fisiológicos son las características intrínsecas del cuerpo humano que pueden ser extraídos de la mano, del rostro, del iris, de la retina, de la oreja, de la voz, de los patrones vasculares, el ADN entre otros. En tanto los rasgos de conducta se refieren a gestos o el comportamiento del cuerpo ante ciertas situaciones, por ejemplo la manera de hablar, de escribir, de caminar, la dinámica del uso del teclado, la firma, entre otros. Recientemente el campo de la biometría ha experimentado con otras características del cuerpo como son los electrocardiogramas (ECG), los rayos X, el análisis de la luna de las uñas, el análisis del movimiento de los labios y la expresión corporal al hablar.

2.1.2 Propiedades de los rasgos biométricos

Los rasgos utilizados en un sistema biométrico deben poseer las siguientes propiedades:

- *Universalidad.* Todos los usuarios del sistema deben de poseer los rasgos que el sistema solicita.
- *Singularidad.* Los rasgos deben ser únicos e irrepetibles entre los usuarios.
- *Cuantificable.* Los rasgos deben ser elementos que puedan ser registrados, cuantificados y procesados
- *Rendimiento.* El nivel de precisión del sistema debe satisfacer las restricciones impuestas en el diseño de la misma.
- *Aceptabilidad.* Los usuarios debe de tener confianza en utilizar el sistema.
- *Evasión o usurpación.* Establece el nivel al que el sistema es capaz de resistir a técnicas fraudulentas, tal como la usurpación de identidad o la alteración de los rasgos para evadir el sistema.

2.2 Sistema biométrico

Un sistema biométrico tiene como objetivo identificar o verificar, de forma automática, la identidad de un usuario a partir del análisis de uno o más rasgos del cuerpo humano. Toda persona registrada en dicho sistema se le conoce como usuario.

Los sistemas de reconocimiento de personas basan su funcionamiento en tres métodos básicos:

- 1) Lo que el usuario sabe.
- 2) Lo que el usuario posee.
- 3) Quién realmente es el usuario.

Los primeros dos métodos se les conocen como enfoques tradicionales y el tercero se refiere a los sistemas biométricos.

Los métodos tradicionales asocian la identidad de un usuario a partir de conoce (por ejemplo el nombre del usuario, la contraseña, el PIN) o a partir de lo que posee (por ejemplo: tarjetas, llaves, pasaporte, licencia de manejo, dispositivo móvil, entre otros). La desventaja de estos métodos es que utilizan recursos que se pueden perder, olvidar, compartir, manipular o robar, que finalmente tiene consecuencias ya sea económicos, accesos ilegales, clonación de tarjetas, por mencionar algunos. En contraparte las técnicas de identificación basadas en la biometría ofrecen una solución de identificación natural y robusta, dado que utilizan los rasgos físicos o de conducta del cuerpo humano que son únicos, permanentes e intransferibles.

2.2.1 Diseño de un sistema biométrico

De manera general, el funcionamiento de un sistema biométrico se basa en dos fases: alta y reconocimiento del usuario tal como se describe en la Figura 2.2.

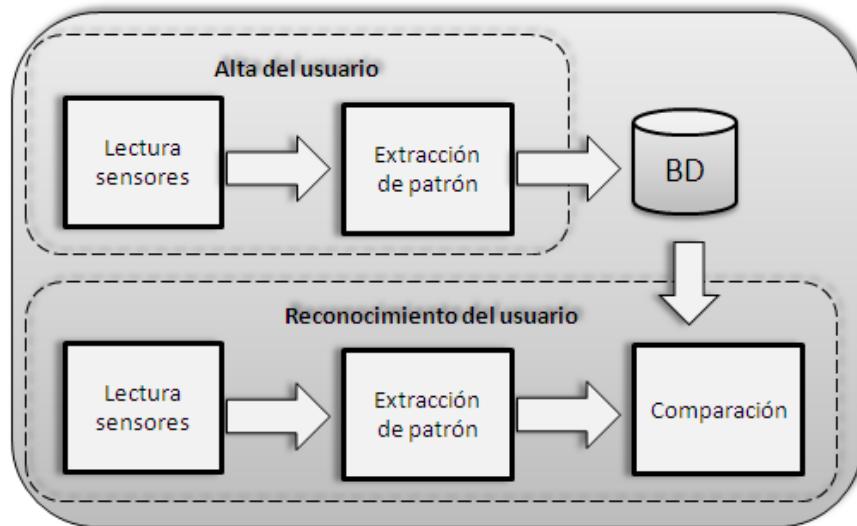


Figura 2.2. Esquema general del funcionamiento de un sistema biométrico.

En la fase de alta, el sistema realiza la adquisición los rasgos biométricos a través de los sensores, seguidamente procesa los datos para obtener los descriptores, finalmente registra los descriptores e identidad del usuario en la base datos (BD).

Mientras tanto en la fase de reconocimiento, el sistema nuevamente lleva a cabo la adquisición de los datos biométricos, luego extrae los descriptores y los compara con los almacenados en la BD para determinar si éste se encuentra registrado en el sistema.

Básicamente un sistema biométrico es un problema de reconocimiento de patrones, por esta razón se encuentra conformado por los módulos que integran un sistema de reconocimiento de patrones. El tema de reconocimiento de patrones se desarrolla en el capítulo 3.

2.2.2 Verificación e identificación

Un sistema biométrico, dependiendo del contexto de su aplicación, puede operar en dos modos: verificación o identificación.

En el modo de verificación, el usuario reclama una identidad y presenta un rasgo biométrico, el sistema compara el rasgo con el almacenado en la BD y realiza una comparación de uno a uno. Si existe un alto grado de similitud entre los patrones, el reclamo se considera *genuino*, de lo contrario el usuario se le conoce como *impostor*. El resultado de éste proceso es un valor verdadero o falso.

En el modo de identificación, el sistema biométrico realiza una comparación del rasgo biométrico con todos los patrones almacenados en la BD, es decir la comparación se realiza de uno a muchos. El resultado del proceso es la identidad del usuario. El término reconocimiento se utiliza de manera indistinta para el proceso de verificación e identificación.

2.3 Variabilidad de los rasgos biométricos

La funcionalidad de un sistema biométrico se basa en dos premisas [12]:

- 1) La unicidad del rasgo biométrico
- 2) La permanencia del rasgo biométrico

La unicidad del rasgo biométrico se refiere a que éstos deben ser características irrepetibles en más de una persona. Mientras que la permanencia del rasgo biométrico hace referencia a que el rasgo debe ser constante y no debe sufrir variación a través del tiempo. Sin embargo, en la práctica, estas premisas no siempre se cumplen por diferentes razones.

Generalmente la unicidad de los rasgos biométricos se ve afectado por factores genéticos. Por ejemplo, existe la posibilidad de la replicación o la similitud de los rasgos entre miembros de una familia como es el caso de la apariencia facial de los gemelos.

La permanencia de los rasgos biométricos tampoco se puede conservar en todo momento, existen factores que las afectan como son: la dinámica del organismo humano (la niñez, adolescencia, juventud y vejez), diferentes condiciones del censado (por ejemplo la iluminación, la temperatura, la humedad), diferentes tipos sensores utilizados en la adquisición de los datos, alteración de los rasgos biométricos (por ejemplo el cambio de la voz por enfermedad respiratoria) y variaciones de la interacción del usuario con el sistema (por ejemplo, la oclusión del iris o del rostro). En estos casos, cada vez que se realiza la adquisición de los rasgos produce variaciones en las muestras.

La variabilidad observada en un conjunto de rasgos biométricos de un usuario se le conoce como variación intra-usuario o variación intra-clase, debido a estas variaciones no puede realizarse un reconocimiento exitoso con dos conjuntos de datos con un grado de variabilidad diferente. Por otro lado no podemos encontrar rasgos biométricos totalmente idénticos, de ahí la diferencia fundamental entre los sistemas biométricos y los sistemas tradicionales. Por ejemplo los sistemas tradicionales que se basan contraseñas, la clave siempre deben coincidir; en tanto los sistemas biométricos validan la identidad de un usuario de acuerdo al grado de semejanza que existe entre los rasgos almacenados y los que se presentan.

2.4 Tipos de errores y su medición

De acuerdo a la variabilidad de los rasgos biométricos, un sistema de reconocimiento automático puede generar dos tipos de errores:

- Error falso positivo (EFP): Cuando una muestra desconocida es declarada erróneamente como conocida.
- Error falso negativo (EFN): cuando una muestra conocida es declarada falsamente como desconocida.

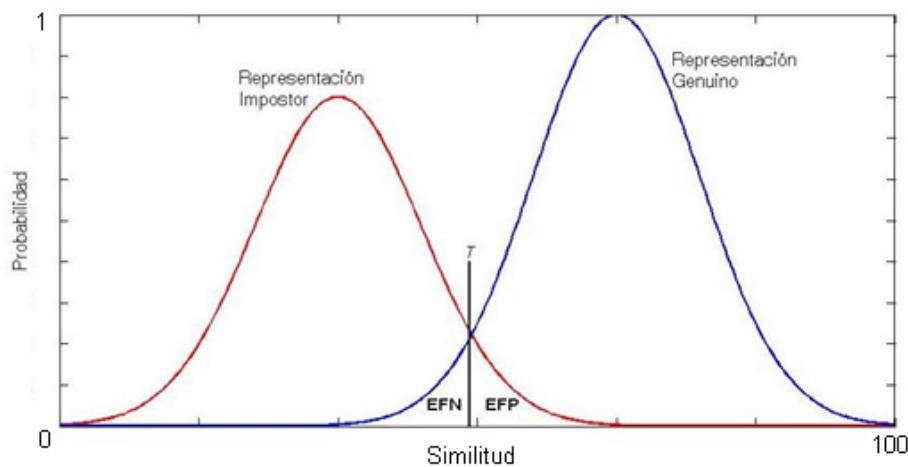


Figura 2.3. Distribución de probabilidad de un usuario impostor y un usuario genuino.

La Figura 2.3, representa la distribución de probabilidad de un usuario impostor y uno genuino. Las dos distribuciones de probabilidad se traslanan en algún punto de la gráfica creando un área común que define la tolerancia o sensibilidad del sistema. eligiendo un valor umbral T situado en el área de tolerancia se definen los EFP y EFN, con un valor bajo el sistema tendrá una mayor tolerancia EFP; si se le asigna un valor alto, el sistema tendrá una menor tolerancia al error de tipo falso Positivo haciéndolo más robusto frente intentos de suplantación de identidad. En aplicaciones como los sistemas de seguridad, se da mayor importancia al error del tipo falso positivo, es decir deben ser robustos frente a estos errores y no dar acceso a personas extrañas o a posibles impostores. Dejando con menor importancia los errores del tipo falso negativo.

Capítulo 3. Reconocimiento de Patrones

Con el paso del tiempo los seres humanos, como demás organismos, han desarrollado importantes mecanismos y habilidades para obtener la información del medio y a partir de ello tomar decisiones. Por ejemplo algunos organismos evalúan el momento de llevar a cabo la migración para conseguir mejores fuentes de alimentos o climas, son capaces de identificar a sus depredadores, pueden marcar y reconocer territorios, entre otros. La capacidad de reconocer diferentes situaciones y tomar decisión pertinentes es una capacidad inherente de los seres vivos, aparentemente esta acción se aprecia como una simple tarea de sentido común, sin embargo implica un proceso cognoscitivo complejo.

El proceso de aprendizaje en los seres vivos ha sido tema de interés en diversas áreas como la psicología, la psicosociología, la biología, la neuropsicología y la teoría de la información. Por otro lado, las ciencias exactas, las ciencias de la computación y la ingeniería se han dado a la tarea de desarrollar teorías y técnicas que intentan emular el proceso cognoscitivo de los seres vivos con el fin de implementarlos en sistemas inteligentes y ser útil en diferentes ámbitos de la actividad humana.

Los seres vivos obtienen la información del medio a través de los sentidos como la vista, el tacto, el gusto, el olfato y la audición; en tanto los sistemas inteligentes la obtienen a través de sensores y presentan los datos en forma conveniente para su procesamiento e interpretación en ordenadores. Los datos registrados son llamados patrones y éstos pueden ser representados como señales, imágenes o tablas de datos.

El reconocimiento de patrones (RP) se define como el proceso de la clasificación automática de un objeto, físico o abstracto, en una o más clases con una intervención mínima del ser humano. La clasificación se basa en el conocimiento a priori o en la información extraída de los patrones [13, 14].

3.1 Conceptos básicos del reconocimiento de patrones

3.1.1 Patrones

Un patrón es la entidad que puede ser identificable de acuerdo a sus características o atributos, éste puede ser abstracto o concreto. Los patrones abstractos son ideas conceptuales mientras que un patrón concreto es la representación física de un objeto, por ejemplo los símbolos (letras, caligrafías), imágenes digitales, imágenes tridimensionales, firmas, señales de voz, electrocardiogramas, electroencefalograma, ondas sísmicas, entre otros.

3.1.2 Clases

Una clase es un estado o categoría de los patrones. Los patrones que pertenecen a la misma clase comparten los mismos atributos y comportamientos. En un sentido abstracto, una clase, se puede interpretar como un modelo o prototipo que idealiza un objeto. Matemáticamente representamos una clase i -ésima con el símbolo ω_i , donde ω_i . El conjunto de todas las clases se le conoce como *espacio de interpretación* [15].

3.1.3 Características

Las características describen a los patrones, de acuerdo al valor de sus atributos, y determinan el *espacio de características* de una clase. En la Tabla 3.1 se presenta una lista de las características comúnmente utilizadas en el reconocimiento de señales e imágenes.

Atributos de señales	Atributos de imágenes
Amplitud	Área de región
Medidas del histograma	Color
Desviación estándar	Intensidad
Morfología de la onda	Momentos invariantes
Cruce por ceros	Histograma
Correlación	Análisis de Fourier
Espectro de frecuencias	Morfología de la imagen

Tabla 3.1. Características comunes extraídas de señales e imágenes.

Los atributos pueden tener valores de tipo numérico, cadenas alfanuméricas o símbolos. Si éstos son de tipo numérico, entonces pueden ser representados mediante vectores. Por ejemplo, la ecuación (3.1), el vector \mathbf{x} representa un vector de características con d características.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} \quad (3.1)$$

Sí los valores son caracteres o símbolos pueden representarse mediante cadenas o una estructura de árbol. Por ejemplo el ADN, está constituido por una cadena de proteínas: adenina (A), guanina (G), citosina (C) y tiamina (T), entonces un gen puede ser descrito por una cadena como la siguiente:

$$GAAGTCCAG \quad (3.2)$$

3.2 Similitud

La noción de similitud, es la pieza angular en el proceso del RP, se refiere a los valores parecidos de un atributo en dos o más objetos. En el proceso de reconocimiento, se evalúa la similitud entre un objeto y un modelo (prototipo) que idealiza a la clase a la que pertenece.

Básicamente existen tres enfoques para evaluar la similitud de los patrones [15]:

- 1) La clasificación
- 2) El análisis de regresión
- 3) La descripción

3.2.1 Clasificación.

La clasificación es el proceso de asignar un patrón desconocido a una clase con base a la evaluación de sus atributos. Por ejemplo, para discernir entre una manzana roja y una pera podemos basarnos en dos atributos: el color y la forma. El vector de características lo define la ecuación (3.3):

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \text{color} \\ \text{forma} \end{bmatrix} \quad (3.3)$$

En su forma más simple, la clasificación se basa en la medición de la distancia euclíadiana de las características del objeto y la de un modelo que idealiza el objeto. Para mejorar los resultados de la clasificación, en ocasiones es conveniente asignar pesos a una característica en particular del objeto, que influye fuertemente en la diferenciación. En este ejemplo podríamos asignar un mayor peso al atributo correspondiente a la forma tridimensional, ya que la forma de una manzana es diferente a la de una pera.

Los errores de clasificación pueden ser por una de las siguientes razones: alto grado de variabilidad de los datos presentados, la insuficiencia de muestras o la mala elección de las características.

3.2.2 Regresión

La tarea de regresión se fundamenta en el proceso cognitivo de la inferencia. La regresión es una herramienta que permite estimar el valor de una variable a partir de valores pasados de la misma o de otras variables. La desviación entre el valor estimado y el real genera un error de aproximación. El objetivo del sistema de regresión es que el error sea lo menor posible.

3.2.3 Descripción

La descripción se basa en el lenguaje formal, busca medir la similitud de los objetos a partir de una regla estructural la cual está conformada de una secuencia

de primitivas. Tanto la clasificación como la regresión evalúan la similitud realizando una medición de distancias; mientras que la descripción realiza la evaluación a partir de reglas sintácticas previamente definidas.

3.3 Paradigmas del reconocimiento de patrones

En el problema del RP existen diferentes paradigmas para la solución de un problema, aunque no existe un consenso de como categorizarlos, una de las aproximaciones se describe en el esquema de la Figura 3.1. Éstos basan su funcionamiento en uno de los enfoques de análisis de similitud revisados previamente, las cuales tambien se les conoce como *hipótesis*.



Figura 3.1. Paradigmas del reconocimiento de patrones

Las hipótesis pueden ser obtenidas de dos maneras: aprendizaje supervisado y aprendizaje no supervisado.

- a) Aprendizaje Supervisado o hipótesis inductiva. Define un conjunto de funciones a partir de datos de entrenamiento o clases previamente definidas. Un patrón desconocido es evaluado en las funciones y de acuerdo al resultado es clasificado en uno o más clases.
- b) Aprendizaje no supervisado o hipótesis deductiva. En este tipo de aprendizaje no existe ninguna clasificación previa del objeto y en algunas ocasiones ni siquiera se han definido las clases.

3.4 Diseño de un sistema de reconocimiento de patrones.

Un sistema de reconocimiento de patrones, independientemente del paradigma que implementa, está conformado por distintos módulos que operan de manera sistemática sobre los patrones. La Figura 3.2 corresponde a la estructura de un sistema típico de un sistema de reconocimiento de patrones [16].

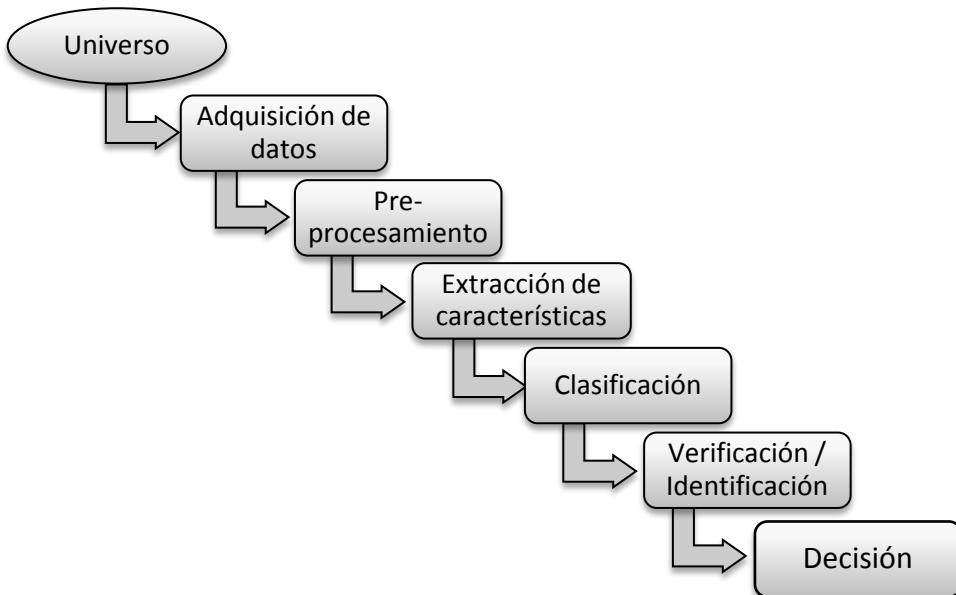


Figura 3.2. Esquema general de un sistema de reconocimiento de patrones.

Módulo de adquisición de datos: El proceso inicia con la adquisición de datos mediante algún sensor que registra las variables físicas y las representa de forma conveniente para su procesamiento y análisis en ordenadores. La calidad de los datos adquiridos depende de las características del sensor tales como la resolución y el diseño; también influyen factores ambientales o las variaciones de la interacción del usuario.

Módulo de preprocesamiento: Esta tarea se realiza con la finalidad de mejorar la calidad de los datos adquiridos. Realiza tareas tales como la normalización, remoción de ruidos aleatorios o eliminación de los datos irrelevantes en las muestras.

Módulo de extracción de características: El objetivo de este módulo es generar información compacta pero significativa de los datos. La información relevante es almacenada en los vectores de características.

Módulo de clasificación: En el proceso de clasificación, los vectores de características son analizados bajo un enfoque de RP para definir las clases y posteriormente asignar un objeto desconocido a una de ellas.

Módulo de postprocesamiento: En esta etapa se busca evaluar el resultado de la clasificación y determinar si un patrón fue asignado a la categoría correcta. Una manera de mejorar los resultados de la clasificación es la utilización de múltiples clasificadores.

3.5 Análisis de componentes principales (Eigenrostros)

En el presente trabajo, los descriptores del rostro se obtienen mediante la implementación de la técnica de Análisis de Componentes Principales (PCA) o eigenrostros, esta técnica se basa en el análisis de la variabilidad de la información de profundidad de los rostros y en la definición de un nuevo subespacio de menor dimensión, en el cual los datos originales son proyectados. Considerando que inicialmente los datos de la superficie del rostro se encuentran

almacenados en una matriz de M elementos, el PCA reduce estos elementos a un vector de N elementos, donde $N \ll M$ [3, 5]. El algoritmo del PCA se describe en la tabla 3.1

Algoritmo: Análisis de Componentes Principales (PCA)

1. Representar cada rostro en un vector.
 2. Definir un rostro promedio.
 3. Normalizar los rostros de entrenamiento respecto al rostro promedio.
 4. Calcular la matriz de covarianza del conjunto de rostros de entrenamiento.
 5. Calcular los componentes principales de la matriz de covarianza.
 6. Definir un nuevo subespacio (espacio de eigenrostros)
 7. Proyectar los rostros de entrenamiento en el nuevo subespacio.
-

Tabla 3.1. Algoritmo de alineación del rostro.

1) Representar los rostros en un vector.

Cada rostro se representa en forma de un vector Γ_i , de longitud N . Donde $N = m \times n$ (m es el número de filas y n el número de columnas de la matriz de datos original).

$$z_i(x, y) = \Gamma_i \quad (3.4)$$

2) Calcular un rostro promedio.

$$\psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i \quad (3.5)$$

donde ψ corresponde al rostro promedio del conjunto de rostros de entrenamiento y M es el número de rostros de entrenamiento.

3) Normalizar los rostros de entrenamiento respecto al rostro promedio.

Normalizamos los rostros sustrayéndole el rostro promedio:

$$\Phi_i = \Gamma_i - \psi \quad (3.6)$$

4) Calcular la matriz de covarianza del conjunto de rostros de entrenamiento.

La matriz de covarianza del conjunto de rostros de entrenamiento se define con la siguiente ecuación:

$$C = \frac{1}{M} \sum_{i=1}^M \Phi_n \Phi_n^T = AA^T \quad (3.7)$$

donde $A = [\Phi_1 \ \Phi_2 \ \dots \ \Phi_i]$.

5) Calcular los componentes principales.

Los componentes principales son los eigenvectores de la matriz de covarianza C . Cada eigenvector lleva asociado un eigenvalor y se definen con la siguiente ecuación.

$$C u_i = \mu_i u_i \quad (3.8)$$

donde u_i es el eigenvector y μ_i es el eigenvalor de la matriz de covarianza C . La matriz de covarianza C contiene $N \times N$ elementos y resulta impráctico procesar tal cantidad de datos, por ello se realiza el cálculo de L envés de C , cuya dimensión es $M \times M$.

$$L = A^T A \quad (3.9)$$

$$A^T A v_i = \mu_i v_i$$

$$AA^T A v_i = A\mu_i v_i$$

$$C A v_i = \mu_i A v_i$$

donde:

$$u_i = A v_i \quad (3.10)$$

6) Definir un nuevo subespacio

En la práctica un pequeño conjunto de eigenvectores es necesario para identificar un rostro, por ello, el nuevo subespacio lo conforman los K eigenvectores que llevan asociado los eigenvalores con valor numérico grande.

$$\mathbf{W} = [u_1, u_2, u_3, \dots, u_K] \quad (3.11)$$

7) Proyectar los rostros de entrenamiento en el espacio de eigenrostros.

El procedimiento final consiste en proyectar los rostros normalizados Φ_i en el sub-espacio \mathbf{W} mediante la ecuación (3.12).

$$\mathbf{X} = \mathbf{W}^T \mathbf{A} \quad (3.2)$$

3.6 Análisis discriminante

El análisis discriminante es un método de clasificación que se basa en la definición de un conjunto de funciones discriminantes y de un conjunto de superficies de decisión. Una función discriminante es el resultado de la combinación lineal de las variables de una clase, mientras que las superficies de decisión se definen a partir de las funciones discriminantes y su objetivo es realizar una partición del espacio de características en regiones mutuamente excluyentes [14].

Un patrón puede ser representado de manera gráfica en el espacio de características siempre que el número de sus atributos sea igual o menor de 3 elementos, de lo contrario se encuentra en un hiperplano haciendo compleja su visualización gráfica. Los patrones que pertenecen a la misma clase tienden a agruparse en la misma región de espacio, tal como se observa en la Figura 3.3, los ejes (x_1, x_2) corresponde a los atributos de la clase, los puntos en color verde son los patrones que pertenecen a la clase ω_1 y los de color rojo pertenecen a la clase ω_2 , la línea en color azul es la función de decisión.

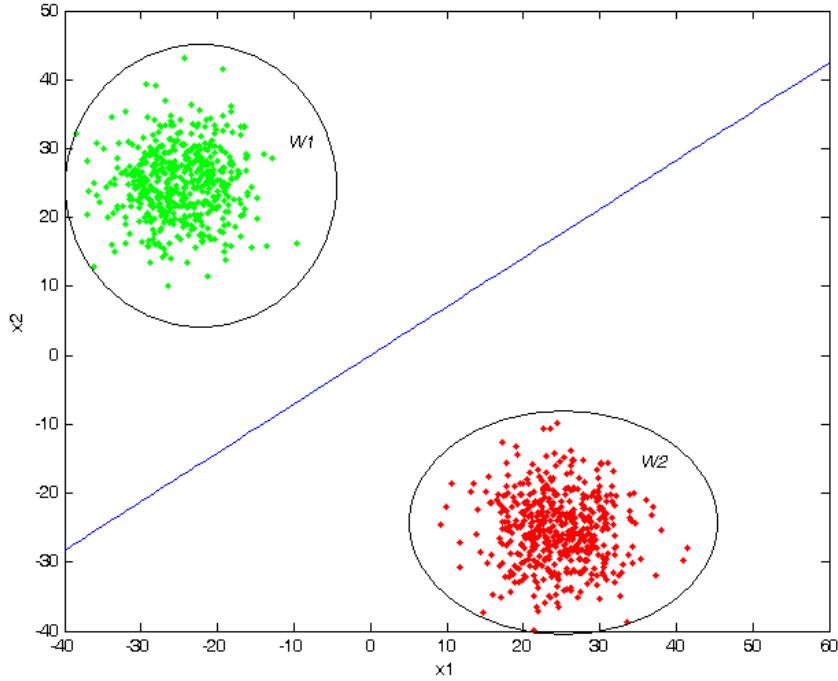


Figura 3.3. Partición del espacio de características.

Considerando el vector de características $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ a clasificar en una de W clases $(\omega_1, \omega_2, \dots, \omega_W)$. El análisis discriminante define W funciones discriminantes $(d_1(\mathbf{x}), d_2(\mathbf{x}), \dots, d_w(\mathbf{x}))$, si \mathbf{x} pertenece a la clase ω_i , entonces:

$$\mathbf{x} \in \omega_i \text{ si } d_i(\mathbf{x}) > d_j(\mathbf{x}); \quad (3.4)$$

donde $i = 1, 2, \dots, W$; $j = 1, 2, \dots, W$; $j \neq i$

Las superficies de decisión se definen a partir de las funciones discriminantes mediante la ecuación (3.5):

$$d_{ij} = d_i(\mathbf{x}) - d_j(\mathbf{x}) = 0 \quad (3.5)$$

donde:

$$\mathbf{x} \in \omega_i \text{ si } d_{ij}(\mathbf{x}) > 0 \quad (3.6)$$

$$\mathbf{x} \in \omega_j \text{ si } d_{ij}(\mathbf{x}) < 0 \quad (3.7)$$

De manera general existente $W(W - 1)/2$ superficies de decisión, sin embargo no todas son significativas sino redundantes, por ello solo $(W - 1)$ superficies son necesarias para separar W clases.

En conclusión la tarea principal del análisis discriminante consiste: primero definir las funciones discriminantes que satisfagan la ecuación (3.4), luego calcular las superficies de decisión de la ecuación (3.5).

3.6.1 Función discriminante lineal

Las funciones discriminantes pueden ser de tipo lineal o no lineal, esto depende de la naturaleza del problema a resolver. En esta sección se describe el caso de las funciones discriminantes lineales [17]. La función discriminante lineal d_j puede representarse en forma vectorial de la siguiente manera:

$$d_j = \mathbf{m}_j^T \mathbf{x} \quad (3.8)$$

Donde \mathbf{m}_j es el prototipo de la j -ésima clase y \mathbf{x} es el vector de características del patrón a clasificar.

Uno de los enfoques básicos para definir las funciones discriminantes lineales es mediante el uso de la distancia euclídea, conocido también como clasificador de la distancia mínima. El clasificador de distancia mínima realiza el cálculo de la distancia euclídea entre el patrón desconocido con los prototipos de las clases, toma una decisión seleccionando la clase en la que resulta una distancia mínima. El prototipo de la clase se refiere al descriptor promedio de ésta y se define de acuerdo a la ecuación (3.9).

$$\mathbf{m}_j = \frac{1}{N_j} \sum_i^{N_j} \mathbf{x}_i \quad \mathbf{x} \in \omega_j; \quad j = 1, 2, \dots, W \quad (3.9)$$

Donde N_j es el número de descriptores que pertenecen a la clase ω_j y W es la cantidad de clases.

Ahora es posible calcular la distancia entre el patrón \mathbf{x} con cada uno de los prototipos:

$$D_j(\mathbf{x}) = \|\mathbf{x} - \mathbf{m}_j\| \quad j = 1, 2, \dots, W \quad (3.10)$$

La ecuación (3.10) asigna el patrón \mathbf{x} la clase ω_i si $D_j(\mathbf{x})$ es la menor distancia. Desarrollando la ecuación (3.10) se puede demostrar que es equivalente a::

$$\|\mathbf{x} - \mathbf{m}_j\|^2 = \mathbf{x}^T \mathbf{x} - 2\mathbf{x}^T \mathbf{m}_j + \mathbf{m}_j^T \mathbf{m}_j \quad (3.11)$$

Se puede prescindir del primer término sin alterar la ecuación debido que es constante para todas las clases:

$$\|\mathbf{x} - \mathbf{m}_j\|^2 = -2\mathbf{x}^T \mathbf{m}_j + \mathbf{m}_j^T \mathbf{m}_j = d_i \quad (3.12)$$

Rescribiendo la ecuación (3.12) obtenemos la función discriminante:

$$d_i(\mathbf{x}) = \mathbf{x}^T \mathbf{m}_j - \frac{1}{2} \mathbf{m}_j^T \mathbf{m}_j \quad (3.13)$$

A partir de las funciones discriminantes se definen las superficies de decisión:

$$d_{ij}(\mathbf{x}) = \mathbf{x}^T (\mathbf{m}_j - \mathbf{m}_i) + \frac{1}{2} (\mathbf{m}_i - \mathbf{m}_j)^T (\mathbf{m}_i + \mathbf{m}_j) = 0 \quad (3.14)$$

donde:

$$\begin{aligned} \mathbf{x} \in \omega_i &\text{ si } d_{ij}(\mathbf{x}) > 0 \\ \mathbf{x} \in \omega_j &\text{ si } d_{ij}(\mathbf{x}) < 0 \end{aligned}$$

La ecuación (3.14) es una superficie bisector respecto al segmento de línea que une a los prototipos \mathbf{m}_i y \mathbf{m}_j . Para $n = 2$, la superficie bisector es una la línea perpendicular, para $n = 3$ es un plano y para $n > 3$ es un hiperplano.

Capítulo 4. Técnicas de Reconstrucción Tridimensional de Objetos

La reconstrucción tridimensional de objetos, uno de los temas de mayor interés en la visión artificial, tiene como objetivo calcular la información espacial y volumétrica de los objetos a partir de un conjunto de imágenes bidimensionales. La reconstrucción tridimensional tiene un vasto campo de aplicación entre las cuales se encuentran el diseño asistido por computadora, en los procesos de manufactura, en la robótica, en aplicaciones médicas, en el entretenimiento, en la seguridad, entre otras [18, 19].

Para llevar a cabo la reconstrucción tridimensional de objetos se han desarrollado distintas técnicas y dispositivos que se aplican según las necesidades y complejidad de los objetos a reconstruir. Comúnmente los métodos de reconstrucción se clasifican: métodos activos y pasivos [20]. Los métodos activos interactúan con el objeto a reconstruir ya sea proyectando un haz de luz blanca, luz láser o bien interactuando de forma mecánica con el objeto a reconstruir. Mientras tanto los métodos pasivos no interactúan de forma directa con el objeto, en vez de ello usan sensores para medir la radiación reflejada por la superficie del objeto a reconstruir, por ejemplo la estereovisión. En la mayoría de los casos, se usa una cámara con lentes el cual es susceptible a los diferentes modelos de distorsión, esto hace necesario realizar un proceso de calibración utilizando algoritmos de optimización.

4.1 Visión estereoscópica

La reconstrucción tridimensional de objetos mediante visión estereoscópica está inspirada en el proceso biológico de visión de los seres humanos, el cual se basa en un sistema binocular o visión estéreo y la estereopsis [21].

La estereopsis es el fenómeno físico de la percepción de profundidad, lejanía o cercanía de los objetos, a partir de dos imágenes con perspectivas diferentes generadas por el ojo derecho e izquierdo. La posibilidad de obtener dos imágenes diferentes de los objetos, es gracias a que nuestros ojos se encuentran separados a una distancia llamada *distancia intraocular* (DIO). La diferencia de las imágenes producida en cada ojo se le denomina *disparidad*.

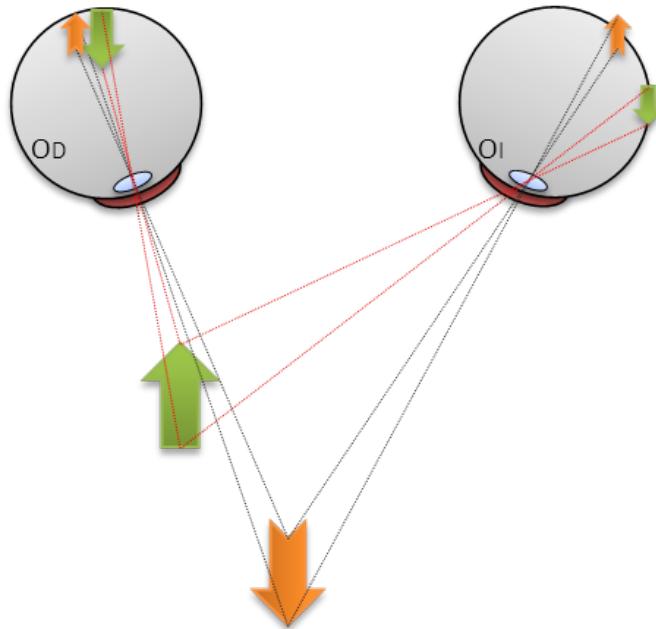


Figura 4.1. Proceso de formación de imágenes en el ojo humano. O_D y O_I son el ojo derecho e izquierdo, respectivamente.

La Figura 4.1 ilustra el proceso de formación de imágenes en el ojo humano. Las imágenes son proyectadas por el cristalino a la retina, donde las imágenes son invertidas y con perspectiva ligeramente diferente. La retina se encuentra en la parte posterior del ojo y está compuesta por dos tipos de células los conos y

bastones, estas células convierten las imágenes en señales eléctricas que son enviadas a la corteza visual del cerebro. Mediante un proceso denominado *fusión binocular* la corteza visual se encarga de integrar las dos imágenes, de esta manera podemos percibir la información de profundidad y nos permite calcular mentalmente las distancias [22].

4.1.1 La visión estereoscópica artificial

La visión estereoscópica artificial es la emulación del proceso de visión humana mediante el uso cámaras. El cálculo de la información tridimensional del objeto se basa en la triangulación y el cómputo de la disparidad de los puntos comunes en dos o más imágenes, los puntos comunes son llamados *puntos conjugados*.

