

**Centro de Investigación Científica y de Educación  
Superior de Ensenada, Baja California**



---

**Programa de Posgrado en Ciencias  
en Ciencias de la Computación**

---

**Métodos de reconocimiento de rostros 3D basados en  
correspondencia de superficies faciales**

Tesis

para cubrir parcialmente los requisitos necesarios para obtener el grado de  
Maestro en Ciencias

Presenta:

**Beatriz Adriana Echeagaray Patrón**

Ensenada, Baja California, México

2016

Tesis defendida por

# **Beatriz Adriana Echeagaray Patrón**

y aprobada por el siguiente Comité

---

Dr. Vitaly Kober

Director del Comité

Dr. Hugo Homero Hidalgo Silva

Dr. Josué Álvarez Borrego



---

Dr. Jesús Favela Vara

Coordinador del Programa de Posgrado en Ciencias de la Computación

---

Dra. Rufina Hernández Martínez

Director de Estudios de Posgrado

Resumen de la tesis que presenta Beatriz Adriana Echeagaray Patrón como requisito parcial para la obtención del grado de Maestro en Ciencias en Ciencias de la Computación.

## **Métodos de reconocimiento de rostros 3D basados en correspondencia de superficies faciales**

Resumen aprobado por:

---

Dr. Vitaly Kober

Director de Tesis

El reconocimiento de rostros es un problema que ha mantenido su relevancia dentro del Reconocimiento de Patrones en los últimos años debido a su gran variedad de aplicaciones, tales como control de acceso, vigilancia, interacción humano-computadora y la creación de identificaciones biométricas en general. Una de las desventajas más importantes de tecnologías actuales en este campo es que los dispositivos usuales de captura de imágenes trabajan con representaciones bidimensionales de los rostros, lo cual proporciona sensibilidad ante variaciones de la imagen de un mismo individuo, principalmente ante cambios de iluminación, expresiones y variaciones de postura. Para hacer frente a estas múltiples variaciones, el presente trabajo de tesis propone un método de reconocimiento de rostros que combina cualidades de distintos enfoques en esta área: métodos basados en modelos estadísticos, métodos de reconocimiento basados en regiones y métodos basados en transformaciones. Con el fin de aprovechar la información geométrica de los rostros, el método parte de nubes de puntos 3D a las cuales se les aplica un análisis de curvatura para obtener descriptores de la forma local del rostro en cada punto; luego, se aplica una transformación del rostro 3D original hacia una malla triangulada en el plano y, finalmente, se consideran segmentaciones que extraen regiones robustas a expresiones.

Palabras Clave: **Reconocimiento de rostros 3D, análisis de curvatura, parametrizaciones conformes.**

Abstract of the thesis presented by Beatriz Adriana Echeagaray Patrón as a partial requirement to obtain the Master of Science degree in Computer Science.

### **3D face recognition methods based on matching on facial surfaces**

Abstract approved by:

---

Dr. Vitaly Kober  
Thesis Director

Face recognition is an important problem that has maintained its relevance in Pattern Recognition in recent years due to its wide range of applications such as access control, surveillance, human-computer interaction, and biometric identification. One of the major drawbacks of current technologies in this field is that the usual imaging devices work with two-dimensional representations of faces. Hence, such methods are sensitive to changes in images of faces, especially to changes in illumination, expressions and posture. To address these multiple variations, this thesis proposes a face recognition method that combines advantages of different approaches in this area; that is, recognition methods based on statistical models, recognition methods based on regions, and recognition methods based on transformations. To take advantage of the geometric information of faces, the proposed method begins with a curvature analysis of the original 3D point clouds in order to obtain local descriptors of the face shape at each point; then, a transformation of the original 3D face is applied to the triangulated mesh in the plane; finally, almost expression-invariant regions are segmented.

**Keywords: 3D face recognition, curvature analysis, conformal parameterization.**

## Dedicatoria

*A mi familia  
y a Luis Bojórquez.*

## Agradecimientos

A mi familia por su amor y apoyo incondicional en cada etapa de mi vida. Todo lo que soy se los debo a ustedes.

A Luis Bojórquez por su compañía y comprensión. Gracias por compartirme tu perspectiva de la vida, por motivarme e inspirarme todo este tiempo.

Al Dr. Vitaly Kober por su ejemplo como investigador, sus enseñanzas y paciencia.

A los miembros de mi comité de tesis, Dr. Hugo Hidalgo y Dr. Josué Álvarez Borrego, por sus consejos y valiosos comentarios.

A todos los compañeros del Departamento de Ciencias de la Computación. Gracias por enseñarme más de la vida, la comida, el café, el vino y la cerveza.

Al Centro de Investigación Científica y de Educación Superior de Ensenada por darme la oportunidad de realizar mis estudios de posgrado.

Al Consejo Nacional de Ciencia y Tecnología por brindarme el apoyo económico para realizar mis estudios de maestría.

*A los que inspiran, a quienes apoyan,  
a quienes acompañan, a quienes motivan.*

*Gracias.*

*Adriana Echeagaray*

# Tabla de contenido

Página

<b>Resumen en español</b>	<b>ii</b>
<b>Resumen en inglés</b>	<b>iii</b>
<b>Dedicatoria</b>	<b>iv</b>
<b>Agradecimientos</b>	<b>v</b>
<b>Lista de figuras</b>	<b>viii</b>
<b>Lista de tablas</b>	<b>x</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Representaciones faciales . . . . .	2
1.2. Problemática del reconocimiento facial tridimensional . . . . .	3
1.3. Investigación previa . . . . .	6
1.3.1. Métodos estadísticos . . . . .	6
1.3.2. Métodos basados en regiones . . . . .	8
1.3.3. Métodos basados en transformaciones . . . . .	9
1.4. Objetivos . . . . .	11
1.4.1. Objetivo general . . . . .	11
1.4.2. Objetivos específicos . . . . .	11
1.5. Limitaciones y suposiciones . . . . .	11
1.6. Organización de la tesis . . . . .	12
<b>2. Fundamentos</b>	<b>13</b>
2.1. Eigenvectores y eigenvalores de una matriz . . . . .	13
2.2. Conceptos básicos de análisis complejo . . . . .	13
2.3. Conceptos básicos en topología y geometría diferencial . . . . .	14
2.3.1. Conceptos sobre variedades en espacios euclídeos. . . . .	15
2.3.2. Curvatura local de superficies . . . . .	16
2.3.2.1. Clasificación HK . . . . .	17
2.3.2.2. Índice de forma . . . . .	18
2.3.2.3. Formas fundamentales de superficies . . . . .	18
2.3.2.4. Discretización del cálculo de curvaturas . . . . .	19
2.4. Triangulación de Delaunay . . . . .	20
<b>3. Parametrización de superficies</b>	<b>23</b>
3.1. Transformaciones conformes . . . . .	23
3.1.1. Mapeos conformes por mínimos cuadrados . . . . .	24
3.1.2. Energía de Dirichlet y energía conforme . . . . .	25
3.2. Parametrización espectral conforme . . . . .	26
3.2.1. Notación . . . . .	26
3.2.2. Análisis espectral de forma . . . . .	27
3.2.3. Discretización del mapeo . . . . .	29
3.2.4. Solución a la parametrización . . . . .	30
3.3. Resumen y comparativa de parametrizaciones conformes . . . . .	31

<b>4.</b>	<b>Métodos clásicos para la correspondencia de superficies faciales</b>	<b>33</b>
4.1.	Análisis de Componentes Principales . . . . .	33
4.2.	Algoritmo iterativo del punto más cercano . . . . .	36
4.3.	Evaluación en el reconocimiento de rostros . . . . .	38
4.3.1.	Medidas de similitud . . . . .	39
4.3.2.	Curva característica de correspondencia acumulada . . . . .	39
4.4.	Resumen . . . . .	40
<b>5.</b>	<b>Metodología propuesta</b>	<b>41</b>
5.1.	Preprocesamiento . . . . .	42
5.2.	Corrección de postura . . . . .	42
5.3.	Segmentación . . . . .	42
5.4.	Análisis de curvatura . . . . .	43
5.4.1.	Localización automática de puntos . . . . .	43
5.5.	Parametrización conforme . . . . .	44
5.5.1.	Normalización . . . . .	44
5.6.	Correspondencia . . . . .	45
5.6.1.	Segmentación de secciones robustas . . . . .	46
5.7.	Resumen . . . . .	46
<b>6.</b>	<b>Experimentos y Resultados</b>	<b>48</b>
6.1.	Base de datos . . . . .	48
6.2.	Experimento 1: Resultados de reconocimiento incluyendo la etapa de corrección de postura . . . . .	50
6.3.	Experimento 2: Reconocimiento automático de rostros con expresión neutral y comparación de resultados con distintas medidas de similitud	51
6.3.1.	Resultados con segmentación circular del rostro . . . . .	52
6.3.2.	Resultados usando la sección circular en torno de la nariz . . . . .	52
6.3.3.	Discusión de resultados . . . . .	53
6.4.	Experimento 3: Reconocimiento de rostros en presencia de expresiones.	54
6.4.1.	Comparación respecto a variaciones de iluminación . . . . .	56
6.5.	Experimento 4: Reconocimiento en presencia de variaciones de postura	57
6.6.	Resumen . . . . .	59
<b>7.</b>	<b>Conclusiones</b>	<b>60</b>
7.1.	Publicaciones en congresos . . . . .	62
7.2.	Trabajo futuro . . . . .	62
	<b>Lista de referencias bibliográficas</b>	<b>63</b>
<b>A.</b>	<b>Comparación de mapeos conformes</b>	<b>66</b>

## Lista de figuras

Figura		Página
1.	Esquema general de los métodos de reconocimiento. . . . .	2
2.	Representaciones faciales comúnmente usadas en reconocimiento de rostros: a) imagen 2D, b) imagen de rango, c) nube de puntos 3D e d) interpolación de nube de puntos. . . . .	3
3.	Ejemplos de modelos 3D de la misma persona bajo distintas condiciones de iluminación. (a) Con iluminación apropiada para el sensor. (b) Con una lámpara de estudio adicional a 1.4m de distancia. (Extraído de Bowyer <i>et al.</i> (2006)). . . . .	4
4.	Ejemplos de imperfecciones en el sensado de rostros 3D. La imagen de la izquierda muestra un ejemplo de información perdida y la de la derecha el ruido a causa del sensor. (Extraído de Bowyer <i>et al.</i> (2006)). . . . .	4
5.	Rostros de la base de datos Gavab <sup>1</sup> con variaciones de pose. (a) Modelo frontal con expresión neutra. En (b) y (c) se muestran rotaciones de la cabeza hacia la izquierda y derecha, respectivamente, junto a su vista con pose normalizada. (Extraído de Berretti <i>et al.</i> (2011)). . . . .	5
6.	Clasificación de técnicas de reconocimiento planteada en Smeets <i>et al.</i> (2012) para rostros 3D en presencia de expresiones. . . . .	6
7.	Modelos deformables para el reconocimiento facial 3D. A partir de un rostro de prueba (a) y uno neutral de la galería (b) se obtiene un modelo deformado (c). El alineamiento de (a) con los modelos (b) y (c) se muestra en (d) y (e) respectivamente. (Extraído de Lu y Jain (2008)). . . . .	7
8.	Extracción de regiones para llevar a cabo reconocimiento facial 3D a partir de un rostro neutral de la galería (a) y distintos rostros de prueba con expresiones (b-d). (Extraído de Chang <i>et al.</i> (2006)). . . . .	8
9.	Impacto de expresiones sobre distancias geodésicas en distintas zonas del rostro. (Extraído de Li <i>et al.</i> (2009)). . . . .	9
10.	Noción de curvatura en 2D. La curvatura en un punto es el inverso del radio del círculo osculante en dicho punto. . . . .	16
11.	Ejemplos representativos superficies en la escala de SI. . . . .	18
12.	La forma tridimensional de la superficie depende de la triangulación. Extraído de Devadoss y O'Rourke (2011). . . . .	21
13.	(a) Una triangulación de Delaunay de 6 puntos (rojo), sus circunferencias circunscritas y sus centros se muestran en negro. (b) Diagrama de Voronoi correspondiente a la triangulación (a). . . . .	21
14.	Transformación conforme. . . . .	24
15.	El mapa $U$ relaciona la superficie facial $\mathcal{X}$ con una triangulación en el plano $\mathcal{U}$ llevando cada triángulo de $\mathbb{R}^3$ a $\mathbb{R}^2$ . . . . .	27

Figura	Página
16. Ángulos usados para definir pesos armónicos en la parametrización. . . . .	29
17. Ejemplo de aplicar ICP sobre nubes de puntos 3D. . . . .	38
18. Esquema general del método de reconocimiento propuesto. . . . .	41
19. Resultado de la segmentación de un rostro. Cada punto contiene información de su índice de forma SI. . . . .	43
20. Normalización de la parametrización basada en la ubicación de los ojos. . . . .	45
21. Gestos de la base de datos CASIA para un mismo individuo adquiridos bajo iluminación de oficina. (a) Sonrisa (b) Risa (c) Enojo (d) Sorpresa (e) Ojos cerrados. Extraído de <a href="http://biometrics.idealtest.org/dbDetailForUser.do?id=8">http://biometrics.idealtest.org/dbDetailForUser.do?id=8</a> . . . . .	49
22. Variaciones de postura bajo iluminación de oficina contenidas en la base de datos CASIA. Extraído de <a href="http://biometrics.idealtest.org/dbDetailForUser.do?id=8">http://biometrics.idealtest.org/dbDetailForUser.do?id=8</a> . . . . .	50
23. Comparación de resultados de distancia euclidiana y coseno utilizando parametrización con frontera libre y curvatura media sobre 10 individuos. . . . .	53
24. Comparación de resultados de distancia euclidiana y coseno utilizando parametrización con segmentación circular y curvatura media sobre 10 individuos. . . . .	54
25. Comparación de resultados de reconocimiento variando la iluminación de los rostros en el conjunto de entrenamiento utilizando parametrización conforme y segmentación circular para 30 individuos. a) Entrenamiento con rostros neutrales y expresiones con iluminación desde la derecha. b) Entrenamiento con rostros neutrales y expresiones con iluminación de oficina. . . . .	57
26. Comparación de distintos métodos para mapeos conformes en un rostro con variación de pose. (Extraído de Zeng y Gu (2009)). . . . .	66
27. Comparación de distintos métodos para mapeos conformes en presencia de oclusiones. (Extraído de Zeng y Gu (2009)). . . . .	67

## Lista de tablas

Tabla		Página
1.	Clasificación de la superficie local en un punto según signos de curvaturas. . . . .	17
2.	Tabla comparativa de técnicas usuales para obtener parametrizaciones. . . . .	32
3.	Medidas de similitud comunes. . . . .	39
4.	Resultados de reconocimiento utilizando segmentación circular de la parametrización junto a curvatura media aplicando PCA y distancia de coseno. . . . .	51
5.	Resultados de reconocimiento utilizando información geométrica 3D sobre conjuntos de rostros con expresión neutra comparando distintas métricas de similitud. . . . .	51
6.	Resultados de reconocimiento utilizando parametrización conforme y curvatura media sobre conjuntos de rostros con expresión neutra comparando distintas métricas de similitud. . . . .	52
7.	Comparación de resultados de reconocimiento a partir de parametrización conforme y curvatura media sobre rostros con expresión neutra aplicando segmentación de la zona nasal y utilizando distintas métricas de similitud. . . . .	53
8.	Resultados de reconocimiento frente a expresiones de sonrisa, risa y enojo utilizando distintos conjuntos de entrenamiento. . . . .	55
9.	Resultados de reconocimiento frente a expresiones de sorpresa y ojos cerrados utilizando distintos conjuntos de entrenamiento. . . . .	56
10.	Resultados de reconocimiento frente a variaciones de postura utilizando segmentación circular de la parametrización conforme y curvatura media para 30 individuos. . . . .	58
11.	Resultados de reconocimiento frente a variaciones de postura utilizando segmentación nasal de la parametrización conforme y curvatura media para 30 individuos. . . . .	58

# Capítulo 1. Introducción

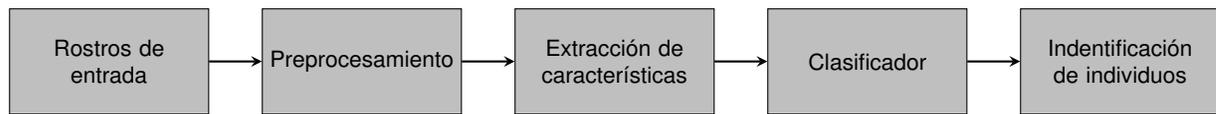
---

El estudio del reconocimiento facial es abordado actualmente en diversas áreas del conocimiento como la psicología, la medicina forense y la antropometría. Desde de la perspectiva de ciencias de la computación, el problema se aborda con el objetivo de automatizar la tarea de reconocimiento a partir de imágenes digitales capturadas mediante sensores que detectan individuos estáticos o en movimiento.

Una de las aplicaciones principales del reconocimiento facial es su uso en biometría, sistemas de reconocimiento automático de patrones que identifican a personas a partir de características físicas o de comportamiento. Entre las características consideradas en estos sistemas están las huellas digitales, el rostro, la forma de las manos y dedos, el iris, la firma y la voz. Tratándose del reconocimiento de rostros, las aplicaciones están enfocadas en la seguridad al cubrir la necesidad de localizar individuos en ambientes públicos, aduanas o aeropuertos; en vigilancia, con el objetivo de detectar intrusos en zonas privadas; en control de acceso, evitando falsificaciones de tarjetas de identificación; en múltiples interfaces hombre-computadora y en videoconferencias al localizar a un individuo en una secuencia de vídeo para poder hacer seguimiento.