La Figura 4.2 describe el arreglo óptico de un sistema de visión estereoscópico, éste requiere de dos cámaras CCD o CMOS que se denotan como cámara 1 y cámara 2, cuyos ejes ópticos son paralelos y se encuentran separados a una distancia b (distancia base, el equivalente al DIO en sistema de visión humano) y tienen una lente de distancia focal f . El sistema de referencia del mundo real se denota con la letra O , éste se encuentra situado en medio de la distancia base. Los puntos de referencia O_1 y O_2 son los sistemas de referencia de la cámara 1 y cámara 2 respectivamente. El objeto bajo análisis se encuentra situado a una distancia z del origen O . La *línea epipolar* es la línea que une los puntos conjugados.

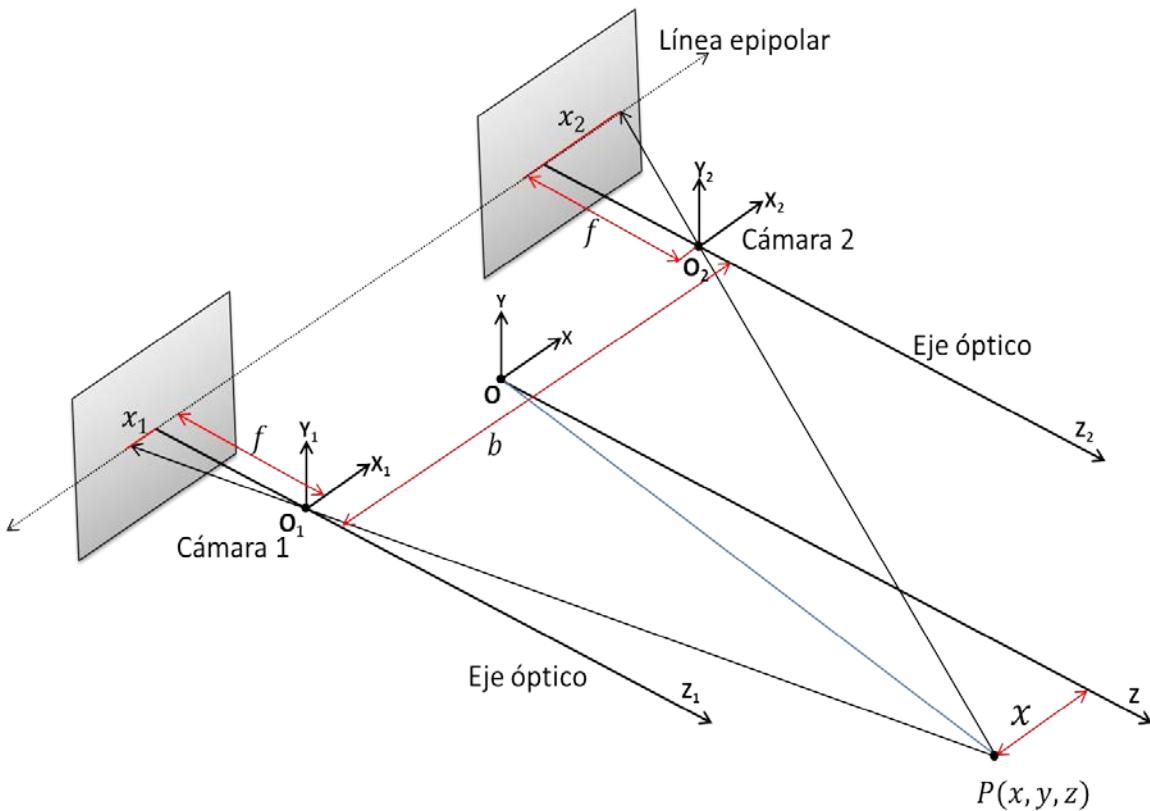


Figura 4.2. Arreglo óptico de un sistema de visión estereoscópica.

Analizando la geometría de la Figura 4.2, obtenemos las ecuaciones (4.1) y (4.2) correspondientes a la posición de los puntos conjugados en el plano imagen.

$$x_1 = \frac{f}{z} \left(x - \frac{b}{2} \right) \quad (4.1)$$

$$x_2 = \frac{f}{z} \left(\frac{b}{2} + x \right) \quad (4.2)$$

La disparidad de los puntos conjugados ($d = x_2 - x_1$) se definen mediante la ecuación (4.3).

$$d = \frac{fb}{z} \quad (4.3)$$

Conociendo el valor de la disparidad d , la distancia focal f y la distancia base b , es posible calcular la distancia z del punto $P(x, y, z)$ al origen O del sistema.

4.1.2 Localización de puntos conjugados.

La detección de los puntos conjugados es uno de los principales problemas dentro del proceso de la visión estereoscópica, éstos pueden o no estar en una de las imágenes. Uno de los métodos de localización es a partir del análisis de la similitud en la vecindad de un punto, utilizando herramientas como la correlación de imágenes [23]. Los puntos conjugados idealmente deben cumplir la siguiente condición:

$$I_1(x_1, y) = I_2(x_2, y) \quad (4.4)$$

Donde $I_1(x_1, y)$ e $I_2(x_1, y)$ corresponden al nivel de intensidad en las coordenadas (x_n, y) del plano imagen formado por la cámara 1 y cámara 2 respectivamente. Sin embargo como los niveles de intensidad no son del todo igual, se busca minimizar la función:

$$\varepsilon_i = \int \int (I_1 - I_2)^2 dx dy \quad (4.5)$$

La restricción se cumple cuando los valores de vecindad de $I_1(x_1, y)$ son aproximadamente iguales a los valores de la vecindad de $I_2(x_2, y)$.

4.2 Proyección de luz estructurada

La técnica de proyección de luz estructurada es la proyección de algún patrón de luz (luz blanca o luz laser) sobre una escena y registrarlos mediante un sistema de adquisición de imágenes (cámara digital CCD o CMOS), para llevar a cabo la reconstrucción tridimensional de objetos [24]. La fuente de iluminación debe proyectar algún patrón de luz como: puntos, una línea, un conjunto de líneas verticales u horizontales, una rejilla entre otros.

4.2.1 Reconstrucción tridimensional mediante proyección de franjas

La reconstrucción tridimensional mediante proyección de franjas es una técnica óptica de campo completo, no invasiva, rápida y versátil para realizar mediciones fuera del plano y obtener la forma 3D de objetos. Este método ha sido ampliamente usado en el campo de la metroología óptica para realizar mediciones de componentes MEMS hasta la medición de objetos relativamente grandes [25, 26].

Las franjas que suelen proyectarse tienen un perfil cosenoidal o binario. Considerando que se proyectan franjas con perfil cosenoidal sobre un plano de referencia, éstas permanecen sin distorsión; sin embargo al colocar un objeto de prueba en el plano de referencia, las franjas se distorsionan o “modulan” de acuerdo a la forma del objeto. Para llevar a cabo la reconstrucción tridimensional es necesario medir el desplazamiento que sufre cada línea respecto a la línea sin deformar y relacionar esta desviación con la información de profundidad del objeto.

Idealmente, el patrón de franjas que registra una cámara digital tiene un perfil cosenoidal, además si ignoramos la no telecentricidad y las aberraciones del sistema óptico podemos modelarlo matemáticamente como [26, 27]

$$I(x, y) = a(x, y) + b(x, y)\cos(2\pi f_o x + \phi(x, y)) \quad (4.6)$$

Donde $a(x, y)$ es la iluminación de fondo, $b(x, y)$ es un factor de modulación o el contraste de las franjas, el término f_o corresponde a la frecuencia espacial de la señal portadora y el término $\phi(x, y)$ es la fase adicional que se encuentra asociado con la forma del objeto. Esta ecuación puede ser comparada con la interferencia de dos haces en interferometría [28]. Para determinar la fase $\phi(x, y)$, pueden usarse técnicas que usan la Transformada de Fourier, el método de Demodulación Directa o el método de Desplazamiento de Fase, entre otros [29].

La Figura 4.3 corresponde al arreglo óptico de un sistema de proyección de franjas, el cual requiere de un sistema de adquisición de imágenes (C), un proyector multimedia (P) y un plano de referencia (R) [30]. Los ejes ópticos de la cámara y el proyector son paralelos, coplanares y son normales al plano de referencia. El eje x y y del sistema recorren la superficie del plano de referencia en dirección horizontal y vertical respectivamente, mientras que el eje z es perpendicular al plano de referencia y paralelo al eje óptico de la lente de la cámara y del proyector. Los puntos C y P se encuentran situados a una distancia d_0 del plano de referencia.

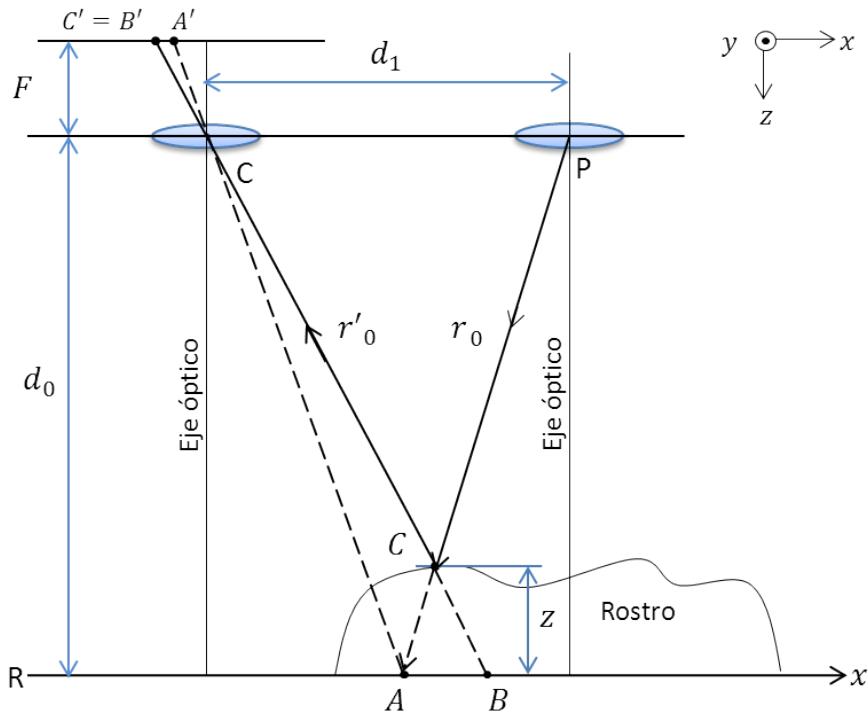


Figura 4.3. Arreglo óptico de un sistema de proyección de luz estructurada.

Se puede observar que antes de colocar un objeto sobre el plano de referencia, el rayo incidente r_0 es reflejado desde el punto A hasta el punto A' del sensor de la cámara. Si se coloca un objeto de altura z sobre el plano de referencia, r_0 ahora se reflejará como r'_0 desde el punto C hasta punto B' del sensor, además el objeto hace que el punto A sufra un desplazamiento hasta el punto B en el plano de

referencia. A partir del cálculo del desplazamiento \overline{AB} se define la relación que permite obtener la variación de la altura de la superficie del objeto bajo análisis. La fase de la franja proyectada en los puntos C y A tiene mismo valor, además el punto C y el punto B son proyectados en el mismo punto del sensor de la cámara. De esta manera es posible expresar la distancia \overline{AB} mediante la ecuación.

$$\overline{AB} = \frac{\phi_B - \phi_C}{2\pi f_0} \quad (4.7)$$

Podemos concluir que la distancia \overline{AB} es proporcional al cambio de fase debido a la presencia del objeto sobre el plano de referencia. A partir de la distancia \overline{AB} , podemos calcular la altura z del objeto. De la Figura 4.3 sí se analizan los triángulos CPO y ABO se obtiene la relación:

$$\overline{AB} = \frac{z d_1}{d_0 - z} \quad (4.8)$$

Al sustituir la ecuación anterior en la ecuación (4.7) se obtiene la profundidad z :

$$z(x, y) = \frac{(\phi_B - \phi_C)d_0}{2\pi f d_1 + (\phi_B - \phi_C)} \quad (4.9)$$

4.3 Métodos de obtención de la fase

Entre los métodos, ampliamente usados, para obtener la fase modulada se encuentran los que usan la Trasformada de Fourier, el método de Desplazamiento de Fase y el método de Detección Espacial de Fase.

4.3.1 Método de la Transformada de Fourier

Este método, propuesto por Takeda y Mutoh en 1983, aplica la Trasformada de Fourier a un patrón de franjas modulado por la superficie de un objeto, posteriormente aplica un filtro pasabanda a la frecuencia portadora, finalmente se realiza el cálculo de la Trasformada Inversa de Fourier obteniendo un conjunto de datos con valores complejos, a partir de estos datos es posible obtener la fase asociada a la forma del objeto bajo análisis [31].

Para realizar la demodulación del patrón de franjas se puede modelar la función de intensidad como [28]:

$$I(x, y) = a(x, y) + \frac{b}{2}(x, y)e^{i2\pi f_0 x}e^{i\phi(x, y)} + \frac{b}{2}(x, y)e^{-i2\pi f_0 x}e^{-i\phi(x, y)} \quad (4.10)$$

Al aplicar la Trasformada de Fourier en términos del eje x , el espectro de frecuencias del eje x esta dado por:

$$\begin{aligned} FT[I(x, y)] &= \int_{-\infty}^{\infty} I(x, y) e^{-i2\pi u_0 x} dx \\ I_F &= A(u, v) + B(u - u_0, v) + B^*(u + u_0, v) \end{aligned} \quad (4.11)$$

De acuerdo a la figura 4.4 (b), el primer término de la ecuación (4.11) corresponde a la iluminación de fondo cuya frecuencia es baja por tanto en el espacio de Fourier aparecen centrados en $u = 0$. Los siguientes dos términos de la ecuación aparecen centrados en $u = u_0$ y su magnitud es simétrica.

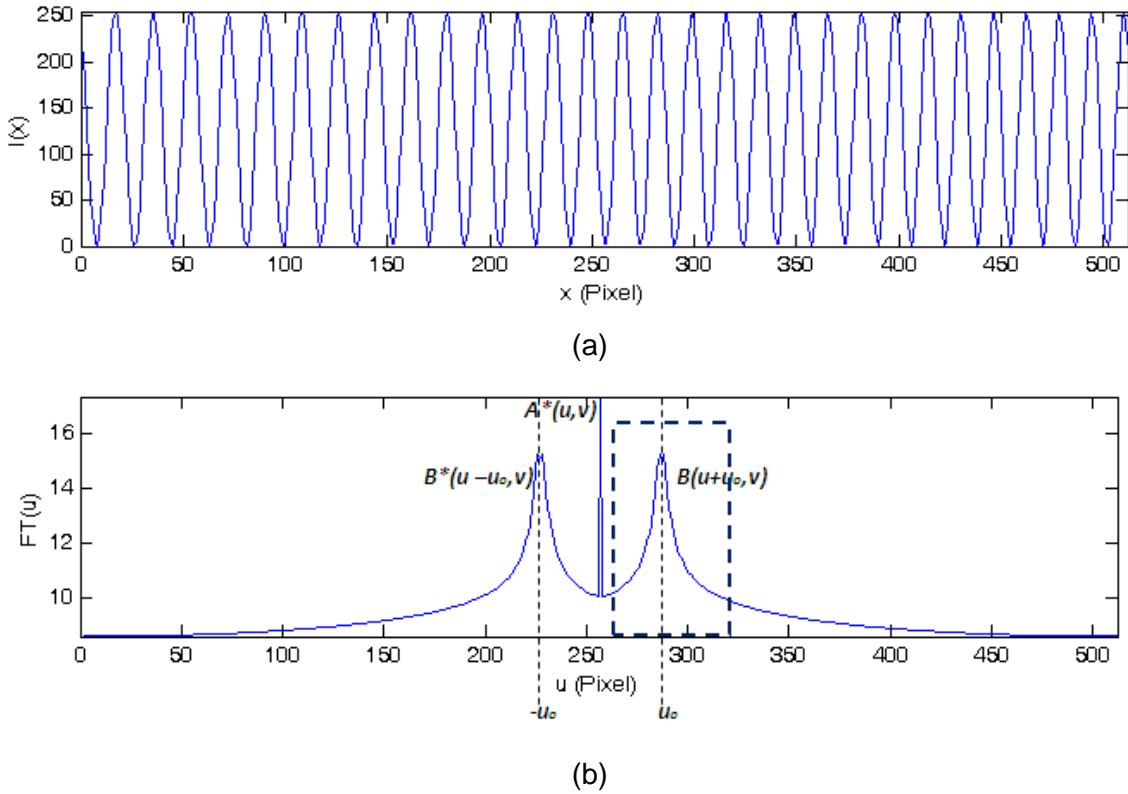


Figura 4.4. (a) Perfil horizontal de una imagen con franjas y (b) perfil horizontal de la Trasformada de Fourier de la imagen (a)

Al aplicar un filtro pasabandas de manera que solo deje pasar la información centrado en u_0 a demás si trasladamos $B(u + u_0, y)$ al origen, obtenemos:

$$I_F = B(u, y) \quad (4.12)$$

Aplicando la Trasformada inversa de Fourier a la ecuación (4.12):

$$c(x, y) = \frac{1}{2}b(x, y) + e^{i\phi(x, y)} \quad (4.13)$$

Aplicando logaritmo a la ecuación:

$$\log [c(x, y)] = \log \left[\frac{1}{2}b(x, y) + e^{i\phi(x, y)} \right] \quad (4.14)$$

Calculamos la fase $\phi(x,y)$ mediante la ecuación 4.16:

$$\phi(x,y) = \tan^{-1} \frac{Im(c(x,y))}{Re(c(x,y))} \quad (4.15)$$

Debido a la naturaleza de la función tangente, la fase $\phi(x,y)$ presenta discontinuidades cada 2π , fase envuelta, por lo que será necesario aplicar algún método que nos permita obtener una fase continua.

4.3.2 Détodo de Desplazamiento de Fase

El método de Desplazamiento de Fase, consiste en proyectar y registrar secuencialmente N patrones de franjas, con una fase inicial ϕ , sobre el plano de referencia y la superficie del objeto bajo análisis. Entre cada patrón de franjas hay un desplazamiento de fase dado por $\nabla\psi = 2\pi/N$ y se requiere por lo menos que $N = 3$, en principio entre mayor sea N , menor es el nivel de ruido aleatorio el cual decrece por un factor $1/\sqrt{N}$ [32-35].

La figura 4.5 corresponde al ejemplo de franjas de perfil cosenoidal, generada por computadora, con un desplazamiento de 4 pasos ($\nabla\psi = \pi/2$).

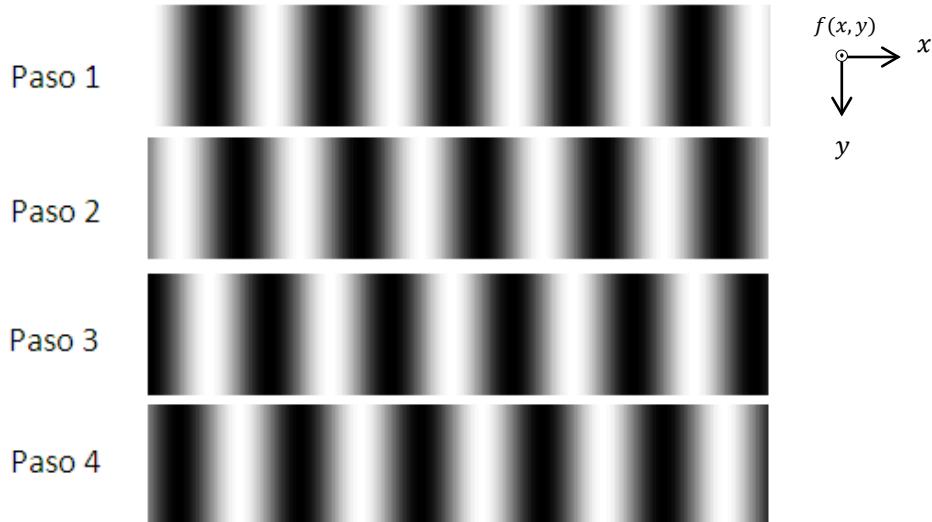


Figura 4.5. Franjas generadas por computadora, con desplazamiento de fase $k = 4$.

El patrón de intensidad registrado por el sistema de captura es descrito por la ecuación (4.16).

$$I_k = a(x, y) + b(x, y)\cos(\phi(x, y) + (k - 1)\nabla\psi) \quad (4.16)$$

donde k es el número secuencial de la captura y $\nabla\psi$ es el desplazamiento de la fase.

Una de las maneras de encontrar la relación entre el mapa de fase $\phi(x, y)$ y la intensidad de cada píxel es mediante la solución de mínimos cuadrados.

$$\phi(x, y) = \tan^{-1} \frac{I_k(x, y)\sin[(k - 1)\nabla\psi]}{I_k(x, y)\cos[(k - 1)\nabla\psi]} \quad (4.17)$$

Para el caso de $N = 3$ la fase está dada por la ecuación:

$$\phi(x, y) = \tan^{-1} \frac{I_3(x, y) - I_2(x, y)}{I_1(x, y) - I_2(x, y)} \quad (4.18)$$

Para el caso de $N = 4$:

$$\phi(x, y) = \tan^{-1} \frac{I_4(x, y) - I_2(x, y)}{I_1(x, y) - I_3(x, y)} \quad (4.19)$$

Para el caso de $N = 5$ (Método de Hariharan) la fase está dada por la ecuación:

$$\phi(x, y) = \tan^{-1} \frac{2 [I_2(x, y) - I_4(x, y)]}{2 I_3(x, y) - I_5(x, y) - I_5(x, y)} \quad (4.20)$$

La fase $\phi(x, y)$ obtenida en cada uno de los casos presenta discontinuidades cada 2π radianes, es decir se encuentra envuelta.

Debido que en el método de desplazamiento de fase requiere de la captura de N patrones de franjas, su uso está limitado para el análisis de objetos estáticos. Para el caso de objetos cuasi-estáticos es necesario el uso de equipos de alta velocidad.

4.3.3 Demodulación directa de fase

El método de demodulación directa de fase, propuesto por Ichioka e Inuiya en 1972, consiste en multiplicar la el patrón de franjas por las funciones seno y coseno con una frecuencia portadora. Posteriormente mediante un filtro pasa bajas, son eliminadas las bajas frecuencias para obtener finalmente dos señales en términos de fase relacionada con la forma del objeto [29].

Recordando la ecuación que describe la función de intensidad de un patrón de franjas:

$$I(x, y) = a(x, y) + b(x, y)\cos(2\pi f_0 x + \phi(x, y)) \quad (4.21)$$

Al multiplicar la ecuación $I(x, y)$ por la función seno y coseno obtenemos:

$$I_s = a(x, y)\sin(2\pi f_0 x) + \frac{1}{2}\sin(4\pi f_0 x + \phi(x, y)) - \frac{1}{2}\sin(\phi(x, y)) \quad (4.22)$$

$$I_c = a(x, y)\cos(2\pi f_0 x) + \frac{1}{2}\cos(4\pi f_0 x + \phi(x, y)) - \frac{1}{2}\cos(\phi(x, y)) \quad (4.23)$$

A partir de estas dos ecuaciones se aplica un filtro pasa bajos para aislar el término de baja frecuencia, el término reaccionado con la forma 3D del objeto:

$$I_{Hs} = -\frac{1}{2}\sin(\phi(x, y)) \quad (4.24)$$

$$I_{Hc} = -\frac{1}{2}\cos(\phi(x, y)) \quad (4.25)$$

Después de aplicar los filtros es posible obtener calcular la fase mediante:

$$\phi(x, y) = \tan^{-1}\left(\frac{I_{Hs}}{I_{Hc}}\right) \quad (4.26)$$

4.4 Técnicas de desenvolvimiento de fase

Las técnicas de obtención de fase analizadas previamente, hacen uso de la función arco tangente la cual está acotada en el rango de $[-\pi, \pi]$ radianes, este hecho implica que la fase tenga saltos cada 2π radianes, a este tipo de fase se le conoce como *fase envuelta* [36].

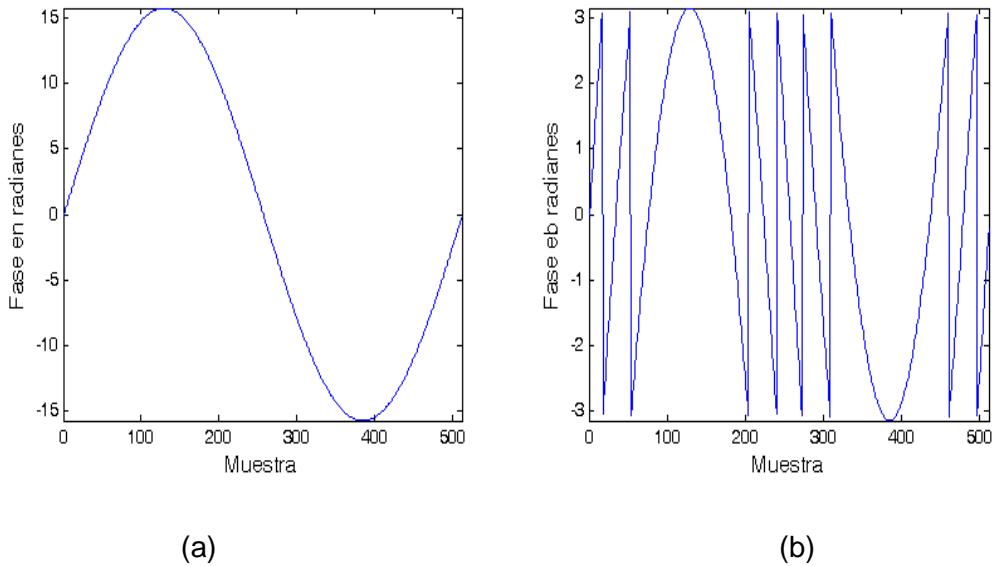


Figura 4.6. Señal original y fase envuelta.

Suponiendo que tenemos una señal discreta $x(n)$ que contiene muestras con amplitudes que exceden el rango de $[-\pi, \pi]$ radianes, tal como se observa en la Figura 4.6 (a). Al aplicar el operador arco tangente a la señal $x(n)$, se obtiene la fase envuelta $\psi(n)$, Figura 4.6 (b). Los saltos de 2π deben de ser removidos para obtener una señal continua. El proceso de eliminar los saltos en la fase envuelta se le denomina *desenvolvimiento de fase* [37].

En los problemas reales, por ejemplo en proyección de franjas, la mayoría de las ocasiones las señales contiene ruido aleatorio, presentan discontinuidades o están submuestreadas. Lo cual disminuye el desempeño de los algoritmos de desenvolvimiento de fase produciendo inconsistencias en los resultados. Debido a

este problema, se han desarrollado numerosos algoritmos de desenvolvimiento de fase con el intento de mejorar los resultados [29, 37].

4.4.1 Desenvolvimiento de fase unidimensional

En este apartado se analizan dos métodos para el desenvolvimiento de fase: el método básico y el método de Itoh.

El método básico, detecta los saltos de 2π radianes y suma o resta, según sea el caso, 2π radianes al elemento bajo análisis [36]. Considerando la fase envuelta $\psi(n)$ de la Figura 4.6 (b) aplicaremos el algoritmo básico de desenvolvimiento de fase el cual se describe a continuación:

1. Calcular la diferencia de fase $\Delta\psi = \psi(i + 1) - \psi(i)$, para $i = 0, \dots, N - 2$, N es el número de muestras.
2. Si la diferencia de fase $\Delta\psi$ es mayor a π , entonces se sustrae 2π radianes a todas las muestras a partir de la muestra $\psi(i + 1)$ hasta $\psi(N)$.
3. Si la diferencia entre las dos muestras es menor a $-\pi$, entonces se suma 2π radianes a todas las muestras a partir de la muestra $\psi(i + 1)$ hasta $\psi(N)$
4. Continuar hasta procesar la última muestra.

Una segundo método de desenvolvimiento de fase es el *Método de Itoh*, el cual se describe a continuación [38].

1. Calcular la diferencia de fase $\Delta\psi(i) = \psi(i + 1) - \psi(i)$, para $i = 0, \dots, N - 2$
2. Calcular la diferencia de fase envuelta $\Delta(i) = \tan^{-1} \frac{\sin(\Delta\psi(i))}{\cos(\Delta\psi(i))}$
3. Inicializar el primer elemento de la fase desenvuelta $\phi(0) = 0$
4. Desenvolver sumando la diferencia de fases : $\phi(i) = \phi(i - 1) + \Delta(i - 1)$, para $i = 1, \dots, N - 1$

4.4.2 Desenvolvimiento de fase de funciones bidimensionales mediante mapas de calidad siguiendo rutas discretas

El algoritmo que se describe a continuación es un método propuesto por A. Arevalio, et al [37] el cual desenvuelve los píxeles según la calidad o bondad que los describe. Los píxeles con mejor calidad se desenvuelven primero y los de menor calidad son desenvueltos al final para minimizar la propagación de errores. La calidad de los píxeles se determina usando la segunda derivada y el proceso de desenvolvimiento sigue una ruta no continua de acuerdo a la bondad de los píxeles.

El algoritmo de desenvolvimiento de fase implica dos tareas: asignar la calidad a cada píxel mediante una función de selección de aptitud y diseñar la ruta de desenvolvimiento de fase que minimice la propagación de errores durante el proceso de desenvolvimiento.

a) Función de selección de aptitud (Mapas de calidad)

Los mapas de calidad, son arreglos de valores que indican la aptitud de cada elemento de un mapa de fase y son usados para guiar el orden del desenvolvimiento de fase. Los mapas de calidad utilizan diferentes criterios para determinar la aptitud de un píxel, los cuales pueden ser: mapas de correlación, mapas de pseudocorrelación, varianza de fase y gradiente de fase [38].

El criterio selección de aptitud que se utiliza en este algoritmo es el de gradiente de fase, el cual se basa en la diferencia de un píxel respecto a sus vecinos. Aquellos puntos cuyo gradiente es menor a 2π radianes son considerados como los más aptos y son los primeros en ser procesados. El gradiente de fase se calcula mediante la segunda derivada del mapa de fase, el cual nos da información de la concavidad y permite detectar posibles inconsistencias.

Para realizar el cálculo de la segunda derivada en un píxel respecto a sus vecinos nos basamos en la Figura 4.7.

$(i - 1, j - 1)$	$(i - 1, j)$	$(i - 1, j + 1)$
$(i, j - 1)$	(i, j)	$(i, j + 1)$
$(i + 1, j - 1)$	$(i + 1, j)$	$(i + 1, j + 1)$

Figura 4.7. Vecindad de un píxel del (i, j) .

El gradiente de fase lo define la ecuación (4.27).

$$D(i, j) = [H^2(i, j) + V^2(i, j) + D_1^2(i, j) + D_2^2(i, j)]^{\frac{1}{2}} \quad (4.27)$$

Dónde: $H(i, j) = \gamma[\varphi(i - 1, j) - \varphi(i, j)] - \gamma[\varphi(i, j) - \varphi(i + 1, j)]$

$$V(i, j) = \gamma[\varphi(i, j) - 1 - \varphi(i, j)] - \gamma[\varphi(i, j) - \varphi(i, j + 1)]$$

$$D_1^2(i, j) = \gamma[\varphi(i - 1, j - 1) - \varphi(i, j)] - \gamma[\varphi(i, j) - \varphi(i + 1, j + 1)]$$

$$D_2^2(i, j) = \gamma[\varphi(i - 1, j + 1) - \varphi(i, j)] - \gamma[\varphi(i, j) - \varphi(i + 1, j - 1)]$$

γ es un operador de desenvolvimiento de fase que elimina cualquier salto de 2π entre dos píxeles consecutivos.

La calidad $R(i, j)$ de un píxel se determina mediante la ecuación (4.28).

$$R(i, j) = \frac{1}{D(i, j)} \quad (4.28)$$

a) Ruta del desenvolvimiento de fase

Para diseñar la ruta del desenvolvimiento de fase se realiza el cálculo de bordes horizontales y verticales a partir del mapa de calidad. Un borde horizontal es la suma de la calidad de los píxeles adyacentes en sentido horizontal. Un borde vertical es la suma de la calidad de los píxeles adyacentes en sentido vertical. En

la Figura 4.8 (a) corresponde al mapa de calidad de una fase envuelta y (b) describe los bordes horizontales (en color celeste) y verticales (en color verde).

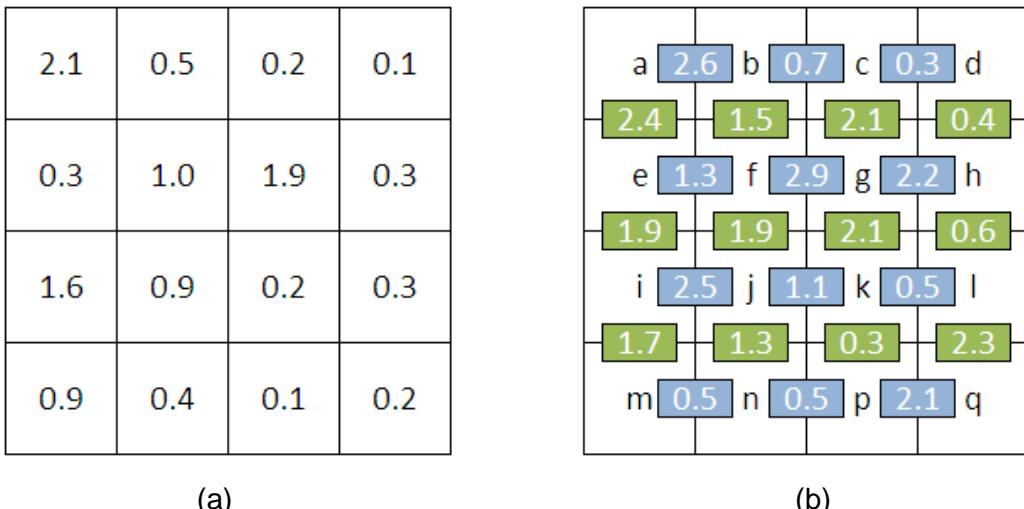


Figura 4.8. (a) Mapa de calidad de la fase envuelta y (b) diseño de ruta de desenvolvimiento de fase, bordes horizontales (en color celeste) y verticales (en color verde).

El primer paso del desenvolvimiento comienza con la localización de los dos píxeles adyacentes que tienen el valor de calidad más alto, estos pueden ser un borde horizontal o vertical, y se desenvuelven creando el primer grupo de fase desenvuelta. Seguidamente se busca el par de píxeles con el segundo valor de calidad más alto y se desenvuelven formando el segundo grupo de fase desenvuelta, se continúa este proceso hasta que se tenga que desenvolver un píxel respecto a un grupo, esto se realiza sumando o restando un múltiplo de 2π al pixel en cuestión. Se continúa de esta manera hasta que se desenvuelvan dos grupos adyacentes, se sumará o restara un múltiplo de 2π radianes al grupo que contenga el menor número de píxeles. A partir de entonces se continuará desenvolviendo grupos adyacentes hasta obtener un solo grupo de fase desenvuelta.

Capítulo 5. Diseño de un Sistema de Reconocimiento Facial a partir del Análisis Tridimensional del Rostro

5.1 Introducción

En este capítulo se describe el diseño de un Sistema de Reconocimiento Facial, el cual es una aplicación biométrica que implementa técnicas de reconocimiento de patrones, análisis estadístico, visión e inteligencia artificial. El sistema está conformado por distintos módulos que operan de manera sistemática sobre la información tridimensional del rostro para reconocer al usuario.

A través del presente capítulo se describe el diseño de los módulos del sistema, se dan detalle de las técnicas implementadas y se presentan los resultados obtenidos.

El esquema de la Figura 5.1 describe el diseño general del sistema.

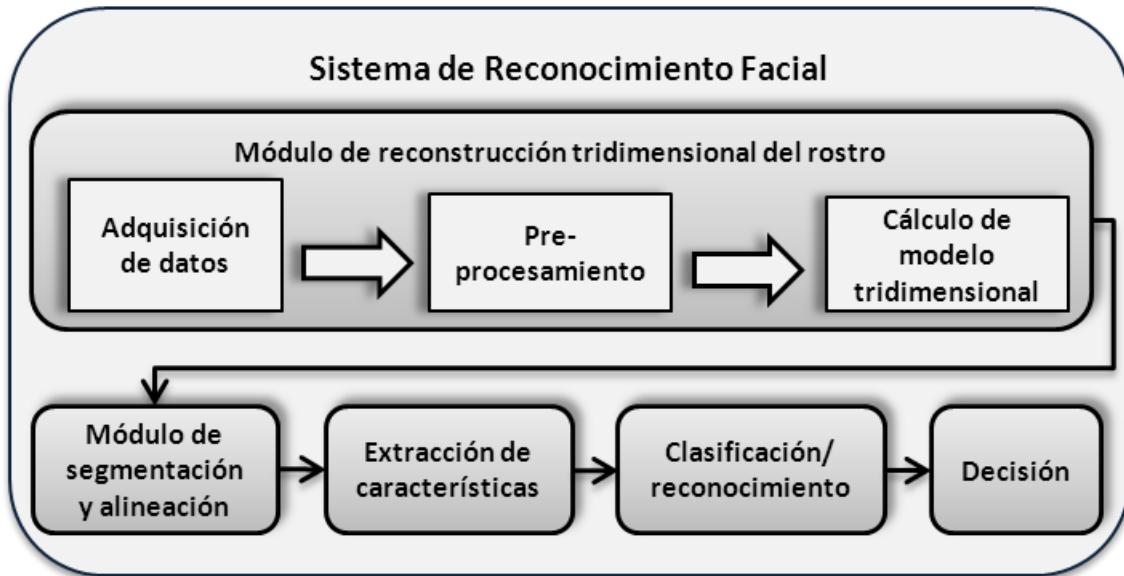


Figura 5.1. Diseño del sistema de reconocimiento facial.

5.1.1 Conceptos claves

En este capítulo se hace uso de los siguientes conceptos, relacionados al reconocimiento de patrones:

Usuario: Es aquella persona que hace uso del sistema de identificación facial. Su información biométrica y sus datos de identidad se encuentran almacenados en el sistema.

Clase: Una clase es definida por un usuario en particular. El sistema tendrá tantas clases como usuarios registrados haya en el sistema. Una clase se denota con ω_i .

Patrón: Los patrones se refieren a los rostros 3D de cada usuario. Un patrón se denota $z(x, y) \in \omega_i$, éste contiene todos los puntos que conforman la superficie del rostro 3D.

Descriptor. Son las características relevantes de los patrones. Un descriptor es un vector columna que se denota con el símbolo x .

Base de datos 3D. Es la colección de los rostros 3D (patrón) almacenados en un disco local de una computadora.

Identidad. Es la información asociada a un usuario, éste puede ser el nombre, el apellido, la edad, la dirección y el teléfono, entre otros.

5.1.2 Aspectos técnicos

El sistema de identificación está integrado por un arreglo óptico para realizar la digitalización del rostro y un conjunto de programas para llevar a cabo los siguientes procesos: reconstrucción tridimensional del rostro, segmentación de la región de interés, la extracción de características, la clasificación y el reconocimiento facial.

El arreglo óptico del sistema de digitalización de rostros consiste en un sistema de proyección de luz estructurada, el cual involucra el uso de un proyector de video LCD con una resolución nativa de 640X480 píxeles, una cámara CMOS monocromática PixeLink® con una resolución máxima de 1280X1024, un plano de referencia que consiste en una superficie plana con un orificio en la parte central donde el usuario asoma el rostro, una computadora portátil con procesador Intel Core Duo® de 1.4 Ghz y 2 Gb en memoria RAM. El diseño del sistema de proyección de luz estructurada se puede observar en la figura 5.2.

Las técnicas y algoritmos que se implementaron en los distintos módulos del sistema fueron programados en el entorno de desarrollo integrado Matlab R2011a®, bajo el sistema operativo Windows Xp®.

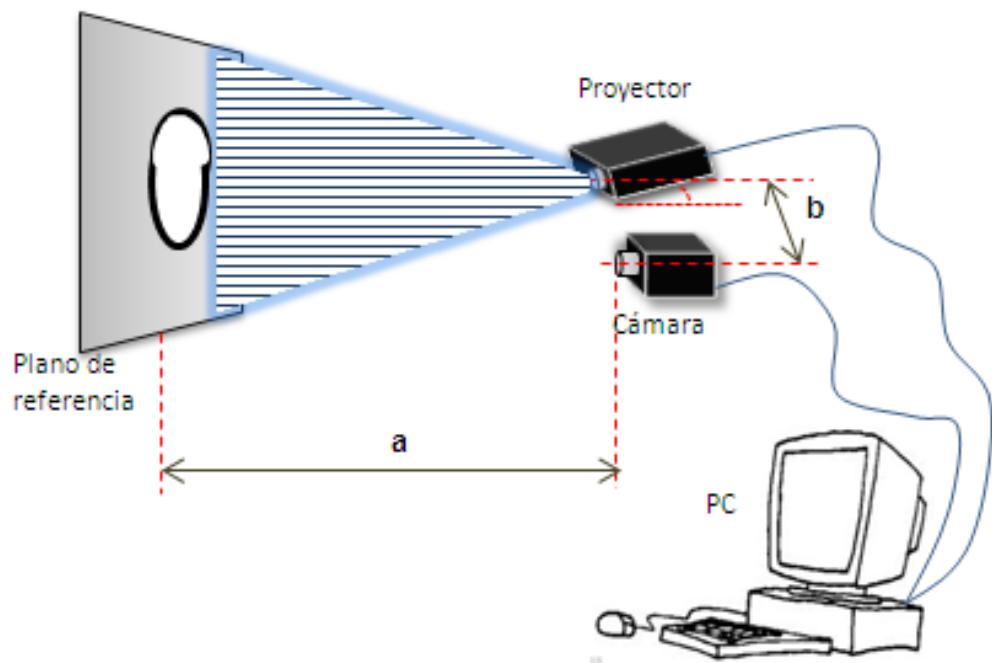


Figura 5.2. Diseño de sistema de adquisición de datos, basado en proyección de luz estructurada.

5.2 Módulo de reconstrucción tridimensional del rostro

El módulo de reconstrucción tridimensional tiene como objetivo obtener el mapa de profundidad del rostro (modelo 3D), éste se encuentra conformado por tres submódulos: submódulo de adquisición de datos, submódulo de pre-procesamiento y submódulo de obtención del modelo tridimensional del rostro, tal como se observa en la Figura 5.3.

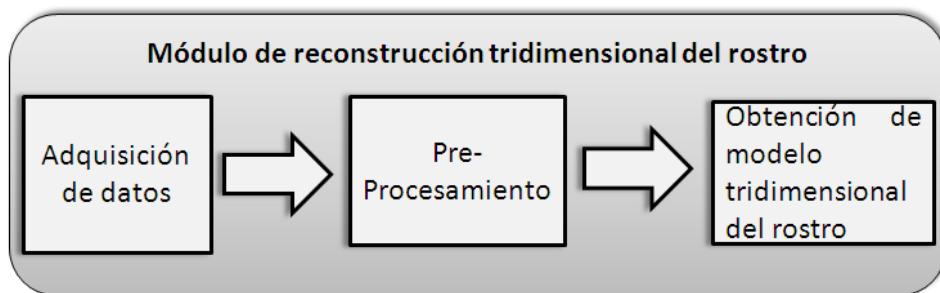
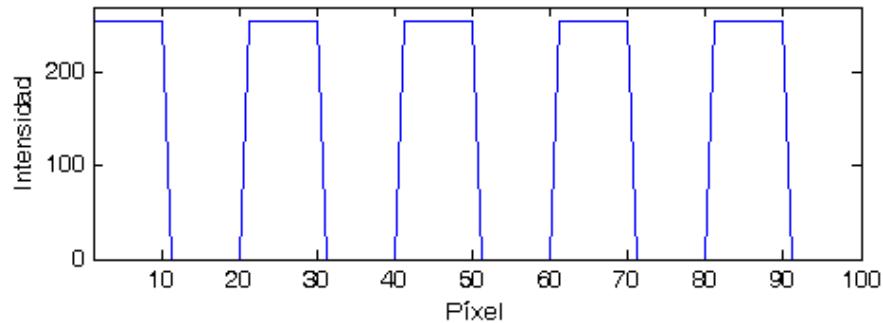


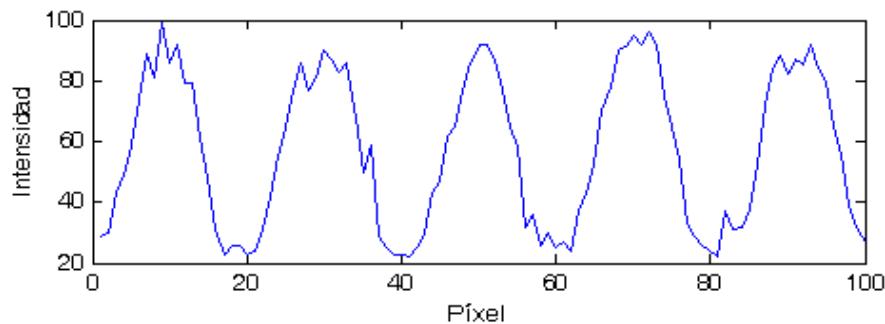
Figura 5.3. Módulo de reconstrucción de tridimensional del rostro.

5.2.1 Submódulo de Adquisición de Datos

El submódulo de adquisición de datos utiliza técnica de proyección de luz estructurada para desplegar un conjunto de patrones de franjas binarias con desplazamiento de fase de cuatro pasos. Proyectar franjas con perfil binario tiene algunas ventajas respecto a las de perfil cosenoidal, ya que presentan uniformidad frente a cambios de intensidad del proyector de video, además el patrón registrado tiende a adoptar el perfil de una señal cosenoidal debido a las aberraciones y no telecentridad de la lente de sistema de captura, tal como se observa en el ejemplo de la Figura 5.4.



(a)



(b)

Figura 5.4 Perfil de patrón de franjas (a) generado por computadora y (b) el perfil registrado mediante una cámara digital.

Para generar un patrón de franjas con perfil binario se considera lo siguiente:

- Las franjas tienen dos posibles valores en la memoria de un ordenador 0 o 1, que equivale a un píxel apagado (oscuro) y otro encendido (brillante).
- Al igual que una señal de tipo senoidal, un patrón de franjas con perfil binario es descrito por un periodo P y una frecuencia f_0 .
- El período de la señal debe ser un múltiplo de 4 píxeles para poder realizar un desplazamiento de cuatro pasos. De esta manera, la señal en el primero y quinto paso se encuentren sustancialmente en la misma posición.

El patrón de franja con perfil binario, generado por computadora, es descrito por la ecuación (5.1).

$$I_k(x, y) = \begin{cases} 0 & \text{si } (x_{i,j} - \nabla\psi) \bmod P < P/2 - 1 \\ 1 & \text{si } (x_{i,j} - \nabla\psi) \bmod P > P/2 - 1 \end{cases}; \quad \nabla\psi = (P/4) * k; \quad k = 1 \dots 4 \quad (5.1)$$

donde k es el número secuencial de patrones de franjas, $\nabla\psi$ es el desplazamiento fase relativo entre cada patrón de franjas y P es el período de la señal. El proceso de proyección y captura de franjas se describe en el pseudocódigo de la Figura 5.5.

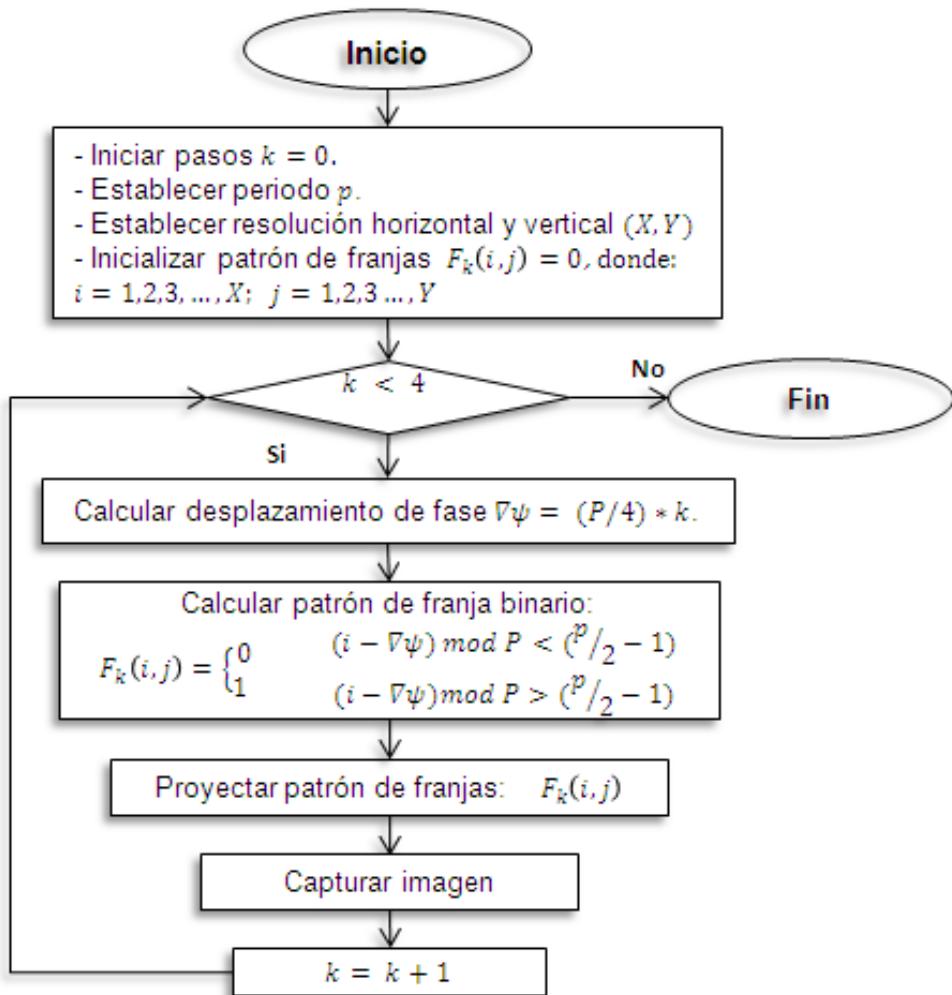


Figura 5.5. Pseudocódigo para generar patrones de franjas con perfil binario.

La Figura 5.6 corresponde a un ejemplo de los patrones de franjas generados por computadora. Donde F_1, F_2, F_3 y F_4 son muestras de los patrones generados, entre cada una de ellas hay un desplazamiento de fase relativo de $\nabla\psi$ píxeles.

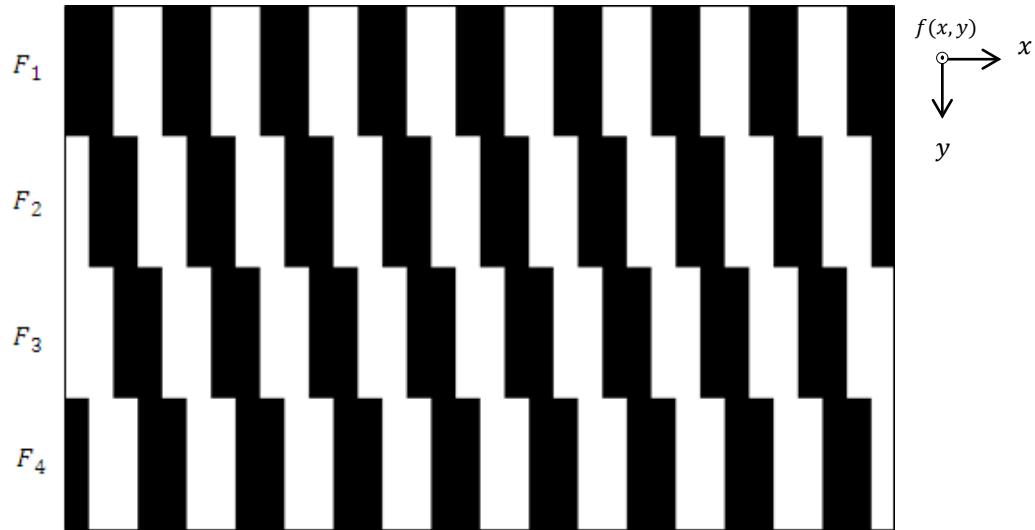
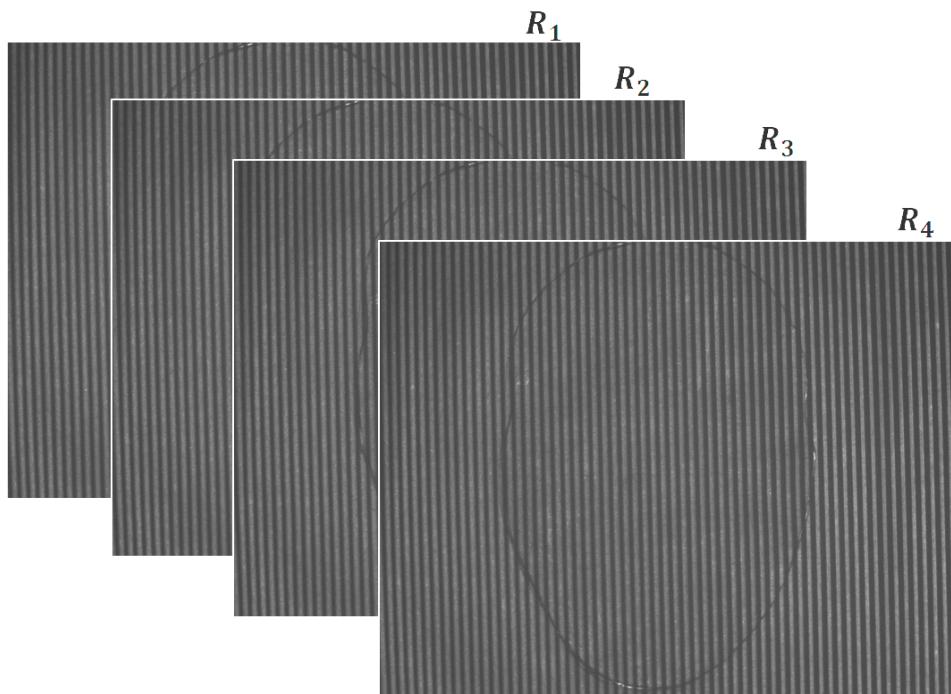


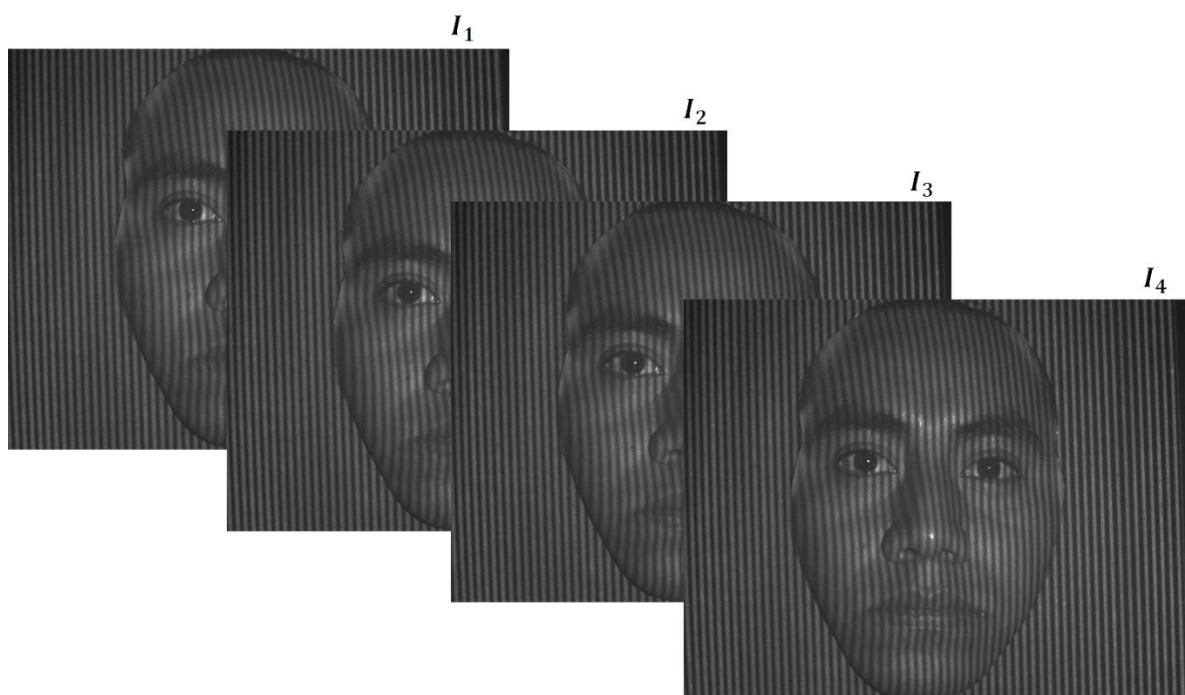
Figura 5.6 Patrones de franjas de perfil binario generados por computadora.

Los patrones de franja de la figura 5.6 son proyectados de manera consecutiva sobre el plano de referencia y la superficie del rostro, mientras son registrados por una cámara digital.

La Figura 5.7 (a) presenta las cuatro imágenes (Ir_1, Ir_2, Ir_3 , e Ir_4) registradas del plano del referencia; la Figura 5.7 (b) corresponde a las imágenes del rostro (Iu_1, Iu_2, Iu_3 , e Iu_4). Las imágenes tienen una resolución de 1280X1024 píxeles y una profundidad de color de 8 bits.



(a)



(b)

Figura 5.7. Patrones de franjas binarias proyectadas y registradas (a) sobre el plano de referencia y (b) sobre el rostro.

5.2.2 Submódulo de Preprocesamiento

Las imágenes registradas en el proceso anterior presenta cambios de intensidad abruptos (ruido aleatorio), éste tipo de variación se interpreta como una señal de alta frecuencia tal como se observa en la figura 5.8 (a). La señal de alta frecuencia puede ser atenuada o eliminada aplicando un filtro pasa bajas en el dominio de la frecuencia o realizando un procesamiento local aplicando un promediado de los píxeles adyacentes a un punto.

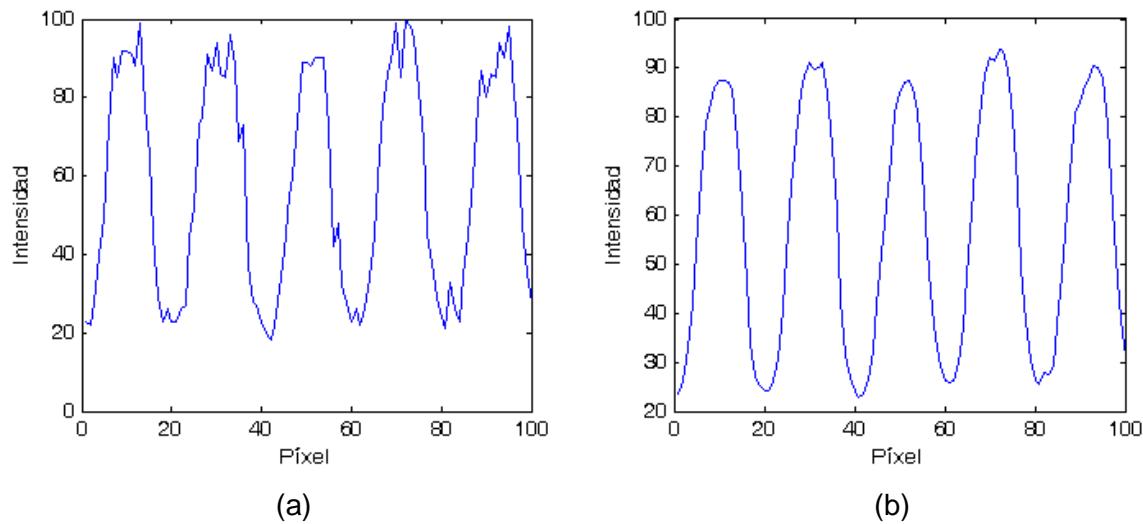


Figura 5.8. Señal (a) original y (b) procesada.

En éste caso el la atenuación del ruido de alta frecuencia se realizó mediante la aplicación de un filtro pasa bajas Butterworth en el dominio de la frecuencia. El proceso se describe en el pseudocódigo de la Figura 5.9. En el apéndice B- 6 se desarrolla el tema del filtro digital de imágenes en el dominio de la frecuencia.

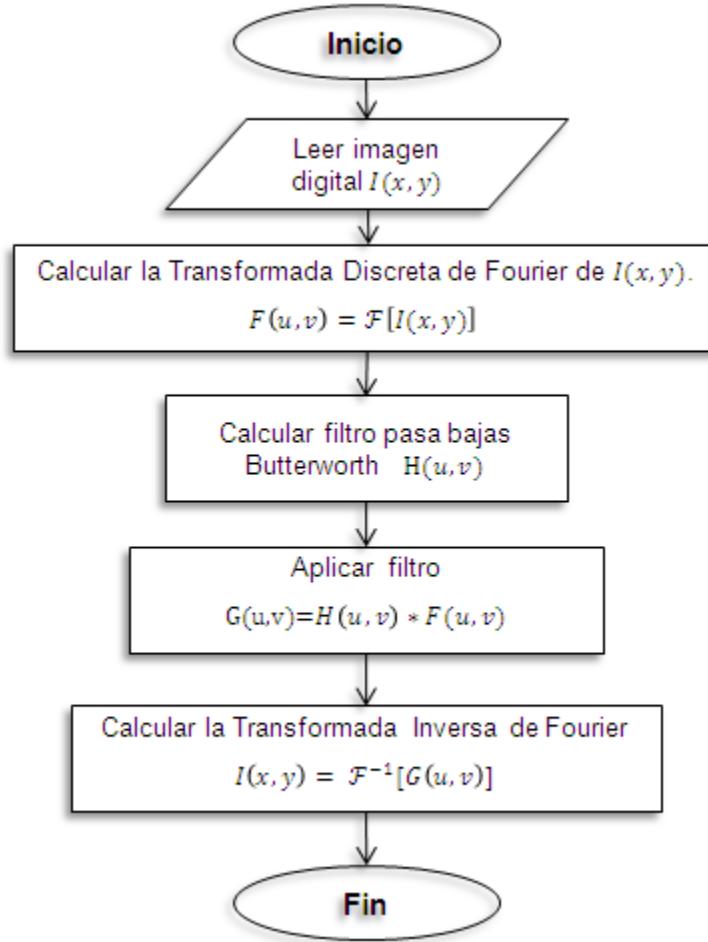
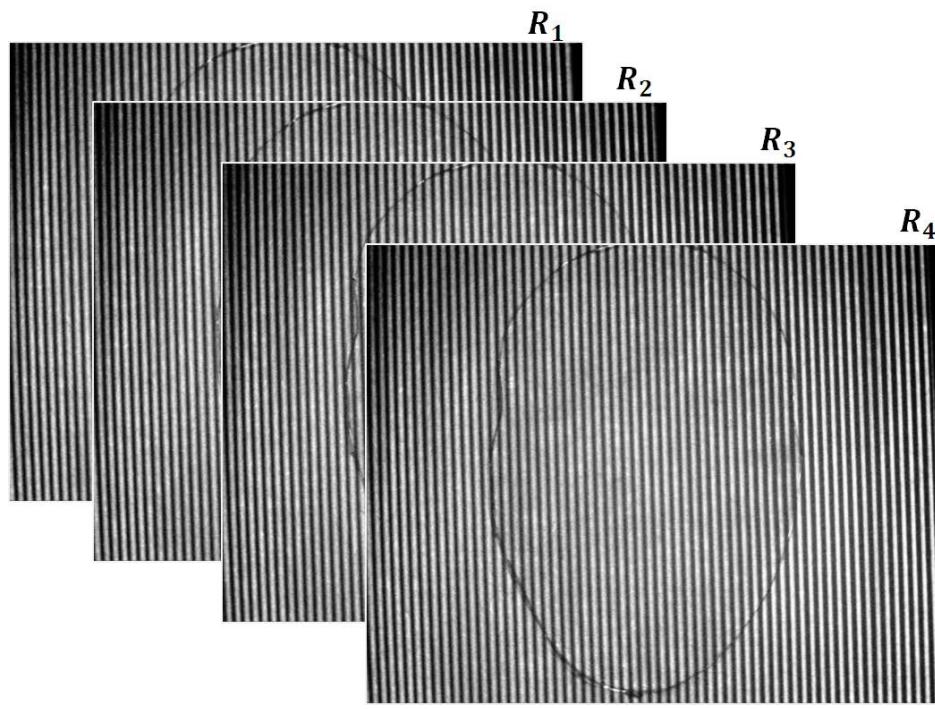
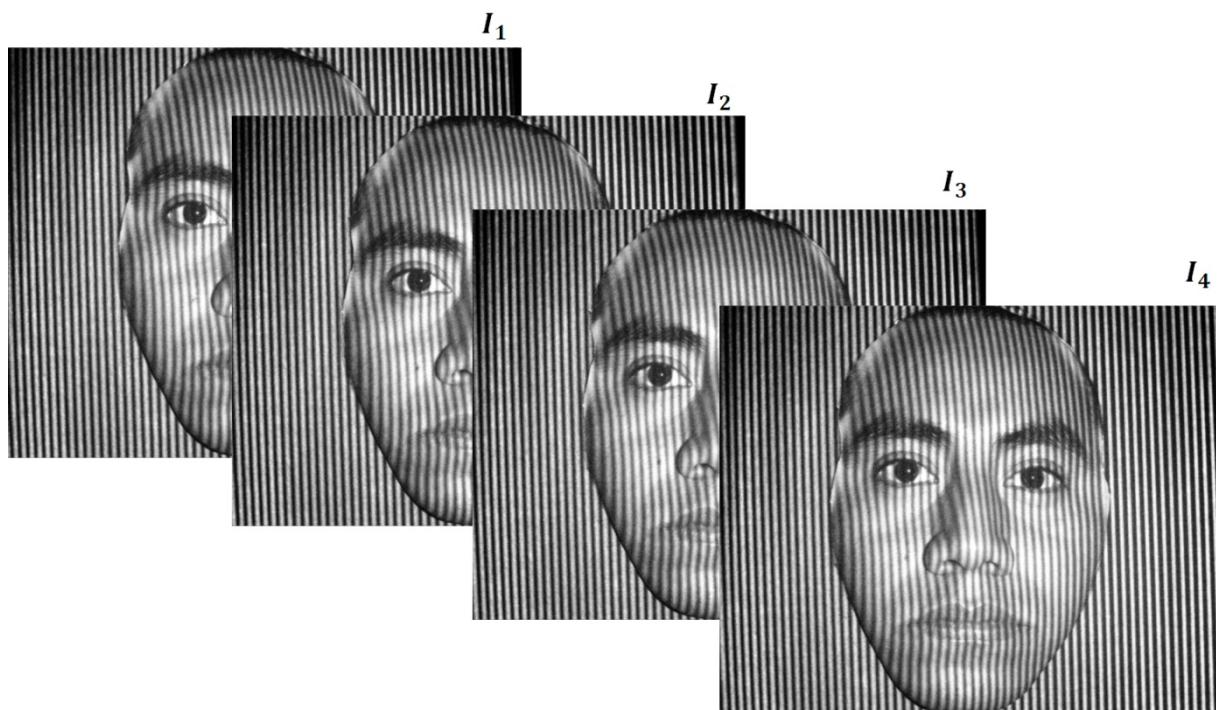


Figura 5.9. Pseudocódigo de filtrado de imagen digital en el dominio de frecuencia

Al aplicar el filtro se atenúan las variaciones aleatorias de los píxeles, se mejora del contraste de las imágenes y el perfil de la señal adopta la forma de una señal cosenoide. Los resultados se pueden observar en la Figura 5.10.



(a)



(b)

Figura 5.10. Resultado después de aplicar un filtro pasa bajas Butterworth a las imágenes registradas. (a) Imágenes del plano de referencia y (b) imágenes del rostro.

5.2.3 Submódulo del cálculo de la información tridimensional

El objetivo de este sub-módulo consiste en obtener el modelo 3D del rostro mediante el método de desplazamiento de fase de cuatro pasos, descrito en el capítulo 4.3.2. Ésta técnica realiza el cálculo de la fase de las franjas con la ecuación (5.2), cabe mencionar que la fase obtenida se encuentra envuelta.

$$\phi(x, y) = \tan^{-1} \frac{I_4(x, y) - I_2(x, y)}{I_1(x, y) - I_3(x, y)} \quad (5.2)$$

En este caso denominamos $\phi_A(x, y)$ la fase asociada al plano de referencia y $\phi_B(x, y)$ la fase de la superficie del rostro:

$$\phi_A(x, y) = \tan^{-1} \frac{R_4(x, y) - R(x, y)}{R(x, y) - R_3(x, y)} \quad (5.3)$$

$$\phi_B(x, y) = \tan^{-1} \frac{I_4(x, y) - I_2(x, y)}{I_1(x, y) - I_3(x, y)} \quad (5.4)$$

A continuación se presentan los resultados obtenidos al aplicar las ecuaciones (5.3) y (5.4) sobre las imágenes registradas. La figura 5.11 corresponde a la fase envuelta relaciona al plano de referencia y la figura 5.12 a la fase envuelta de la superficie del rostro.

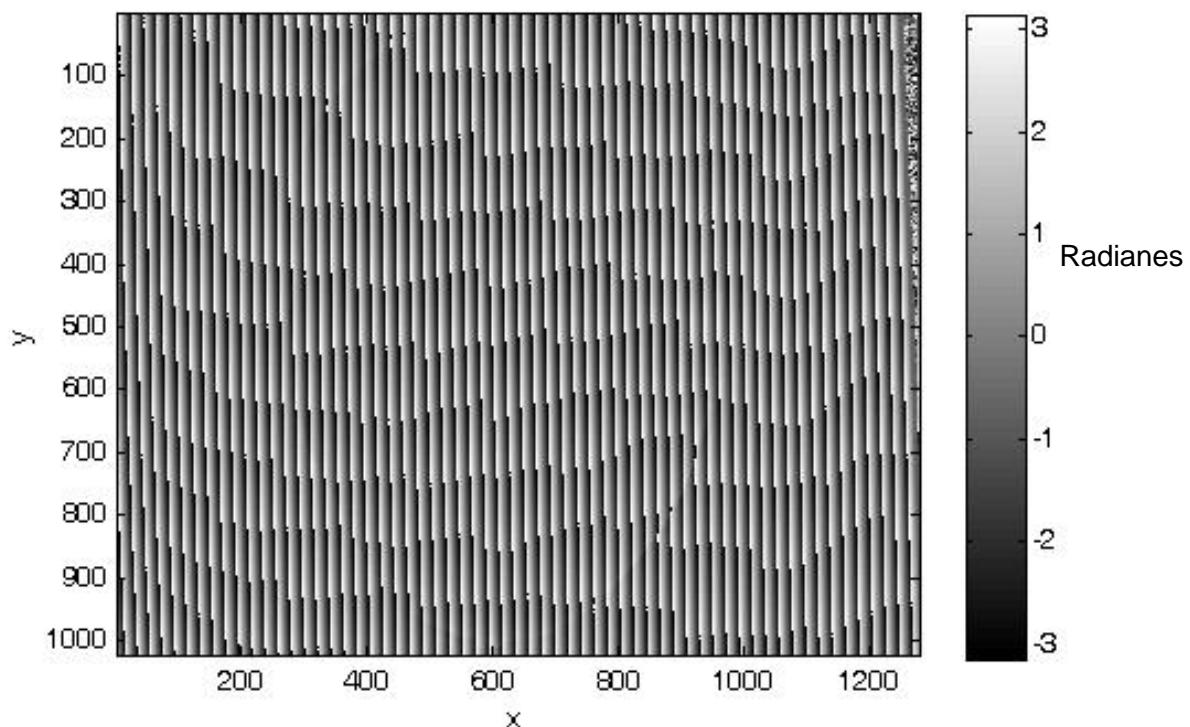


Figura 5.11. Fase envuelta asociada al plano de referencia, $\phi_A(x, y)$

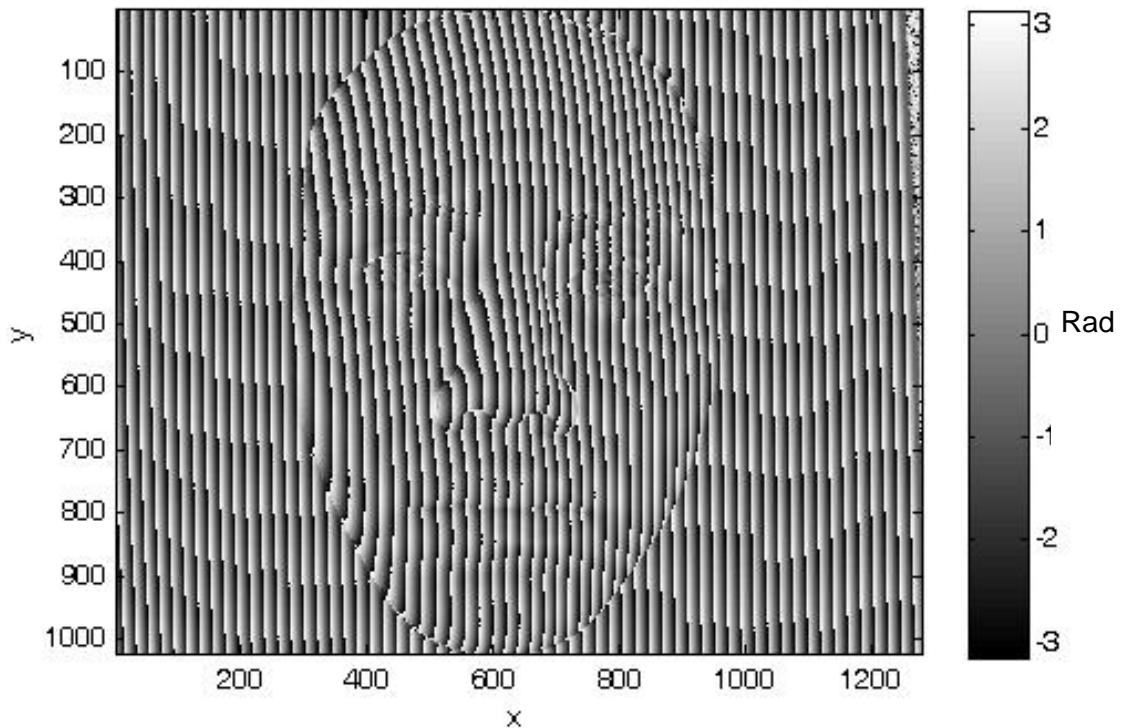


Figura 5.12. Fase envuelta asociada a la superficie del rostro, $\phi_B(x, y)$

Para obtener una fase continua se aplicó el algoritmo propuesto por M. Arevalio, et al [37], el cual se basa en el cálculo de mapas de calidad y a partir ella se diseña una ruta discreta de desenvolvimiento. Este algoritmo se expuso en el capítulo 4.4.2.