Dentro del proceso de reconocimiento, inicialmente se considera una base de datos (denominada galería) y un conjunto de rostros de prueba que sirven de entrada al algoritmo. Luego, el sistema de reconocimiento puede operar de dos maneras: bajo verificación o identificación (referida como reconocimiento). La verificación es una comparación uno a uno que trata de relacionar un rostro de prueba con un rostro registrado en la galería que se sospecha es la identidad correspondiente. Por otro lado, el reconocimiento o identificación es una correspondencia de uno a muchos en la que se compara un rostro con múltiples individuos de la galería con el fin de asignar una identidad al rostro de prueba. Dentro del problema de reconocimiento también puede considerarse el caso en que algunos de los rostros de prueba no pertenecen a ningún individuo registrado en la base de datos.

En este trabajo de tesis se considera el problema de reconocimiento, el cual se aborda siguiendo el procedimiento representado en la figura 1. Los distintos enfoques en la literatura para extraer y comparar datos de superficies faciales adquieren sus respectivas ventajas



**Figura 1: Esquema general de los métodos de reconocimiento.**

y desventajas de la complejidad estimada para cada una de estas etapas. Las principales diferencias entre las investigaciones se establecen al considerar la representación facial que utilizan, la manera en que se hace frente a las variaciones del rostro de un mismo individuo y de tomar en cuenta si la información extraída se da de entrada a un análisis estadístico para reducir la dimensión de los datos.

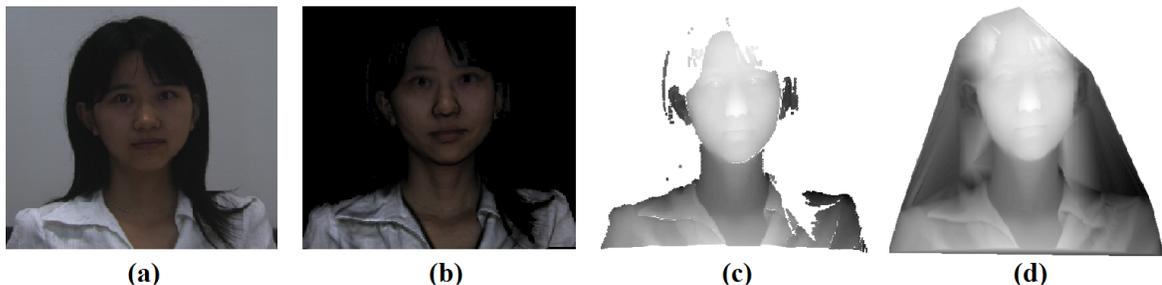
### **1.1. Representaciones faciales**

La mayoría de la investigación y sistemas comerciales de reconocimiento de rostros utiliza imágenes de intensidad, a las cuales se les referirá como imágenes 2D. Una representación que ha ganado popularidad en los últimos años son las imágenes de profundidad, también denominadas imágenes de rango, las cuales se obtienen a través de sensores 3D. Recientemente se ha considerado también un enfoque multimodal, el cual combina una imagen de intensidad (con o sin color) con la forma tridimensional de manera que se sobrepone la imagen 2D a una superficie 3D del mismo rostro considerando a la primera como información de textura sobrepuesta a la forma. Para este último enfoque, la fusión de información 2D y 3D se puede hacer en las primeras etapas del reconocimiento o en las etapas finales considerando independiente cada modalidad; con la primera alternativa se mezcla información a nivel de píxeles y con la segunda a nivel de características, por lo que se considera que la fusión temprana de ambas modalidades contiene mayor información característica de los rostros de entrada. Además, el considerar una fusión temprana de los datos permite acelerar el reconocimiento a través de la construcción de un clasificador de rechazo que descarte una gran cantidad de rostros de la galería desde las primeras etapas (Mian *et al.* (2007)).

A pesar de la gran variedad de técnicas que se han desarrollado, el problema de reconocimiento de rostros sigue abierto y aún presenta varios retos, especialmente en ambientes no controlados. La iluminación, por ejemplo, es un aspecto que usualmente no

se controla al usar interfaces hombre-computadora: la persona puede estar cerca de una ventana, una lámpara, con cierta iluminación desde el techo, etc. Sin embargo, es un factor que puede llegar a afectar drásticamente la apariencia de los objetos en general y, en consecuencia, verse reflejado en serias disminuciones en la tasa de reconocimiento. Por otro lado, considerar las múltiples variaciones que la imagen de una misma persona puede llegar a tener vuelve complicado el reconocimiento automático. Basta mencionar factores como su peinado, barba, bigote, gestos, maquillaje, rotación de la cabeza, distancia entre la cara y el sensor que se utilice, variaciones de color en imágenes del mismo individuo, la escala, posición de la fuente de iluminación, oclusiones, entre otros.

Con el fin de hacer frente a estos factores, en este trabajo se utiliza información tridimensional extraída de imágenes de profundidad. En la figura 2 se muestra cómo se pasa de una imagen de rango a una nube de puntos tridimensionales para después realizar una interpolación y obtener una superficie tridimensional sin color y con información de forma.

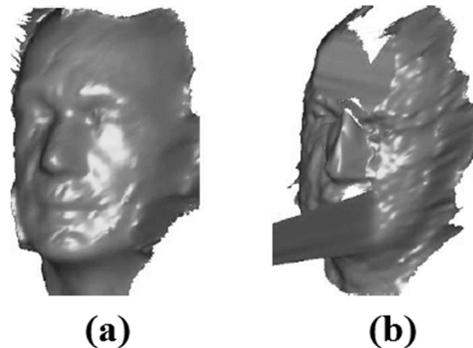


**Figura 2: Representaciones faciales comúnmente usadas en reconocimiento de rostros: a) imagen 2D, b) imagen de rango, c) nube de puntos 3D e d) interpolación de nube de puntos.**

## 1.2. Problemática del reconocimiento facial tridimensional

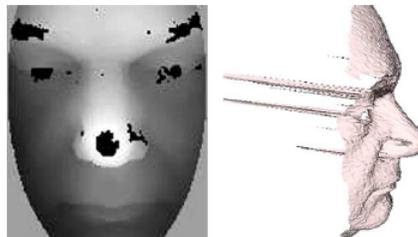
La relevancia del reconocimiento de rostros a través de superficies faciales tridimensionales radica en que la información extraída de ellas aporta un mayor poder de discriminación entre los individuos comparada con la que proviene de imágenes en dos dimensiones, además de que cubre parcialmente el problema de la dependencia a las condiciones en que se adquirieron las imágenes respecto a la iluminación y el maquillaje. Sin embargo, aunque es cierto que una forma 3D per se es independiente a la iluminación en el sentido que la forma existe sin importar cómo se ilumine, el sensado de los objetos generalmente

no es independiente a la iluminación debido a las características del sensor 3D (Bowyer *et al.* (2006)). Además, como se muestra en la figura 3, la iluminación puede introducir ruido a la información sensada.



**Figura 3: Ejemplos de modelos 3D de la misma persona bajo distintas condiciones de iluminación. (a) Con iluminación apropiada para el sensor. (b) Con una lámpara de estudio adicional a 1.4m de distancia. (Extraído de Bowyer *et al.* (2006)).**

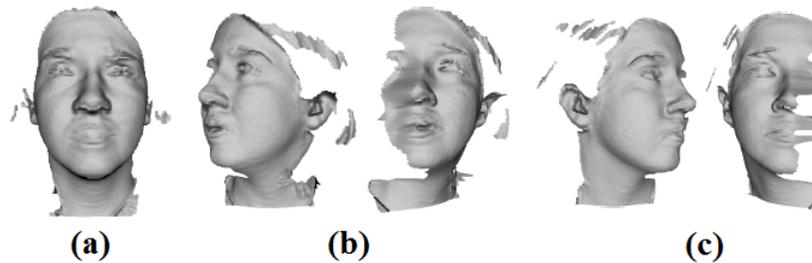
Es importante mencionar que a pesar de buscar una iluminación apropiada, los sensores 3D algunas veces proporcionan superficies que contienen ruido e incluso información faltante (véase figura 4). Estos son errores propios del sensor y no pueden ser controlados, de ahí la importancia de tener una etapa de preprocesamiento en la metodología para tratar estos errores.



**Figura 4: Ejemplos de imperfecciones en el sensado de rostros 3D. La imagen de la izquierda muestra un ejemplo de información perdida y la de la derecha el ruido a causa del sensor. (Extraído de Bowyer *et al.* (2006)).**

Por otro lado, la información tridimensional sigue estando sujeta a variaciones de postura: cuando se trabaja con sólo un sensor 3D, la superficie facial no es totalmente independiente de la pose, ya que el sensor posee un rango de visión limitado. Ésta característica reduce significativamente el rendimiento de algoritmos de reconocimiento si no se toma en cuenta que se está trabajando con una superficie parcial de los rostros

(véase figura 5). Sin embargo, el trabajar con esta representación facial tiene la ventaja de que es posible construir una etapa de normalización que corrija la orientación original de la cabeza (Berretti *et al.* (2011)).



**Figura 5: Rostros de la base de datos Gavab<sup>1</sup> con variaciones de pose. (a) Modelo frontal con expresión neutra. En (b) y (c) se muestran rotaciones de la cabeza hacia la izquierda y derecha, respectivamente, junto a su vista con pose normalizada. (Extraído de Berretti *et al.* (2011)).**

Además, para comparar los resultados de la literatura se debe considerar que se utilizaron diferentes tamaños para las bases de datos y distintos niveles de dificultad en la variedad de los rostros y no sólo distinto diseño en los experimentos. Por ejemplo, en Xu *et al.* (2004) reportan un 96.1 % de reconocimiento usando una base de datos de 30 personas, el cual bajó a 72.4 % al incrementarla a 120 personas. En Chang *et al.* (2006) se reporta una menor disminución en el rendimiento al incrementar el tamaño de la base de datos, pero se menciona que la disminución es mayor cuando se agregan rostros con expresiones en lugar de rostros neutros. Adjunto a esto, usualmente se reporta un incremento en el porcentaje de reconocimiento al usar más de una imagen por persona en la galería.

Acerca de la resolución, no se cuenta con un estándar sobre la densidad del muestreo ni la profundidad necesarias para el reconocimiento 3D, las investigaciones van de manejar cientos a decenas de miles de puntos. Para afrontar el problema de reconocimiento en términos de correspondencia de superficies con grandes cantidades de puntos una solución atractiva es extraer información cuantitativa que permita discernir entre individuos, la cual puede concernir a puntos, curvas o superficies; otra opción es aplicar técnicas que reduzcan la dimensión de los datos, esto es, embeber la información extraída en un análisis estadístico. Estas características en las técnicas marcan una pauta para su clasificación y la estimación de su robustez puede valorarse a través de su manejo de

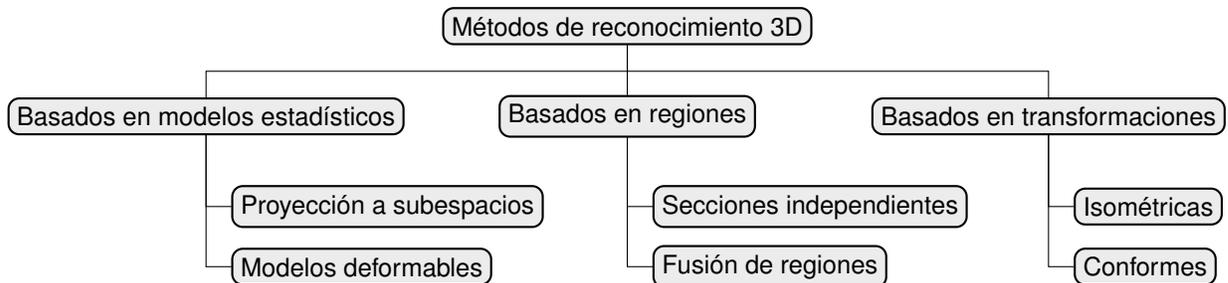
<sup>1</sup>[www.gavab.etsii.urjc.es/recursos.html#GavabDB](http://www.gavab.etsii.urjc.es/recursos.html#GavabDB)

expresiones y variaciones de postura, el cual puede efectuarse mediante una etapa de normalización o la transformación a una representación invariante a pose.

### 1.3. Investigación previa

En esta sección se presentan los enfoques que se han desarrollado en reconocimiento de rostros dirigiendo el estudio a hacer frente a expresiones y oclusiones. Al mismo tiempo, se mencionan algunos trabajos de investigación relevantes que se llevaron a cabo en los últimos años y marcan una pauta para el presente trabajo de tesis.

De acuerdo con la clasificación de Smeets *et al.* (2012), las técnicas de reconocimiento 3D en presencia de expresiones se pueden catalogar en tres grandes grupos: métodos basados en modelos estadísticos, en regiones y basados en transformaciones (véase figura 6). Los métodos estadísticos construyen un modelo de las deformaciones causadas por expresiones usando un conjunto de rostros con gestos no neutrales, los basados en transformaciones asumen cierta invariancia de su representación facial frente a estas deformaciones y los métodos basados en regiones toman partes del rostro considerando qué tan rígidas son durante expresiones. La literatura indica que el rendimiento en el reconocimiento puede incrementarse al mezclar distintas clases de métodos y diferentes modalidades respecto a la representación facial.



**Figura 6: Clasificación de técnicas de reconocimiento planteada en Smeets *et al.* (2012) para rostros 3D en presencia de expresiones.**

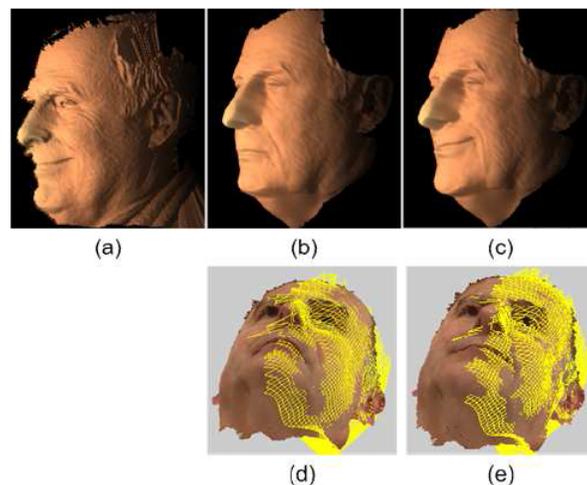
#### 1.3.1. Métodos estadísticos

Una de las formas de llevar a cabo el reconocimiento es a través de la construcción de un modelo estadístico de las deformaciones de los rostros durante expresiones. Se caracterizan por requerir de un paso de entrenamiento para construir el modelo sobre el

cual se van a proyectar los rostros, de manera que el rendimiento depende de incorporar datos suficientemente representativos en este proceso.

Un punto importante al aplicar estos métodos es que, antes de efectuar el análisis, los rostros deben ser estandarizados y alineados si se busca hacer frente a los cambios de postura (Al-Osaimi *et al.* (2009)).

El Análisis de Componentes Principales (PCA, por sus siglas en inglés) es una de las técnicas estadísticas que ha sido ampliamente utilizada en algoritmos de reconocimiento. Su uso más frecuente va dirigido a la reducción de la dimensión de datos a través de una proyección que minimiza distancias en términos de mínimos cuadrados. Otro uso popular de esta técnica dentro del marco de reconocimiento de rostros 3D es la corrección de la orientación en el espacio 3D de cada superficie facial (Mian *et al.* (2007); Lei *et al.* (2014)).



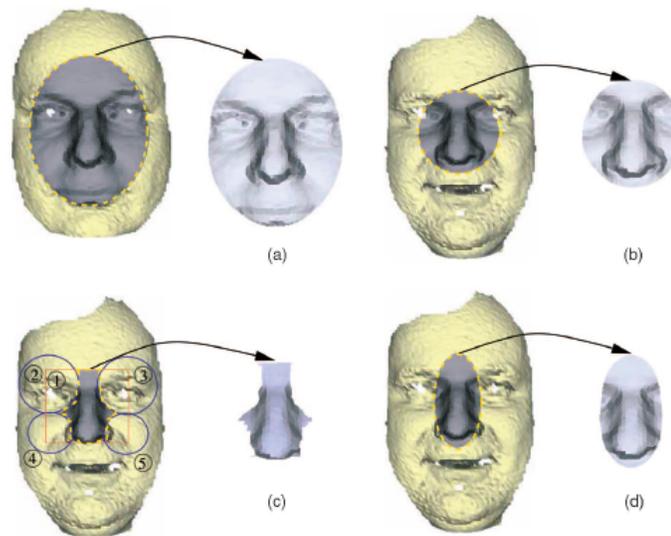
**Figura 7: Modelos deformables para el reconocimiento facial 3D. A partir de un rostro de prueba (a) y uno neutral de la galería (b) se obtiene un modelo deformado (c). El alineamiento de (a) con los modelos (b) y (c) se muestra en (d) y (e) respectivamente. (Extraído de Lu y Jain (2008)).**

Otra técnica usada como base en reconocimiento de rostros es el algoritmo iterativo del punto más cercano (ICP, por sus siglas en inglés), el cual se usa para minimizar iterativamente la distancia euclidiana entre conjuntos de puntos sobre superficies tridimensionales (Chang *et al.* (2006)). Un inconveniente del método es que requiere de una buena alineación inicial de los conjuntos de entrada para llegar a obtener buenos resultados. Además, en caso de que se considere un parámetro de escala, este debe ser parecido inicialmente o estar acotado desde un principio para asegurar la convergencia del método,

de otra manera, los puntos de un conjunto convergen a un pequeño subconjunto del otro (Du *et al.* (2010)). Además de utilizarse para la alineación, es la base de algoritmos que deforman modelos de la galería (véase figura 7) para realizar la correspondencia (Lu y Jain (2008)).