La Figura 5.13 corresponde a la fase desenvuelta asociada al plano de referencia y la Figura 5.14 a la fase desenvuelta asociada a la superficie del rostro.

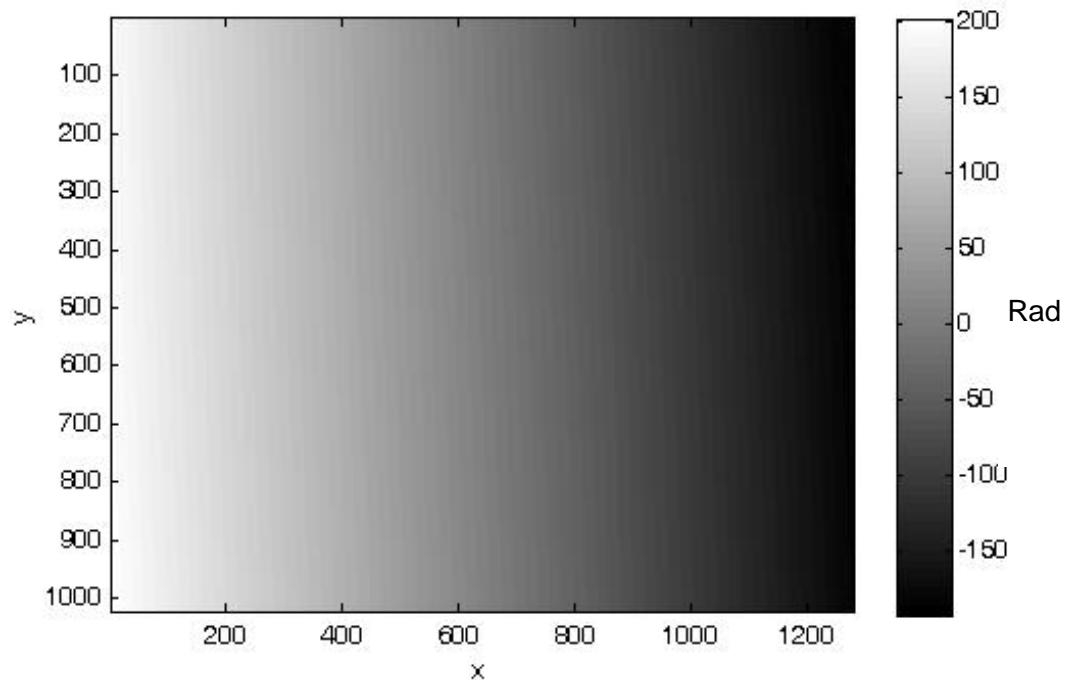


Figura 5.13. Fase desenvuelta asociada al plano de referencia.

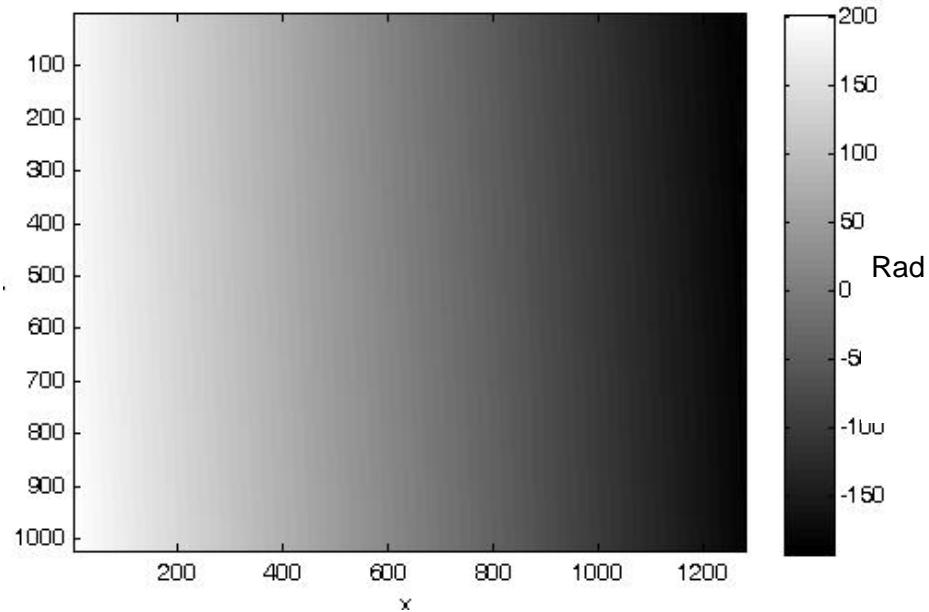


Figura 5.14. Fase desenvuelta asociada a la superficie del rostro.

Realizando el cálculo de la diferencia entre la fase del plano de referencia y la del rostro, $\phi(x, y) = (\phi_B - \phi_A)$, obtenemos el modelo tridimensional del rostro (modelo 3D), según se observan en la Figura 5.15.

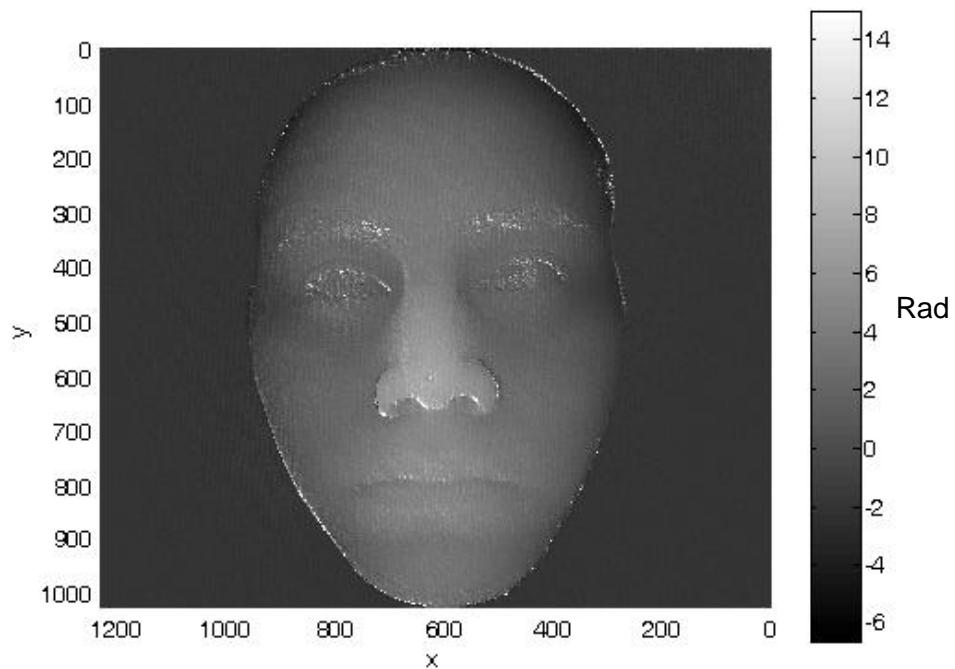


Figura 5.15 modelo tridimensional del rostro, vista del plano x, y

El modelo 3D de la figura 5.15 se encuentra en un espacio de tres dimensiones. Los ejes (x, y) corresponden a las coordenadas espaciales de los píxeles de las imágenes registradas, el eje z es la proyección de la fase ϕ del pixel situado en la posición (x, y) .

$$z(x, y) = \phi(x, y) \quad (5.5)$$

El modelo 3D del rostro recuperado presenta variaciones aleatorias que deben ser removido mediante un filtro pasa bajos en el espacio de frecuencias. Para realizar la remoción del ruido se aplica un filtro pasa bajas Butterworth en el dominio de frecuencias. El resultado final se presenta en las siguientes figuras:

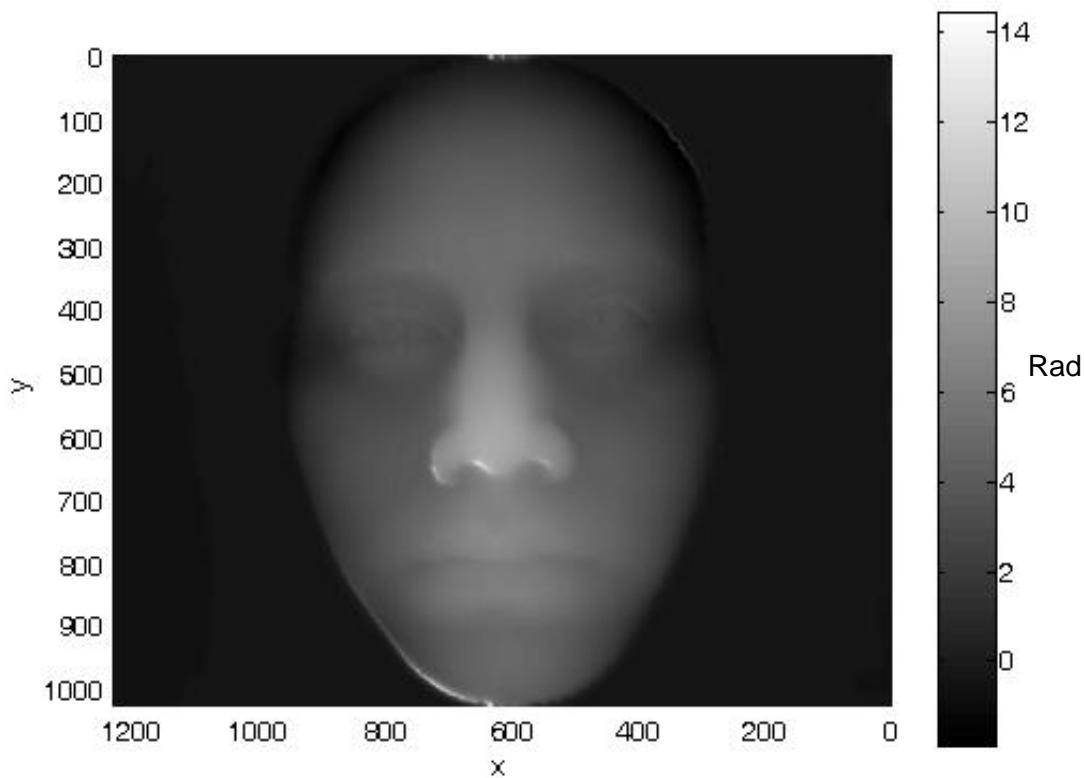


Figura 5.16. Modelo tridimensional del rostro. Vista del plano (x, y)

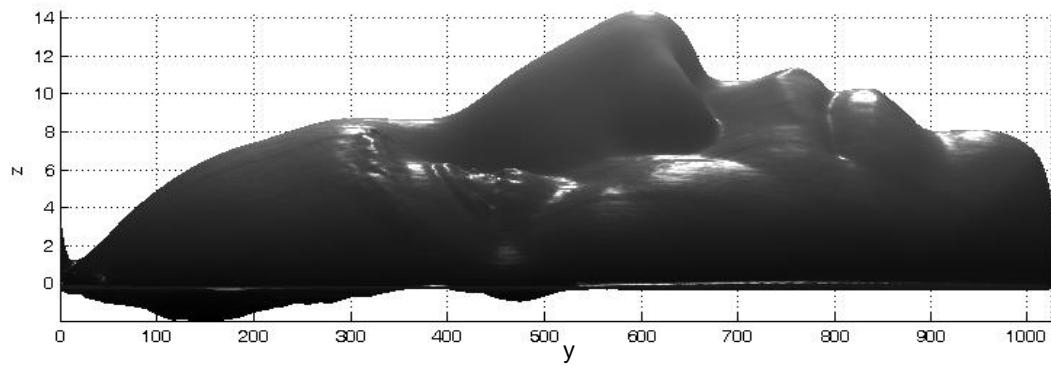


Figura 5.17. Modelo tridimensional del rostro. Vista del plano (y, z).

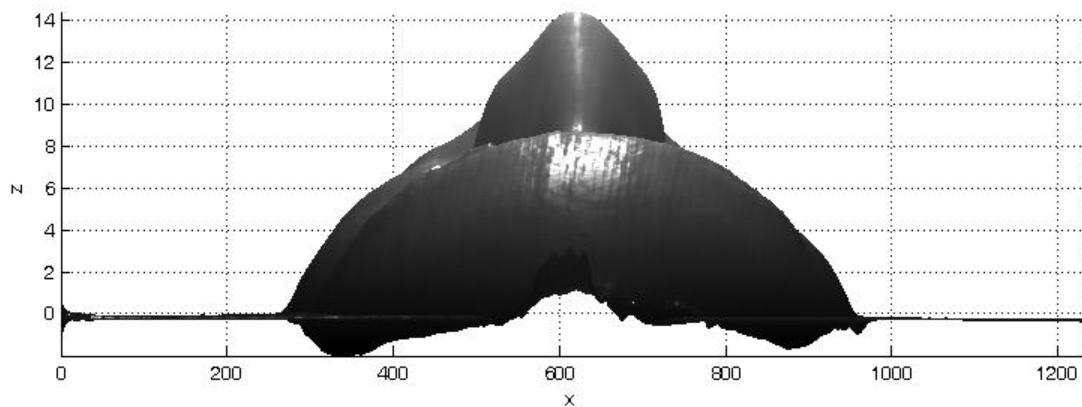


Figura 5.18. Modelo tridimensional del rostro. Vista del plano (x, z)

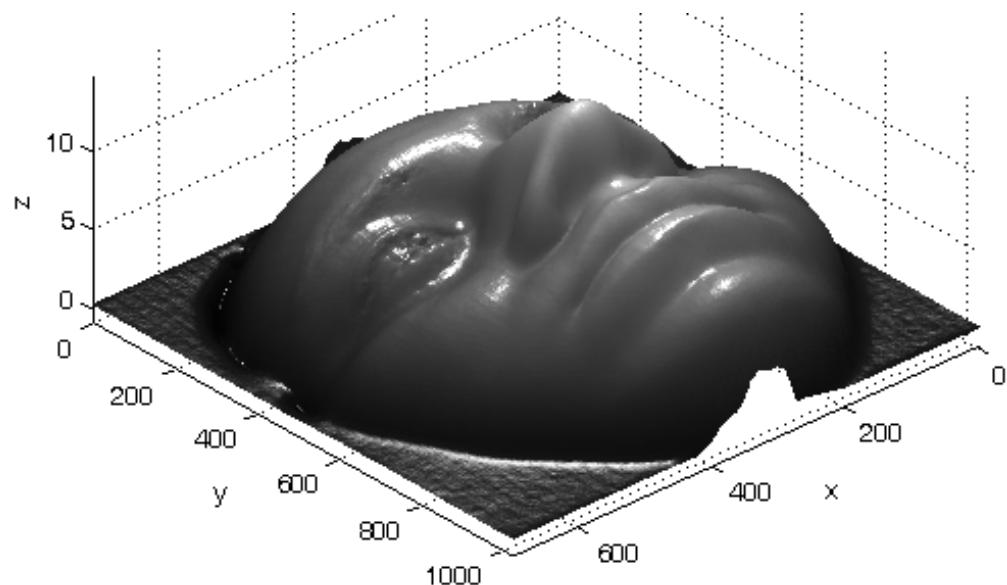


Figura 5.19. Modelo tridimensional del rostro, vista arbitraria.

5.3 Módulo de alineación y segmentación

El proceso de alineación consiste en corregir la orientación del rostro respecto a un rostro modelo, esta tarea es crucial para obtener buen rendimiento del módulo de clasificación. Después de normalizar la posición del rostro, se lleva a cabo el proceso de segmentación que tiene como objetivo determinar la región de la frente, los ojos y la nariz. A continuación se describen estos procesos.

5.3.1 Alineación

Una de las restricciones del Módulo de Clasificación y Reconocimiento, es que los modelos 3D de los rostros deben estar sustancialmente orientados en la misma posición, desafortunadamente éstos tienen ligeras rotaciones por lo que es necesario normalizarlos a una misma posición.

La alineación se lleva a cabo mediante la aplicación del algoritmo ICP (Iterative Closest Point). Éste minimiza la distancia entre los puntos de un rostro objeto y los de un rostro de referencia a partir del cálculo iterativo de la matriz de transformación geométrica de uno de estos [39, 40]. El algoritmo ICP minimiza la siguiente ecuación:

$$d(R, t) = \frac{1}{N} \|(Rp + t) - y\|^2 \quad (5.6)$$

donde R es la matriz de rotación, t es la matriz de translación, y es la nube de puntos de referencia y p es la nube de puntos objeto (nube de puntos a alinear).

El modelo 3D del rostro se encuentra conformado por cientos de puntos, realizar el cálculo de la matriz de transformación geométrica con tal cantidad de datos es un proceso lento e ineficiente. En este trabajo se realizó el cálculo de la matriz de transformación únicamente con el perfil vertical del rostro y los lagrimales de los ojos, en vez de utilizar todos los puntos. Posteriormente la matriz de transformación calculada se aplica a todos los puntos de la superficie del rostro, mediante este proceso se logra optimizar el cálculo y el tiempo de procesamiento.

Para definir el perfil vertical del rostro se requiere localizar dos puntos claves: la base de la nariz (región situada entre los ojos) y el domus de la nariz (región más prominente de la nariz), adicionalmente se localizan los lagrimales izquierdo y derecho. Los puntos clave son localizados mediante el análisis de la superficie del rostro y la clasificación de curvaturas de las medianas y gaussianas.

El algoritmo que se propone para realizar la alineación del rostro es el siguiente:

Algoritmo de alineación del rostro

1. Realizar el análisis de la curvatura de la superficie del rostro.
 2. Localizar los puntos clave del rostro mediante la clasificación de superficies medianas y gaussianas (Clasificador HK).
 3. Definir el perfil vertical del rostro.
 4. Calcular la matriz de trasformación geométrica mediante el algoritmo ICP, a partir del perfil del rostro.
 5. Aplicar la matriz de trasformación geométrica a todos los puntos de la superficie del rostro.
-

Tabla 5.1. Algoritmo de alineación del rostro.

A continuación se describe la implementación del algoritmo propuesto.

1) Análisis de curvatura de la superficie del rostro.

El objetivo del análisis de curvatura es determinar las curvaturas mediana (H) y gaussiana (K) de la superficie del rostro [41].

Considerando que los puntos del rostro pertenecen a una superficie S , cuyos elementos son doblemente diferenciables, $f = U \rightarrow \mathbb{R}$, definido en un conjunto abierto $U \subseteq \mathbb{R}^2$, donde \mathbb{R} es el conjunto de los números reales, tenemos:

$$S = \{(x, y, z) | (x, y) \in U; (f(x, y) = z | z) \in R\} \quad (5.7)$$

Para cada elemento $(x_0, y_0, f(x_0, y_0)) \in S$ se consideran dos tipos de curvaturas: Curvatura Mediana (H) y Curvatura Gaussiana (K):

$$H(x_0, y_0) = \frac{(1 + f_y^2)f_{xx} - 2f_x f_y f_{xy} + (1 + f_x^2)f_{yy}}{2(1 + f_x^2 + f_y^2)^{\frac{1}{3}}} \quad (5.8)$$

$$K = \frac{f_{xx}f_{yy} - f_{xy}^2}{(1 + f_x^2 + f_y^2)^2} \quad (5.9)$$

donde $f_x, f_y, f_{xy}, f_{xx}, f_{yy}$ son la primera y la segunda derivada parcial de $f(x, y)$.

Debido que los elementos de la superficie del rostro son discretos, las derivadas parciales son estimadas mediante la aproximación de un polinomio biquadrático:

$$\begin{aligned} g_{ij}(x, y) = & a_{ij} + b_{ij}(x - x_i) + c_{ij}(y - y_i) + d_{ij}(x - x_i)(y - y_i) + e_{ij}(x - x_i)^2 \\ & + f_{ij}(y - y_i)^2 \end{aligned} \quad (5.10)$$

donde $i = 1..M, j = 1..N$ (M es el número de filas y N es el número de columnas de la matriz que contiene los datos de S), los coeficientes $a_{ij}, b_{ij}, c_{ij}, d_{ij}, e_{ij}, f_{ij}$ son calculados mediante mínimos cuadrados en la vecindad de $f(x_i, y_j)$. Realizando las derivadas parciales de $g_{ij}(x, y)$ obtenemos una aproximación de las derivas parciales de $f(x, y)$:

$$\begin{aligned} f_x(x_i, y_j) &= b_{ij} \\ f_y(x_i, y_j) &= c_{ij} \\ f_{xy}(x_i, y_j) &= d_{ij} \\ f_{xx}(x_i, y_j) &= 2e_{ij} \\ f_{yy}(x_i, y_j) &= 2f_{ij} \end{aligned} \quad (5.11)$$

Sustituyendo los valores de la ecuación (5.11) en las ecuaciones (5.8) y (5.9), se obtienen la curvatura H y K respectivamente.

La Figura 5.20 ejemplifica el proceso de análisis de superficies de un rostro donde la figura (a) corresponde a la superficie a analizar, (b) es curvatura H y la (c)

corresponde a la curvatura K; se observa que los valores de ambos mapas son relativamente pequeños.

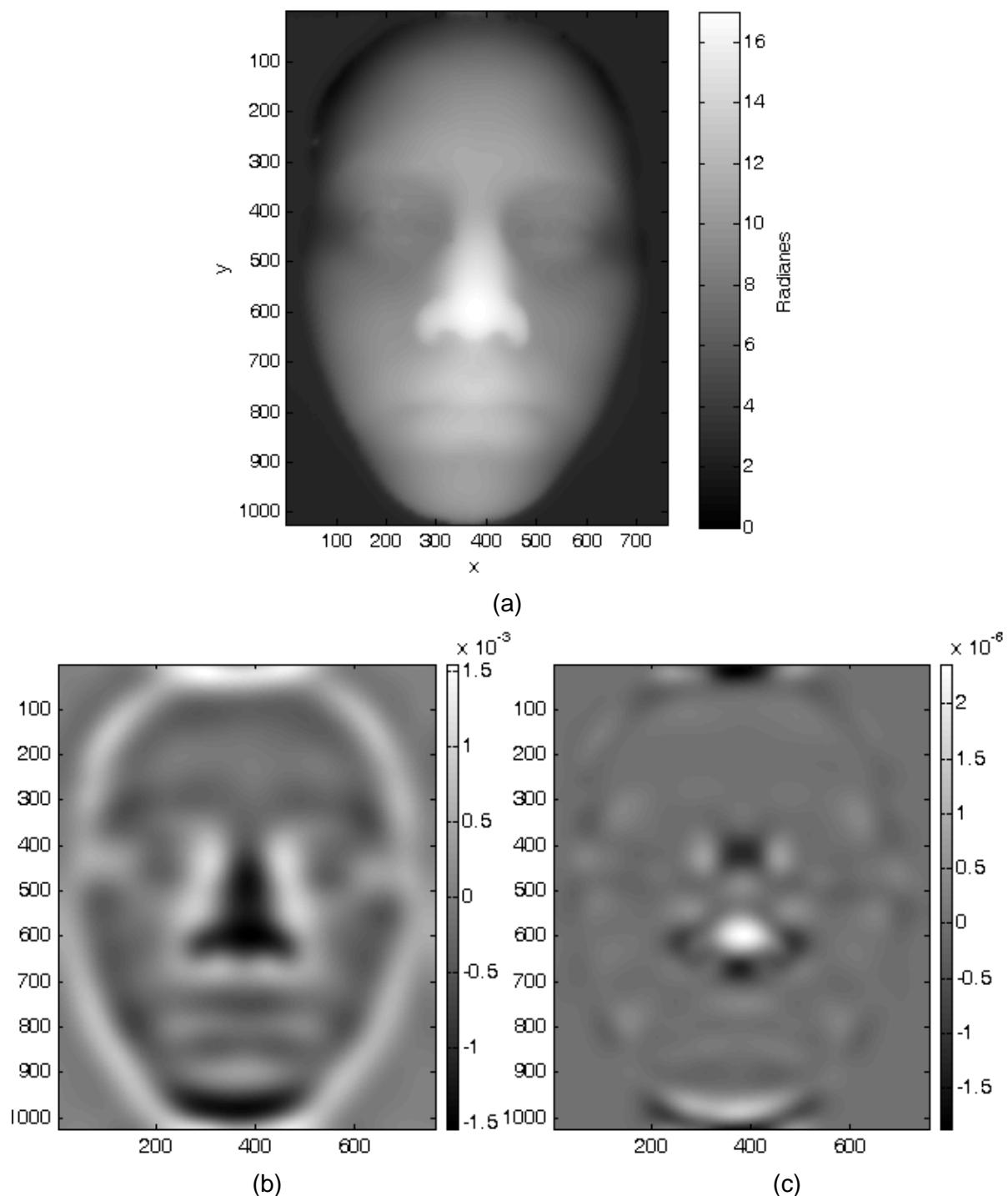


Figura 5.20. Mapa de la Curvatura H y K. (a) Superficie del rostro a analizar, (b) curvatura H y (c) curvatura K.

2) Localizar los puntos clave del rostro mediante la clasificación HK

El objetivo del clasificador HK, es localizar la posición de cuatro regiones claves del rostro: los lagrimales derecho e izquierdo, la base de la nariz y el domus de la nariz, descritos en la Figura 5.21.

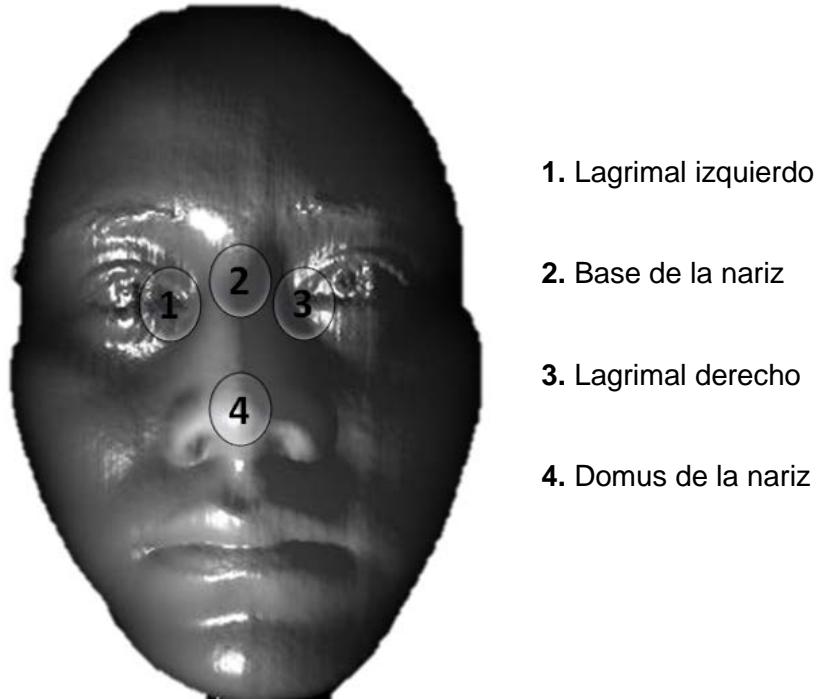


Figura 5.21. Puntos clave del rostro a localizar.

El clasificador HK se basa en el análisis de los signos de las curvaturas H y K . Según sea el caso, los elementos de la superficie S son clasificadas en uno de los siguientes tipos:

	$K < 0$	$K = 0$	$K > 0$
$H < 0$	Hiperbólica cóncava	Cilíndrica convexa	Elíptica convекса
$H = 0$	Hiperbólica simétrica	Plano	Impossible
$H > 0$	Hiperbólica convexa	Cilíndrica cóncava	Elíptica cóncava

Tabla 5.2. Clasificador HK.

Además el clasificador HK define descriptores precisos de los elementos del rostro que son de nuestro interés tal como se muestra en la Tabla 5.3.

Región de interés	Tipo de superficie
Lagrimal derecho	Elíptica cóncava
Base de la nariz	Hiperbólica simétrica
Lagrimal izquierdo	Elíptica cóncava
Domus de la nariz	Elíptica convexa

Tabla 5.3 Descriptores de las regiones de interés.

Para llevar a cabo la clasificación HK, se establece un umbral $\pm Th$ y $\pm Tk$ los cuales permiten filtrar únicamente las curvaturas con valor absoluto grande. Los umbrales son calculados mediante las ecuaciones (5.12) y (5.13):

$$Th = \pm \frac{\max(H)}{p} \quad (5.12)$$

$$Tk = \pm \frac{\max(K)}{p} \quad (5.13)$$

donde $\max(H)$ y $\max(K)$ es valor máximo de las curvaturas H y K, respectivamente, mientras que p define la proporción de los valores máximos a considerar.

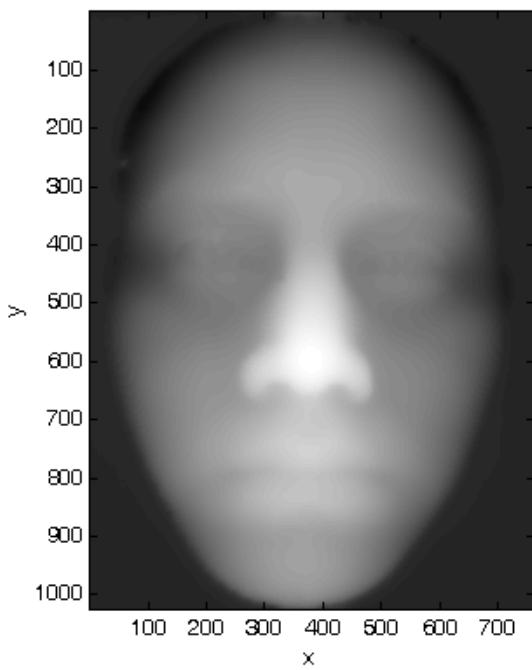
A partir de las condiciones de la tabla 5.3 y los umbrales de las ecuaciones 5.12 y 5.13, se generan mapas binarios correspondientes para cada tipo de superficies. Las imágenes binarias resultantes se abrevian como:

$$ECC = \text{Superficie Elíptica Cónica} \quad (5.14)$$

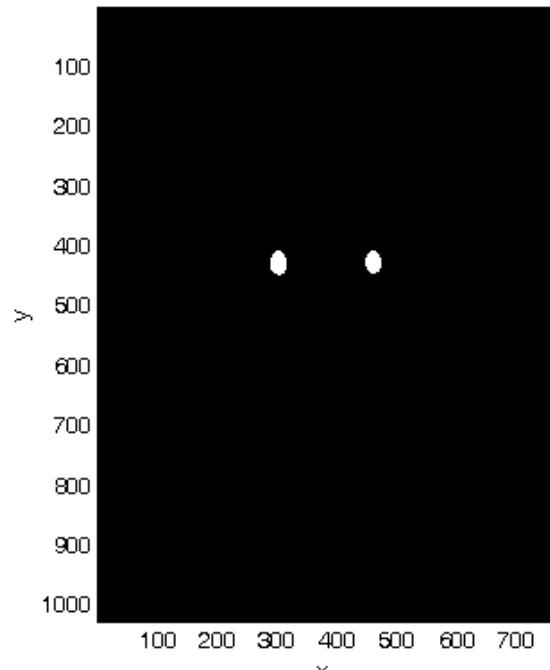
$$HIS = \text{Superficie Hiperbólica Simétrica} \quad (5.15)$$

$$ECV = \text{Superficie Elíptica Convexa} \quad (5.16)$$

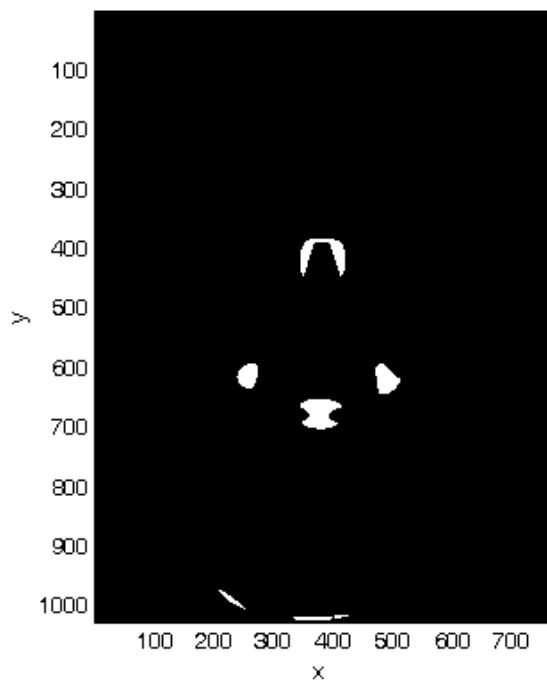
La Figura 5.22 presenta los el resultado de la clasificación HK de la superficie de un rostro.



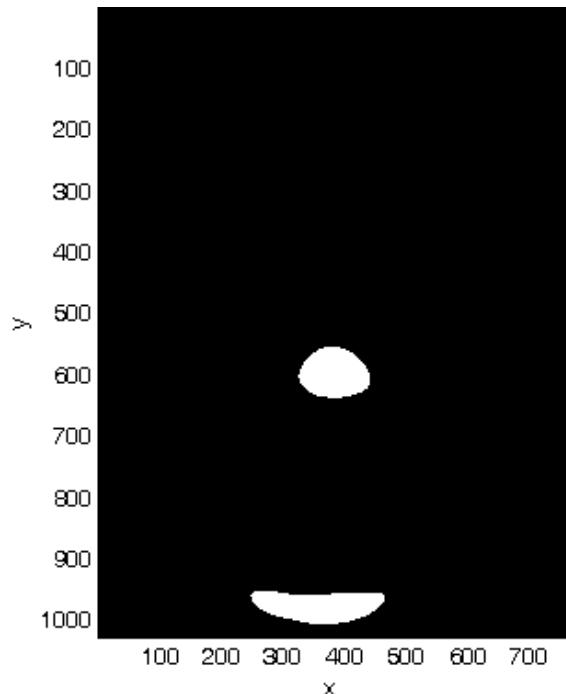
(a)



(b)



(c)



(d)

Figura 5.22. Mapa binario de los descriptores del rostro. (a) Rostro a analizar, (b) superficie elíptica cóncava (lagrimales), (c) superficie hiperbólica simétrica (base de la nariz) y (d) superficie elíptica convexa (domus de la nariz).

Es preciso describir las regiones de interés respecto a las coordenadas de un solo punto. Localizamos las coordenadas en este orden: primero el domus de la nariz, seguido de la base de la nariz y finalmente las coordenadas de los lagrimales.

La coordenada del domus de la nariz corresponde al elemento de mayor magnitud del producto punto de la superficie ECV y el modelo 3D bajo análisis. Describimos su posición espacial (xd, yd, zd) mediante:

$$domus_nariz(xd, yd, zd) = \max [(ECV) (S)] \quad (5.17)$$

Siguiendo el orden de búsqueda, ahora se localizan las coordenadas de la base de la nariz. La búsqueda se limita en un área rectangular (R1) sobre el mapa binario *HIS*. El rectángulo R1 se define a partir de las coordenadas del domus de la nariz y se sitúa en la parte superior de éste, tal como se muestra en la Figura 5.23.

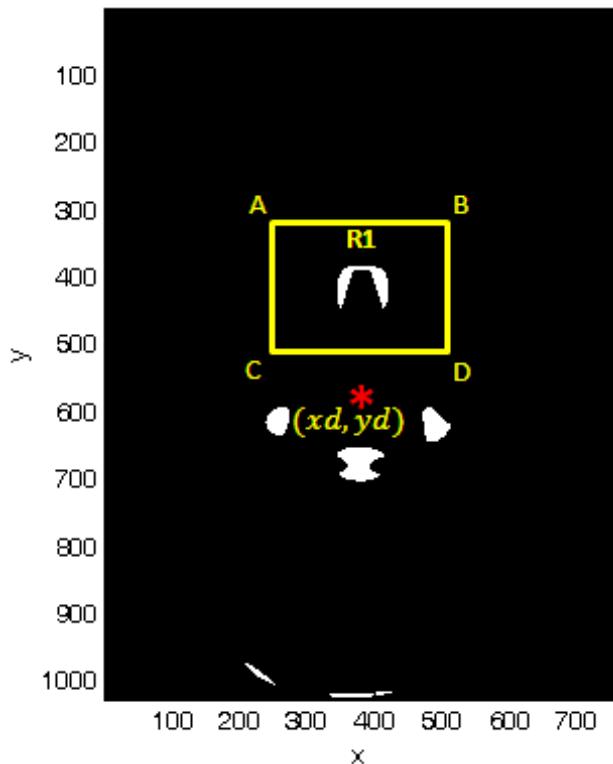


Figura 5.23. Área de búsqueda de la base de la nariz sobre la superficie *HIS*.

La región rectangular se determina con la siguiente ecuación:

$$R1 = HIS(A:D) \quad (5.18)$$

Donde A es la coordenada superior derecha y D es la coordenada inferior izquierda de R1. Las equinas del rectángulo R1, se determinan sumando y/o restando unidades preestablecidas a la coordenada del domus, tal como se describe a continuación:

- Esquina superior derecha: $A(xd - m, yd - n)$
- Esquina superior izquierda: $B(xd + m, yd - n)$
- Esquina inferior derecha: $C(xd - m, yd - o)$
- Esquina inferior izquierda: $D(xd + m, yd - o)$

Las coordenadas de base de la nariz se determina calculando el centroide de R1.

$$\text{base_nariz}(xb, yb) = \text{centroide}(R1) \quad (5.19)$$

Finalmente, las coordenadas de los lagrimales se localizan siguiendo un proceso similar al anterior, en este caso se realiza la búsqueda en el mapa binario *ECC*. Se definen dos áreas de búsqueda R2 y R3, para los lagrimales derecho e izquierdo respectivamente. Los rectángulos se definen a partir de las coordenadas de la base de la nariz, tal como se puede ver en la Figura 5.24.

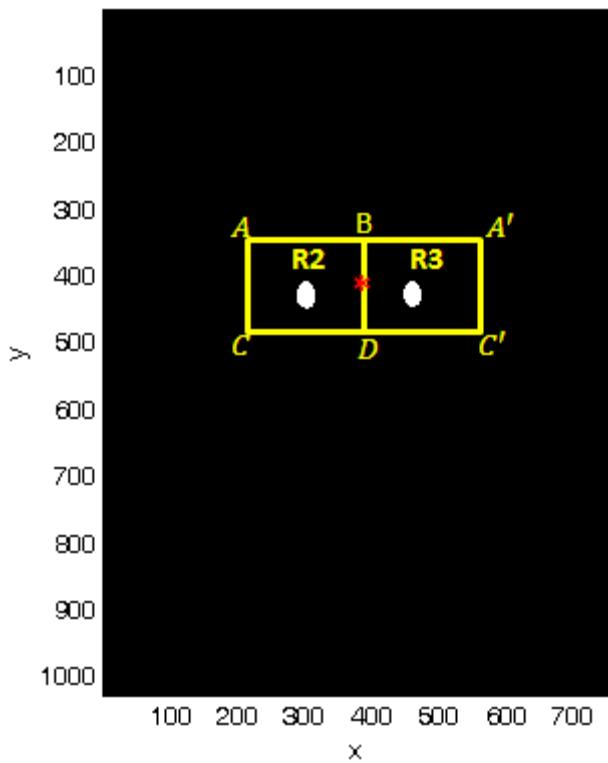


Figura 5.24. Área de búsqueda de los lagrimales en la superficie *ECC*.

Las regiones R2 y R3 se definen mediante las ecuaciones (5.20) y (5.21), respectivamente:

$$R2 = ECV(A, D) \quad (5.20)$$

$$R3 = ECV(B, C') \quad (5.21)$$

donde *A* es la esquina superior derecha y *D* la esquina inferior izquierda del rectángulo R2; *B* es la esquina superior derecha y *C'* la esquina inferior izquierda del rectángulo R3. Los rectángulos R2 y R3 se definen de la siguiente manera:

- Rectángulo *R2*
 - Esquina superior derecha: $A = (xb - m, yb - n)$
 - Esquina superior izquierda derecha: $B = (xb, yb - n)$
 - Esquina inferior derecha: $C = (xb - m, yb + n)$
 - Esquina inferior izquierda: $D = (xb, yb + n)$

- Rectángulo $R3$
- Esquina superior derecha: $B = (xb, yb - n)$
- Esquina superior izquierda: $A' = (xb + m, yb - n)$
- Esquina inferior derecha: $D = (xb, yb + n)$
- Esquina inferior izquierda: $C' = (xb + m, yb + n)$

Realizando el cálculo del centroide de los rectángulos de $R2$ y $R3$ obtenemos las coordenadas de los lagrimales:

$$Lagrimon_derecho(xoi, oyi) = \text{centroide}(R2) \quad (5.22)$$

$$lagrimon_izquierdo(xod, yod) = \text{centroide}(R3) \quad (5.23)$$

Los resultados se presentan en la Figura 5.25, en ella han proyectado las coordenadas de los puntos clave sobre la superficie del rostro.

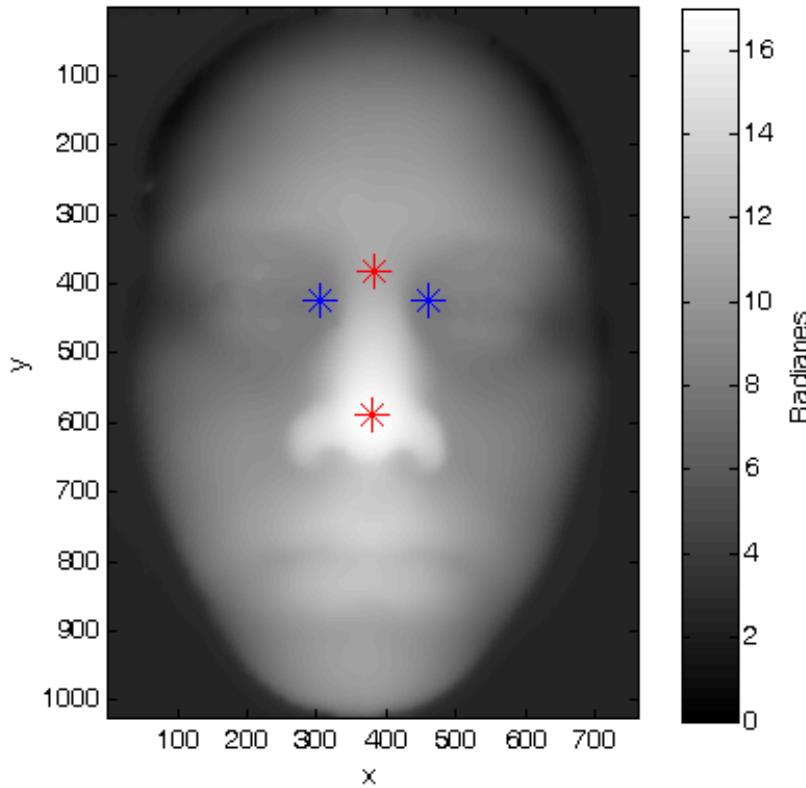


Figura 5.25. Posición de los puntos claves proyectados sobre el modelo 3D del rostro.

3) Definir el perfil vertical del rostro.

En este proceso, el objetivo es definir el perfil vertical del rostro que pasa a través del dorso de la nariz y divide el rostro en dos regiones simétricas: izquierda y derecha.

El siguiente algoritmo describe el proceso para definir el perfil del rostro:

Algoritmo para definir el perfil del rostro

1. Proyectar la base y el domus de la nariz en el plano (x, y) .
 2. Calcular la línea recta que pasa en los puntos proyectados en el paso (a). El cálculo de la línea recta se realiza respecto al eje y .
 3. Definir el perfil del rostro mediante la proyección de la línea recta sobre la superficie del rostro.
-

Tabla 5.4. Algoritmo de definición del perfil del rostro.

La proyección de los puntos de interés sobre el plano (x, y) consiste en considerar únicamente las coordenadas en el eje x y y de los puntos de interés.

El cálculo de la línea recta que pasa por los puntos antes mencionados, se realiza mediante la ecuación (5.24):

$$x = \frac{(y_2 - y_1)}{(x_2 - x_1)}(y - y_1) + x_1 \quad (5.24)$$

donde las variables (x_1, y_1) corresponden a las coordenadas de la base de la nariz, (x_2, y_2) son las coordenadas del domus de la nariz, la variable y es la coordenada del perfil que recorren el rostro en dicho sentido.

La proyección de la línea recta sobre la superficie lo define la ecuación (5.25).

$$\text{Perfil}(x, y, z) = \text{linea_recta}(\text{base_nariz}, \text{domus_nariz}) \quad (5.25)$$

Los resultados de la definición del perfil vertical del rostro se presentan en la Figura 5.26. La figura (a) corresponde a la vista del plano (y, z) del perfil y (b) es la superposición del perfil sobre la superficie del rostro.

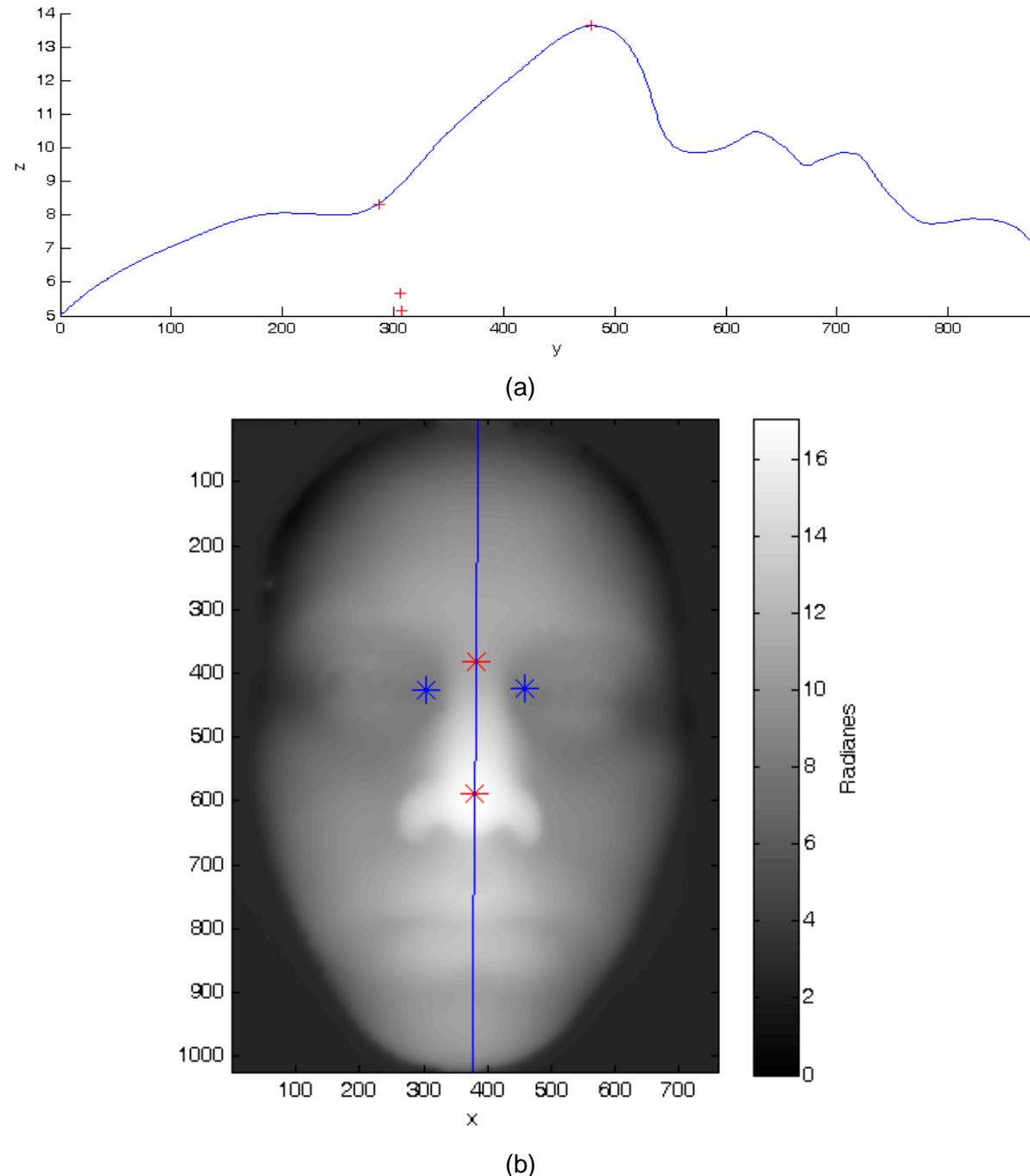


Figura 5.26. (a) Perfil vertical del rostro y (b) perfil sobrepuerto sobre el modelo 3D del rostro.

4) Cálculo de la matriz de trasformación a partir del perfil vertical del rostro.

El cálculo de la matriz de trasformación se realiza mediante el algoritmo ICP, el cual minimiza la distancia entre el perfil vertical de un rostro objeto y la de un rostro modelo. Posteriormente la matriz de trasformación calculada se aplica a todos los puntos que conforman la superficie del rostro objeto.

Cabe mencionar que en este trabajo se utilizó un script que implementa el algoritmo ICP en el entorno de desarrollo Matlab®. El script fue escrito por W. Jakob [42] y está disponible en la página web de Matlab® Central. El script tiene la siguiente estructura:

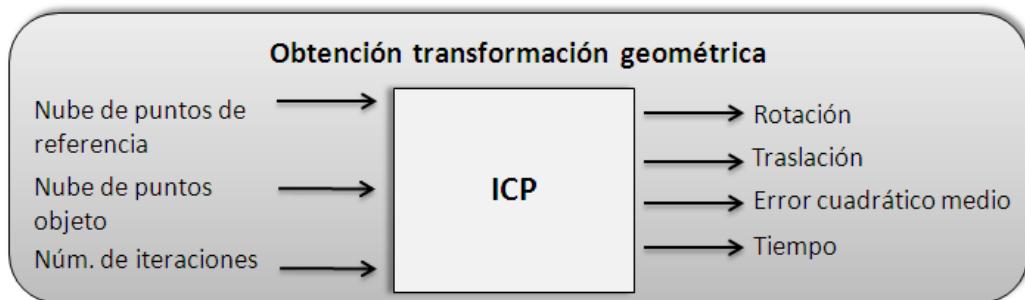


Figura 5.27. Estructura del algoritmo ICP.

El cálculo de los parámetros de la matriz de trasformación geométrica se realiza con la siguiente instrucción:

$$[\mathbf{R}, \mathbf{T}, \mathbf{ECM}, \mathbf{t}] = ICP(\text{perfil_referencia}, \text{perfil_objeto}, \text{nun_iteración}) \quad (5.26)$$

donde *perfil_referencia* son las coordenadas del perfil de referencia, *perfil_objeto* son a las coordenadas del perfil a alinear y *nun_iteración* el el número máximo de iteraciones a realizar. Los parámetros que retorna el script son la matriz de rotación (\mathbf{R}), vector de traslación (\mathbf{T}), vector del error cuadrático medio \mathbf{ECM} y vector del tiempo de ejecución (\mathbf{t}).

Las imágenes de la Figura 5.28 corresponde a las posiciones iniciales de los perfiles de referencia y objeto en color rojo y azul respectivamente, los puntos aislados que se observan pertenecen a las coordenadas de los lagrimales. En la imagen se observan dos perspectivas diferentes de los perfiles: (a) vista del plano (y, z), (b) vista del plano (x, z).

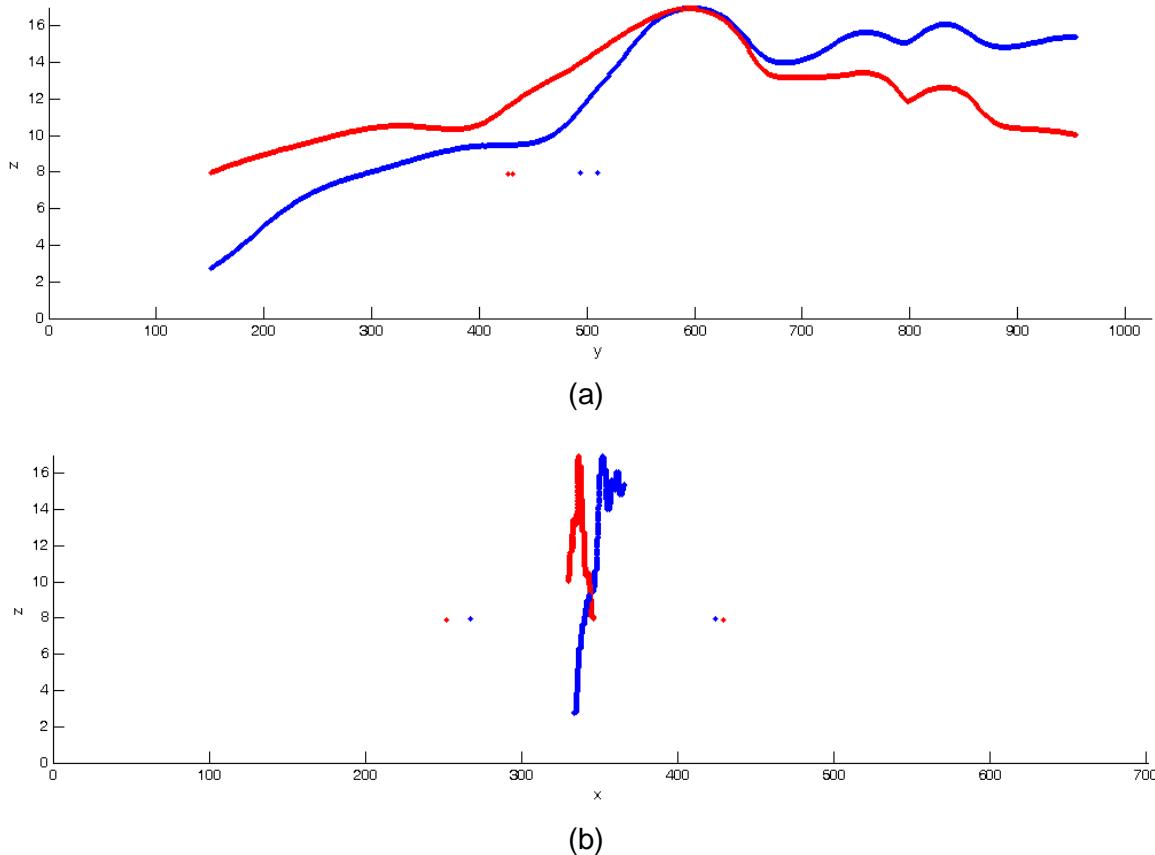


Figura 5.28. Perfil del rostro de referencia y objeto en color rojo y azul respectivamente, antes del registro. (a) Vista del plano (y, z), (b) vista del plano (x, z)

Antes de aplicar el algoritmo ICP, se realizó el cálculo del error cuadrático medio (ECM) entre los perfiles, para el caso de este ejemplo, el ECM inicial fue de 19.1169. Después de aplicar el algoritmo ICP con 8 iteraciones se obtuvieron los resultados descritos en la tabla 5.3:

Iteración	ECM	Tiempo (seg)
1	4.0008	0.0331
2	3.7819	0.0681
3	3.7781	0.0983
4	3.7703	0.1233
5	3.7452	0.1475
6	3.7283	0.1793
7	3.7328	0.2018
8	3.7328	0.2289

Tabla 5.5. ECM y tiempo de ejecución de las iteraciones del algoritmo ICP.

En la tabla 5.5 se observa que el EMC converge en la sexta iteración, en las subsecuentes iteraciones permanece constante. El grado de acoplamiento entre ambos perfiles es del 80.47 %, debido a que no puede haber dos rostros idénticos (propiedad de unicidad de los rasgos biométricos). A pesar de ello el registro realizado proporciona una buena aproximación para alinear los perfiles

Los resultados gráficos del registro se pueden apreciar en la Figura 5.29. En color rojo el perfil de referencia y en color azul el perfil objeto. En la imagen se observan dos perspectivas diferentes: a) vista del plano (y, z) y b) vista del plano(x, z).

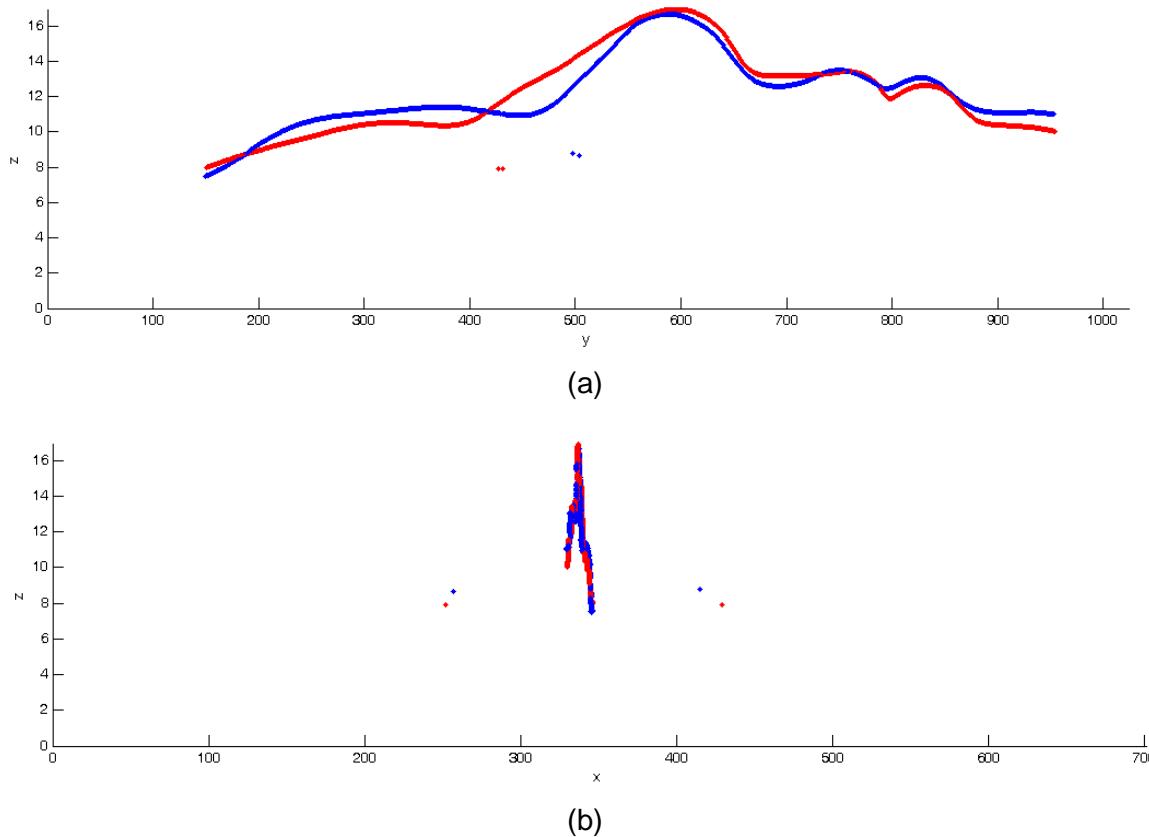


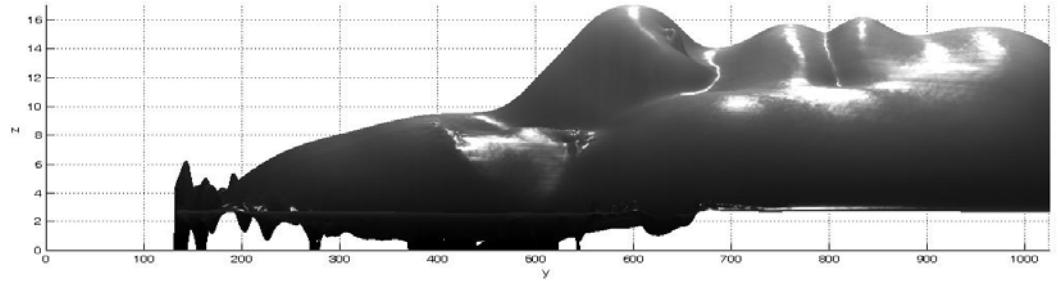
Figura 5.29. Perfil del rostro de referencia y objeto en color rojo y azul respectivamente, después del registro. (a) Vista del plano (y, z), (b) vista del plano (x, z).

5) Aplicar la matriz de trasformación geométrica a la superficie del rostro 3D.

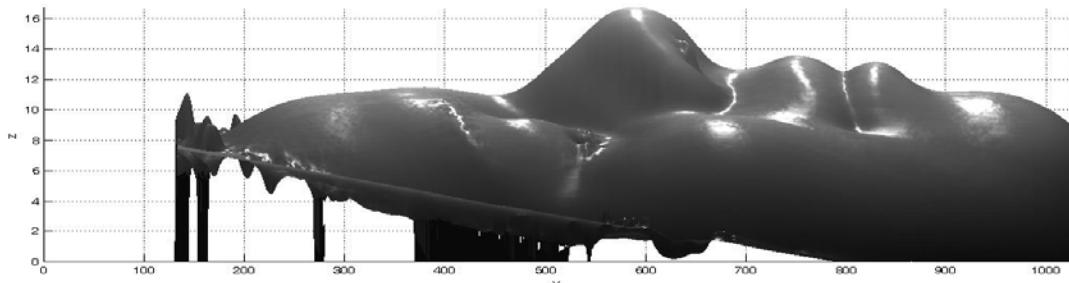
La matriz de transformación que retorna el algoritmo ICP se aplica en el registro de todos los puntos del rostro objeto mediante la siguiente ecuación:

$$rostro_registro(x, y, z) = (rostro_objeto(x, y, z))(\mathbf{R})(\mathbf{T}) \quad (5.28)$$

Los resultados gráficos de la alineación de la superficie se puede observar en la Figura 5.30, donde (a) es el modelo 3D antes de la alineación y (b) es el modelo 3D después de la alineación.



(a)



(b)

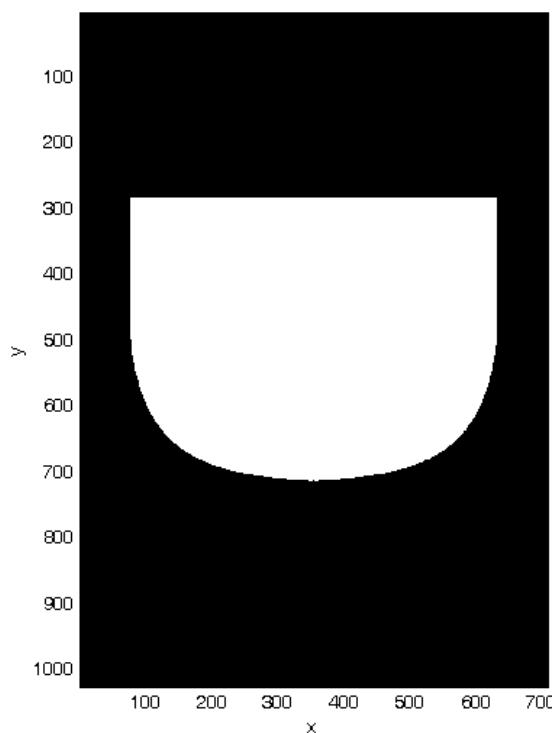
Figura 5.30. Superficie del rostro vista desde el plano (y, z): (a) antes del registro, (b) después del registro.

5.3.2 Segmentación

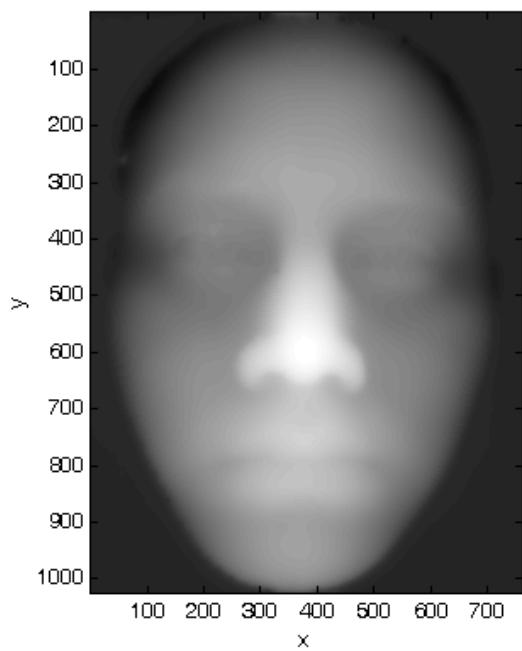
Para realizar la segmentación de la frente, los ojos y el dorso de la nariz se hace uso de una máscara binaria donde la región válida del rostro se le asigna el valor de uno en otro caso se le asigna un cero, Figura 5.31.

$$z_{seg}(x, y) = [z_{ob}(x, y). m_2(x, y)] \quad (5.29)$$

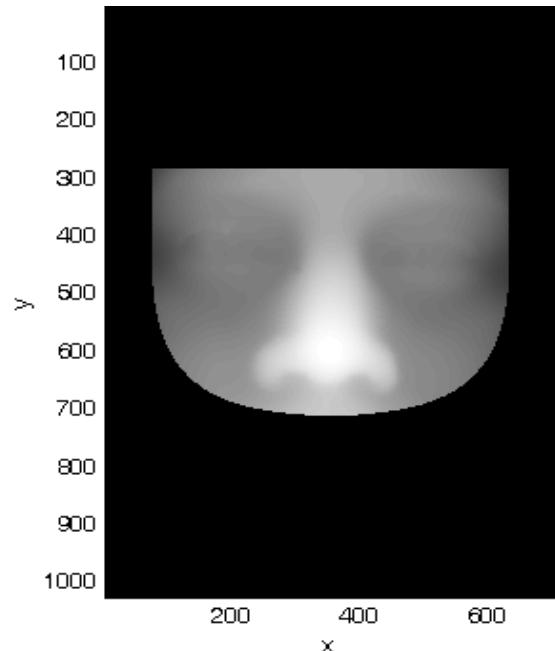
Donde $z_{seg}(x, y)$ es rostro objeto segmentado, $z_{ob}(x, y)$ es el rostro objeto y m_2 es la máscara binaria.



(a)



(b)



(c)

Figura 5.31. Proceso de segmentación de la región del rostro. (a) Máscara binaria, (b) modelo 3D antes de realizar la segmentación, (c) modelo 3D después de realizar la segmentación.

5.4 Módulo de Extracción de Características

Los descriptores del rostro se obtienen mediante la implementación de la técnica de Análisis de Componentes Principales (PCA) o eigenrostros [3, 7, 41, 5], esta técnica se basa en el análisis de la variabilidad de la información de profundidad de los rostros y en la definición de un nuevo subespacio de menor dimensión, en el cual los datos originales son proyectados. Este tema fue desarrollado en la sección 6.5. La figura 5.32 describe pseudocódigo del PCA.

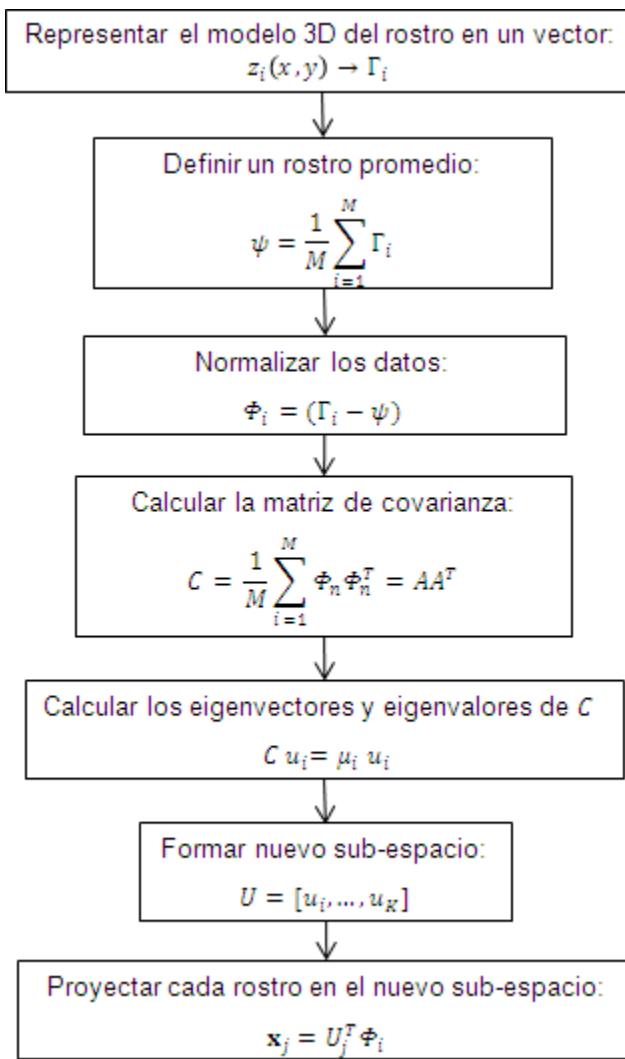


Figura 5.32. Pseudocódigo del Análisis de Componentes Principales.

A continuación se presentan los resultados significativos del proceso del PCA.

1) Representar el modelo 3D en un vector.

Para realizar el PCA sobre un conjunto de rostros de entrenamiento, se necesita representar los modelos 3D en un vector fila Γ_i de N elementos ($N = (m)(n)$, donde n es el número de filas y m el número de columnas de la matriz de datos original). Los vectores Γ_i son concatenados en la matriz L , la cual contiene $N \times M$ elementos, M es el número de modelos 3D de la base de datos. Este proceso se describe gráficamente en la Figura 5.33.

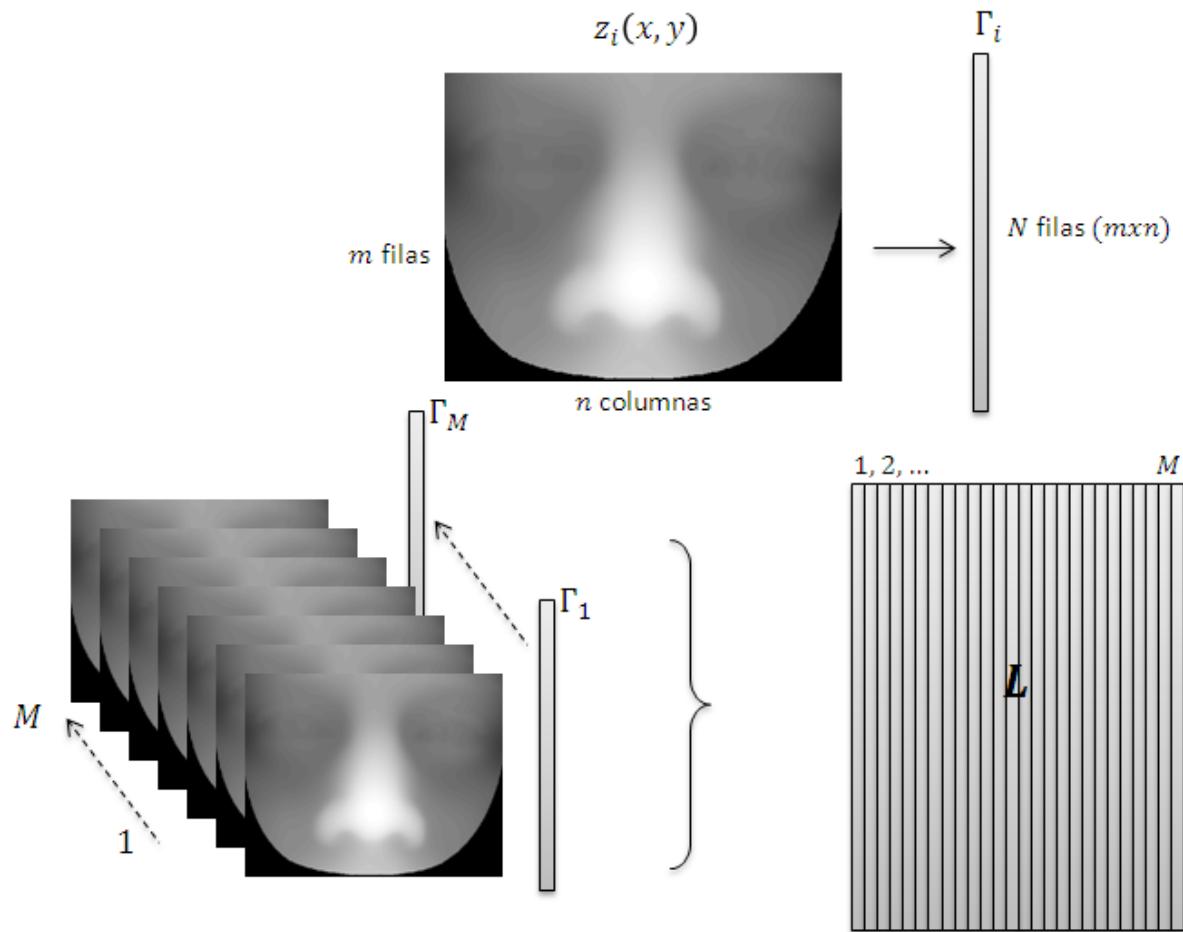


Figura 5.33. Proceso de conversión de matriz de datos en un vector fila.