### 1.3.2. Métodos basados en regiones

Como su nombre lo indica, estos métodos seleccionan regiones del rostro que presenten menor impacto bajo variaciones típicas de los rostros. La selección de las zonas a extraer depende de las variaciones a las que se quiera hacer frente. Al manifestarse las expresiones, se seleccionan regiones que no cambian tanto como la boca y mejillas, por ejemplo, la cuenca de los ojos y la sección alrededor de la nariz. Al evaluar la capacidad de discriminación de distintas zonas en presencia de oclusiones los resultados dependen del tipo de oclusión considerada (Alyuz *et al.* (2008)), ya sea oclusión de un solo ojo, uso de lentes o bufanda, oclusión de la boca o simplemente cabello sobre la cara; como cada región tendrá un rendimiento diferente en cada caso, generalmente se recomienda hacer una fusión de los resultados individuales.



**Figura 8: Extracción de regiones para llevar a cabo reconocimiento facial 3D a partir de un rostro neutral de la galería (a) y distintos rostros de prueba con expresiones (b-d). (Extraído de Chang *et al.* (2006)).**

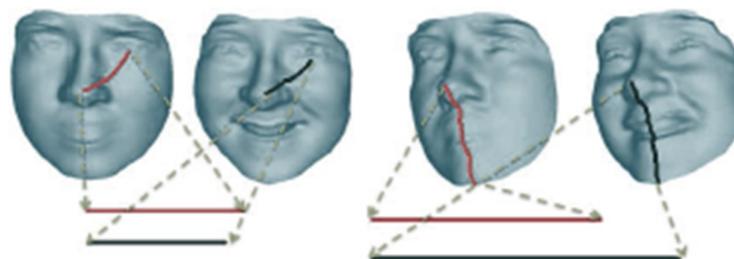
La opción más popular es considerar la región alrededor de la nariz (Chang *et al.* (2006); Faltemier *et al.* (2008); Queirolo *et al.* (2010)), pues contiene información distintiva

de los individuos y ha demostrado ser la más robusta. En la figura 8 se muestra un ejemplo de la extracción de información de esta zona. Como se ilustra, usualmente la información contenida en la galería proviene de un rostro con expresión neutra y cubre una mayor área en comparación con las zonas que se toman como prueba, las cuales pueden proceder de rostros con gestos notables pero que cubren un área menor bien definida que se espera sea invariante a deformaciones por expresiones.

Algunas investigaciones utilizan métodos basados en regiones para hacer reconocimiento facial centrándose en el principio de que el rendimiento se puede mejorar al tener más imágenes de referencia del rostro a reconocer dentro de la galería. Con base en esta idea, se aplican técnicas que descomponen el rostro de un individuo en múltiples observaciones para después evaluar su capacidad de reconocimiento separadamente y combinar sus resultados usando métodos estadísticos tradicionales como los anteriormente mencionados (Mian *et al.* (2007)).

### 1.3.3. Métodos basados en transformaciones

Como se mencionó anteriormente, las técnicas de reconocimiento de rostros no solo utilizan superficies faciales o regiones específicas de estas, sino también puntos o curvas. Comúnmente, la finalidad de extraer esta información en cada individuo es calcular las distancias entre dichos puntos o la longitud de ciertas curvas para identificar a través de ellas a cada individuo.



**Figura 9: Impacto de expresiones sobre distancias geodésicas en distintas zonas del rostro. (Extraído de Li *et al.* (2009)).**

Algunos trabajos suponen que al considerar rostros con expresiones el cambio de distancias sobre las superficies es insignificante, pero en realidad la información extraída puede variar mucho entre las superficies faciales de un mismo individuo (véase figura 9). Los cambios más notables en distancias corresponden a la zona inferior del rostro por movimientos de la boca.

Este problema originó otro enfoque para lidiar con expresiones que se centra en convertir las superficies faciales de la base de datos a una representación invariante a deformaciones. El objetivo es llevar el rostro a una forma invariante mediante el mapeo de la superficie facial original a una forma rígida, de manera que, al transformar el rostro del mismo individuo con alguna variante en su expresión, el cambio en la representación obtenida sea insignificante.

Una idea popular bajo el enfoque de transformaciones es considerar que las expresiones pueden ser aproximadas por transformaciones isométricas, es decir, aquellas que preservan distancias sobre una superficie. Luego, el construir una representación invariante a transformaciones isométricas brinda la capacidad de ser invariante a expresiones faciales bajo esta suposición.

Además de utilizar transformaciones isométricas, es de gran utilidad obtener representaciones de los rostros aplicando transformaciones conformes, esto es, transformaciones que conservan ángulos de la forma tridimensional original. El interés detrás de su uso es que permiten llevar un rostro 3D a una representación plana, lo cual puede aprovecharse para aplicar técnicas de reconocimiento de rostros 2D ampliamente estudiadas.

Así, una de las ventajas de los métodos basados en transformaciones es que permiten combinar varias técnicas para realizar el reconocimiento. Por ejemplo, el trabajo de Drira *et al.* (2010) utiliza curvas radiales y posteriormente aplica una transformación que las haga isométricas; en Kakadiaris *et al.* (2007) se realiza un mapeo de las superficies faciales 3D a dos dimensiones y después se emplean wavelets para extraer características y en Bronstein *et al.* (2003) se emplea un enfoque multimodal en que se usa una transformación isométrica para lidiar con expresiones faciales y se considera después la información de textura sobrepuesta.

## **1.4. Objetivos**

### **1.4.1. Objetivo general**

Realizar un estudio comparativo entre métodos de reconocimiento de rostros 3D basados en superficies faciales con el fin de desarrollar un método de reconocimiento híbrido con el que se pueda garantizar cierto grado de confiabilidad respecto a distintas métricas de desempeño.

### **1.4.2. Objetivos específicos**

1. Estudiar la clasificación, ventajas y desventajas de las técnicas actuales de reconocimiento de rostros en 2D y 3D.
2. Estudiar el problema de correspondencia de superficies desde diferentes enfoques.
3. Estudiar transformaciones conformes.
4. Diseñar un algoritmo de reconocimiento de rostros basado en métodos de correspondencia de superficies faciales y mapeos conformes.
5. Realizar un estudio comparativo entre el método propuesto y los existentes.

## **1.5. Limitaciones y suposiciones**

- El uso de superficies faciales desprovistas de datos de color disminuye en gran medida el volumen de datos a emplear en el reconocimiento, lo cual se espera contribuya a mejorar la eficiencia y tiempo de ejecución del método propuesto.
- Se propone excluir las características de la región comprendida entre la boca y la parte inferior de la mandíbula para reforzar la invariancia a expresiones faciales, además de hacer uso de datos provenientes de zonas desprovistas de vello facial para reducir el efecto negativo producido por barba, bigote, etc.
- Con objeto de que el método propuesto sea robusto ante cambios de posición y orientación de la cabeza, se pretende incluir una etapa de extracción automática

de puntos que permita realizar una corrección de pose. Estos puntos deberán ser distinguibles y estar situados en zonas que no varíen mucho con expresiones faciales.

- Se propone efectuar la correspondencia con base en un mapeo conforme de la representación original del rostro a una malla triangulada en el plano, ya que este enfoque preserva ángulos de la superficie facial original al mismo tiempo que simplifica el problema de correspondencia.
- Los experimentos se diseñarán bajo el supuesto de que todos los rostros de prueba pertenecen a los individuos de la galería. Además, se utilizará una base de datos que presente variabilidad respecto a expresiones faciales y rotaciones del rostro con el fin de poder evaluar la metodología propuesta frente a estas variaciones.

## **1.6. Organización de la tesis**

A continuación se detalla la organización del resto del presente trabajo de tesis. El capítulo 2 introduce los fundamentos de los algoritmos de correspondencia con los que se realizó este trabajo. El capítulo 3 forma la base teórica del enfoque de reconocimiento 3D a través de parametrización de superficies. En el capítulo 4 se presentan los algoritmos de correspondencia que se tomaron en cuenta para el diseño de los experimentos.

El capítulo 5 presenta la metodología propuesta mostrando la información extraída y las transformaciones por las que pasan las superficies faciales para buscar su correspondencia. En el capítulo 6 se describen los resultados obtenidos con la metodología propuesta. Finalmente, se exponen las conclusiones de este trabajo de investigación al igual que las oportunidades de trabajo a futuro en el capítulo 7.

## Capítulo 2. Fundamentos

---

### 2.1. Eigenvectores y eigenvalores de una matriz

Dada una matriz cuadrada  $A$  de tamaño  $n$  con entradas reales, es decir,  $A \in \mathbb{R}^{n \times n}$ , se dice que un escalar  $\lambda$  es un valor propio o eigenvalor de  $A$  si existe alguna matriz no nula  $X \in \mathbb{R}^{n \times 1}$  tal que

$$AX = \lambda X. \quad (1)$$

Se dice entonces que la matriz columna  $X$  es un vector propio o eigenvalor de  $A$  correspondiente al eigenvalor  $\lambda$ .

Un resultado importante es que si la matriz  $A$  es compleja y simétrica, entonces sus eigenvalores son reales.

### 2.2. Conceptos básicos de análisis complejo

**Función compleja.** Sea una función compleja  $f(z)$  con  $z = x + iy$ . Se sabe que  $f(z)$  puede descomponerse en una suma de dos funciones reales  $u$  y  $v$  de dos variables de manera que

$$f(z) = f(x, y) = f(x + iy) = u(x, y) + iv(x, y). \quad (2)$$

**Función holomorfa.** Sea  $\Omega$  un conjunto abierto en el plano complejo,  $z_0 \in \Omega$  y sea  $f : \Omega \rightarrow \mathbb{C}$ .  $f$  se dice complejo derivable o complejo diferenciable en  $z_0$  si

$$\lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0} \text{ existe.} \quad (3)$$

Cuando tal límite existe, a este se le llama la derivada de  $f$  en  $z_0$  y se denota por  $f'(z_0)$ . Si  $f$  es diferenciable en todo  $\Omega$ ,  $f$  se llama holomorfa en  $\Omega$ .

Si la función  $f$  es derivable en un punto  $z_0 = x_0 + iy_0$ , entonces se satisfacen las ecuaciones de Cauchy-Riemann:

$$\frac{\partial u}{\partial x}(x_0, y_0) = \frac{\partial v}{\partial y}(x_0, y_0) \quad \text{y} \quad \frac{\partial u}{\partial y}(x_0, y_0) = -\frac{\partial v}{\partial x}(x_0, y_0). \quad (4)$$

### 2.3. Conceptos básicos en topología y geometría diferencial

Para realizar un análisis de forma, en este trabajo se utilizan conceptos del marco teórico de geometría diferencial de superficies (cálculo en variedades diferenciables). En particular, en este trabajo se toma en cuenta esta teoría aplicada a un espacio euclídeo,  $\mathbb{R}^3$ .

La definición formal de superficie se manejará a través del concepto de variedad, un objeto geométrico que generaliza la noción de curva (1-variedad) y superficie (2-variedad) a cualquier dimensión. Para describir formalmente una variedad, se introducen primeramente conceptos básicos de topología.

Un **espacio topológico** es un conjunto  $\mathcal{E}$  junto con una colección  $\mathcal{T}$  de subconjuntos de  $\mathcal{E}$  (denominada topología) que satisfacen:

1. El conjunto vacío y  $\mathcal{E}$  pertenecen a  $\mathcal{T}$ .
2. La intersección de cualquier subcolección finita de conjuntos de  $\mathcal{T}$  pertenece a  $\mathcal{T}$ .
3. La unión arbitraria de conjuntos de  $\mathcal{T}$  pertenece a  $\mathcal{T}$ .

**Homeomorfismo.** Un homeomorfismo es una biyección entre dos espacios topológicos mediante una aplicación biyectiva que es continua y cuya inversa es continua. Si existe un homeomorfismo entre ellos, los dos espacios topológicos se dicen homeomorfos.

**Variedad.** En términos generales, una variedad de dimensión  $n$  es un espacio que se parece localmente a  $\mathbb{R}^n$ , de manera que puede verse como un conjunto de parches  $n$ -dimensionales.

Esto es, una variedad topológica de dimensión  $n \geq 0$  es un espacio topológico  $\mathcal{M}$  en el que para cada  $p \in \mathcal{M}$  existe una vecindad abierta  $\mathcal{U}_p \subset \mathcal{M}$  homeomorfo a un abierto  $\mathcal{V}$  de  $\mathbb{R}^n$  mediante una aplicación  $\phi_p : \mathcal{U}_p \rightarrow \mathcal{V}_p \subset \mathbb{R}^n$ .

### 2.3.1. Conceptos sobre variedades en espacios euclídeos.

**Espacio topológico de Hausdorff.** Un espacio topológico  $\mathcal{X}$  se dice Hausdorff si dados  $x, y \in \mathcal{X}$ ,  $x \neq y$ , existen abiertos  $\mathcal{U}, \mathcal{V} \in \mathcal{X}$ ,  $x \in \mathcal{U}$ ,  $y \in \mathcal{V}$  tales que  $\mathcal{U} \cap \mathcal{V} = \emptyset$ . Esto es, dados dos puntos distintos del espacio, hay abiertos disjuntos conteniendo a cada uno de ellos.

Un espacio topológico  $\mathcal{X}$  se dice que tiene una base numerable si existe una colección numerable  $\mathcal{B}$  de conjuntos abiertos de  $\mathcal{X}$  tal que cada elemento no vacío del espacio topológico es la unión de elementos de  $\mathcal{B}$ . Todo espacio métrico es Hausdorff.

**Variedad topológica.** Formalmente, decimos que  $\mathcal{M}$  es una variedad topológica de dimensión  $n$  si  $\mathcal{M}$  es un espacio topológico de Hausdorff con base numerable y tal que, para cada uno de sus puntos, existe una vecindad homeomorfa a un subconjunto abierto  $\mathcal{U} \subset \mathbb{R}^n$ .

**Superficie.** Una superficie es una variedad bidimensional, es decir, un objeto topológico que es localmente homeomorfo al plano euclídeo  $\mathbb{R}^2$ .

El homeomorfismo local entre una superficie y el plano euclídeo implica que para cada punto de una superficie hay una vecindad centrada en dicho punto que es homeomorfa a un disco abierto de  $\mathbb{R}^2$ . Esto permite construir un sistema de coordenadas local bidimensional en torno a cualquier punto de la superficie.

**Parametrización de superficies.** Puesto que una superficie en  $\mathbb{R}^3$  es una variedad de dimensión dos, para toda vecindad en ella las coordenadas de cualquier punto  $(x, y, z)$  pueden escribirse en función de dos parámetros  $(u, v)$  que funcionan como sistema de coordenadas propio de la superficie:

$$(x(u, v), y(u, v), z(u, v)) = r(u, v), \quad (5)$$

lo cual se denomina parametrización de la superficie.

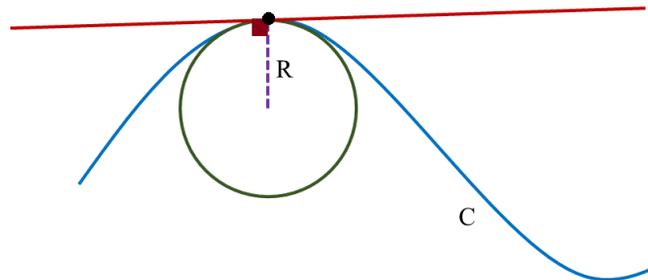
### 2.3.2. Curvatura local de superficies

Extraer información de curvatura de superficies tiene la ventaja de ser una propiedad local invariante a transformaciones rígidas (traslaciones y rotaciones).

De manera intuitiva, la curvatura es la medida por la cual un objeto geométrico dentro de un espacio euclídeo se desvía de ser plano, proviene de nociones de geometría en 2D que mide qué tanto se “dobla” una curva en cierto punto.

Para explicar geoméricamente el concepto se muestra primero el caso de una curva en 2D. Bajo este esquema, considérese un punto sobre la curva y el vector tangente que pasa por dicho punto; a medida que el punto se desplaza a lo largo de la curva, la dirección del vector tangente cambia; se dice entonces que entre más rápido cambia su dirección, la curvatura es más grande.

Dado un punto  $p$  sobre una curva  $C$ , existe una circunferencia cuyo centro se encuentra sobre la normal a la curva en  $p$  y que posee la misma curvatura que  $C$  en  $p$  (véase figura 10). Así, la curvatura de  $C$  en  $p$  se define como el inverso del radio de dicha circunferencia.