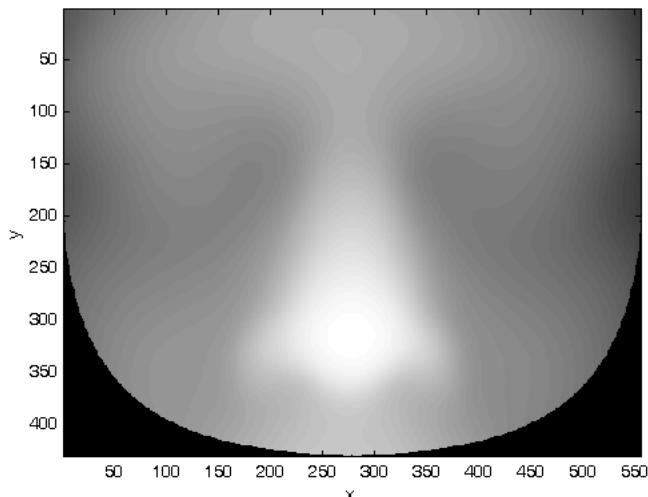
2) Normalizar los modelos 3D.

A partir del Γ_i se define un rostro promedio y se normalizan los rostros mediante las ecuaciones (5.30) y (5.31) respectivamente.

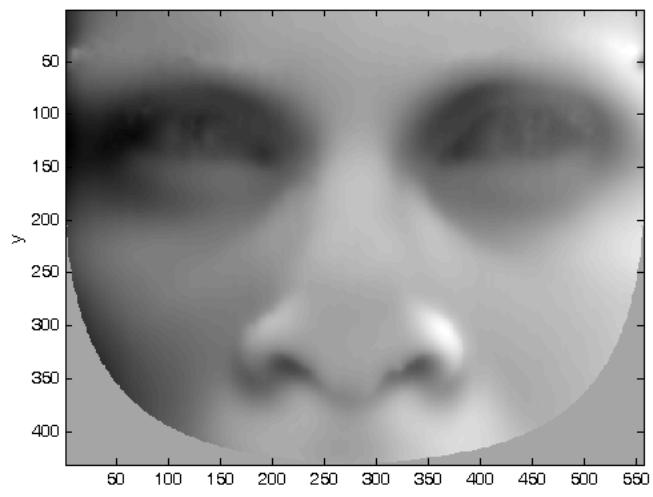
$$\psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i \quad (5.30)$$

$$\Phi_i = (\Gamma_i - \psi) \quad (5.31)$$

En la figura 5.34 se presenta un rostro del conjunto de entrenamiento y la diferencia entre el rostro promedio y uno de prueba.



(a)



(b)

Figura 5.34. (a) Rostro promedio del conjunto de entrenamiento y (b) diferencia entre el rostro promedio y un rostro de prueba.

3) Cálculo de la matriz de covarianza.

La matriz de covarianza es definida por la ecuación (5.32):

$$C = \frac{1}{M} \sum_{i=1}^M \Phi_n \Phi_n^T = AA^T \quad (5.32)$$

donde $A = [\Phi_1 \Phi_2 \dots \Phi_i]$

El producto de AA^T define una matriz de $N \times N$ elementos, sin embargo para casos prácticos resulta poco eficiente procesar tal cantidad de información, por ello el cálculo de la matriz de covarianza se realiza con la ecuación (5.33).

$$C' = A^T A \quad (5.33)$$

donde C' es una matriz con $M \times M$ elementos, ($M \times M$) \ll ($N \times N$).

4) Calcular los componentes principales

Los componentes principales son los eigenvectores de la matriz de covarianza C . Cada eigenvector lleva asociado un eigenvalor y se definen con la ecuación (5.34).

$$C u_i = \mu_i u_i \quad (5.34)$$

donde u_i es el eigenvector y μ_i es el eigenvalor. El eigenvector u_i se define con la ecuación (5.36)

$$u_i = A v_i \quad (5.36)$$

Los componentes principales definen una matriz de $M \times N$ elementos, además están ordenados de menor a mayor variabilidad.

5) Definir un nuevo subespacio

El nuevo sub-espacio \mathbf{W} lo conforman los K eigenvectores que llevan asociado los eigenvalores con valor numérico grande.

$$\mathbf{W} = [u_1, u_2, u_3, \dots, u_K] \quad (5.37)$$

donde $u_1, u_2, u_3, \dots, u_K$ son los eigenvectores seleccionados.

Se puede expresar los eigenvalores en términos de energía y representarlos en una gráfica tal como se observa en la Figura 5.35, en ella la distribución de la energía encuentra ordenada de manera ascendente, de manera que las energías más altas se encuentran en los últimos eigenvalores.

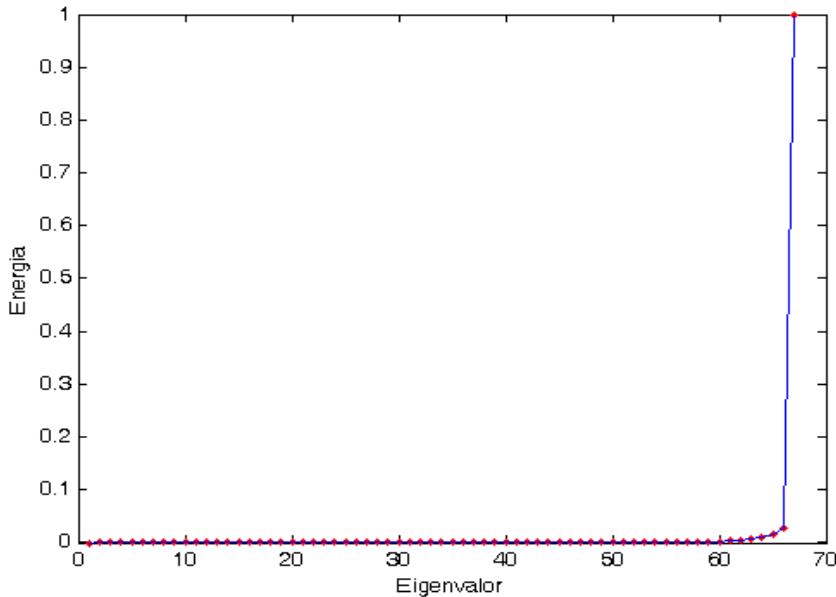


Figura 5.35. Energía de los eigenvalores

La totalidad de los eigenvalores representan el 100% de la energía, que es igual al 100% la variabilidad de los rostros 3D. De manera que para elegir los K eigenvalores se considera solo el 99.8% de la energía, el cual se encuentra distribuido en los últimos 19 eigenvalores.

6) Proyectar los rostros en el nuevo sub-espacio

El procedimiento final consiste en proyectar los rostros normalizados Φ_i en el sub-espacio \mathbf{W} mediante la ecuación (5.38).

$$\mathbf{X} = \mathbf{W}^T \mathbf{A} \quad (5.38)$$

donde cada vector $\mathbf{x}_i \in \mathbf{X}$ es el descriptor del i -ésimo modelo 3D de la base de datos de rostros 3D.

Con este procedimiento se logra reducir la cantidad de datos originales del rostro 3D. En nuestro caso, inicialmente el rostro $z_i(x, y)$ estaba representado en una matriz con $431 * 557$ elementos, en el nuevo sub-espacio el mismo rostro se representa en un vector con 19 elementos, es decir, se han conservado únicamente los datos que retienen las máximas variaciones

5.5 Módulo de clasificación/reconocimiento

El objetivo del Módulo de Clasificación es definir un conjunto de funciones de decisión que permita clasificar el patrón \mathbf{x} a una de las clases válidas.

El proceso de clasificación y reconocimiento se realiza mediante la técnica de Análisis Discriminante Lineal, esta técnica determina un conjunto de funciones discriminantes y un conjuntos de funciones de decisión. En el proceso de reconocimiento, los descriptores a clasificar son evaluados en cada una de las funciones de decisión. Este tema fue revisado en la sección 3.6.

Para fines prácticos, el Módulo de Clasificación y Reconocimiento hace uso de la función *Classify* del entorno de desarrollo Matlab, ésta función implementa la técnica de Análisis Discriminante para asignar un descriptor a una o más clases válidas [43]. La función *classify* retorna los coeficientes la función de decisión de tipo lineal descrita por la ecuación (5.39).

$$d_{ij}(\mathbf{x}) = \mathbf{x}^T (\mathbf{m}_j - \mathbf{m}_i) + \frac{1}{2} (\mathbf{m}_i - \mathbf{m}_j)^T (\mathbf{m}_i + \mathbf{m}_j) \quad (5.39)$$

expresado en forma vectorial se tiene:

$$d_{ij}(\mathbf{x}) = \mathbf{x}^T L + K \quad (5.40)$$

donde $d_{ij}(\mathbf{x})$ es la función de decisión entre las clases ω_i y ω_j , $i \neq j$; \mathbf{m}_i y \mathbf{m}_j son los prototipos de las clases ω_i y ω_j respectivamente. El descriptor \mathbf{x} es evaluado en cada una de las funciones $d_{ij}(\mathbf{x})$ y es asignado a la clase ω_i si $d_{ij}(\mathbf{x}) > 0$ de lo contrario es asignado en la clase ω_j .

Después de realizar la clasificación del descriptor \mathbf{x} a una de las clases ω_i , se realiza el cálculo del error cuadrático medio (ECM) de la asignación del vector de características \mathbf{x} a la clase ω_i . La medición del error se basa en el cálculo de la distancia euclídea entre el rostro objeto y el prototipo de la clase a la que fue asignada.

$$ECM(\mathbf{x}) = \sqrt{\frac{(\mathbf{pm} - \mathbf{px})^2}{N}} \quad (5.41)$$

donde $E(x)$ es el error promedio, \mathbf{px} es el rostro a clasificar, \mathbf{pm} es el prototipo de la clase a la que fue asignado y N es la longitud del vector de $(\mathbf{pm} - \mathbf{px})^2$.

5.5.1 Decisión

La validación de la asignación del descriptor \mathbf{x} a la clase ω_i , se lleva a cabo comparando el ECM de asignación con un umbral previamente establecido. Si el error es mayor al umbral, el usuario es rechazado por el sistema. Sin embargo se pueden producir dos tipos de errores: Erro Falso Negativo (EFN), si la persona aun estando registrado en el sistema produce un ECM mayor al umbral y un Error Falso Positivo (EFP) si la persona que no se encuentra registrada en el sistema produce un ECM menor al umbral, por lo que el sistema acepta erróneamente al usuario.

Para definir el umbral de rechazo, se realizó la clasificación de tres grupos de prueba los cuales se denominan: TEST1, TEST2 y TEST3, con las siguientes características.

- TEST1. Es un conjunto de rostros tomados de manera aleatoria de la base de datos de rostros 3D.
- TEST2. Es un conjunto de rostros cuya identidad se encuentra registrada en el sistema y tiene muestras registradas en la base de datos de rostros 3D. Los elementos de este grupo son muestras nuevas no almacenadas en el sistema.
- TEST3. Es un conjunto de rostros cuya identidad del usuario y rostros 3D no se encuentra registrada en el sistema de reconocimiento facial, (conjunto desconocido).

La tabla 5.6, describe los resultados de la clasificación de los rostros de cada grupo de prueba, la primera columna es el identificador del rostro a clasificar, la

segunda columna corresponde al identificador de la clase asignada y la tercera columna indica el error de clasificación.

TEST1			TEST2			TEST3		
ID_Objeto	ID_Enc	ECM	ID_Objeto	ID_Enc	ECM	ID_Objeto	ID_Enc	ECM
0004	0004	0.11	0025	0025	0.0345	0000	0013	0.4906
0007	0007	0.1134	0021	0021	0.0995	0003	0014	0.6426
0010	0010	0.1015	0033	0033	0.0814	0003	0014	0.6410
0014	0014	0.0746	0035	0035	0.1880	0008	0013	0.6394
0015	0015	0.0419	0037	0037	0.0953	0008	0013	0.5918
0017	0017	0.0468	0030	0030	0.0506	0018	0042	0.3681
0020	0020	0.1243	0040	0040	0.2730	0018	0042	0.3925
0020	0020	0.226	0041	0041	0.0837	0024	0010	0.5385
0021	0021	0.1732	0042	0042	0.0595	0027	0013	0.5871
0015	0015	0.1185	0043	0043	0.1855	0027	0013	0.5829
0026	0026	0.1145	0047	0047	0.3109	0048	0010	0.7973
0031	0031	0.0945	0032	0010	0.4723	0048	0030	0.7436
0031	0031	0.1207	0037	0037	0.0696	0049	0015	0.6626
0040	0040	0.0885	0022	0022	0.1290	0002	0010	1.0000
0043	0043	0.0502	0020	0020	0.2145	0002	0010	1.0635

Máximo	0.226
Mínimimo	0.0419
Promedio	0.1106
Desv. Est.	0.0468

Máximo	0.4723
Mínimimo	0.0345
Promedio	0.1565
Desv. Est.	0.1207

Máximo	1.0635
Mínimimo	0.3681
Promedio	0.6494
Desv. Est.	0.1927

Tabla 5.6. Resultados de la ejecución de módulo de reconocimiento sobre los grupos: Test1, Test2 y Test3.

La Figura 5.36 corresponde a la gráfica del análisis estadístico básico del ECM, podemos observar el comportamiento de los errores máximos y mínimos, el error promedio y la desviación estándar de cada grupo de prueba.

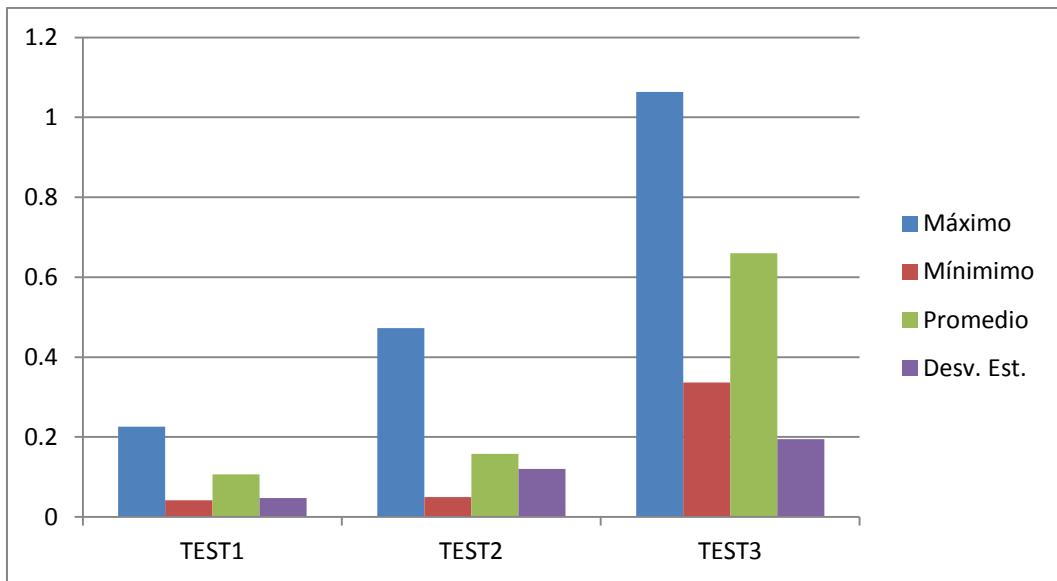


Figura 5.36. Análisis estadístico del ECM de clasificación de los grupos de prueba.

A partir de las observaciones podemos definir un umbral de rechazo para que las siguientes clasificaciones se puedan llevar a cabo de manera automática. Se decidió tomar como umbral el valor del error mínimo del grupo TEST3, el cual tiene asociado un valor mayor en comparación con la mayoría de los parámetros de los grupos TEST1 Y TEST2 a excepción del valor máximo del grupo TEST2.

$$Umbral = 0.3681 \quad (5.42)$$

De esta manera en el proceso de autentificación un rostro de prueba se considera que pertenece a una la clase si el ECM es menor al umbral:

$$ECM < Umbral \quad (5.43)$$

5.6 Resultados

En esta sección se discute de manera general los resultados obtenidos en los procesos de mayor relevancia de este trabajo, los cuales son: la obtención de la información de profundidad del rostro, la alineación del rostro a partir del perfil de un rostro modelo y la tolerancia de reconocimiento con diferentes expresiones faciales.

5.6.1 Obtención del modelo tridimensional de los rostros mediante proyección de franjas.

La Figura 5.37 corresponde a una primera aproximación del modelo 3D del rostro de un usuario obtenido mediante el método de proyección de franjas con perfil binario. En la imagen se observa que en la superficie del modelo 3D presenta ruido aleatorio que se interpretan como señales de alta frecuencia.

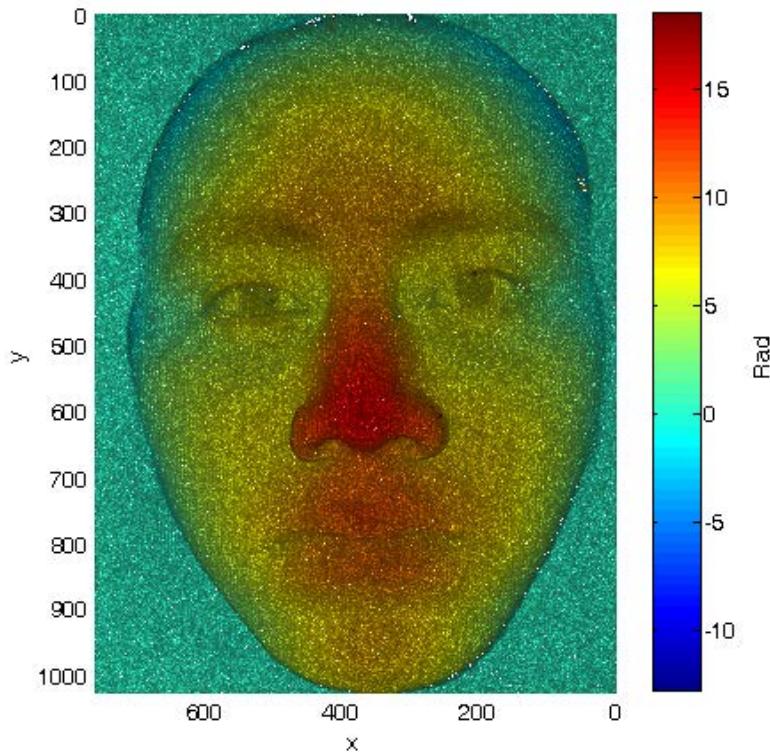


Figura 5.37. Primera aproximación del modelo 3D del rostro. Vista del plano (x, y)

Para realizar la remoción del ruido, de la superficie del modelo 3D, se aplica un filtro pasa bajas Butterworth en el dominio de frecuencias. Los resultados gráficos del filtro aplicado, se presentan en las Figuras 5.38 – 5.40, en ella se observa una superficie con variación suave donde se han suprimido las variaciones aleatorias.

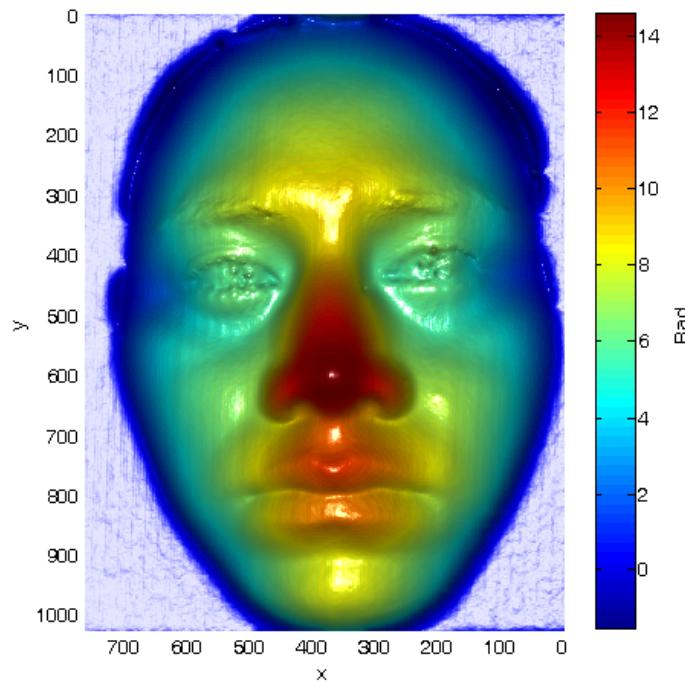


Figura 5.38. Modelo 3D después de realizar la remoción del ruido, vista del plano (x, y)

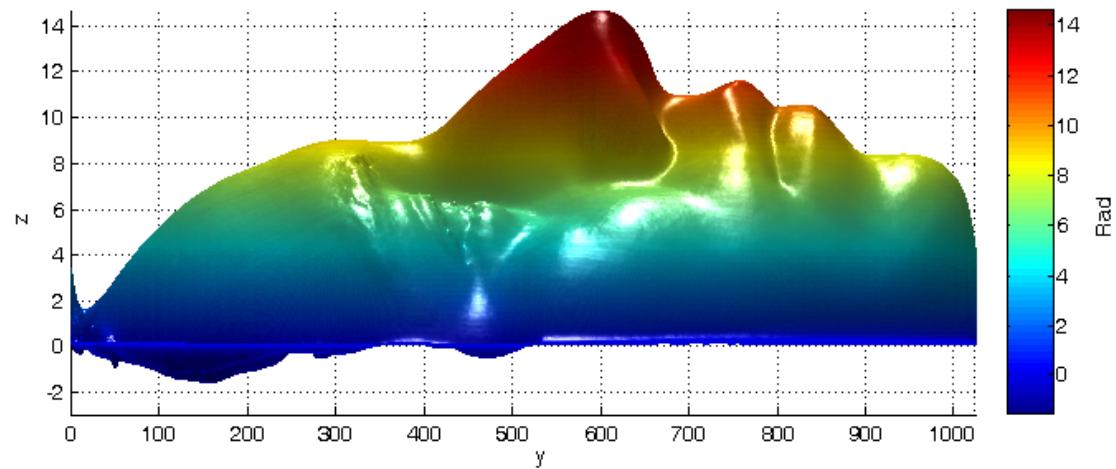


Figura 5.39. Modelo 3D después de realizar la remoción del ruido, vista del plano (x, y).

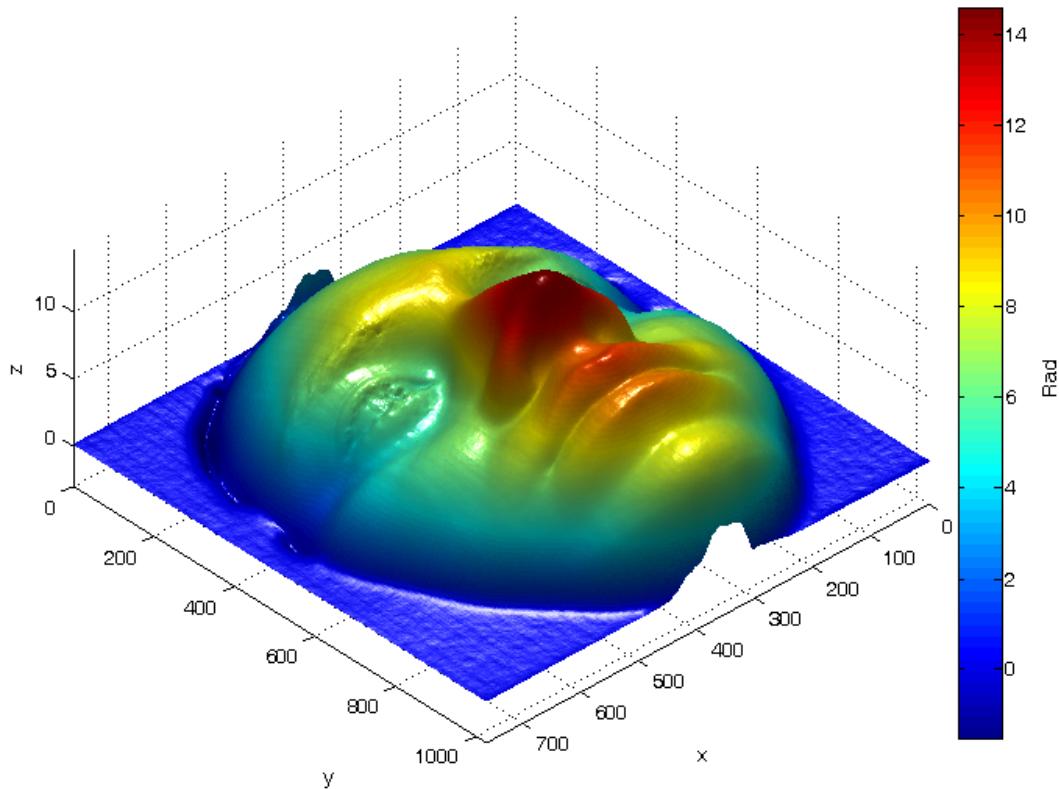


Figura 5.40. Modelo 3D después de realizar la remoción del ruido, vista arbitraria.

Uno de los aspectos importantes a considerar en este proceso es el tiempo que se requiere desde la adquisición de las imágenes del rostro hasta la obtención del modelo 3D. El tiempo promedio en cada uno de estos procesos se presenta en la Tabla 5.7, estos datos se obtuvieron al realizar la reconstrucción tridimensional de diez rostros de diferentes usuarios.

T. promedio de adquisición (s)	T. promedio de reconstrucción (s)	T. promedio de remoción de ruido	T. promedio total (s)
1.3	2.4971	0.6886	4.48567

Tabla 5.7. Tiempo promedio del proceso de reconstrucción 3D del rostro.

Mediante el proceso descrito anteriormente, se registraron 172 rostros correspondientes a 47 usuarios con un promedio de 4 rostros por persona. Dentro de esta colección, algunos presentan distintas expresiones faciales y ligeras rotaciones sobre los ejes coordenados. La Figura 5.41 corresponde a una muestra de los modelos 3D registrados en el sistema de reconocimiento, originalmente estas imágenes fueron registradas a una resolución de 1024 X 711 píxeles y están codificados en niveles de grises para su visualización en 2D.

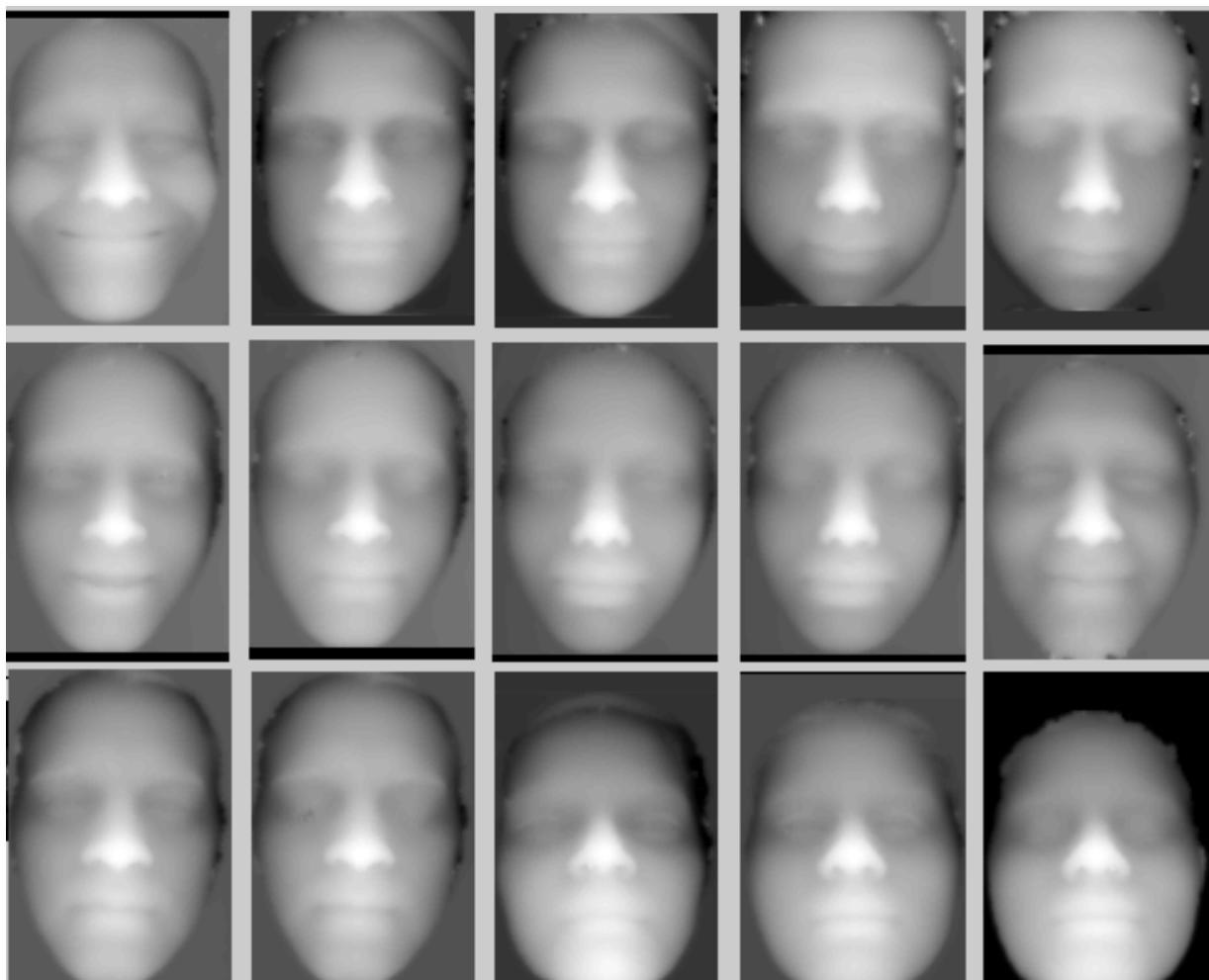


Figura 5.41. Muestra de los modelos tridimensionales de los rostros registrados.

5.6.2 Alineación del rostro a partir del perfil de un rostro modelo.

La segunda contribucion del trabajo de tesis es la alineación de los rostros a partir del perfil y los lagrimales de un rostro modelo. Mediante este enfoque se optimiza la velocidad del procesamiento. En la Figura 5.42 (a) se puede observar el rostro modelo y (b) el perfil del mismo.

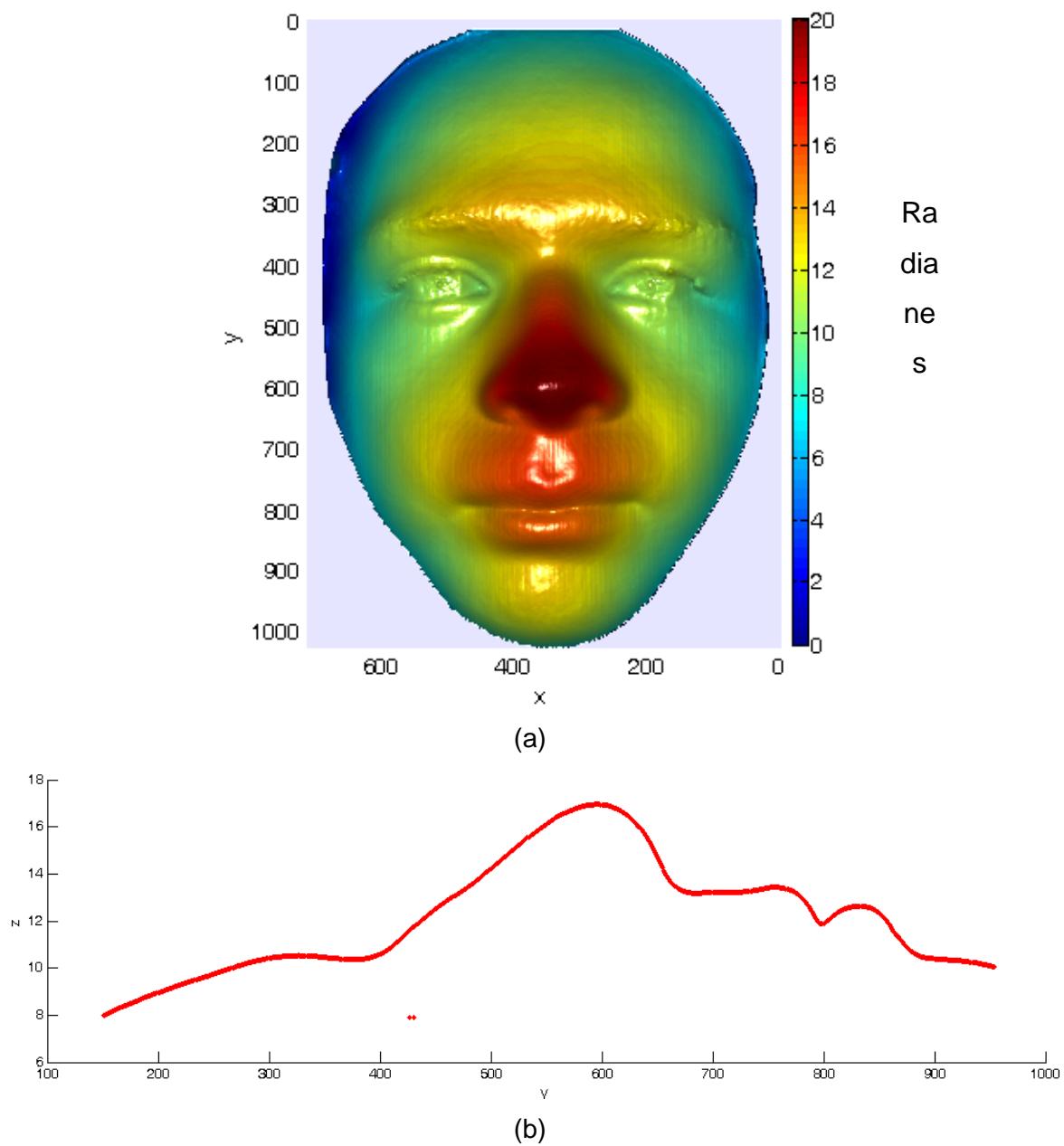


Figura 5.42. (a) Rostro modelo y (b) perfil del rostro.

A continuación se presenta el resultado de la alineación de tres rostros que presentan rotaciones sobre alguno de los ejes.

a) Ejemplo 1. Rostro con rotación respecto al eje x , frente sobresaliente.