**Figura 10: Noción de curvatura en 2D. La curvatura en un punto es el inverso del radio del círculo osculante en dicho punto.**

Para mostrar la idea de curvatura de superficies debe considerarse un punto  $p$  sobre una superficie  $S$  y un plano  $\Pi \in \mathbb{R}^3$  que pasa por  $p$  y que contiene al vector normal  $N$ . La intersección de  $S$  con  $\Pi$  resulta en una curva que pasa por  $p$ ; viéndola como una curva sobre el plano  $\Pi$ , puede asignársele un valor de curvatura con la noción previa en 2D.

Al repetir el proceso con distintos planos se obtienen distintos valores de curvatura

para la curva de intersección con  $S$  en  $p$ . Las curvaturas principales de una superficie,  $k_{max}$  y  $k_{min}$  son los valores máximo y mínimo de estas curvaturas. Éstas curvaturas principales contienen la información del cambio de la superficie en las dos direcciones extremas, por lo cual se utilizan para caracterizar objetos tridimensionales; sin embargo, no suelen utilizarse directamente, sino que se calculan descriptores de forma a través de ellas. Algunos de los descriptores de forma más utilizados son los siguientes:

- Curvatura gaussiana (K)

$$K = k_{max}k_{min} \quad (6)$$

- Curvatura media (H)

$$H = \frac{k_{max} + k_{min}}{2} \quad (7)$$

- Índice de forma (Shape Index, SI)

$$SI = \frac{1}{2} - \tan^{-1} \frac{(k_{max} + k_{min})}{\pi(k_{max} - k_{min})} \quad (8)$$

### 2.3.2.1. Clasificación HK

La relevancia de las curvaturas media ( $H$ ) y gaussiana ( $K$ ) es que permiten asignar una clase de forma local a cada punto por medio de la tabla 1. Este catálogo de formas se conoce como clasificación HK.

**Tabla 1: Clasificación de la superficie local en un punto según signos de curvaturas.**

Curvatura media H	Curvatura gaussiana K	Forma local
0	0	Plano
Positivo	0	Cóncavo cilíndrico
Negativo	0	Convexo cilíndrico
Positivo	Positivo	Cóncavo elíptico
Negativo	Positivo	Convexo elíptico
Cualquier signo	Negativo	Hiperbólico

### 2.3.2.2. Índice de forma

El índice de forma (SI, por sus siglas en inglés) es un descriptor de forma invariante a escala calculado a partir de las curvaturas principales (Mahmood *et al.* (2013)).

Esta medida asigna un valor dentro del intervalo  $[0, 1]$  a cada punto en la superficie a través de la ecuación 8 (Szeptycki *et al.* (2010)). A excepción del plano donde el SI es indeterminado (puesto que  $k_{max} = k_{min}$ ), cada forma posible de una superficie corresponde a un único valor de SI. Algunas formas representativas se muestran en la figura 11. Como puede observarse, los valores pequeños corresponden a puntos en zonas cóncavas y los más grandes a puntos en zonas convexas.

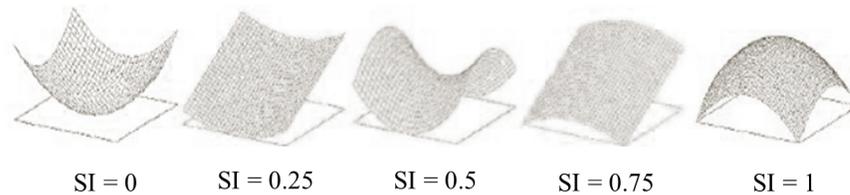


Figura 11: Ejemplos representativos superficies en la escala de SI.

### 2.3.2.3. Formas fundamentales de superficies

Tratándose de superficies parametrizadas, la descripción geométrica local se realiza a través de su primera y segunda forma fundamental considerando una vecindad de la superficie parametrizada por las variables  $u$  y  $v$ .

La primera forma fundamental se obtiene a través de la siguiente expresión:

$$I(u, v) = Edu^2 + 2Fdudv + Gdv^2, \ni \quad (9)$$

$$\begin{aligned} E &= \|x_u\|^2 = \left| \frac{\partial x}{\partial u} \right|^2, \\ F &= x_u \cdot x_v = \frac{\partial x}{\partial u} \cdot \frac{\partial x}{\partial v}, \\ G &= \|x_v\|^2 = \left| \frac{\partial x}{\partial v} \right|^2. \end{aligned} \quad (10)$$

Los coeficientes de la primera forma fundamental permiten calcular longitudes de curvas sobre la superficie, el ángulo entre dos curvas y áreas sobre una región específica.

La segunda forma fundamental  $II(u, v)$  está relacionada con el vector  $n$  normal a la superficie en un punto y las segundas derivadas respecto a las variables de la parametrización:

$$II(u, v) = edu^2 + 2fdudv + gdv^2, \quad (11)$$

donde

$$\begin{aligned} e &= n \cdot \frac{\partial^2 r}{\partial u^2} = -n_u \cdot r_u \\ f &= n \cdot \frac{\partial^2 r}{\partial u \partial v} = -n_u \cdot r_v = -n_v \cdot r_u \\ g &= n \cdot \frac{\partial^2 r}{\partial v^2} = -n_v \cdot r_v. \end{aligned} \quad (12)$$

La relación entre las medidas de curvatura y las formas fundamentales se establece por medio de los coeficientes definidos en 10 y 12 a través de las ecuaciones

$$\begin{aligned} H &= \frac{eG - 2fF + gE}{2(EG - F^2)} \\ K &= \frac{eg - f^2}{EG - F^2} \end{aligned} \quad (13)$$

#### 2.3.2.4. Discretización del cálculo de curvaturas

En cuanto a superficies discretas, el análisis de curvatura puede realizarse a través de cálculos basados en la aproximación de la segunda derivada en cada punto utilizado (Colombo *et al.* (2006)). Para este enfoque, se considera el rostro como una superficie  $S$  definida por una función real dos veces diferenciable  $f : U \rightarrow \mathbb{R}$ , definida en un conjunto abierto  $U \subseteq \mathbb{R}^2$  tal que

$$S = \{(x, y, z) | (x, y) \in U; z \in \mathbb{R}; f(x, y) = z\}. \quad (14)$$

Así, para cada punto  $(x, y, f(x, y))$  las curvaturas media ( $H$ ) y gaussiana ( $K$ ) se definen como (Do Carmo (1976)):

$$H(x, y) = \frac{(1 + f_y^2)f_{xx} - 2f_x f_y f_{xy} + (1 + f_x^2)f_{yy}}{2(1 + f_x^2 + f_y^2)^{3/2}}, \quad (15)$$

$$K(x, y) = \frac{f_{xx}f_{yy} - f_{xy}^2}{(1 + f_x^2 + f_y^2)^2}, \quad (16)$$

donde  $f_x, f_y, f_{xy}, f_{xx}, f_{yy}$  son las primeras y segundas derivadas de  $f$  en  $(x, y)$ .

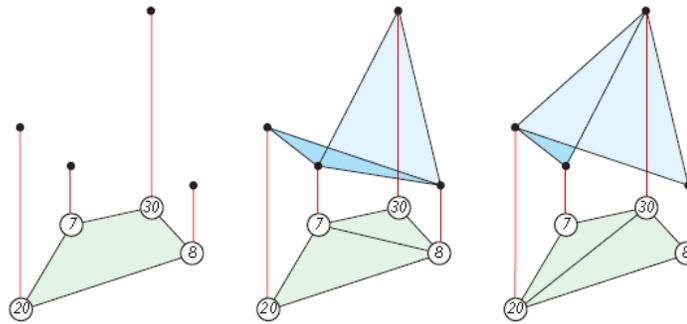
#### 2.4. Triangulación de Delaunay

En este trabajo se aborda el problema de encontrar correspondencias entre superficies faciales a partir de nubes de puntos. Para cada punto  $q$  sobre un rostro de referencia, el objetivo es encontrar un punto correspondiente en el rostro de prueba  $P = \{p_1, p_2, \dots, p_N\}$  a través de cierta distancia. Este es un problema ampliamente estudiado debido su variedad de aplicaciones, por lo cual existen diversas técnicas que abordan el problema.

Una idea básica es calcular la distancia de  $q$  hacia todos los puntos en  $P$  y elegir el punto con la menor distancia, lo cual se conoce como búsqueda de fuerza bruta o búsqueda exhaustiva. Sin embargo, es preferible representar a  $P$  en una estructura que permita búsquedas más eficientes. Una posible solución es determinar previamente regiones de puntos en  $P$  que sean cercanos entre ellos, de manera que al buscar un punto correspondiente a  $q$  se realicen cálculos solamente sobre puntos agrupados en determinada región. En este trabajo se adopta este enfoque a través del cómputo de una triangulación de las nubes de puntos.

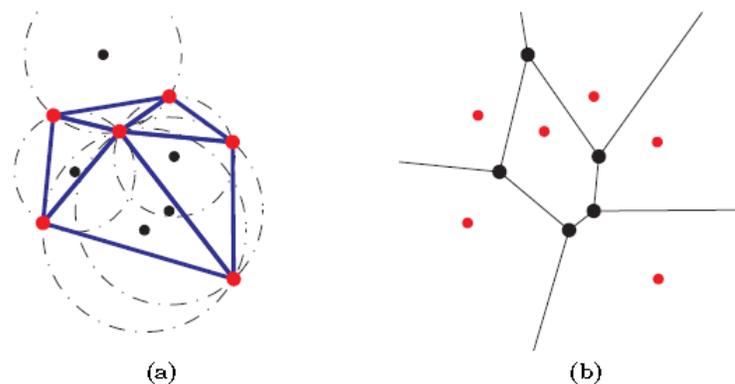
En la figura 12 se muestra un ejemplo de dos triangulaciones sobre el mismo conjunto de puntos, donde puede verse que la primer triangulación describe una región cóncava y la segunda triangulación una región convexa. Debido a que es de interés conservar información geométrica tridimensional de los rostros, en este trabajo se utilizó el criterio de Delaunay para establecer las triangulaciones.

La idea en dos dimensiones de una triangulación de Delaunay consiste en que tres



**Figura 12: La forma tridimensional de la superficie depende de la triangulación. Extraído de Devadoss y O'Rourke (2011).**

puntos forman un triángulo válido si se cumple la condición de que la circunferencia circunscrita al triángulo no contiene otros puntos del conjunto. Esta triangulación no es necesariamente única para cada conjunto de puntos y puede no existir, por ejemplo, para un conjunto de puntos alineados; sin embargo, se opta por utilizarla debido a que posee una estructura asociada de gran utilidad.



**Figura 13: (a) Una triangulación de Delaunay de 6 puntos (rojo), sus circunferencias circunscritas y sus centros se muestran en negro. (b) Diagrama de Voronoi correspondiente a la triangulación (a).**

Para cada triangulación de Delaunay existe un diagrama asociado, llamado diagrama de Voronoi, que conecta los centros de las circunferencias circunscritas. En la figura 13 se muestra una triangulación de Delaunay en 2D y su respectivo diagrama de Voronoi.

Cabe mencionar que cada punto en la triangulación tiene asociada una celda de Voronoi: el área para el cual dicho punto es el más cercano de entre todos los que forman la triangulación. Esta información es la que permite realizar las búsquedas de puntos más cercanos con mayor facilidad.

La región de Voronoi,  $Vor(q)$ , de un punto  $q_i$  en la triangulación  $\mathcal{T}$  está definida por (Devados y O'Rourke (2011))

$$Vor(q_i) = \{x \in \mathbb{R}^2 : \|x - q_i\| \leq \|x - q_j\| \quad \forall q_j \in \mathcal{T}\}. \quad (17)$$

Estos conceptos pueden ser extendidos a dimensiones mayores. Es de particular importancia en 3D, donde la triangulación consiste en tetraedros (cada uno definido ahora por 4 puntos circunscritos en una esfera).

## Capítulo 3. Parametrización de superficies

---

Un enfoque particular para el estudio de superficies en  $\mathbb{R}^3$  es aprovechar que estas tienen una estructura bidimensional inherente bajo la orientación de geometría diferencial de superficies suaves. Así, el estudio de superficies tridimensionales se hace a través de parametrizaciones, lo cual puede ser visto como el proceso de embeber una superficie 3D con topología de disco a un dominio plano. Generalmente, los trabajos en este campo adoptan términos pertenecientes a geometría conforme (transformaciones que preservan ángulos) donde la teoría respalda la existencia de mapeos conformes biyectivos.

En el dominio discreto, la aplicación de transformaciones conformes resulta en una distorsión de ángulos y áreas, por lo que las técnicas desarrolladas están enfocadas a minimizar dichas deformaciones. En particular, los métodos para calcular mapeos conformes surgen de la aproximación de ecuaciones diferenciales provenientes de geometría conforme y de las distintas formas de medir la distorsión angular (mediante el cálculo directo de las diferencias entre los ángulos correspondientes en las mallas discretas o la optimización de una función que la mida).

En las secciones siguientes se muestran las bases de transformaciones conformes, se presenta el panorama de técnicas para su cálculo y se exponen los fundamentos de la parametrización espectral conforme utilizada en este trabajo.

### 3.1. Transformaciones conformes

Los mapeos conformes tienen gran importancia en el análisis complejo y la geometría riemanniana. En general, una transformación conforme es una función que preserva ángulos. Para describirlas formalmente, se considera el caso más común entre dominios del plano complejo.

Una función  $f : \Omega \subset \mathbb{C} \rightarrow \mathbb{C}$  se llama conforme o transformación conforme en un punto  $p \in \Omega$  si  $f$  preserva ángulos entre curvas diferenciables que se intersecan en  $p$ , como se muestra en la figura 14.

Lo anterior es equivalente a que si  $f$  es holomorfa en  $\Omega$  y si  $f'(p) \neq 0$ , entonces  $f$  es conforme en  $p$ . Un resultado importante es que el recíproco de esta definición también es

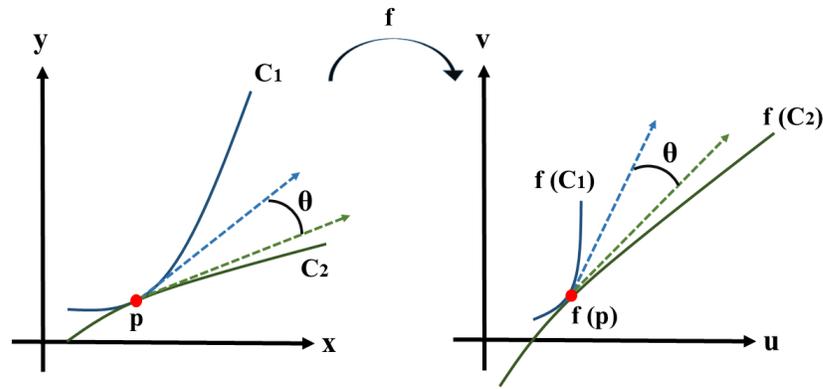


Figura 14: Transformación conforme.

cierto; en consecuencia, si  $f$  es conforme en  $p$ , entonces se verifican las ecuaciones de Cauchy-Riemann  $u_x(p) = v_y(p)$  y  $u_y(p) = -v_x(p)$ , donde  $f(p) = u(p) + iv(p)$  y  $p = x + iy$ . Diferenciando estas ecuaciones con respecto a  $x$  y  $y$  respectivamente se obtiene que  $u_{xx}(p) = 0$  y  $u_{yy}(p) = 0$ . Luego, cualquier función real que satisfaga la ecuación de Laplace ( $\nabla^2 f = 0$ ) es llamado armónico, por lo tanto, la parte real e imaginaria de cualquier función conforme es armónica.

La aproximación de las ecuaciones de Cauchy-Riemann han originado varios métodos para el cálculo de parametrizaciones conformes sobre superficies discretas. A continuación se describe el trabajo de Lévy *et al.* (2002), el cual sigue esta metodología.

### 3.1.1. Mapeos conformes por mínimos cuadrados

Como se mencionó anteriormente, dada una superficie 3D discreta  $\mathcal{S}$  y un mapeo suave  $f : \mathcal{X} \rightarrow \mathcal{U}$ ,  $f$  se dice conforme en  $p \in \mathcal{X}$  si y sólo si se satisfacen las ecuaciones de Cauchy-Riemann. Esto implica que  $f$  es conforme si

$$\frac{\partial f}{\partial x}(p) + i \frac{\partial f}{\partial y}(p) = 0 \quad (18)$$

se cumple para todo  $p \in \mathcal{X}$ .

En la práctica, no se puede cumplir dicha condición de conformalidad en toda la

superficie triangulada, por lo cual se construye el mapeo conforme en el sentido de mínimos cuadrados, esto es, buscando minimizar la expresión 18 para todo punto en  $\mathcal{X}$  a través de la función

$$C(\mathcal{X}) = \sum_{\tau \in \mathcal{X}} \left\| \frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} \right\|^2 A(\tau), \quad (19)$$

donde  $\tau$  es un triángulo en la malla  $\mathcal{X}$  y  $A(\tau)$  es su área.

El método de mapeos conformes por mínimos cuadrados (Lévy *et al.* (2002)) parte de una triangulación en  $\mathbb{R}^3$  y considera una base ortonormal local para cada triángulo de manera que éstos quedan representados por sus vértices en dicha base como  $(x_1, y_1)$ ,  $(x_2, y_2)$  y  $(x_3, y_3)$ . Luego, la minimización de la ecuación 19, resolverá el problema de buscar la correspondencia de cada uno de estos vértices a sus coordenadas en la parametrización. Una característica importante del trabajo de Lévy *et al.* (2002) es que se fija previamente la correspondencia de al menos un par de puntos de la malla original, de manera que el problema se resuelve a través de un sistema matricial que considera una división entre las correspondencias desconocidas entre puntos libres y las coordenadas que se han fijado.

### 3.1.2. Energía de Dirichlet y energía conforme

Otro enfoque para el cálculo de parametrizaciones es a través de la energía de Dirichlet, también llamada energía armónica (ecuación 20), la cual puede verse como una medida de la distorsión que la parametrización genera. Para un mapa  $U : \mathcal{X} \subset \mathbb{R}^3 \rightarrow \mathcal{U}$  sobre la superficie  $\mathcal{X}$ , su energía de Dirichlet,  $E_D$ , está definida como

$$E_D = \frac{1}{2} \int_{\mathcal{X}} |\nabla U|^2 dA, \quad (20)$$

donde  $dA$  es un elemento de área de la superficie original  $\mathcal{X}$  y  $\nabla U$  es el gradiente de  $U$ .

Un mapa que minimiza la energía de Dirichlet (bajo ciertas condiciones de frontera dadas) es armónico porque satisface  $\Delta U = 0$  (donde  $\Delta$  es el operador de Laplace-Beltrami, i.e., el laplaciano de la superficie original  $\mathcal{X}$ ).

De nuevo, el trabajar con mallas discretas provoca que estas funciones deban ser aproximadas. La discretización de la energía de Dirichlet es sencilla de calcular, sin embargo,

requiere fijar condiciones de frontera. El hecho de que se suponga que las parametrizaciones tendrán una frontera dada es una desventaja para realizar reconocimiento de rostros debido a que es habitual que el rostro se encuentre ocluido por cabello provocando que la frontera de la representación bidimensional varíe de un modelo a otro aún tratándose del mismo individuo. Por otro lado, el constreñir el contorno del rostro a una forma particular, un círculo por ejemplo, crea distorsiones para zonas internas.

El punto relevante para este trabajo es que si se alcanza el mínimo valor de  $E_D$ , entonces el mapeo es conforme. Dicho valor crítico corresponde al área de la imagen del mapeo,  $\mathcal{A}(U) = \int_{\mathcal{X}} \det(U) dA$ . Bajo esta idea, es conveniente medir la diferencia entre  $E_D$  y el área de la imagen de  $U$ ,  $\mathcal{A}(U)$ , mediante la energía conforme definida como

$$E_C(U) = E_D(U) - \mathcal{A}(U), \quad (21)$$

de manera que el mapa  $U$  es conforme si y sólo si  $E_C$  es cero.

### 3.2. Parametrización espectral conforme

En este trabajo de tesis se utilizó el método de parametrización espectral conforme desarrollado por Mullen *et al.* (2008) para el cálculo de mapeos conformes. Esta técnica busca la minimización de la ecuación 21 considerando una discretización de la energía de Dirichlet y resolviendo el problema de manera matricial a través de análisis espectral.

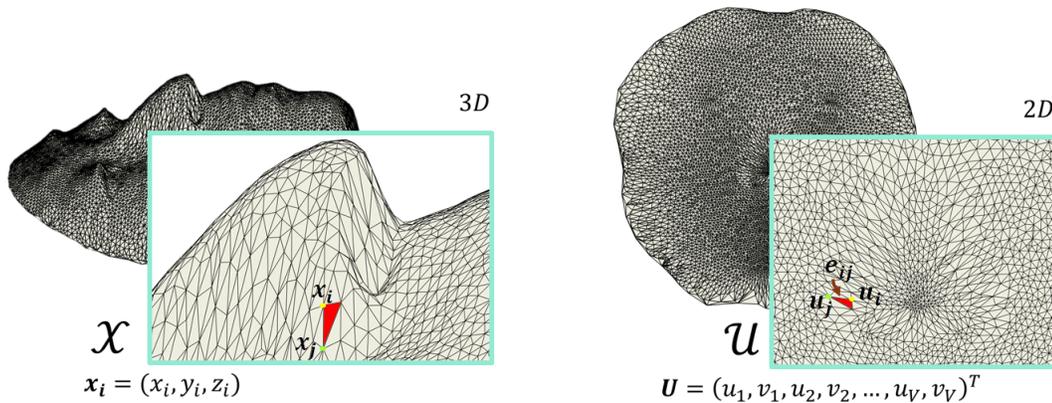
A continuación se describe primeramente en qué consiste el enfoque de análisis espectral, posteriormente se presenta la discretización del problema y su representación matricial y, finalmente, se muestra la formulación de la solución al problema de parametrización conforme.

#### 3.2.1. Notación

El mapa  $U$  representa una correspondencia entre la superficie facial  $\mathcal{X}$  y la triangulación en el plano  $\mathcal{U}$  relacionando cada triángulo de  $\mathbb{R}^3$  a  $\mathbb{R}^2$  como se muestra en la figura 15. Se denota con  $E$  al conjunto de aristas, con  $V$  al número de vértices en  $\mathcal{X}$  y con  $V_b$  al número de vértices que forman la frontera.

El  $i$ -ésimo nodo en la superficie original  $\mathcal{X}$  se denota por  $\mathbf{x}_i = (x_i, y_i, z_i)$  y la posición del nodo correspondiente en  $\mathcal{U}$  se representa como  $\mathbf{u}_i = (u_i, v_i)$ . Por otra parte,  $e_{ij}$  denota el arista en  $\mathcal{U}$  entre los vértices  $\mathbf{u}_i$  y  $\mathbf{u}_j$ .

Así, el rostro  $\mathcal{X}$  queda representado por la concatenación de las coordenadas de la parametrización en 2D por el vector columna  $U = (u_1, v_1, u_2, v_2, \dots, u_V, v_V)^T$ .



**Figura 15:** El mapa  $U$  relaciona la superficie facial  $\mathcal{X}$  con una triangulación en el plano  $\mathcal{U}$  llevando cada triángulo de  $\mathbb{R}^3$  a  $\mathbb{R}^2$ .

### 3.2.2. Análisis espectral de forma

El análisis espectral involucra matrices que provienen de la discretización de un operador continuo (típicamente lineal) obtenido a partir de determinada estructura topológica (Lévy y Zhang (2010)). Una vez que se construye la representación discreta del operador, estas matrices representan relaciones entre parejas de elementos sobre una malla, denotan los elementos adyacentes, etc.

Luego, dependiendo del problema, se hace una descomposición de la matriz previamente construida. Para resolver el problema de parametrización, se realiza una eigendescomposición, es decir, se calculan los eigenvalores y eigenvectores de la matriz. Tratándose de superficies faciales, se necesitan calcular eigenvalores de matrices muy grandes, por lo que se requieren métodos adicionales para hacer más eficiente el proceso.

Los primeros métodos desarrollados sobre parametrización de superficies al plano bajo el esquema de análisis espectral consistían en dos etapas: primero, se mapeaban los vértices de la frontera de la superficie hacia una región en 2D y posteriormente el resto de

los vértices se obtenían resolviendo un sistema lineal de la forma  $Lu = 0$ ,  $Lv = 0$ , donde  $L$  es la representación matricial de la triangulación de la superficie original. En este trabajo, dada una malla tridimensional, su representación  $L$  se construye a partir del siguiente operador:

$$L_{i,j} = \begin{cases} -\sum_{e_{ij} \in E} w_{ij} & i = j, \\ w_{ij} & i \neq j \wedge e_{ij} \in E, \\ 0 & \text{de otra manera.} \end{cases} \quad (22)$$

Los pesos  $w_{ij}$  se definen para cada arista  $e_{ij}$  en la malla. Bajo esta construcción se produce una matriz dispersa, simétrica y con entradas reales donde un renglón toma la forma

$$i \rightarrow [0 \dots w_{ij_1} \dots 0 \dots w_{ij_2} \dots 0 \dots \sum_{e_{ij} \in E} w_{ij} \dots 0 \dots w_{ij_3} \dots 0]. \quad (23)$$

Una de las principales diferencias entre las aplicaciones del análisis espectral en parametrizaciones es la manera de definir los valores de cada  $w_{ij}$ , cuya elección influye en las distorsión y la biyectividad de la parametrización. Por ejemplo, si los pesos son positivos y la matriz  $L$  es simétrica, la parametrización se garantiza biyectiva. La forma más utilizada de definir los pesos es a través de los llamados pesos armónicos que definen los elementos de la matriz a través de

$$w_{ij} = \frac{\cot(\alpha_{ij}) + \cot(\beta_{ij})}{2}, \quad (24)$$

donde  $\alpha_{ij}$  y  $\beta_{ij}$  son ángulos opuestos en los dos triángulos que comparten el arista  $e_{ij}$  como se muestra en la figura 16. Estos valores se derivan de una descripción mediante elemento finito de la energía armónica, por lo que pretende reducir la distorsión angular de la parametrización (Sheffer *et al.* (2006)). El inconveniente de su uso es que si la triangulación contiene ángulos obtusos, los pesos pueden ser negativos resultando en una parametrización no biyectiva. Un punto importante es que se ha probado que si la

triangulación satisface el criterio de Delaunay la parametrización obtenida será biyectiva, es por ello que en el presente trabajo se utilizan triangulaciones de Delaunay sobre las superficies faciales tridimensionales.

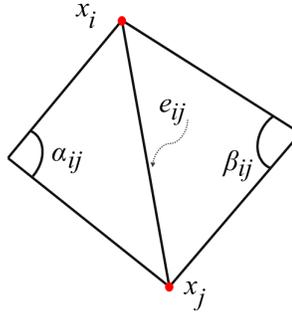


Figura 16: Ángulos usados para definir pesos armónicos en la parametrización.

### 3.2.3. Discretización del mapeo

Para adaptar los conceptos previos de energía de Dirichlet (ecuación 20) y energía conforme (ecuación 21) a superficies trianguladas, solo es necesario discretizar el mapa  $U$  considerando que transforma cada triángulo en  $\mathcal{X}$  a un triángulo en  $\mathcal{U}$ . En éste caso, la energía de Dirichlet puede expresarse como

$$E_D(U) = \sum_{e_{ij}} \frac{1}{4} (\cot(\theta_{ij}) + \cot(\theta_{ji})) (u_i - u_j)^2, \quad (25)$$

donde  $\theta_{ij}$  y  $\theta_{ji}$  son dos ángulos opuestos al arista que une a  $x_i$  y  $x_j$  en la malla  $\mathcal{X}$ . Como la ecuación anterior es cuadrática en las coordenadas de  $\mathcal{U}$ , puede ser descrita de forma matricial como

$$E_D(U) = \frac{1}{2} U^t L_D U, \quad (26)$$

donde  $L_D$  es la matriz simétrica y dispersa de  $2V \times 2V$  que contiene entradas correspondientes a la suma de cotangentes en la ecuación 25.

Por otro lado, el área de la parametrización  $\mathcal{U}$  puede ser calculada sumando las áreas  $\mathcal{A}_{\mathcal{T}}$  de cada triángulo  $\mathcal{T}$  en  $\mathcal{U}$  a través de la ecuación  $\mathcal{A}_{\mathcal{T}}(U) = \sum_{e_{ij} \in \mathcal{T}} \frac{1}{2} (u_i v_j - u_j v_i)$ . Como los aristas internos se cancelan mutuamente, el área total puede ser calculada en términos

de únicamente las coordenadas de los vértices de la frontera, de manera que

$$\mathcal{A}(U) = \sum_{e_{ij} \in \partial\mathcal{U}} \frac{1}{2} (u_i v_j - u_j v_i), \quad (27)$$

donde el arista  $e_{ij}$  se encuentra en la frontera  $\partial\mathcal{U}$ .

Luego, podemos definir una matriz  $A$  tal que  $\mathcal{A}(U) = \frac{1}{2} U^t A U$ . Esta matriz, también de  $2V \times 2V$ , es simétrica y extremadamente dispersa, pues sólo las entradas que corresponden a vértices en la frontera son diferentes de cero.

Finalmente, la energía conforme  $E_C = E_D - \mathcal{A}$  tiene una representación cuadrática como

$$E_C(U) = \frac{1}{2} U^t L_C U, \quad (28)$$

donde  $L_C = L_D - A$ . En conclusión, el mapa discreto  $U$  es conforme cuando minimiza la energía cuadrática discreta  $E_C(U)$ .

### 3.2.4. Solución a la parametrización

Para resolver el problema de minimizar la expresión 28, se sigue el enfoque de Mullen *et al.* (2008) donde se utiliza la parametrización  $U$  que soluciona el problema de encontrar el eigenvector correspondiente al eigenvalor más pequeño distinto de cero que satisface

$$L_C U = \lambda B U, \quad (29)$$

donde  $B$  es una matriz diagonal de  $2V \times 2V$  de la forma

$$B_{i,j} = \begin{cases} 1 & e_{ij} \in \partial\mathcal{U} \\ 0 & \text{de otra manera.} \end{cases} \quad (30)$$

El trabajo de Mullen *et al.* (2008) busca la disminución de la sensibilidad a la irregularidad del muestreo en la triangulación original al asignar un peso en el funcional de área  $\mathcal{A}$  y la energía de Dirichlet  $L_D$ ; con este propósito, cada triángulo es pesado con el inverso de su área original  $|\mathcal{T}|$  en  $\mathcal{X}$ .

La ventaja de esta metodología es que, a diferencia de otros métodos lineales de frontera libre, esta no requiere fijar correspondencias previas como el enfoque de mínimos cuadrados, reduciendo así la distorsión típica que dicho proceso genera.

### **3.3. Resumen y comparativa de parametrizaciones conformes**

En general, el propósito de la parametrización de una superficie triangulada es obtener un mapeo entre la malla tridimensional y una triangulación en 2D tal que se asocie cada triángulo de la malla original con un triángulo en el dominio plano. Una característica importante para ciertas aplicaciones es que dicho mapeo sea biyectivo, al menos localmente, para evitar que los triángulos se traslapen al emplear la transformación.

Para este trabajo se adoptó un enfoque en que no se cambia la topología del rostro. Específicamente, se considera que las superficies faciales no tienen agujeros ni fronteras internas y que ésto no cambia en presencia de expresiones, por ejemplo, al abrir la boca.

Al trabajar sobre mallas discretas, un enfoque muy utilizado para el cálculo de parametrizaciones es aproximar transformaciones armónicas o conformes. La desventaja de utilizar funciones armónicas es que estas técnicas requieren fijar la frontera mientras que los mapeos conformes pueden ser calculados sin esta restricción. Respecto a transformaciones conformes, un enfoque popular es el de aproximar las ecuaciones de Cauchy-Riemann por medio de mínimos cuadrados, metodología que no es sensible a cambios de resolución, no necesita restricciones de la frontera, que no manifiesta un impacto relevante frente a la oclusión de la frontera y es un método lineal (Wang *et al.* (2007)); sin embargo, tiene la restricción de que requiere de fijar puntos interiores en la parametrización para evitar soluciones triviales (constantes) del sistema matricial y para calcular a partir de dichas correspondencias el resto del mapeo.

En el trabajo de Wang *et al.* (2007), los mapeos conformes por mínimos cuadrados muestran superioridad ante aproximaciones de mapas armónicos al considerar su aplicación a correspondencia de superficies 3D. En el apéndice A se muestran comparaciones mostradas por Zeng y Gu (2009) de distintas transformaciones sobre rostros en presencia de variación de postura y oclusiones.

Las transformaciones conformes utilizadas en este trabajo se calcularon mediante el esquema de parametrización espectral conforme propuesta en Mullen *et al.* (2008). Se utiliza este método que no requiere de fijar correspondencias, maneja una forma de frontera libre y se resuelve a través de la solución de un problema de eigenvalores en lugar de buscar la solución a un sistema lineal directamente. Además, busca la minimización de una energía conforme con pesos que se presume insensible a la irregularidad en el muestreo de la malla original.

A continuación se muestra una tabla que compara las técnicas presentadas para calcular parametrizaciones. Se sabe que los métodos que asumen fija la frontera en el dominio plano usualmente son más rápidos que los que determinan la frontera como parte de la solución, sin embargo, estos últimos presentan una menor distorsión, lo cual resulta atractivo para aplicaciones de reconocimiento.

**Tabla 2: Tabla comparativa de técnicas usuales para obtener parametrizaciones.**

	<b>Mapeos armónicos</b>	<b>Mapeos conformes</b>	<b>Mapeos conformes por mínimos cuadrados</b>	<b>Parametrización espectral conforme</b>
Cambios de resolución	Insensible	Insensible	Insensible	Insensible
Restricción de frontera	Sí	No	No	No
Oclusión de fronteras	Difícil de manejar	Sin impacto significativo	Sin impacto significativo	Sin impacto significativo
Uso de puntos interiores fijos	No usa	Usa 2 puntos	Usa al menos 2	No usa
Complejidad	Lineal	No lineal (con aproximaciones lineales)	Lineal	Lineal

## Capítulo 4. Métodos clásicos para la correspondencia de superficies faciales

---

### 4.1. Análisis de Componentes Principales

Una de las técnicas más populares para tratar el problema de reconocimiento de rostros es el análisis de componentes principales (PCA, por sus siglas en inglés). En general, PCA es un método estadístico que transforma linealmente un conjunto de datos a otros denominados componentes principales; cada componente contiene la varianza de los datos originales: el primer componente contiene la máxima varianza, el segundo contiene la segunda máxima varianza y así sucesivamente. Puede verse como una compresión de información, motivo por el cual es un método aprovechado para realizar reducción de dimensiones.