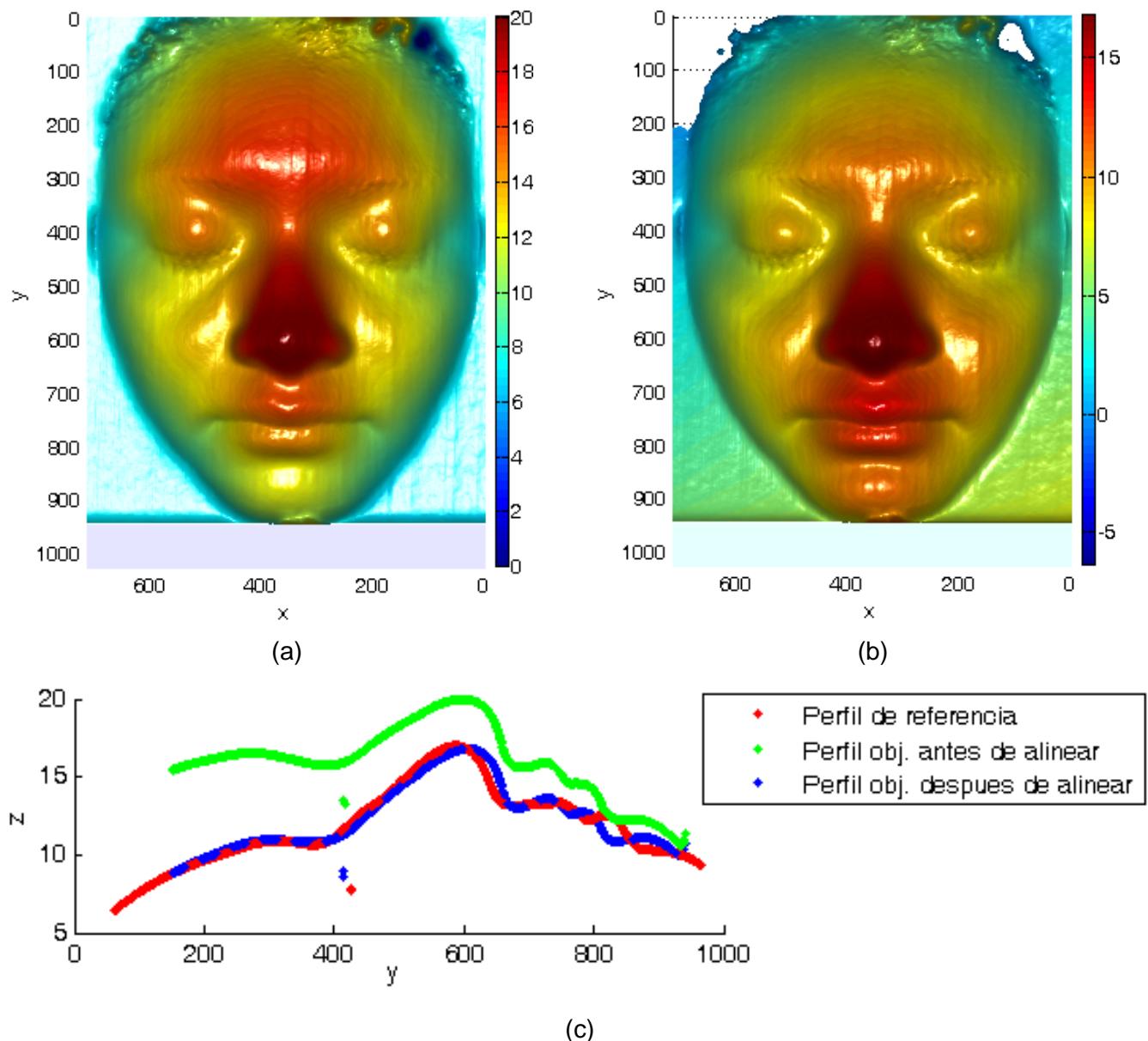


Figura 5.43 (a) Rostro antes de alinear, (b) rostro después de alinear y (c) perfiles antes y después de la alineación.

b) Ejemplo 2. Rostro con rotacion respecto al eje x , barbilla sobresaliente.

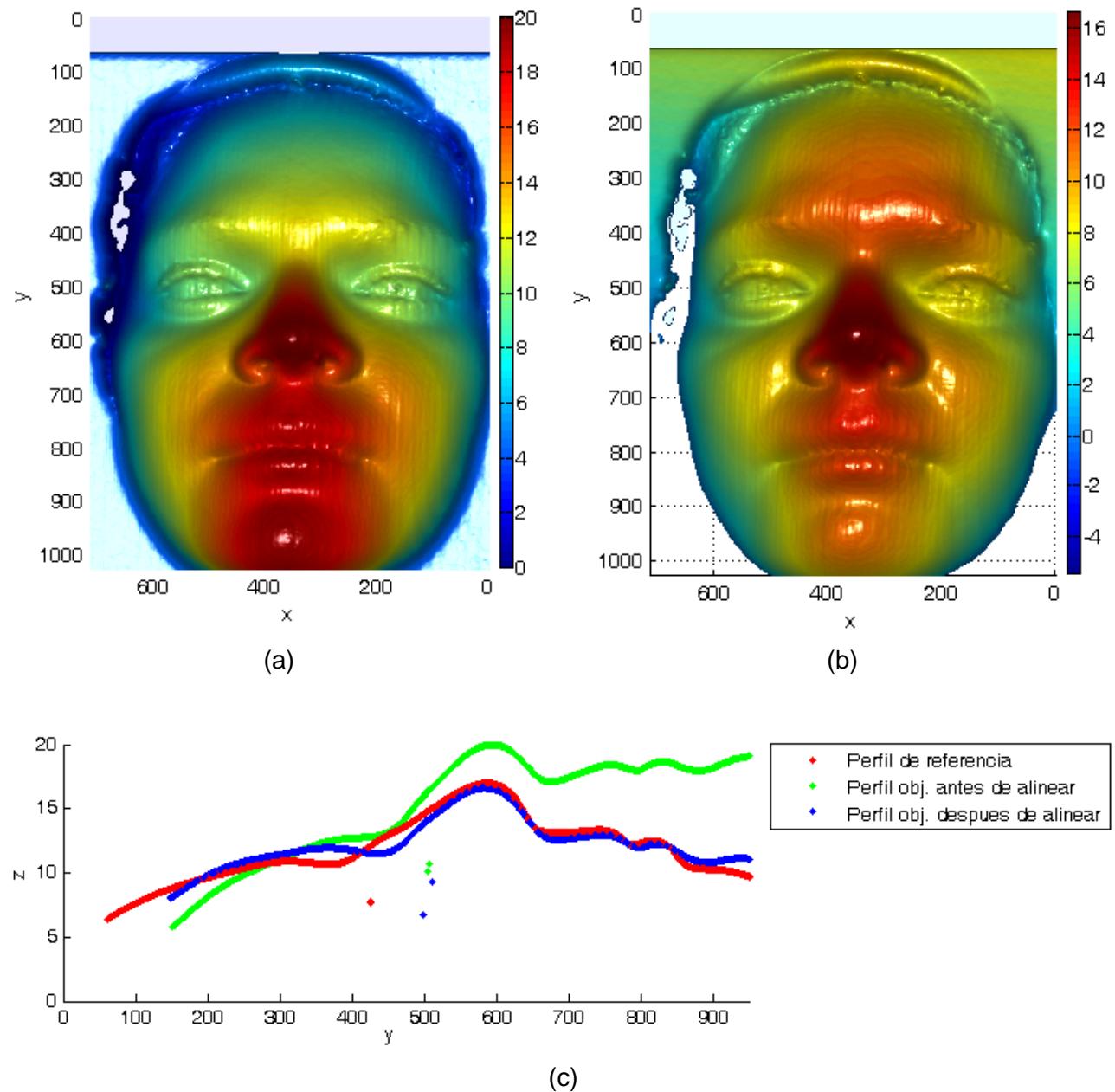


Figura 5.44 (a) Rostro antes de alinear, (b) rostro después de alinear y (c) perfiles de antes y después de la alineación.

c) Ejemplo 3. Rotacion respecto al eje y.

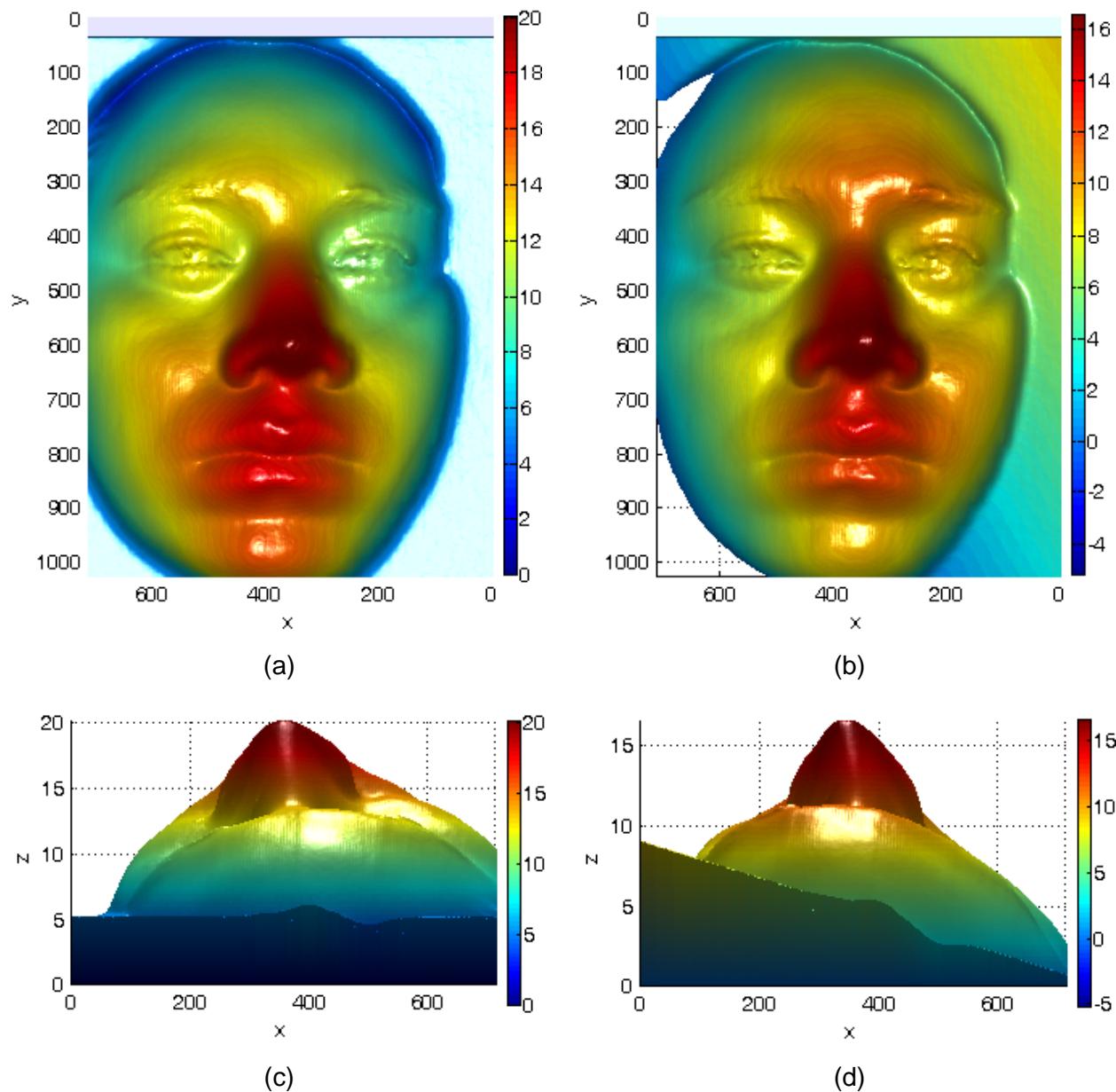


Figura 5.45 (a) Rostro antes de alinear, vista (x, y) , (b) rostro después de alinear, vista (x, y) , (c) rostro antes de alinear, vista (x, z) y (d) rostro después de alinear, vista (x, z) .

La Tabla 5.8 corresponde al análisis del error cuadrático medio (ECM) de la alineación de los perfiles modelo y referencia, de los ejemplos presentados.

	ECM antes de la alineación	ECM después de la alineación	Acercamiento (%)	Núm. de Iteraciones	Tiempo (s)
Ejemplo 1	10.952	1.0658	90.27	6	0.2059
Ejemplo 2	22.2891	1.7124	92.32	6	0.2566
Ejemplo 3	16.7606	1.1518	93.13	6	0.2091

Tabla 5.8. Datos de la alineación de los perfiles de modelo y objeto

La Tabla 5.9 corresponde al ECM de la alineación de la superficie del rostro objeto respecto al del rostro modelo.

Ejemplo	ECM antes de la alineación	ECM después de la alineación	Acercamiento (%)	Tiempo (s)
1	3.7851	0.8533	77.46	0.3812
2	4.3035	1.2762	70.35	0.3875
3	4.2876	1.1014	74.31	2.9567

Tabla 5.9. Datos de la alineación de la superficie del rostro objeto

En la Tabla 5.10 se presenta análisis del tiempo de cada una de las etapas del proceso de alineación. En ella se establece el tiempo necesario para realizar el análisis de curvatura de la superficie del rostro, el tiempo para realizar la clasificación HK y localizar los puntos de interés, el tiempo de cálculo de la matriz de transformación (MT), el tiempo para alinear la superficie del rostro y finalmente el tiempo total del proceso de alineación.

Ejemplo	Análisis de curvatura (s)	Clasificación HK (s)	Cálculo de la MT (s)	Aplicación de MT (s)	Tiempo total (s)
1	2.2168	0.3089	0.2059	0.3812	3.1128
2	2.3654	0.2793	0.2566	0.3875	3.2887
3	2.1447	0.2554	0.2091	0.3475	2.9567
T. Promedio	2.2423	0.2812	0.2239	0.3721	3.1194

Tabla 5.10. Datos del tiempo de procesamiento en cada etapa del proceso de alineación.

5.6.3 Reconocimiento a partir de la señal demodulada.

El proceso de reconocimiento se realiza sobre los rostros normalizados y segmentados de acuerdo a los procesos descritos en la sección 5.4. La figura 5.46 corresponde a una muestra de los modelos 3D antes y después de ser normalizados a una misma posición y segmentados reteniendo únicamente la región de interés.

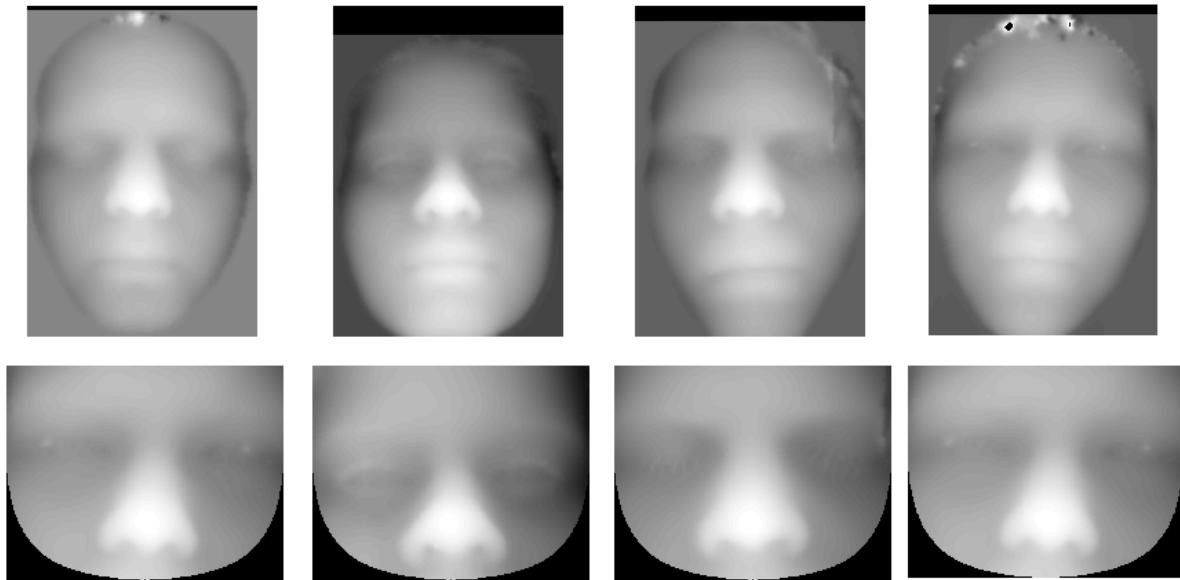


Figura 5.46 Modelos faciales antes y después de realizar el proceso de normalización y segmentación.

La validacion de la asignacion de un rostro a una clase se realiza de acuerdo al umbral establecido en la seccion 5.5.1.

$$umbral = 0.3681 \quad (5.44)$$

En las pruebas de identificación, se consideraron dos grupos de prueba: Grupo A, Conformado por 9 usuarios no registrados en el sistema; y el Grupo B, conformado por todos los usuarios registrados en el sistema (47 usuarios).

En las pruebas realizadas con el Grupo A, el sistema rechazó el 100% de los usuarios debido que generan un ECM mucho mayor al umbral establecido. En el caso del Grupo B, 46 de los 47 usuarios fueron identificados de manera

satisfactoria y un usuario fue rechazado erróneamente, es decir produjo un error *falso negativo*, por haber generado un error mayor al umbral establecido. En la tabla 5.11 se resumen los resultados obtenidos.

	Usuarios	Identificación		Rechazo		Falso positivo		Falso negativo	
		Núm.	%	Núm.	%	Núm.	%	Núm.	%
Grupo A	9	0	0	9	100	0	0	0	0
Grupo B	47	46	97.87	0	0	0	0	1	2.13

Tabla 5.11. Resultados de la clasificación de usuarios registrados y no registrados en el sistema de reconocimiento facial.

a) Resultados del Grupo A.

La tabla 5.12 corresponde a los resultados de la clasificación del Grupo A. Se debe tomar en cuenta que el usuario es asignado a una clase incorrecta, además el ECM es mucho mayor que el umbral de aceptación, por lo tanto el tipo de resultado es *negativo*, es decir el usuario no es identificado por el sistema.

ID_Obj	ID_Enc	ECM	Resul.	T. Bsq.
0000	0013	0.4906	Negativo	0.2487
0003	0014	0.6426	Negativo	0.2545
0003	0014	0.641	Negativo	0.229
0008	0013	0.6394	Negativo	0.2296
0008	0013	0.5918	Negativo	0.2292
0018	0020	0.7288	Negativo	0.2968
0018	0042	0.3925	Negativo	0.2885
0024	0047	0.3367	Negativo	0.2672
0027	0013	0.5871	Negativo	0.2487
0027	0013	0.5829	Negativo	0.2327
0048	0010	0.7973	Negativo	0.2741
0048	0030	0.7436	Negativo	0.2343
0049	0015	0.6626	Negativo	0.4846
0002	0010	1.0000	Negativo	0.2432
<u>002</u>	10	1.0635	Negativo	0.2407
Tiempo promedio de búsqueda				0.2668

Tabla 5.12. Resultados de la clasificación del grupo de los usuarios no registrados en el sistema.

b) Resultados del Grupo B.

La tabla 5.13 corresponde a los resultados de la clasificación del Grupo B, usuarios registrados en el sistema. En este caso, solo un usuario generó un ECM mayor que el umbral, sin embargo la asignación a la clase es correcta por lo tanto el error generado es del tipo *Falso Negativo*.

ID_Obj	ID_Enc	ECM	Resul.	T. Bsq.	ID_Obj	ID_Enc	ECM	Resul.	T. Bsq.
0003	0003	0.0177	Positivo	0.2746	0029	0029	0.059	Positivo	0.4134
0004	0004	0.1651	Positivo	0.3501	0026	0026	0.3164	Positivo	0.3072
0005	0005	0.106	Positivo	0.2901	0031	0031	0.1008	Positivo	0.2949
0006	0006	0.1349	Positivo	0.2583	0032	0032	0.1777	Positivo	0.3107
0007	0007	0.2268	Positivo	0.2485	0033	0033	0.0319	Positivo	0.268
0008	0008	0.1132	Positivo	0.2463	0034	0034	0.0664	Positivo	0.3477
0009	0009	0.2696	Positivo	0.2642	0035	0035	0.1529	Positivo	0.3803
0010	0010	0.1353	Positivo	0.2584	0036	0036	0.3174	Positivo	0.3562
0011	0011	0.2578	Positivo	0.2631	0037	0037	0.2493	Positivo	0.277
0012	0012	0.2288	Positivo	0.2636	0038	0038	0.5946	FN	0.2877
0013	0013	0.1155	Positivo	0.2516	0039	0039	0.0359	Positivo	0.3609
0014	0014	0.1493	Positivo	0.3875	0040	0040	0.1124	Positivo	0.3997
0015	0015	0.173	Positivo	0.3436	0041	0041	0.1414	Positivo	0.4066
0016	0016	0.1722	Positivo	0.277	0042	0042	0.033	Positivo	0.3006
0017	0017	0.0937	Positivo	0.2579	0043	0043	0.2668	Positivo	0.3228
0018	0018	0.1405	Positivo	0.2745	0048	0048	0.1782	Positivo	0.2474
0020	0020	0.1611	Positivo	0.3279	0044	0044	0.137	Positivo	0.279
0021	0021	0.2883	Positivo	0.3186	0045	0045	0.0833	Positivo	0.2705
0001	0001	0.1098	Positivo	0.3271	0046	0046	0.2282	Positivo	0.2784
0023	0023	0.1195	Positivo	0.3465	0047	0047	0.1603	Positivo	0.3001
0025	0025	0.16	Positivo	0.4623	0050	0050	0.141	Positivo	0.3006
0027	0027	0.0606	Positivo	0.2541	0002	0002	0.1132	Positivo	0.2746
0030	0030	0.0675	Positivo	0.2714	0022	0022	0.199	Positivo	0.5294
0028	0028	0.1256	Positivo	0.363					

Tabla 5.13. Resultados de la clasificación de los usuarios registrados en el sistema.

Finalmente en la tabla 5.14, se presenta el tiempo promedio requerido desde la digitalización de los rostros hasta el proceso de reconocimiento. Cabe recordar que todo el proceso descrito se ha realizado en un equipo portátil con condiciones limitadas de procesamiento. Las posibles mejoras en cuanto al tiempo de procesamiento se discuten en la sección de trabajos a futuro (sección 6.2).

T. prom. de adquisición y reconstrucción	4.4857 s
T. prom. de alineación	3.1194 s
T. prom de búsqueda	0.2668 s
Tiempo prom. total	7.8719 s

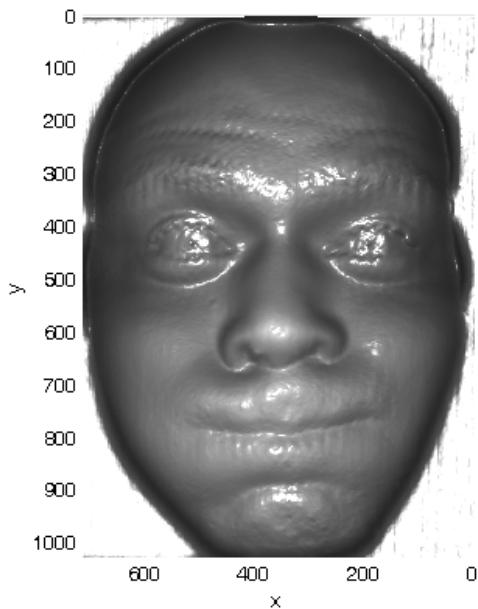
Tabla 5.14. Tiempo promedio de la respuesta del sistema.

5.6.4 Tolerancia de reconocimiento frente a gestos y expresiones faciales.

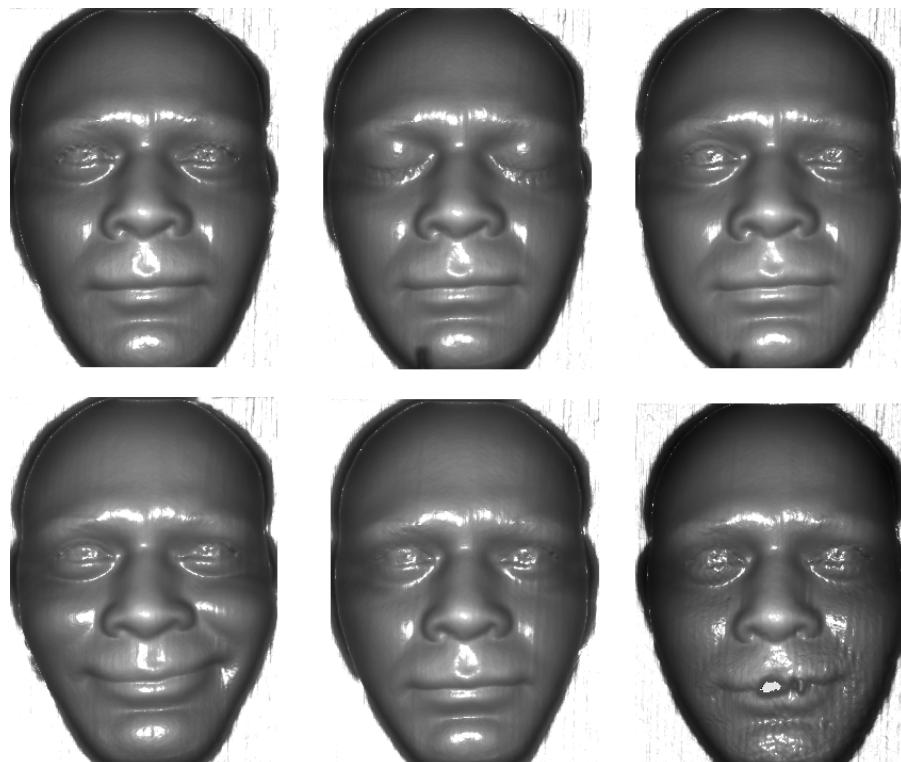
El sistema desarrollado tiene la capacidad clasificar de manera positiva a personas con distintas expresiones faciales, esto debido a que el análisis del rostro se realiza únicamente con regiones que sufre variaciones mínimas frente a estos cambios como son la frente, los ojos y la nariz.

A continuación se presentan 4 ejemplos del proceso de reconocimiento bajo distintas expresiones faciales. Las Figuras 5.47 - 5.50, corresponden a rostros con diferentes expresiones faciales, las imágenes etiquetadas con el inciso (a) corresponde al rostro a identificar (rostro objeto) y la imagen etiquetadas con el inciso (b) corresponde al conjunto de rostros de entrenamiento de dicho usuario. Posterior a la presentación de las figuras, se presentan los resultados y comentarios del proceso de clasificación.

a) Ejemplo 1. ID usuario: 0047



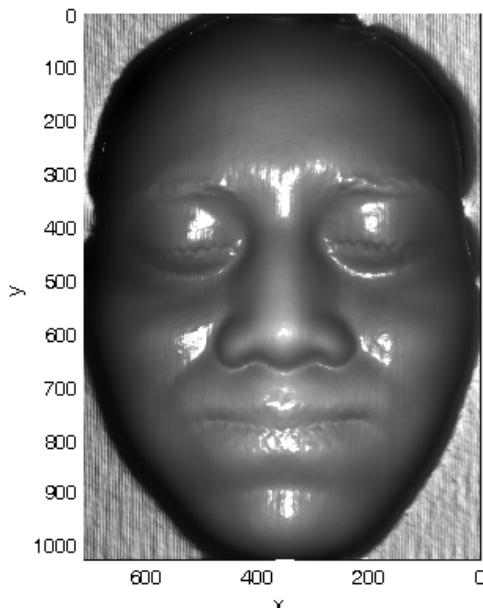
(a)



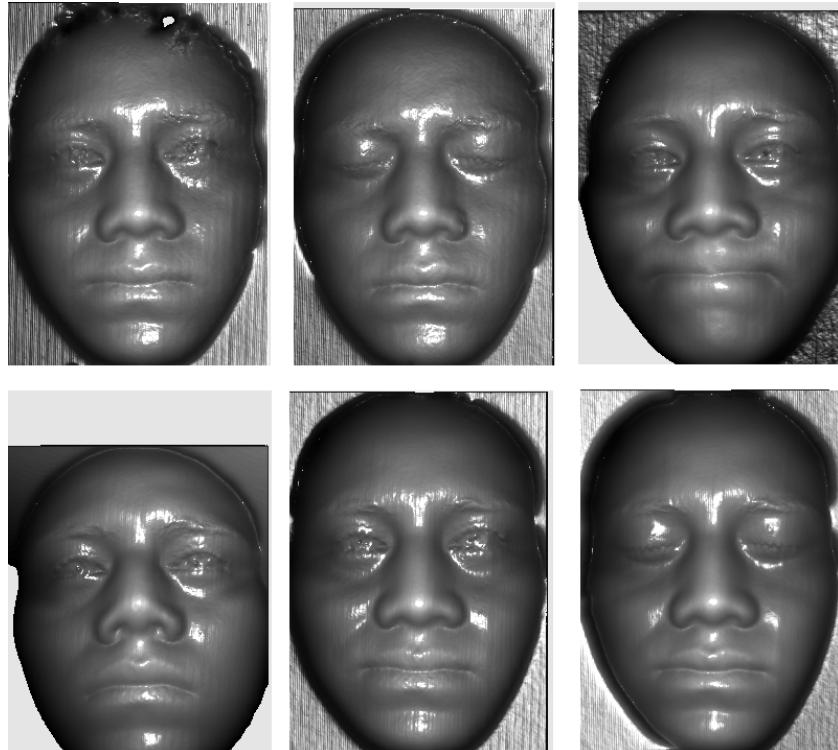
(b)

Figura 5.47. ID usuario: 0047. (a) Rostro objeto y (b) conjunto de rostros de entrenamiento.

b) Ejemplo 2. ID usuario: 0022



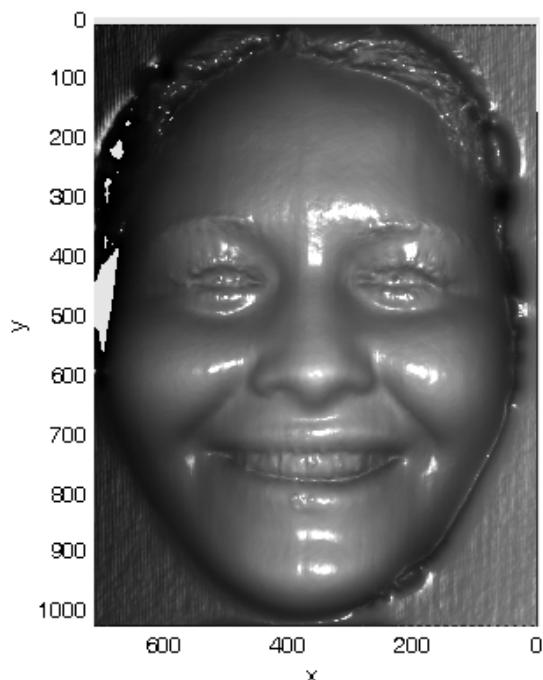
(a)



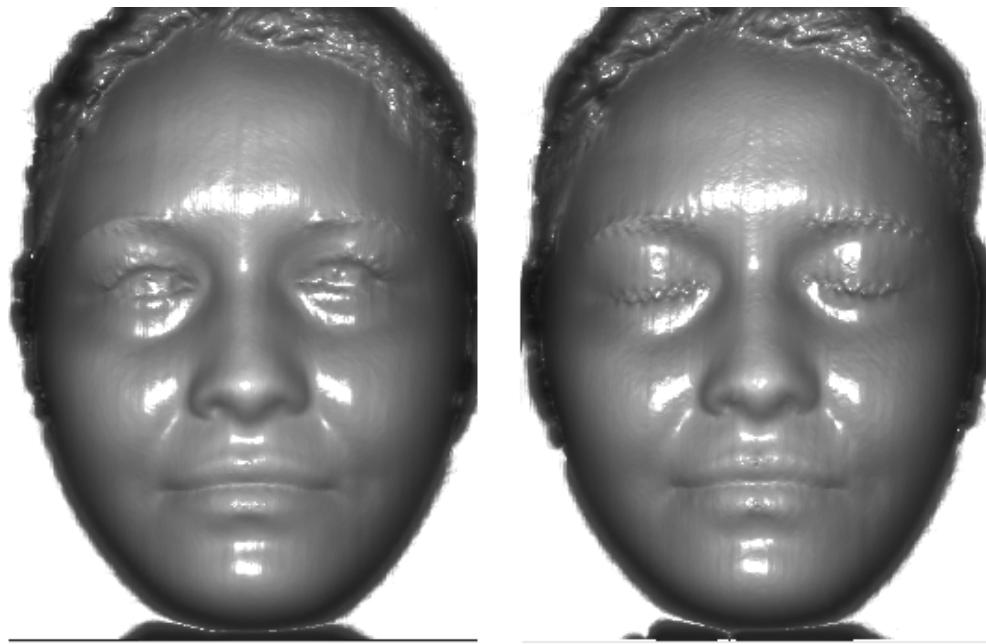
(b)

Figura 5.48. ID usuario: 0022. (a) Rostro objeto y (b) conjunto de rostros de entrenamiento.

c) Ejemplo 3. ID usuario: 0031.



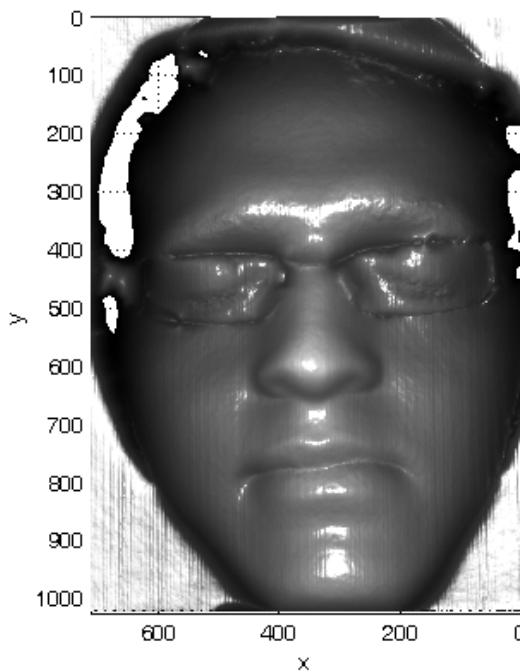
(a)



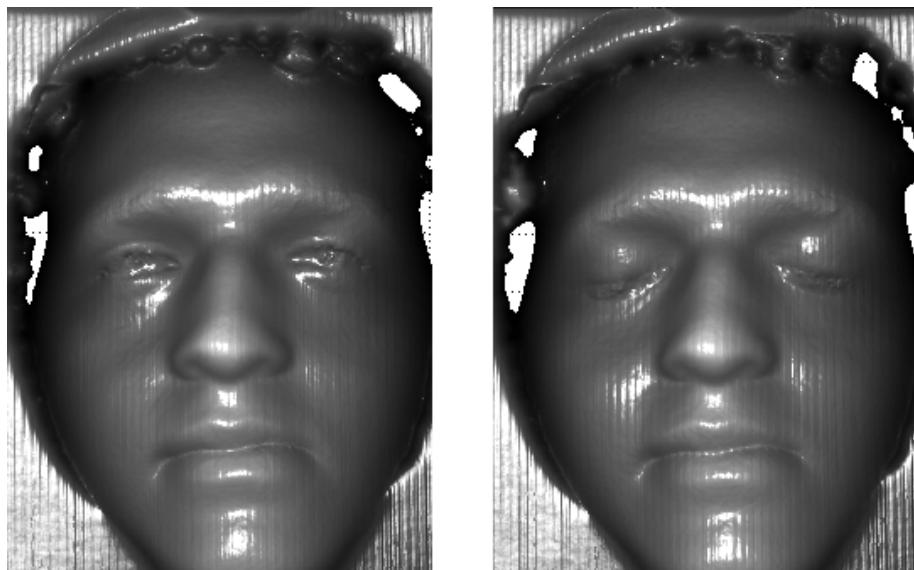
(b)

Figura 5.49. ID usuario: 0031. (a) Rostro objeto y (b) conjunto de rostros de entrenamiento.

d) Ejemplo 4. ID usuario: 0004.



(a)



(b)

Figura 5.50. ID usuario: 0004. (a) Rostro objeto y (b) conjunto de rostros de entrenamiento.

Los resultados de la clasificación de los rostros de las de los ejemplos 1–4, se presentan en la Tabla 5.15. Cabe recordar que el error umbral de rechazo es $ECM = 0.3682$.

Ejemplo	ID_Obj	ID_Enc	ECM	Tipo de resultado	Núm. de rostros de entrenamiento
1	0022	0022	0.2547	Positivo	20
2	0047	0047	0.3109	Positivo	6
3	0031	0031	0.1809	Positivo	2
4	0004	0004	0.2279	Positivo	3

Tabla 5.15. Resultados de la clasificación de rostros con distintas expresiones faciales.

En el ejemplo 1, se observa que en el conjunto de rostros de entrenamiento uno de ellos presenta errores de reconstrucción en la parte de la mejilla, mandíbula y labios, debido que al realizar la digitalización del rostro del usuario, éste hizo algún movimiento alterando los resultados de la reconstrucción. A pesar la expresión facial en el rostro de prueba y la variación en el conjunto de rostros de entrenamiento el resultado de la clasificación es positivo.

En el ejemplo 2, el aspecto importante a destacar es que uno de los rostros de entrenamiento tiene una ligera rotación; aún con esta variación en los rostros de entrenamiento y la expresión en el rostro de prueba, el resultado de la clasificación es positivo.

Para Ejemplo 3, el usuario solo cuenta con dos imágenes de entrenamiento, a pesar de ello fue uno de los que generó un ECM más bajo.

En el ejemplo 4, hay dos aspectos importantes a destacar: 1) el número de rostros de entrenamiento es de dos muestras y 2) el rostro de prueba fue digitalizado con anteojos, mientras que los de entrenamiento fueron digitalizados sin éste. A pesar de la combinación de las variaciones entre el rostro objeto y los rostros de entrenamiento, se realizó una identificación exitosa.

Capítulo 6. Conclusión y trabajo a futuro

6.1 Conclusión

En el presente trabajo de tesis se ha presentado el desarrollo de un sistema de reconocimiento facial a partir del análisis de la información de profundidad del rostro, el cual se encuentra conformado por distintos módulos que implementan diferentes algoritmos para realizar el procesamiento y análisis de los datos del rostro.

La Tabla 6.1 describe los módulos del sistema de identificación facial así como las técnicas y algoritmos que implementan.