Esta técnica ha sido ampliamente utilizada en el marco de reconocimiento de rostros usando imágenes de intensidad, se ha aplicado a bases de datos grandes con el fin de realizar reconocimiento obteniendo un buen rendimiento cuando las imágenes eran capturadas bajo condiciones uniformes.

En el marco de la representación bidimensional de los rostros, donde una imagen consiste en una matriz cuadrada de  $N \times N$  valores de intensidad, un rostro se representa como un vector de longitud  $N^2$  que se forma por la concatenación de todos los valores de los píxeles en la imagen. Luego, este vector se puede considerar como un elemento del espacio vectorial de todas las imágenes cuya representación como vector es de tamaño  $N^2$  sin importar el contenido de la imagen.

El reconocimiento de rostros basado en PCA considera que hay un subespacio que corresponde al conjunto de imágenes de rostros. La metodología consiste en buscar los vectores de la base de este subespacio de rostros, a los cuales se les denomina eigenfaces, de manera que cualquier imagen de un rostro pertenece a este subespacio y se puede escribir como una combinación lineal de eigenfaces. Los coeficientes de dicha combinación lineal conforman un vector de características que representa al rostro en cuestión, de manera que se puede construir una base de datos que consiste en estos

vectores de características de los individuos conocidos.

La metodología inicia considerando que se tienen  $M$  imágenes de rostros de tamaño  $N \times N$ , pertenecientes a distintos individuos. Sea  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M$  un conjunto de imágenes de rostros de la base de datos en su representación como vector, donde  $\mathbf{x}_i \in \mathbb{R}^d$  con  $d = N \times N$ .

Se espera que, al verlos como puntos, los rostros de un mismo individuo se encuentren aglomerados en el espacio de rostros  $\mathbb{R}^d$  (altamente dimensional), de forma que una técnica de clasificación sea capaz de separar los rostros por identidades.

Es conveniente considerar la media de las imágenes como el origen para hacer un análisis de su variación, por ello se calcula la media del conjunto de imágenes de entrenamiento

$$\mu = \frac{1}{M} \sum_{i=1}^M \mathbf{x}_i, \quad (31)$$

y se calculan los rostros centrados  $\varphi_i = x_i - \mu$ , es decir, los vectores que indican por cuánto difiere cada rostro de la media. Después se forma la matriz  $D$  de tamaño  $d \times M$  cuyas columnas son los rostros centrados.

Los vectores de la base del subespacio de rostros se obtienen a partir de los eigenvec-tores asociados a los eigenvalores de la matriz de covarianza de  $D$

$$C = \frac{1}{M} \sum_{i=1}^M \varphi_i \varphi_i^T = \frac{1}{M} D D^T. \quad (32)$$

Para manejar la gran dimensionalidad de  $C$  (de tamaño  $d \times d$ ), se realiza el siguiente procedimiento: dada la matriz de covarianza  $C$ , su descomposición por eigenvectores es

$$C \mathbf{v}_i = D D^T \mathbf{v}_i = \lambda \mathbf{v}_i. \quad (33)$$

Como la matriz  $D D^T$  es muy grande, se considera la descomposición

$$D^T D \mathbf{r}_i = \lambda \mathbf{r}_i \quad (34)$$

y la expresión resultante de multiplicar esta ecuación por  $D$

$$DD^T D\mathbf{r}_i = D\lambda\mathbf{r}_i. \quad (35)$$

Esto significa que si  $\mathbf{r}_i$  es un eigenvector de  $D^T D$ , entonces  $D\mathbf{r}_i$  es un eigenvector de  $C = DD^T$ . La relevancia de este procedimiento es que si se tienen  $M = 50$  imágenes de  $256 \times 256$  píxeles, su representación como vector sería de tamaño  $d = 65,536$  y calcular la matriz  $D^T D$  de tamaño  $2,500 \times 2,500$  es preferible a manejar la matriz de covarianza  $DD^T$  de  $65,536 \times 65,536$ .

Para calcular los eigenvectores suele utilizarse descomposición en valores singulares. Los eigenvectores obtenidos se normalizan para obtener un sistema de vectores  $\Phi$  que formen una base ortonormal tal que

$$\forall \phi_i, \phi_j \in \Phi, \quad \phi_i \cdot \phi_j = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases} \quad (36)$$

En la práctica suelen tomarse sólo los eigenvectores normalizados que corresponden a los  $k$  eigenvalores más grandes (usualmente de 20 a 50 eigenvalores), puesto que proveen la información más útil para caracterizar la variación entre los rostros. Finalmente, se construye la matriz de proyección de tamaño  $d \times k$

$$\Phi_{PCA} = [\phi_1 \phi_2 \dots \phi_k], \quad (37)$$

a través de la cual se obtiene una representación de los rostros en un espacio de dimensión menor. Un rostro  $\mathbf{x}_i$  pasa de tener una representación en  $\mathbb{R}^d$  a  $\mathbb{R}^k$  bajo la proyección

$$\mathbf{x}'_i = (\mathbf{x}_i - \mu)\Phi_{PCA}. \quad (38)$$

A través de esta proyección es que se obtienen los vectores de características de la base de datos (denominado conjunto de entrenamiento) y de los rostros nuevos que se quieran identificar (conjunto de prueba). Para realizar el proceso de clasificación se

compara el vector de características de un individuo desconocido con los contenidos en la base de datos mediante alguna función de distancia, de manera que aquél vector de la base de datos que obtenga la mínima distancia corresponde al individuo más parecido al que se desea identificar.

Es posible crear un clasificador que determine que el rostro no pertenece a ninguna identidad conocida si la distancia de su vector de características a los de la base de datos es mayor a cierto umbral, sin embargo, en este trabajo se considera que todo rostro de prueba pertenece a un individuo conocido.

#### 4.2. Algoritmo iterativo del punto más cercano

La técnica del punto iterativo más cercano es una técnica usada ampliamente para el alineamiento utilizando información geométrica proveniente de un escáner 3D (Rusinkiewicz y Levoy (2001)). El algoritmo inicia con dos mallas tridimensionales y una idea inicial de cómo se encuentran orientadas una con respecto de la otra para después calcular una transformación entre ellas que se refina iterativamente minimizando una métrica de error.

Se han desarrollado variantes del algoritmo modificando las siguientes etapas base.

1. **Selección** de un conjunto de puntos de una o ambas mallas.

Para esta etapa se pueden utilizar todos los puntos disponibles, una selección aleatoria o puntos con el mayor gradiente.

2. Buscar la **correspondencia** de los puntos seleccionados.

Se determina el punto más cercano en el rostro de referencia para cada punto en la superficie a ser transformada, generando así pares de puntos correspondientes entre las superficies. El punto más cercano se determina tomando aquél que tenga la mínima distancia euclidiana al punto en cuestión. El proceso de buscar el punto más cercano puede acelerarse usando un k-d árbol.

3. Asignar **pesos** a las parejas de puntos correspondientes.

4. **Rechazo** de ciertos pares de puntos.

5. Asignar una **transformación rígida**.

Se determinan parámetros de una transformación rígida y la transformación es aplicada al rostro problema. Dicha transformación consiste en aplicar una matriz de rotación  $R$  y un vector de traslación  $T$ .

6. **Minimizar** la métrica de error.

Se calcula el error correspondiente a la transformación y se busca minimizarlo iterativamente. El alineamiento iterativo se detiene al alcanzar un número máximo de iteraciones o cuando la diferencia del error respecto a una iteración anterior no alcanza cierto umbral.

La métrica de error define una función a minimizar en cada iteración del algoritmo. Usualmente, se elige la raíz cuadrada del error cuadrático medio (RMSE, por sus siglas en inglés), suponiendo que la distancia RMSE es pequeña entre rostros del mismo individuo en comparación a la distancia con un rostro de un individuo distinto.

A continuación se muestran dos variantes muy utilizadas para la minimización de la métrica de error. Se parte de considerar que los puntos  $p_i, i \in \{1, \dots, N\}$  en el rostro de prueba están en correspondencia con un conjunto de puntos  $q_i$  del rostro de referencia.

*Minimización punto a punto.* Suma el cuadrado de las distancias de los puntos originales a los de prueba que han sido transformados. El error en este caso se puede expresar como

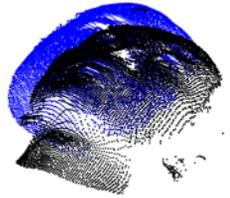
$$E = \sum_{i=1}^N \|Rp_i + T - q_i\|^2. \quad (39)$$

*Minimización punto a plano.* Suma las distancias de los puntos originales a los planos tangentes en los cuales los puntos de prueba residen. Puede expresarse por

$$E = \sum_{i=1}^N [(Rp_i + T - q_i) \cdot n_i]^2, \quad (40)$$

donde  $n_i$  denota la normal del plano tangente sobre el punto  $p_i$  al punto  $q_i$  en la superficie de referencia.

Una de las ventajas de utilizar este algoritmo para tratar el problema de correspondencia de superficies faciales es que permite hacer el emparejamiento entre una región parcial del rostro contra una cara completa; además, como se ilustra en la figura 17, el algoritmo es robusto a incluir pequeñas zonas que no correspondan al rostro.



**Figura 17: Ejemplo de aplicar ICP sobre nubes de puntos 3D.**

Una desventaja de este método bajo el marco de reconocimiento de rostros en 3D es que el algoritmo se queda estancado en un óptimo local cuando hay grandes variaciones de postura en los rostros (Smeets 2012).

#### **4.3. Evaluación en el reconocimiento de rostros**

Como se mencionó anteriormente, el proceso de reconocimiento de rostros parte usualmente de una división de la base de datos en una galería de referencia y un conjunto de rostros de prueba, siendo que los rostros de la galería sirven de entrenamiento para ciertos métodos estadísticos como análisis de componentes principales. Para asegurarse que los conjuntos de entrenamiento y de prueba sean representativos de cada identidad, suele hacerse un muestreo aleatorio; para evitar que la elección de las muestras produzca sesgo, se repite el proceso de entrenamiento y prueba con distintas muestras aleatorias y, al final, se toma el promedio de los errores individuales en cada experimento.

El enfoque del algoritmo propuesto es hacia propiedades geométricas de manera que, para llevar a cabo el reconocimiento, a cada rostro se le asocia un vector de características que contenga información discriminante obtenida de las superficies faciales originales. Así, si el vector de características de un rostro a reconocer es similar al que caracteriza un individuo (obtenido del conjunto de entrenamiento), se le asigna dicha identidad conocida. La representación de los rostros a través de un vector facilita las comparaciones entre ellos simplemente asignando un valor de similitud.

### 4.3.1. Medidas de similitud

Un algoritmo de reconocimiento requiere de un valor escalar para llevar a cabo una decisión en el proceso de identificación. En la tabla 3 se enlista las medidas más comunes. Los vectores  $p_i$  y  $q_i$  representan puntos pertenecientes a un rostro de prueba y de referencia, respectivamente, entre los cuales se quiere establecer una correspondencia.

**Tabla 3: Medidas de similitud comunes.**

Medida de similitud	Definición
Distancia euclidiana	$D(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$
Distancia de Mahalanobis	$D(p, q) = \sqrt{\sum_{i=1}^n \sum_{j=1}^n \frac{(p_i - q_i)^2}{\sigma_{ij}^2}}$
Distancia raíz media cuadrática	$D(p, q) = \sqrt{\frac{1}{n} \sum_{i=1}^n (p_i - q_i)^2}$
Distancia de coseno	$D(p, q) = 1 - \frac{\sum_{i=1}^n p_i q_i}{\sqrt{\sum_{i=1}^n p_i^2} \sqrt{\sum_{i=1}^n q_i^2}}$

### 4.3.2. Curva característica de correspondencia acumulada

El objetivo principal de calcular métricas de similitud entre un rostro prueba y los de la galería es asignar una identidad en base a la mejor correspondencia. En este trabajo se considera el problema de reconocimiento o identificación de rostros, no sólo de verificación de identidad; además, se realizan correspondencias entre rostros en un conjunto cerrado, es decir, se asume que cada rostro de prueba pertenece a una identidad en la galería. Para realizar una evaluación de este tipo de metodología, es recomendable tomar en cuenta no solo el porcentaje de identificaciones correctas, sino también los casos en que la identidad correcta se encuentra dentro de las mejores  $n$  correspondencias.

Considerando una galería  $\mathcal{G}$  y un conjunto de prueba  $\mathcal{P}$ , se tiene que cada prueba  $p \in \mathcal{P}$  posee un valor de distancia hacia cada rostro conocido  $q \in \mathcal{G}$ . Se dice entonces que el rostro  $p$  tiene un rango  $k$  si la identidad correcta pertenece a la  $k$ -ésima mejor correspondencia, es decir, a la  $k$ -ésima menor distancia.

Usando este concepto, el número de pruebas con rango  $n$  o menor se puede expresar como

$$C(n) = |\{p \in \mathcal{P} : \text{rango}(p) \leq n\}|. \quad (41)$$

Así, la razón de identificación de rango  $n$ ,  $P_{Id}(n)$ , es la fracción de rostros prueba con

rango  $n$  o menor, es decir, el cociente

$$P_{Id}(n) = \frac{C(n)}{|\mathcal{G}|}. \quad (42)$$

A la tasa de identificación en rango  $n = 1$ ,  $P_{Id}(1)$ , también se conoce como la tasa de identificación correcta o tasa de mejor correspondencia.

La representación de la razón de identificación  $P_{Id}$  como función del rango se conoce como curva característica de correspondencia acumulada (CMC, por sus siglas en inglés), siendo la tasa de reconocimiento de rango 1 el valor de la curva CMC más utilizado (Bowyer *et al.* (2006)). Esta curva es no decreciente y para el rango máximo,  $|\mathcal{P}|$ , la curva tiene valor 1. El rango 1 resume el desempeño de la técnica, pero una descripción más amplia suele hacerse a través de incluir el rango 5, 10 y 20 al manejarse bases de datos extensas. Lo que hay que tomar en consideración es que una buena calidad en una técnica de reconocimiento se ve reflejada en valores altos de la curva para rangos bajos.

#### 4.4. Resumen

En este capítulo se describieron métricas de desempeño comúnmente utilizadas en reconocimiento de rostros. Se presentan también técnicas tradicionales de reconocimiento de rostros tridimensionales basados en métodos estadísticos que se incorporaron a la metodología propuesta y sirvieron de base para realizar una comparación con el rendimiento del método. Las técnicas consisten en emplear una proyección de los rostros a un subespacio produciendo una reducción en la dimensión de los datos y la segunda permite mejorar la correspondencia entre rostros de prueba y de referencia de manera iterativa.

## Capítulo 5. Metodología propuesta

---

En este capítulo se describe el algoritmo de reconocimiento de rostros 3D propuesto que combina técnicas estadísticas tradicionales de reconocimiento de rostros con análisis de curvatura y parametrizaciones conformes.

Cabe resaltar que se propone el uso de nubes de puntos 3D (sin información de color) como entrada al algoritmo de reconocimiento con el objeto de eliminar la dependencia a factores que no afectan la forma tridimensional de las superficies faciales tales como el color de piel y uso de maquillaje.

De manera general, el algoritmo propuesto sigue el procedimiento ilustrado a continuación.

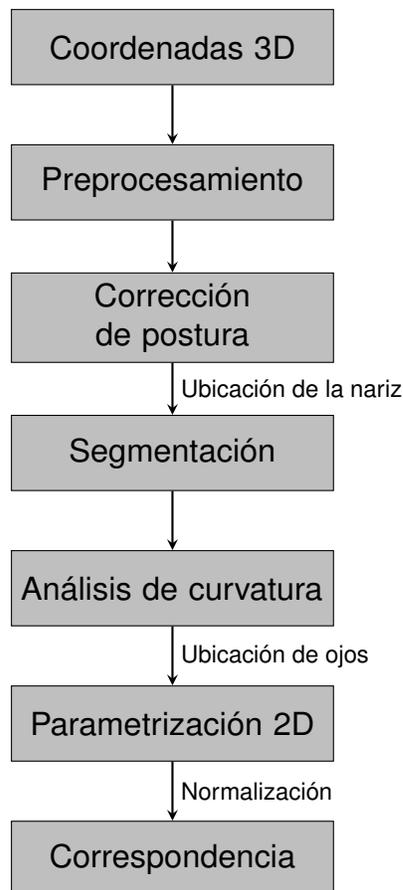


Figura 18: Esquema general del método de reconocimiento propuesto.

### **5.1. Preprocesamiento**

El algoritmo parte de extraer una nube de puntos 3D de los archivos contenidos en la base de datos. Primeramente, se aplica un filtro de mediana que tiene por objetivo eliminar el ruido de la nube de puntos. Después, se establecieron conexiones entre los puntos definiendo una triangulación de Delaunay. Luego, se aplica un filtro laplaciano para suavizar la superficie triangulada y eliminar zonas que no hayan sido muestreadas, es decir, rellenar agujeros internos en el rostro.

### **5.2. Corrección de postura**

Esta etapa tiene el fin de homogenizar la postura de las superficies faciales a través de una rotación y una traslación definida por el método ICP (descrito en la sección 4.2) tomando como referencia un rostro frontal único para todos los rostros. Se consideró la variante de punto a plano utilizando un árbol de búsqueda y el rechazo del 30% de las peores correspondencias entre los puntos en cada iteración.

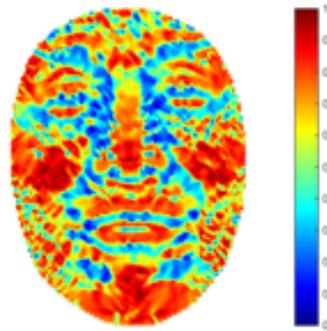
Esta alineación es de gran relevancia para la segmentación inicial del rostro, puesto que espera que la nariz quede ubicada en el centro del rostro para que se realice correctamente la segmentación que a continuación se detalla.

### **5.3. Segmentación**

Esta etapa permite excluir información extra de hombros, orejas, cabello y cuello que no son relevantes para el reconocimiento.

Se automatizó la tarea de segmentación a través de la ubicación de la punta de la nariz de acuerdo al diccionario de formas que brinda el índice de forma SI presentado en la sección (2.3.2.2). Primeramente, se calcula el valor de SI para cada punto en el rostro; después, se extraen los puntos con valores de SI en el intervalo  $[0.9, 1]$  para considerarlos como candidatos y, finalmente, se selecciona aquel con la coordenada  $z$  mayor.

Para realizar el corte del rostro, se seleccionan los puntos sobre la superficie facial que se encuentran a una distancia menor o igual a 100 de la punta de la nariz. Esto puede



**Figura 19: Resultado de la segmentación de un rostro. Cada punto contiene información de su índice de forma SI.**

verse como centrar en la nariz una esfera de radio 100 y conservar únicamente los puntos que encuentran dentro de ella.

Por último, se traslada la superficie de tal manera que la punta de la nariz quede ubicada en el origen.

#### **5.4. Análisis de curvatura**

Una de las ventajas de utilizar superficies faciales en lugar de imágenes de intensidad recae en la posibilidad de calcular descriptores de curvatura que diferencian formas sobre los rostros y permiten distinguir entre mejillas, nariz, frente y barbilla. Estos descriptores son de gran relevancia al hacer frente a expresiones faciales, puesto que, si se pretende excluir zonas fuertemente afectadas por gestos, las imágenes 2D tradicionales presentan un escaso gradiente de intensidad en dichas áreas imposibilitando su segmentación automática.

Para conservar información local de las superficies se aplicó el algoritmo de clasificación HK (sección 2.3.2.1) que caracteriza cada punto sobre las superficies tridimensionales de acuerdo a la forma en un entorno local de los mismos.

##### **5.4.1. Localización automática de puntos**

Además de crear un mapa general de las formas sobre las superficies faciales, el análisis de curvatura realizado se utilizó para ubicar la parte interna de los ojos y refinar la selección de la punta de la nariz.

Se utilizaron umbrales para la curvatura media ( $H$ ) y la curvatura gaussiana ( $K$ ) de cada punto con el fin de seleccionar aquellos sobre las zonas específicas buscadas: la parte interna de los ojos y la punta de la nariz. El umbral para la curvatura media fue  $Thr_H = 0.00005$  y  $Thr_K = 0.0025$  para la curvatura gaussiana. El procedimiento para cada localizar estos tres puntos fue el siguiente:

- Ojos:**
1. Seleccionar puntos con  $H < Thr_H$  y  $K > Thr_K$ .
  2. Extraer puntos con valores del índice de forma SI en el intervalo  $[0.28, 0.32]$ .
  3. Seleccionar los que se encontraran entre los percentiles 40 y 80 para su coordenada  $x$  y entre los percentiles 25 y 75 para su coordenada en  $y$ .
  4. Buscar una pareja de puntos con valor igual en la coordenada  $y$  (aprovechando la simetría del rostro).
- Nariz:**
1. Seleccionar puntos con  $H < Thr_H$  y  $K > Thr_K$ .
  2. Elegir el punto cuyo valor en la coordenada  $z$  sea el mayor.

Cabe señalar que la ubicación automática de puntos aprovecha la simetría y características generales del rostro, por ejemplo, que los ojos se encuentren alineados y que la nariz esté al centro del rostro. Más aún, en rostros que presentan variaciones de postura, la localización de estos puntos depende fuertemente de que la etapa de corrección de pose previa se haya efectuado correctamente.

## 5.5. Parametrización conforme

Esta etapa consiste en aplicar el método de parametrización espectral conforme descrito en la sección 3.2 a cada uno de los rostros. De ésta manera, se transforman las triangulaciones tridimensionales a triangulaciones en el plano que buscan conservar los ángulos originales.

### 5.5.1. Normalización

Una vez que se tiene una representación plana de los rostros, se realiza una normalización basada en la ubicación de los puntos correspondientes a la parte interna de los ojos y a la punta de la nariz.

Primero, se determina el ángulo que forma respecto al eje  $x$  el segmento de recta entre los puntos que caracterizan a los ojos; luego, se aplica una rotación en el plano por dicho ángulo. Así, se considera que el segmento de recta que une a los ojos queda totalmente horizontal. En la figura 20 se ilustra este procedimiento.



(a) Ubicación de los ojos sobre la parametrización. (b) Aplicación de la rotación en el plano.

**Figura 20: Normalización de la parametrización basada en la ubicación de los ojos.**

Después, se aplica una traslación de manera que el punto correspondiente a la punta de la nariz quede ubicado en el origen del plano. Posteriormente, se seleccionan los puntos que se encuentren dentro de un círculo centrado en la nariz, de manera que el rostro queda representado en un círculo unitario después de realizar un escalamiento.

## 5.6. Correspondencia

Para realizar el proceso de correspondencia se utilizaron vectores que concatenan la información geométrica de los rostros. Las características consideradas fueron las coordenadas 2D obtenidas de la parametrización en conjunto con valores de curvatura media en cada punto.

La clasificación se llevó a cabo a través de la aplicación de las técnicas de análisis de componentes principales (PCA) y búsqueda de vecinos cercanos (NN, por sus siglas en inglés) de manera independiente y combinada. Los resultados se muestran en el siguiente capítulo.

### 5.6.1. Segmentación de secciones robustas

El proceso de correspondencia se realizó considerando tres distintas modalidades del rostro. Primeramente, se considera el rostro completo con frontera libre que se obtiene directamente de la parametrización; después se considera la segmentación en forma de círculo unitario explicada anteriormente y, por último, se realiza el proceso de correspondencia limitándolo a la información geométrica correspondiente a un círculo que abarca únicamente la región nasal (con el fin de evitar la influencia de expresiones faciales en el reconocimiento).

### 5.7. Resumen

A continuación se muestra un resumen de las etapas de la metodología propuesta:

**Preprocesamiento:** Tiene por objetivo eliminar el ruido y realizar una interpolación de la malla tridimensional de modo que no se tengan “hoyos” en la superficie.

**Corrección de pose:** Se aplica una transformación a la nube de puntos de manera que el rostro tenga orientación frontal.

**Segmentación:** Consiste en localizar de forma automática la punta de la nariz en cada superficie facial y segmentar a partir de este punto la región comprendida por el rostro eliminando información extra de orejas, cuello y hombros.

**Análisis de curvatura:** En esta etapa se calculan descriptores de la forma del rostro en cada punto.

**Localización automática de ojos:** Se ubican las esquinas internas de los ojos y se refina la selección de la punta de la nariz con base en medidas de curvatura.

**Parametrización conforme:** Se transforma la triangulación tridimensional del rostro original a una malla triangulada en 2D que busca conservar ángulos en el plano.

**Normalización:** Se aplica una rotación de los rostros en el plano con base en los puntos caracterizados como esquinas internas de los ojos y se hace una traslación de manera que la punta de la nariz sea el origen.

**Cálculo de correspondencias:** Se lleva a cabo una clasificación de los rostros a través de su representación como vectores de características combinando las técnicas de PCA y búsqueda de vecinos cercanos.

**Extracción de regiones invariantes a expresiones:** Los experimentos se llevan a cabo considerando también distintas zonas del rostro que se espera sean invariantes a expresiones.

## Capítulo 6. Experimentos y Resultados

---

En este capítulo se describen los experimentos realizados para la evaluación de la metodología propuesta. Los experimentos están diseñados para evaluar el desempeño de distintas combinaciones de información geométrica junto a diferentes métodos de correspondencia.

Se hace la comparación entre utilizar las coordenadas 2D provenientes de la parametrización descrita en el capítulo 3.2 de manera aislada y junto a curvatura media. Es de interés conocer su robustez frente a expresiones y variaciones de postura de manera separada, así como evaluar su rendimiento al añadir dos tipos de segmentación a cada experimento. La comparación de los resultados se hace ante el uso tradicional de coordenadas 3D de cada vértice del rostro original.

Todas las técnicas para el algoritmo de reconocimiento se implementaron en Matlab 2014b. La implementación de la parametrización se basa en el Matlab Mesh Toolkit<sup>1</sup> para calcular la parametrización espectral conforme; este toolbox usa la función *eigs* de las librerías de Matlab para resolver el problema de eigenvalores y eigenvectores asociados bajo la función *embedSCP*. La correspondencia se hizo a través de análisis de componentes principales (PCA) y el cálculo del vecino más cercano (NN). Para aplicar la técnica PCA se separó la base de datos en conjuntos de entrenamiento y conjuntos de prueba considerando en su implementación el tomar todos los componentes principales posibles ( $M - 1$ , donde  $M$  es el número de imágenes de entrenamiento).

En las siguientes secciones se describen la base de datos utilizada, los conjuntos de entrenamiento y prueba formados para cada experimento y los resultados de la evaluación variando la complejidad respecto a la inclusión de modelos con expresiones y variaciones de postura.

### 6.1. Base de datos

Los rostros considerados en los experimentos provienen de la base de datos pública CASIA<sup>2</sup>. La selección de esta base de datos fue a partir de considerar que esta contiene

---

<sup>1</sup><http://www.dgp.toronto.edu/~rms/software/matlabmesh/>

<sup>2</sup><http://biometrics.idealtest.org>

modelos faciales 3D con una variación considerable respecto a expresiones y variaciones de postura en comparación a otras bases de datos. Una característica atractiva es que provee información acerca de qué tipo de expresión se manifiesta en cada modelo, en qué rango se encuentra la variación de postura de la cabeza y una posición aproximada de la fuente de iluminación.

Las expresiones consideradas fueron sonrisa, enojo, risa, sorpresa y ojos cerrados (figura 21). Las variaciones de postura extraídas fueron voltear a la derecha e izquierda con una rotación de 20 a 30 grados , girar la cabeza hacia arriba y abajo de 20 a 30 grados e inclinar la cabeza de 20 a 30 grados hacia la izquierda y derecha (véase figura 22). Respecto a la variación en iluminación, los rostros que se incluyeron fueron adquiridos bajo luz de oficina, iluminación desde arriba, abajo, izquierda y derecha.

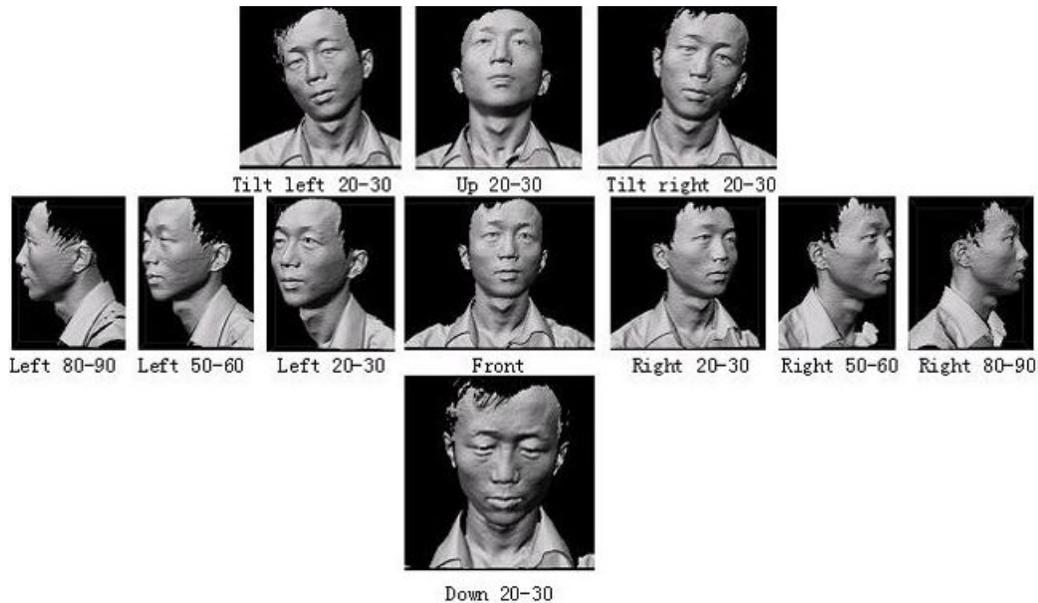
Se formaron grupos de 5, 10 y 30 identidades con sus respectivos rostros neutrales, con expresiones y con variación de postura e iluminación para verificar el rendimiento de la metodología propuesta frente al número de individuos.

En general, se consideraron dos grupos de entrenamiento. El primero consiste en 5 modelos con expresión neutra y postura frontal, uno para cada variación de iluminación contemplada. El segundo conjunto de entrenamiento consiste en 5 modelos adquiridos bajo iluminación de oficina y postura frontal, incluyendo un rostro perteneciente a cada expresión antes mencionada.



**Figura 21: Gestos de la base de datos CASIA para un mismo individuo adquiridos bajo iluminación de oficina. (a) Sonrisa (b) Risa (c) Enojo (d) Sorpresa (e) Ojos cerrados. Extraído de <http://biometrics.idealtest.org/dbDetailForUser.do?id=8>.**

Se formó un tercer conjunto de entrenamiento que sirvió para verificar el rendimiento del método propuesto al incluir variaciones de postura. Las características de estos modelos se describen en el experimento 4.



**Figura 22: Variaciones de postura bajo iluminación de oficina contenidas en la base de datos CASIA. Extraído de <http://biometrics.idealtest.org/dbDetailForUser.do?id=8>.**

Es preciso señalar que el no usar color en las imágenes 3D contribuye a la independencia del reconocimiento respecto a ciertas condiciones de adquisición de las imágenes, en particular, si se busca liberarlo de la iluminación controlada.

## **6.2. Experimento 1: Resultados de reconocimiento incluyendo la etapa de corrección de postura**

Antes de evaluar el rendimiento del método propuesto mediante conjuntos de entrenamiento grandes y con variaciones de expresión e iluminación considerables, se optó por verificar la aportación de la corrección de postura previa al cálculo de la parametrización conforme.

Se hizo una evaluación con 10 individuos utilizando la parametrización conforme junto a información de curvatura media y tomando una segmentación circular del rostro. Todos los modelos utilizados en este caso fueron adquiridos con iluminación de oficina. En la tabla 4 se muestra un ejemplo de los resultados de utilizar un modelo neutral contra otro neutral (neutral vs. neutral), dos modelos neutrales contra un gesto de sonrisa (neutrales vs. sonrisa) y dos modelos neutrales contra una expresión de enojo (neutrales vs. enojo).

Debido a que los modelos sin corrección de pose obtuvieron mejores resultados en

**Tabla 4: Resultados de reconocimiento utilizando segmentación circular de la parametrización junto a curvatura media aplicando PCA y distancia de coseno.**

Variaciones	Con normalización de pose	Sin normalización de pose
Neutral vs. neutral	30	60
Neutrales vs. sonrisa	50	60
Neutrales vs. enojo	30	50

este experimento preliminar, se optó por no considerar una normalización de postura en experimentos posteriores. Esto permitió que el método probara su robustez ante cambios de postura y no ante modelos con información incompleta producto de la etapa de corrección de pose.

### 6.3. Experimento 2: Reconocimiento automático de rostros con expresión neutral y comparación de resultados con distintas medidas de similitud

Este experimento se llevó a cabo con la finalidad de seleccionar una función de distancia que tuviese buen rendimiento en conjunto con el método propuesto.

El conjunto de entrenamiento utilizado en este experimento se formó con 5 modelos con expresión neutra y postura frontal (con variaciones en la iluminación con que fueron adquiridas). La prueba se hizo frente a un rostro adquirido con iluminación desde la derecha, expresión neutra y postura frontal (diferente a los modelos del conjunto de entrenamiento).

**Tabla 5: Resultados de reconocimiento utilizando información geométrica 3D sobre conjuntos de rostros con expresión neutra comparando distintas métricas de similitud.**

No. de individuos	Euclidiana	Coseno
	PCA+NN	PCA+NN
5 individuos		
Vértices 3D	100	100
Vértices 2D + curvatura media	100	100
10 individuos		
Vértices 3D	100	100
Vértices 2D + curvatura media	80	100
30 individuos		
Vértices 3D	93.33	96.66
Vértices 2D + curvatura media	86.66	93.33

La correspondencia se hizo bajo Análisis de Componentes Principales (PCA) y búsqueda del vecino más cercano (NN) mediante distancias euclidiana y de coseno. Se analizó el rendimiento de la combinación de información de curvatura junto a información 3D original y la parametrización conforme. El análisis del reconocimiento de coordenadas 3D originales se muestra en la tabla 5 y el del método propuesto en las tablas 6 y 7 para segmentaciones circular y de la zona nasal respectivamente.