Módulo	Técnica / Algoritmo
Módulo de reconstrucción tridimensional	<ul style="list-style-type: none">- Sistema de proyección de luz estructurada.- Proyección de franjas binarias mediante desplazamiento de fase.- Desenvolvimiento de fase mediante mapas de calidad- Fundamentos de procesamiento de imágenes: filtros en el dominio de frecuencias.
Módulo de alineación y segmentación	<ul style="list-style-type: none">- Análisis de superficies- Alineación de nubes de puntos mediante ICP- Mascaras binarias
Módulo de extracción de características	<ul style="list-style-type: none">- Análisis de componentes principales (Eigenrostros)
Módulo de clasificación y reconocimiento	<ul style="list-style-type: none">- Análisis discriminante.

Tabla 6.1. Técnicas y algoritmos implementados en el sistema de reconocimiento facial.

El sistema de reconocimiento facial es, entonces, una aplicación biométrica que integra un conjunto de herramientas de diferentes disciplinas tal como la óptica, el, reconocimiento de patrones, la visión artificial, geometría diferencial, y la estadística.

Las contribuciones de este trabajo son las siguientes: la aplicación de la técnica de proyección de franjas para el cálculo de la nube de puntos relacionada al rostro, el uso de la fase de la señal demodulada para realizar el reconocimiento, la alineación del rostro a partir del perfil de un rostro modelo así como la tolerancia de gestos y expresiones faciales en el proceso de reconocimiento.

La obtención de la información de profundidad del rostro se lleva a cabo mediante la técnica de proyección de franjas, el cual es un proceso rápido y sobre todo seguro ya que no afecta la integridad del usuario, esto da posibilidad de obtener el modelo tridimensional de otras partes del cuerpo humano con fines médicos o en el área de la industria en general. En las revisiones bibliográficas realizadas se encuentran trabajos que realizan el cálculo de la información de profundidad a partir de la proyección de luz láser o mediante sistemas de visión estereoscópica, ambos métodos resultan ser procesos lentos y en el caso del uso de la luz láser conlleva riesgos a daños permanentes en los ojos.

La segunda contribución del trabajo es referente al proceso de alineación del rostro, el cual se basa en la alineación respecto al perfil de un rostro modelo. Este proceso permite normalizar los rostros a la misma posición con pocos puntos en vez de usar todos los puntos del rostro. Mediante este proceso se logra optimizar la carga computacional y el tiempo de procesamiento. De esta manera, en las pruebas realizadas se logró reducir el tiempo de procesamiento hasta un 99% respecto a la alineación con todos los puntos de la superficie del rostro.

El proceso de reconocimiento se basa en el análisis de la señal demodulada de un conjunto de franjas proyectadas. El usar este tipo de señal resulta ser robusto frente a intentos de manipulación y/o suplantación de la identidad; además, el sistema, ha demostrado ser fiable y seguro para realizar la identificación efectiva de los usuarios. En las pruebas de identificación se logró la identificación positiva en el 97.87% de los casos, mientras que el 2.13% fueron casos de falso negativo. Cabe destacar que en los sistemas de verificación es preferible obtener resultados del tipo falso negativo que los resultados de tipo falso positivo; en muchas de las

ocasiones los resultados de falsos negativos se solucionan con una nueva adquisición de los datos biométricos del usuario.

Es importante recalcar que el sistema desarrollado es capaz de clasificar, de manera positiva, a personas con distintas expresiones faciales del usuario. Esto se debe, en gran medida, a que el análisis del rostro se realiza únicamente en las regiones que sufre variaciones mínimas frente a estos cambios como son la frente, los ojos y la nariz.

Finalmente, los sistemas de identificación biométrica basados en el análisis de la tridimensional del rostro han demostrados ser fiables, son fáciles de usar, tiene un alto grado de aceptación entre los usuarios y por lo general son estables. Los sistemas basados en el análisis de información tridimensional han demostrado ser mucho más robusto frente a intentos de manipulación o suplantación de los usuarios, esto en gran medida, por la complejidad de la reconstrucción tridimensional del rostro y los algoritmos implementados en el análisis de los datos.

6.2 Trabajo a futuro

Es primordial que los sistemas de reconocimientos biométricos sean fiables y robustos desde la adquisición de los datos hasta el reconocimiento del usuario. Por esta razón, adicional al trabajo desarrollado, se tiene contemplado realizar mejoras tanto en el paradigma de reconocimiento biométrico, en los algoritmos de clasificación, el software y el hardware del sistema.

En cuanto al paradigma de reconocimiento biométrico se prevé implementar un sistema multibiométrico, que combine el análisis de la profundidad del rostro con otros rasgos del cuerpo tales como la estatura, el peso o el tono de la voz por mencionar algunos. Esto permite diseñar un sistema de reconocimiento robusto frente a intentos de suplantación de usuarios.

Respecto a los algoritmos de clasificación, en un futuro, se pretende explorar y realizar pruebas con diferentes herramientas como son las redes neuronales, algoritmos evolutivos, lógica difusa y algoritmos de agrupamiento (clustering) para evaluar el algoritmo idóneo.

Las mejoras en cuanto a las mejoras del software, queda pendiente programar los distintos algoritmos en lenguajes no restrictivos (software libre) como son C++ o phyton. Por otra parte se pretende diseñar una arquitectura de la base de datos de tipo entidad-relación, que permita la administración y garantice la seguridad de la información de los usuarios.

Respecto al hardware del sistema, se pretende desarrollar un sistema de escaneo compacto y portable, que faciliten su calibración e instalación en distintos ambientes. Cabe la posibilidad de sustituir el proyector de video por alguna fuente de luz infrarroja y la cámara CCD por una cámara sensible a radiaciones en infrarrojas, para disminuir los efectos de la luz blanca sobre los ojos del usuario. Además, en un futuro, se busca desarrollar los algoritmos en sistemas embebidos o implementar computación paralela para disminuir el tiempo de procesamiento.

Finalmente, las aplicaciones del sistema desarrollado, en principio, puede ser usado desde el control de entrada/salida en el ámbito empresarial o residencial, en el control de acceso virtual a recursos informáticos hasta el control de acceso físico en áreas restringidas solo a personal autorizada. Explorando otros campos de aplicación, se puede aprovechar el sistema de reconstrucción 3D para el desarrollo de herramientas para el análisis y diagnósticos médicos, creación de prótesis, herramientas forenses, realidad virtual, computación gráfica, impresión 3D de rostros, aplicaciones comerciales que usen la información tridimensional del rostro como credencial o membresía, por mencionar algunos.

Apendice A. Manipulación de matrices

Un vector columna \mathbf{x} y su traspuesta \mathbf{x}^T se define como:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} \quad (\text{A.1})$$

$$\mathbf{x}^T = [x_1 \ x_2 \ \dots \ x_d] \quad (\text{A.2})$$

Una matriz \mathbf{X} de tamaño $(i \times j)$ y su traspuesta \mathbf{X}^T se define como:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1j} \\ x_{21} & x_{22} & \dots & x_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ x_{i1} & x_{i2} & \dots & x_{ij} \end{bmatrix} \quad (\text{A.4})$$

$$\mathbf{X}^T = \begin{bmatrix} x_{11} & x_{21} & \dots & x_{i1} \\ x_{12} & x_{22} & \dots & x_{i2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1j} & x_{2j} & \dots & x_{ij} \end{bmatrix} \quad (\text{A.3})$$

Una matriz \mathbf{X} de tamaño $(d \times d)$ es una matriz cuadrada. Si el elemento $x_{ij} = x_{ji}$ la matriz es simétrica, si los elementos $x_{ij} = -x_{ji}$ es una matriz antisimétrica.

Una matriz identidad \mathbf{I} es aquella matriz cuadrada cuya diagonal tiene valores de unos.

$$\mathbf{X} = \begin{bmatrix} 1 & x_{12} & \dots & x_{1j} \\ x_{21} & 1 & \dots & x_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ x_{i1} & x_{i2} & \dots & 1 \end{bmatrix} \quad (\text{A.4})$$

A.1 Producto interno de dos vectores

El producto interno o producto punto de dos vectores \mathbf{x} y \mathbf{y} se denota:

$$\mathbf{x}^T \mathbf{y} = \sum_{i=1}^d x_i y_i = \mathbf{y}^T \mathbf{x} \quad (\text{A.5})$$

El producto interno es útil en el cálculo de la o similitud entre vectores, usualmente se utiliza en el cálculo de la norma o el ángulo entre dos vectores:

$$\|\mathbf{x}\| = \sqrt[2]{\mathbf{x}^T \mathbf{x}} \quad (\text{A.6})$$

$$\cos \theta = \frac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|} \quad (\text{A.7})$$

A.2 Producto de un escalar y un vector

El producto de un escalar por un vector es la multiplicación de cada elemento del vector por el escalar, el resultado es un nuevo vector:

$$a\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} = \begin{bmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_d \end{bmatrix} \quad (\text{A.8})$$

A.3 Vector ortogonal

Dos vectores son ortogonales si su producto interno es igual a cero. En un espacio dimensional significa que estos vectores son perpendiculares y el ángulo θ entre ambos es de 90° .

$$\mathbf{x}^T \mathbf{y} = \sum_{i=1}^d x_i y_i = 0 \quad (\text{A.9})$$

A.4 Vector normal

Es aquel vector cuya norma es igual a la unidad. Un vector cuya norma inicialmente es mayor a 1, puede ser normalizado de la siguiente manera:

$$\mathbf{Y} = \frac{\mathbf{x}}{\|\mathbf{x}\|} \quad (\text{A.10})$$

A.5 Matriz ortogonal

Una matriz \mathbf{M} es ortogonal si

$$\mathbf{MM}^T = \mathbf{M}^T\mathbf{M} = \mathbf{I} \quad (\text{A.11})$$

A.6 Determinante de una matriz

La determinante de una matriz es una función que reduce la matriz a un número escalar. La determinante de una matriz se denota $|\mathbf{M}|$ o $\det(\mathbf{M})$. La determinante de una matriz \mathbf{M} de (2×2) es:

$$|\mathbf{M}| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \quad (\text{A.12})$$

La determinante de una matriz de $(d \times d)$ donde $d > 2$, se obtiene eliminando de manera recursiva las filas y columnas para crear matrices de (2×2) y luego se aplica el proceso analizado previamente, finalmente se restan los determinantes.

A.7 Producto matricial

Para realizar el producto entre dos matrices \mathbf{A} y \mathbf{B} , se considera la siguiente restricción: el número de filas de \mathbf{B} matriz debe ser igual al número de columnas de \mathbf{A} . El producto \mathbf{C} resultante tendrá tantas filas como \mathbf{A} y tantas columnas como \mathbf{B} . Los elementos de la nueva matriz se hallan aplicando el producto interno entre cada fila de \mathbf{A} y la columna de \mathbf{B} .

A.8 Eigenvectores y eigenvalores

Considerando una matriz \mathbf{M} de tamaño ($d \times d$), es posible calcular los eigenvectores y eigenvalores mediante la siguiente ecuación:

$$\mathbf{M}\mathbf{x} = \lambda\mathbf{x} \quad (\text{A.13})$$

donde λ es el eigenvector y \mathbf{x} es el eigenvalor de \mathbf{M} . Esta ecuación puede ser reescrita como:

$$(\mathbf{M} - \lambda\mathbf{I})\mathbf{x} = 0 \quad (\text{A.14})$$

donde \mathbf{I} es una matriz identidad.

Una de las maneras de resolver esta ecuación es a partir de la siguiente ecuación característica:

$$(\mathbf{M} - \lambda\mathbf{I}) = \lambda^d + a_1\lambda^{d-1} + \cdots + a_{d-1}\lambda + a_d = 0 \quad (\text{A.15})$$

Para cada d raíz de λ_j se calcula un conjunto de ecuaciones lineales. Finalmente la traza de la matriz \mathbf{M} es la suma de los eigenvalores y la determinante de esta matriz es el producto de estos

$$\text{tr}[\mathbf{M}] = \sum_{i=1}^d \lambda_i \quad (\text{A.16})$$

$$|\mathbf{M}| = \prod_{i=1}^d \lambda_i \quad (\text{A.17})$$

Si la matriz \mathbf{M} es una matriz diagonal, los eigenvalores son elementos no ceros de la diagonal principal y los eigenvectores es el vector unidad paralelo a los ejes coordenados.

Apendice B. Procesamiento digital de imágenes

Una imagen está compuesta por una componente que depende de la iluminación que incide sobre la escena y otra relacionada a la reflectancia que corresponde a la luz reflejada por los objetos. Ambas componentes determinan la intensidad que se cuantifica en los fotodetectores los cuales establecen los niveles de intensidad de los niveles de gris en una imagen digital. La intensidad es detectada por un dispositivo fotosensible cuya respuesta depende de la longitud de onda del espectro electromagnético. Entre los fotodetectores destacan los denominados Dispositivos de Acoplamiento de Carga (Charged Couple Devices . CCD), los cuales son construidos disponiendo en forma de matriz los elementos señalados. Cada uno de estos elementos genera un valor que representa la intensidad de la imagen en ese punto, denominado píxel (picture element).

La disposición matricial de los elementos fotosensibles genera una matriz de valores que representan intensidades de la escena 3D formando lo que se conoce como imagen digital. La imagen digital es entonces una estructura de datos representada como una matriz de $M \times N$ elementos, donde cada elemento representa a un píxel.

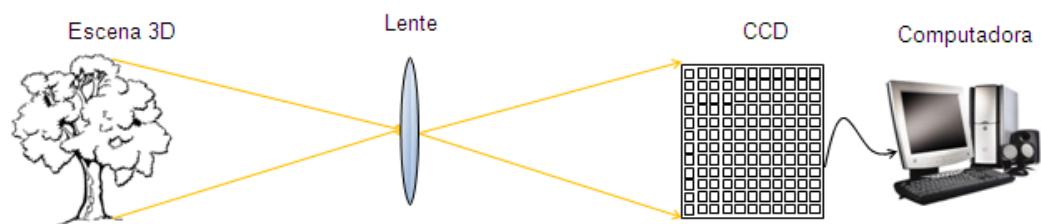


Figura B.1 Imagen digital registrada por un dispositivo CCD (Charged Couple Devices).

Así pues, el término imagen digital se refiere a una función bidimensional $f(x, y)$ con M filas y N columnas, donde cada elemento corresponde a un valor de intensidad cuantizada.

La Figura B.2 muestra la convención de ejes generalmente utilizadas para representar una imagen digital.

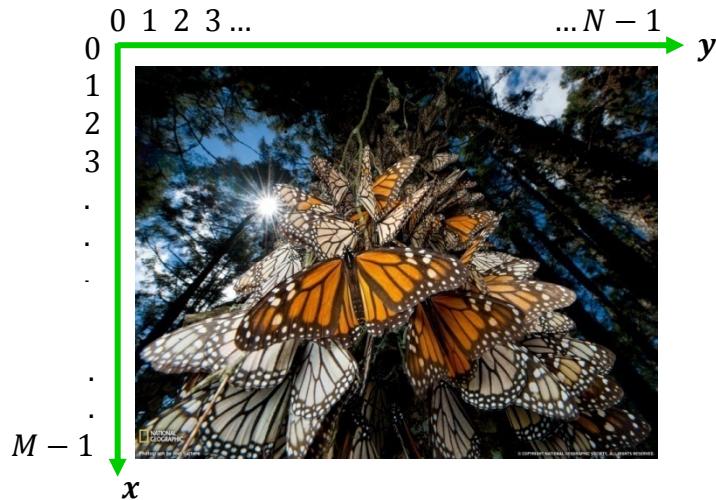


Figura B.2. Coordenada convencional de los píxeles en una imagen digital

B.1 Resolución de una imagen

La resolución de una imagen digital es la cantidad de píxel por unidad de área. Suele medirse en términos de "píxel por pulgada" (ppi) y de ella depende tanto la calidad de la representación como el tamaño que ocupa en memoria el archivo gráfico. La Figura B.3 corresponde a una imagen con distintos niveles de resolución.

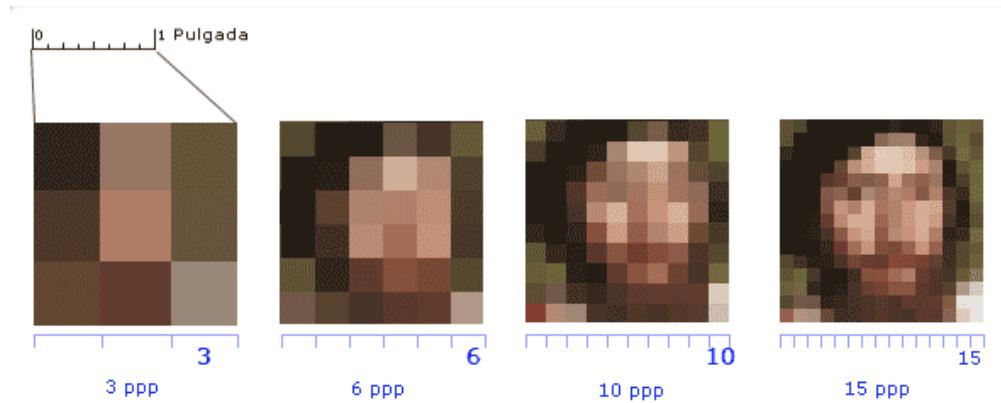


Figura B.3. Resolución por píxeles por pulgada

B.2 Profundidad de color

La profundidad de color o profundidad de bits, es el número de bits necesarios para representar el color de un píxel de una imagen digital. Por ejemplo, un píxel con profundidad 1 bit tiene dos valores posibles 0 o 1, por lo tanto, puede representar dos tonos: negro y blanco; un píxel con profundidad 8 tiene 2^8 tonalidades y así sucesivamente.

B.3 Frecuencia espacial de una imagen digital

Si exploramos el perfil de una imagen digital, por ejemplo la primera fila $I(x, 0)$, podemos encontrar que ciertos valores de I pueden repetirse más de una vez, esta comportamiento se le llama frecuencia espacial. Las zonas con pequeños cambios o con transiciones graduales se denominan áreas de bajas frecuencias, mientras que las regiones con cambios bruscos y rápidos, se le conocen como zonas de altas frecuencias (por ejemplo, los bordes o los ruidos).

Con base a lo anterior una imagen puede ser modelada a partir de funciones armónicas usando técnicas de análisis de Fourier, es decir podemos transformar la información de una línea de exploración en una serie de funciones cosenoide de amplitud y frecuencia apropiada. Por otro lado, el filtraje espacial de imágenes, básicamente, consistente en la eliminación de ciertas frecuencias presentes en una imagen, éstos pueden ser de alguno de los siguientes tipos:

- 1) *Filtros pasa bajos*: enfatizan las bajas frecuencias, suavizando las imágenes y atenuando los datos ruidos. En cuyo caso el nivel de gris de cada píxel se calculará en base al valor de los niveles de gris de los píxeles vecinos, reduciendo la variabilidad espacial de la imagen. Ello produce un borroneado de los bordes, perdiéndose la nitidez de la imagen, pero ganando homogeneidad.
- 2) *Filtros pasa altos*: enfatizan las altas frecuencias, para mejorar o afilar las características lineales como los bordes o contornos de los objetos. Realizan por tanto el efecto contrario a los filtros pasa bajos, eliminando las bajas frecuencias.

B.4 Filtro pasa bajo ideal

El filtro pasa bajos ideal es aquel que suprime todas las altas frecuencias en la Trasformada de Fourier que están a una distancia mayor de D_0 , que se encuentra definido desde el centro de la Trasformada hacia la periferia de la imagen. Un filtro pasa bajos ideal está definido por la siguiente función:

$$H = \begin{cases} 1 & \text{si } D(u, v) \leq D_0 \\ 0 & \text{si } D(u, v) > D_0 \end{cases} \quad (\text{B.1})$$

Donde D_0 es la distancia desde el centro de la transformada de Fourier a un punto a un punto distante del centro.

B.5 Filtro pasa bajo Butterworth

La función de trasferencia del filtro pasa bajos Butterworth es de orden n y suprime las frecuencias que se encuentran más allá de una distancia D_0 medida desde el origen de la trasformada de Fourier. El filtro se define con la siguiente ecuación.

$$H(u, v) = \frac{1}{1 + [D(u, v)/D_0]^{2n}} \quad (\text{B.2})$$

Las distancias de los puntos $D(u, v)$ se pueden calcular con la ecuación de la distancia euclíadiana entre dos puntos:

$$D(u, v) = \sqrt{(u - M/2)^2 + (v - N/2)^2} \quad (\text{B.3})$$

B.6 Transformada de Fourier

La transformada de Fourier de una función continua se define con la siguiente ecuación:

$$F(u) = \int_{-\infty}^{\infty} f(x)e^{-2i\pi ux} dx \quad (\text{B.4})$$

La ecuación anterior nos permite pasar una función del dominio espacial al dominio de la frecuencia. La transformada inversa nos permite regresar al dominio espacial y se define con la siguiente ecuación.

$$f(x) = \int_{-\infty}^{\infty} F(u)e^{2i\pi ux} du \quad (\text{B.5})$$

La transformada de Fourier para funciones continuas de dos variables independientes x y y se define con la ecuación (B.6). La transformada inversa de Fourier para dos variables se define con la ecuación (B.7).

$$F(u, v) = \iint_{-\infty}^{\infty} f(x, y)e^{-2i\pi(ux+vy)} dx dy \quad (\text{B.6})$$

$$f(x, y) = \iint_{-\infty}^{\infty} F(u, v)e^{2i\pi(ux+vy)} du dv \quad (\text{B.7})$$

B.7 Transformada Discreta de Fourier

La transformada discreta de Fourier para una función unidimensional $f(x)$ se define con la siguiente ecuación:

$$F(u) = \sum_{x=0}^{M-1} f(x)e^{-2i\pi ux} dx \quad (\text{B.8})$$

dónde $f(x)$ es una función discreta y $F(u)$ es la transformada discreta de Fourier.

La Trasformada inversa es:

$$f(x) = \sum_{u=0}^{M-1} F(u)e^{2i\pi ux} dx \quad (\text{B.9})$$

La transformada discreta de la función bidimensional $f(x, y)$ se define con la ecuación (B.10). La transformada inversa es definida por la ecuación (B.11).

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-2i\pi(ux/M + vy/m)} dx dy \quad (\text{B.10})$$

$$I(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{2i\pi(ux/M + vy/m)} du dv \quad (\text{B.11})$$

Apendice C Análisis Multivariante

Un evento aleatorio multivariado consiste en m mediciones con n variables, donde $n > 2$. La cantidad de variables del evento determina la dimensión un evento [44]. Los datos generados en un evento multivariado pueden ser representados en una matriz de m filas y n columnas. Donde m son las mediciones realizadas y n son las variables del evento.

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m,1} & x_{m,2} & \cdots & x_{m,n} \end{bmatrix} \quad (\text{C.1})$$

La reducción de variables es posible debido a que en muchos de los casos la variabilidad de los datos puede ser representado por k componentes principales, en vez de n variables ($k \ll n$). Es mucho más factible buscar información relevante en k componentes que buscar la misma información en n variables.

Las técnicas de análisis multivariante se clasifica en dos grupos: técnicas de dependencia y técnicas de interdependencia [45]. Las técnicas de dependencia involucra un conjunto de variables que son explicadas a partir del comportamiento de otras variables denominadas independientes, algunos de los métodos que se encuentran dentro de esta clasificación son: regresión múltiple, análisis discriminante múltiple, análisis conjunto, análisis de correlaciones canónicas, y modelos de ecuaciones estructurales.

Las técnicas de interdependencia usa el análisis simultáneo de todas las variables sin distinción entre dependientes e independientes. Los métodos dentro de esta clasificación se encuentran: análisis de componentes principales, análisis factorial, análisis de conglomerados, escalamiento multidimensional y análisis de correspondencia

La técnica de análisis multivariante que es de interés en este trabajo de tesis es el *Análisis de Componentes Principales* (ACP), clasificado como una técnica de interdependencia.

C.1 Análisis de componentes principales

El análisis de componentes principales, PCA por sus siglas en inglés, es una técnica estadística introducida por Karl Pearson en 1901, sin embargo debido a la complejidad del cálculo no fue usada ampliamente hasta la presencia de ordenadores en los principales laboratorios de investigación [46]. El método estadístico de PCA permite explicar un evento multivariado con un número reducido de nuevas variables que se les llaman *componentes principales* (CP). Los CP se obtienen a partir de la matriz de covarianza de los datos, éstos forman una base ortonormal, no existe correlación entre estos y están ordenados de manera que los primeros CP representan la máxima varianza de las variables originales. Suponiendo que tenemos un evento con p variables y n observaciones, mediante el análisis de componentes principales es posible explicar este evento con un nuevo conjunto de k componentes principales, siendo $k < p$ [47].

El porcentaje de variabilidad que recoge cada componente principal suele ser muy distinto. Si entre las variables existe una fuerte correlación es posible describir los datos con unos pocos componentes principales; por el contrario, correlaciones pequeñas indican de poca información redundante, por lo tanto es necesario considerar la mayor cantidad posible de componentes principales para explicar el evento.

C.2 Derivación de componentes principales

Suponiendo que tenemos una matriz \mathbf{X} de tamaño $(m \times n)$ donde $\mathbf{X} \in \mathcal{R}^{m \times n}$, m es el número de mediciones y n es el número de variables,

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{m,1} & x_{m,2} & \cdots & x_{m,n} \end{bmatrix} \quad (\text{C.2})$$

es posible modelar \mathbf{X} de la siguiente forma:

$$\mathbf{X} = \mathbf{T}\mathbf{P}' \quad (\text{B.3})$$

Donde \mathbf{P} es la matriz traspuesta de los eigenvectores de la matriz de covarianza de \mathbf{X} y \mathbf{T} es la proyección de la matriz \mathbf{X} sobre los eigenvectores:

$$\mathbf{T} = \mathbf{X}\mathbf{P} \quad (\text{B.4})$$

La matriz $\mathbf{P} \in \mathcal{R}^{n \times r}$ mientras que $\mathbf{T} \in \mathcal{R}^{m \times r}$. La matriz \mathbf{P} está ordenada de manera que la primera columna representa la máxima variabilidad, la segunda columna corresponde a la segunda variabilidad más alta y así sucesivamente hasta la última columna que representa la variabilidad más baja de los datos.

La matriz de datos \mathbf{P} se obtiene a partir de la combinación lineal de la matriz de covarianza de las variables. La matriz de covarianza con la siguiente ecuación:

$$\Sigma = \frac{\mathbf{\tilde{X}}'\mathbf{\tilde{X}}}{(m - 1)} \quad (\text{C.5})$$

donde $\mathbf{\tilde{X}}$ es la son los datos normalizados de \mathbf{X} respecto a su valor promedio. Se puede comprobar que:

$$\mathbf{P}\Lambda\mathbf{P}' = \Sigma \quad (\text{C.6})$$

Donde Λ es una matriz diagonal que corresponde a los eigenvalores de la matriz de covarianza..

$$\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_r, \dots, 0, \dots 0) \quad (\text{C.7})$$

$$\lambda_1 > \lambda_2 > \lambda_r \geq 0$$

Los eigenvalores son indicadores de la variación de los eigenvectores, los que tiene un valor alto representa la máxima variación de los datos.

El cálculo de \mathbf{T} se realiza con la siguiente ecuación:

$$\mathbf{T}_k = \mathbf{X}\mathbf{P}_k \quad (\text{B.8})$$

La proporción de la varianza total de \mathbf{T}_k está dada por $v_k/vr * 100$, donde

$$v_k = \sum_{i=1}^k \lambda_i \quad y \quad v_r = \sum_{i=1}^r \lambda_i \quad (\text{B.9})$$

Bibliografía

- [1] J. W. Kane y M. M. Sternheim, Física, Barcelona: Reverte, 2007.
- [2] S.-T. Bow, Pattern recognition: Application to large data-set problems., Pennsylvania: Electrocal engineering and electronics, 1984.
- [3] M. Turk y A. Petland, «Eigenfaces for recognition,» *Journal of Cognitive Neuroscience*, vol. 3, nº 1, pp. 71-86, 1991.
- [4] E. Cabello Pardos, «Técnicas de reconocimiento facial mediante redes neuronales,» Departamento de tecnología fotonica, facultad de informática., Madrid, 2004.
- [5] X. Chenghua, W. Yunhong, T. Tieniu y Q. Long, «A new attempt to face recognition using 3D-eigenfaces,» de *Asian Conference on Computer Vision ACCV*, Jeju Island, Korea, 2004.
- [6] S. Hwanjong, Y. Ukil, L. Sangyoun y S. Kwanghoon, «3D Face Recognition Based on Facial Shape Indexes with Dynamic Programming,» de *International Conference, ICB 2006*, Hong Kong, China, 2006.
- [7] Y. Xue, L. Jianming y Y. Takashi, «A method of 3D face recognition based on principal component analysis algorithm,» de *Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on*, 2005.
- [8] C. Zhang, «A survey of recent advances in face detection,» Microsoft corporation, 2010.
- [9] K. Saeed y T. Nagashima, Biometrics and Kansei Engineering, New York: Springer, 2012.
- [10] A. K. Jain, P. Flynn y A. A. Ross, Handbook of biometrics, New York: Springer, 2008.
- [11] W. Zhao, R. Chellappa, P. J. Phillips y A. Rosenfeld, «Face Recognition: A Literature Survey,» *ACM Computing Surveys*, vol. 35, nº 4, pp. 399-458, 2003.
- [12] J. Wayman, Introduction to biometrics, New York: Springer, 2011.
- [13] M. Narasimha Murty y V. Susheela Devi, Pattern Recognition. An Algorithmic Approach, Springer, 2011.
- [14] R. C. Gonzalez y R. E. Woods, Digital image processing, Ney Jersey: Pearson Pretince Hall, 2008.
- [15] J. P. Marques de Sá, Pattern recognition. Concepts, methods and applications, Oporto, Portugal: Springer, 2001.

- [16] R. O. Duda, P. E. Hart y D. G. Stork, *Pattern clasification*, A Wiley International Publication, 2001.
- [17] H. C. Andrews, *Introduction to mathematical techniques in pattern recognition*, Canada: Wiley-Interscience, 1972.
- [18] G. Sánchez Torres y J. William Branch, «Revisión general de los modelos de reconstrucción tridimensionales desde imágenes de rango,» *Revista Avances en Sistemas e Informática*, vol. 5, nº 2, pp. 23-31, 2008.
- [19] M. Mazaheri y M. Momeni, «3d modeling using structured light pattern and photogrammetric epipolar geometry,» *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.*, vol. XXXVII, nº Part B3b, pp. 87-90, 2008.
- [20] G. Pajares Martinsanz y M. Santos Peñas, *Inteligencia Artificial e ingeniería del conocimiento*, Madrid, España: Alfaomega, 2006.
- [21] B. Jähne y H. Haussecker, *Computer Vision and Applications: A Guide for Students and Practitioners*, Academic Press, 2000.
- [22] M. Quirós Hernández, *Tecnologías de la Información Geográfica (TIG): cartografía, fotointerpretación, teledetección y SIG*, Salamanca, España: Universidad de Salamanca, 2011.
- [23] G. Pajares y J. M. de la Cruz, *Visión por computador*, Madrid, España: Alfaomega, 2002.
- [24] R. Klette, K. Schluns y A. Koschan, *Computer Vision. Three-Dimensional Data from Images*, Springer, 1998.
- [25] S. Siva Gorthi y P. Rastogi, «Fringe Projection Techniques: Whither we are?,» *Optics and Lasers in Engiering*, vol. 48, nº 2, pp. 133-140, 2009.
- [26] Y. Fu y Q. Luo, «Fringe projection profilometry based on a novel phase shift method,» *OPTICS EXPRESS*, vol. 19, nº 22, 2011.
- [27] D. M. Quintero Oliveros, R. Villamizar Mejía y E. Yesith Mendoza, «Reconstrucción tridimensional de deformaciones mecánicas superficiales en platinas, basada en la proyección de franjas,» *Revista de la Facultad de Ingenieria. Universidad de Antioquia*, nº 59, 2011.
- [28] T. Mitsuo, H. Ina y S. Kobayashi, «Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry.,» *J. Opt. Soc. Am.*, vol. 72, nº 1, pp. 156-160, 1982.
- [29] F. J. Cuevas de la Rosa, «Aplicación de redes neuronales en la reconstrucción tridimensional de objetos,» Tesis, Centro de Investigaciones en Óptica, León,

Guanajuato, México, 2000.

- [30] M. Cerca, B. Barrientos García, J. García Márquez y C. Hernandez Bernal, «Obtención de relieve digital mediante proyección de luz estructurada en modelos analógicos de extensión,» *Boletín de la Sociedad Geológica Mexicana*, vol. LIX, nº 1, pp. 101-113, 2007.
- [31] Y. Gong y S. Zhang, «Ultrafast 3-D shape measurement with an off-the-shelf DLP projector,» *OPTICS EXPRESS*, vol. 18, nº 19, 2010.
- [32] Y. Hu, J. Xi, J. Chicharo y Z. Yang, «Improved three-step phase shifting profilometry using digital fringe pattern projection,» de *Computer Graphics, Imaging and Visualisation, 2006 International Conference on*, 2006.
- [33] P. S. Huang, Q. Hu, . F. Jin y F.-P. Chiang, «Color-encoded digital fringe projection technique for high-speed three-dimensional surface contouring,» *Optical Engineering*, vol. 38, nº 6, pp. 1065-1071, 1999.
- [34] J.-i. Kato, «Fringe Analysis,» de *Handbook of optical metrology: principles and applications*, Toru Yoshizawa, 2009, pp. 449-509.
- [35] L. Kinell, «Multichannel method for absolute shape measurement using projected fringes,» *Optics and Lasers in Engineering*, vol. 41, nº 1, pp. 57 - 71, 2004.
- [36] S. A. Karout, M. A. Gdeisat, D. R. Burton y M. J. Lalor, «Residue vector, an approach to branch-cut placement in phase unwrapping: theoretical study,» *APPLIED OPTICS*, vol. 46, nº 21, pp. 4212 - 4727, 2007.
- [37] M. Arevalillo Herráez, D. R. Burton, M. J. Lalor y M. A. Gdeisat, «Fast two-dimensional phase-unwrapping algorithm based on sorting by reliability following a noncontinuous path,» *APPLIED OPTICS*, vol. 41, nº 35, pp. 7437 - 7444, 2002.
- [38] D. C. Ghiglia y M. D. Pritt, *Two-dimensional phase unwrapping: theory, algorithms, and software*, John Wiley & Sons, INC, 1998.
- [39] P. J. Besl y N. D. McKay, «A method for registration of 3-D shapes,» *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 14, nº 2, pp. 39-256, 1992.
- [40] B. Ben Amor, M. Ardabilian y L. Chen, «Enhancing 3D Face Recognition By Mimics Segmentation,» *Proceedings of the Sixth International Conference on Intelligent Systems Design and Applications*, vol. 3, pp. 150 - 155 , 2006.
- [41] A. Colombo, C. Cusano y R. Schettini, «3D face detection using curvature analysis,» *Pattern Recognition. The Journal of the Pattern Recognition Society*, vol. 39, nº 3, pp. 445 - 455, 2006.
- [42] M. Central, «Matlab Central,» 30 05 2010. [En línea]. Available: <http://www.mathworks.com/matlabcentral/fileexchange/27804-iterative-closest-point>.

[Último acceso: 3 05 2013].

- [43] MathWorks, «Documentation center. Discriminant analysis,» MathWorks, 2013. [En línea]. Available: <http://www.mathworks.com/help/stats/classify.html>. [Último acceso: 12 5 2013].
- [44] O. Hernández Rodríguez, Temas de análisis estadístico multivariado, San José: Editorial de la Universidad de Costa Rica, 1998.
- [45] A. Rial Boubeta y M. J. Varela, Estadística práctica para la investigación de ciencias de la salud, La Coruña: Netbiblo, 2008.
- [46] R. Serrano Gallego, Introducción al análisis de datos experimentales: tratamiento de datos en bioensayos, Universitat Jaume, 2003.
- [47] R. Álvarez Cáceres, Estadística multivariante y no paramétrica con SPSS. Aplicaciones a las ciencias de la salud., Madrid, España: Díaz de Santos S.A, 1995.
- [48] D. Tulsiani, «A Fringe Projection System for Measurement of Condensing Fluid Films in Reduced Gravity,» 2005.
- [49] L. Ling, «Phase unwrapping,» de *Wiley Encyclopedia of Biomedical Engineering*, Milwaukee, Wisconsin, John Wiley & Sons, Inc., 2006, pp. 1-12.
- [50] A. Bivaldi, «Two-dimensional phase unwrapping by quad-tree decomposition,» *APPLIED OPTICS*, vol. 40, nº 8, pp. 1187 - 1194, 2001.
- [51] G. Cong y X. Jiang, «Fully automatic face recognition framework based,» *Machine Vision and Applications* , vol. 24, nº 3, pp. 537-549, 2012.
- [52] A. R. Arun, N. Karthik y K. J. Anil, Handbook of multibiometrics, New York: Springer, 2006.