### 6.3.1. Resultados con segmentación circular del rostro

Los resultados utilizando coordenadas tridimensionales mostraron que la distancia de coseno obtiene un mejor grado de reconocimiento (tabla 5). Para el caso de la metodología propuesta, en la tabla 6 puede verse que, en general, la distancia coseno también obtiene mejores resultados en comparación a la euclidiana. El aplicar una segmentación circular del rostro mejora el grado de reconocimiento a pesar de que se aumente el número de individuos en el experimento.

**Tabla 6: Resultados de reconocimiento utilizando parametrización conforme y curvatura media sobre conjuntos de rostros con expresión neutra comparando distintas métricas de similitud.**

No. de individuos	Euclidiana	Coseno
	PCA+NN	PCA+NN
5 individuos		
Parametrización 2D	100	100
Parametrización + curvatura media (frontera libre)	60	80
Parametrización + curvatura media + segmentación circular	40	100
10 individuos		
Parametrización 2D	90	80
Parametrización + curvatura media (frontera libre)	40	60
Parametrización + curvatura media + segmentación circular	60	90
30 individuos		
Parametrización 2D	60	53.33
Parametrización + curvatura media (frontera libre)	23.33	36.66
Parametrización + curvatura media + segmentación circular	60	83.33

### 6.3.2. Resultados usando la sección circular en torno de la nariz

Con el fin de seleccionar la distancia adecuada para la correspondencia usando segmentación de zonas invariantes del rostro, se comparó el grado de reconocimiento de la combinación de la parametrización, curvatura media y una segmentación de la zona nasal de los rostros. Los resultados se muestran en la tabla 7.

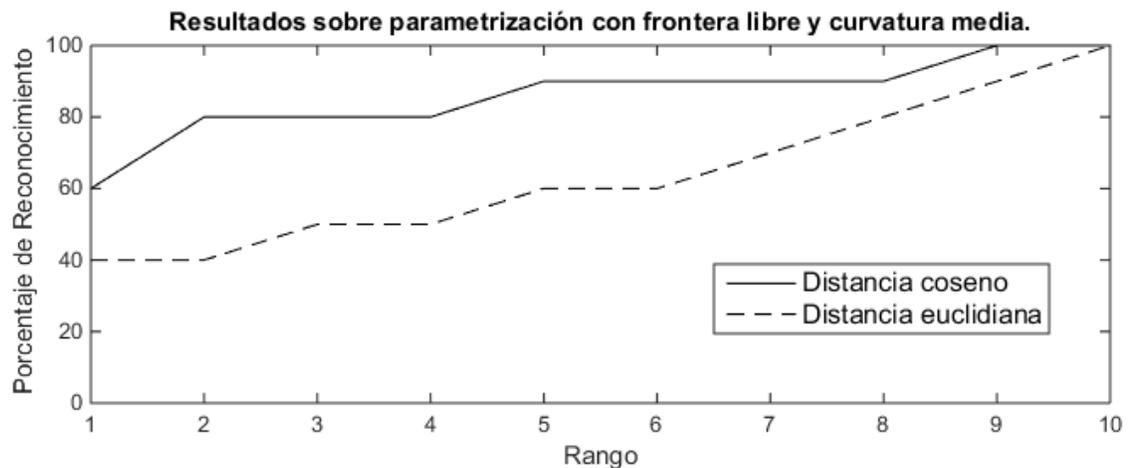
**Tabla 7: Comparación de resultados de reconocimiento a partir de parametrización conforme y curvatura media sobre rostros con expresión neutra aplicando segmentación de la zona nasal y utilizando distintas métricas de similitud.**

No. de individuos	Euclidiana	Coseno
	PCA+NN	PCA+NN
5 individuos	80	80
10 individuos	60	80
30 individuos	60	73.33

El bajo rendimiento de los resultados utilizando una segmentación nasal respecto a la segmentación circular se explica por el hecho de que, para varios individuos, la nariz pudo haberse localizado erróneamente. Sin embargo, dichos resultados fueron superiores a los obtenidos por la parametrización con frontera libre.

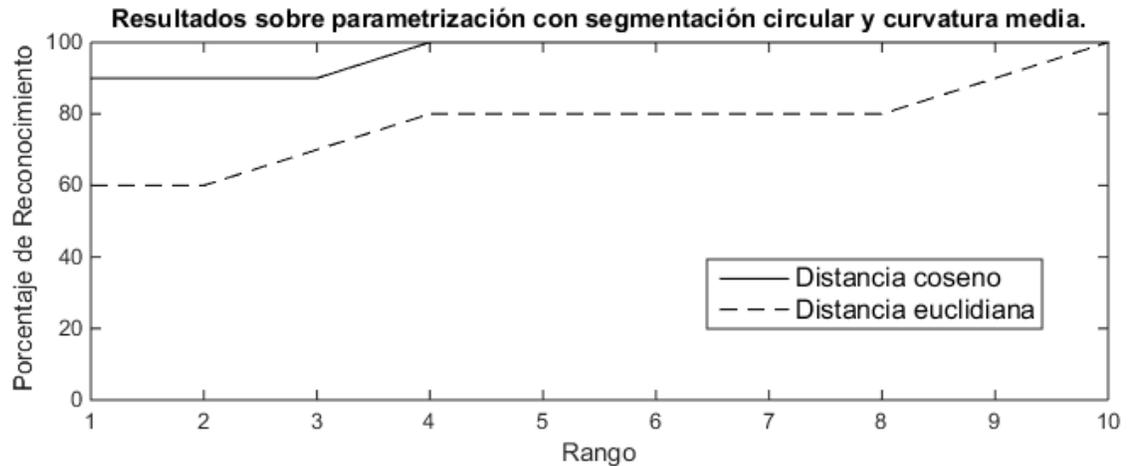
### 6.3.3. Discusión de resultados

A pesar de que los resultados no fueron muy buenos bajo la elección del conjunto de entrenamiento utilizado, el experimento logró mostrar la superioridad en cuanto a capacidad de reconocimiento de la distancia de coseno. A continuación, en las figuras 23 y 24 se muestran las curvas de correspondencia acumulada al utilizar frontera libre y segmentación circular del rostro para las dos distancias utilizando 10 individuos



**Figura 23: Comparación de resultados de distancia euclidiana y coseno utilizando parametrización con frontera libre y curvatura media sobre 10 individuos.**

Bajo este conjunto de entrenamiento de expresiones neutras, al utilizar una segmentación circular la distancia coseno encuentra la identidad correcta para todos los individuos



**Figura 24:** Comparación de resultados de distancia euclidiana y coseno utilizando parametrización con segmentación circular y curvatura media sobre 10 individuos.

considerando las 4 identidades más cercanas (figura 24), mientras que la distancia euclidiana no logra alcanzar el 100 % de reconocimiento en este caso. Con base en los resultados de este experimento, se seleccionó la distancia coseno para experimentos posteriores.

#### **6.4. Experimento 3: Reconocimiento de rostros en presencia de expresiones.**

Este experimento se llevó a cabo con la finalidad de probar el rendimiento de la metodología propuesta ante variaciones de expresiones en los rostros. Se seleccionó la información que tuvo mayor capacidad de discriminación en experimentos anteriores: los vértices 3D originales, la parametrización conforme junto a curvatura media aplicando segmentación circular del rostro y segmentación de la región nasal.

La evaluación se hizo frente a los siguientes conjuntos de entrenamiento:

**Neutral (N).** Consiste en 6 modelos con expresión neutra y postura frontal, de los cuales 2 fueron adquiridos con iluminación de oficina y el resto con iluminación desde arriba, abajo, izquierda y derecha.

**Expresiones (E).** Se formó a partir de 5 modelos con expresiones (sonrisa, enojo, risa, sorpresa y ojos cerrados) adquiridas bajo iluminación de oficina y postura frontal.

**Neutral + expresiones (N+E)** Se combinaron los rostros contenidos en los conjuntos de entrenamiento anteriores.

Las pruebas se hicieron evaluando cada conjunto de entrenamiento frente a un rostro manifestando una expresión adquirida bajo iluminación desde la derecha y con postura frontal. Los mejores resultados se obtuvieron frente a las expresiones de sonrisa, enojo y sorpresa, siendo la primera la que obtuvo la mejor tasa de reconocimiento.

El hecho de que la información tridimensional original no obtuviera resultados favorables indica que las expresiones en la base de datos afectan de manera considerable la forma de los rostros. También, la fuerza de las expresiones es muy variable, puesto que los resultados varían mucho de una expresión a otra.

Como puede observarse en la tabla 8, el uso de la parametrización conforme obtuvo resultados ligeramente superiores para la expresión de enojo y sonrisa al utilizar un conjunto de entrenamiento que incluyera modelos neutrales.

**Tabla 8: Resultados de reconocimiento frente a expresiones de sonrisa, risa y enojo utilizando distintos conjuntos de entrenamiento.**

No. de individuos	Sonrisa			Risa			Enojo		
	N	E	N+E	N	E	N+E	N	E	N+E
5 individuos									
Vértices 3D	100	40	100	100	60	100	100	40	80
Parametrización + curvatura media + segmentación circular	80	60	80	60	60	80	100	40	100
Parametrización + curvatura media + segmentación nasal	100	60	100	80	80	100	100	80	100
10 individuos									
Vértices 3D	80	50	80	80	40	80	80	60	80
Parametrización + curvatura media + segmentación circular	70	40	70	40	50	50	70	30	70
Parametrización + curvatura media + segmentación nasal	90	40	90	30	20	40	90	60	90
30 individuos									
Vértices 3D	86.66	53.33	90	60	43.33	63.33	53.33	40	60
Parametrización + curvatura media + segmentación circular	73.33	16.66	70	43.33	20	46.66	60	26.66	66.66
Parametrización + curvatura media + segmentación nasal	70	26.66	70	33.33	33.33	36.66	46.66	36.66	46.66

Se esperaba que la segmentación nasal tuviera un mejor rendimiento en todos los casos fundamentando que la zona en torno a la nariz tiene menos variabilidad ante expresiones faciales. A pesar que esta zona no obtuvo una buena capacidad de reconocimiento al considerar 30 identidades distintas, se obtuvieron resultados aceptables e incluso

mejores que las coordenadas 3D al considerar 10 individuos. La principal causa de esta diferencia es que la punta de la nariz no haya sido correctamente identificada, llevando a segmentaciones de zonas con mucho impacto frente a expresiones, como mejillas o la boca, o que cubrieran zonas con poca capacidad de discriminación entre individuos, por ejemplo, la frente o la barbilla.

**Tabla 9: Resultados de reconocimiento frente a expresiones de sorpresa y ojos cerrados utilizando distintos conjuntos de entrenamiento.**

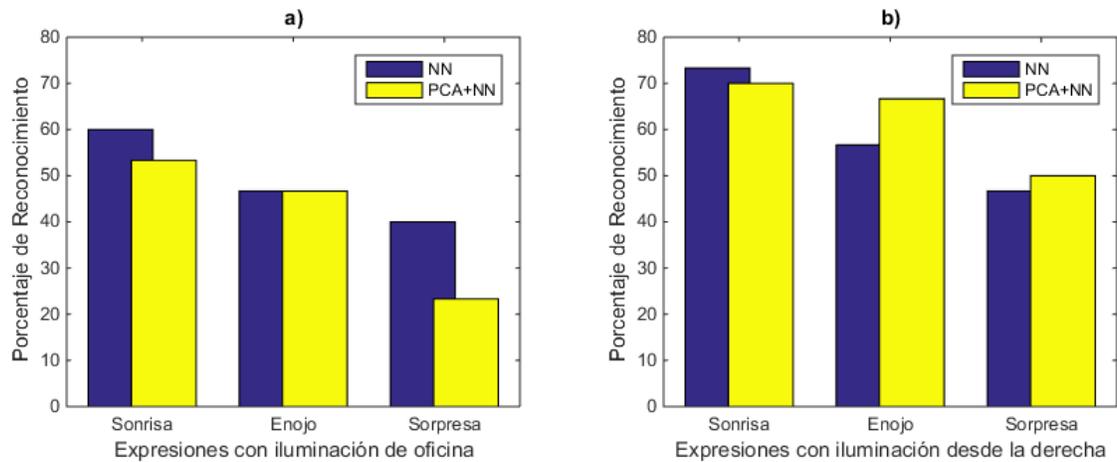
No. de individuos	Sorpresa			Ojos cerrados		
	N	E	N+E	N	E	N+E
<b>5 individuos</b>						
Vértices 3D	100	40	100	100	60	100
Parametrización + curvatura media						
+ segmentación circular	100	80	100	60	40	40
Parametrización + curvatura media						
+ segmentación nasal	80	40	80	80	60	80
<b>10 individuos</b>						
Vértices 3D	90	40	80	80	50	80
Parametrización + curvatura media						
+ segmentación circular	60	50	60	50	30	50
Parametrización + curvatura media						
+ segmentación nasal	80	40	80	80	50	80
<b>30 individuos</b>						
Vértices 3D	56.66	50	76.66	70	50	73.33
Parametrización + curvatura media						
+ segmentación circular	50	16.66	50	70	16.66	50
Parametrización + curvatura media						
+ segmentación nasal	60	26.66	63.33	53.33	40	53.33

#### 6.4.1. Comparación respecto a variaciones de iluminación

Con el fin de explorar el impacto de las diferencias de iluminación en el reconocimiento en presencia de expresiones, se repitió el experimento utilizando dos conjuntos de entrenamiento: los rostros neutrales en conjunto con los de expresiones con iluminación de oficina y los rostros neutrales en conjunto con los de expresiones con iluminación desde la derecha. La prueba se hizo ante las expresiones que obtuvieron los mejores resultados en el experimento pasado (sonrisa, enojo y sorpresa) considerando que no se tuviera la misma iluminación en el modelo de prueba que en el conjunto de entrenamiento.

Como puede apreciarse en la figura 25, se obtuvieron mejores resultados utilizando como entrenamiento rostros adquiridos bajo iluminación de oficina y dejando como prueba los adquiridos bajo iluminación desde la derecha. Además, en este caso, los resultados

obtenidos mediante el cálculo del vecino más cercano fueron mejores que su combinación con análisis de componentes principales en el caso de la expresión sonrisa.



**Figura 25: Comparación de resultados de reconocimiento variando la iluminación de los rostros en el conjunto de entrenamiento utilizando parametrización conforme y segmentación circular para 30 individuos. a) Entrenamiento con rostros neutrales y expresiones con iluminación desde la derecha. b) Entrenamiento con rostros neutrales y expresiones con iluminación de oficina.**

#### 6.5. Experimento 4: Reconocimiento en presencia de variaciones de postura

Este experimento se llevó a cabo con la finalidad de probar el rendimiento de la metodología propuesta ante variaciones de postura de los individuos. Los modelos con variaciones de pose se seleccionaron de manera que manifestaran una expresión neutra y hubiesen sido adquiridos bajo iluminación de oficina.

La evaluación se hizo considerando el conjunto neutral del experimento anterior agregando los siguientes modelos de variación de postura al conjunto de entrenamiento:

**Vertical.** Consiste en un modelo del individuo volteando hacia arriba y otro modelo volteando hacia abajo ambos con un ángulo de 20 a 30 grados.

**Horizontal.** Está formado por un modelo del individuo volteando hacia la izquierda y hacia la derecha ambos con un ángulo de 20 a 30 grados.

**Inclinación.** Consiste en un modelo del individuo inclinando su cabeza hacia la izquierda y otro hacia la derecha ambos en un ángulo de 20 a 30 grados.

Para comparar los resultados obtenidos con los de experimentos anteriores, se optó por utilizar de nuevo la parametrización conforme en conjunto con información de curvatura media aplicando segmentación circular del rostro y segmentación del área nasal. Las pruebas se hicieron evaluando cada conjunto de entrenamiento frente a una variación de postura a la vez.

**Tabla 10: Resultados de reconocimiento frente a variaciones de postura utilizando segmentación circular de la parametrización conforme y curvatura media para 30 individuos.**

Conjunto de entrenamiento	Horizontal		Vertical		Inclinación	
	Izquierda	Derecha	Arriba	Abajo	Izquierda	Derecha
<b>Parametrización + Segmentación circular</b>						
Neutral + Horizontal	-	-	93.33	93.33	93.33	93.33
Neutral + Vertical	3.33	96.66	-	-	100	100
Neutral + Inclinación	3.33	96.66	100	100	-	-
<b>Vértices 3D</b>						
Neutral + Horizontal	-	-	93.33	93.33	96.66	93.33
Neutral + Vertical	23.33	93.33	-	-	100	100
Neutral + Inclinación	23.33	93.33	100	100	-	-

En general, los resultados de reconocimiento en modelos con variación de postura fueron muy buenos. En las tablas 10 y 11 puede observarse que la variante de giro horizontal de la cabeza hacia la izquierda es la variación que obtuvo el peor rendimiento. Sin embargo, debido a que los resultados utilizando información tridimensional original también fueron bajos frente a este modelo, se pudo verificar que el problema había sido una mala segmentación de los rostros por haber ubicado erróneamente la punta de la nariz sobre el área de la boca. Excluyendo este caso, los resultados de reconocimiento para modelos con variación de postura fueron todos superiores al 93%. Un punto importante es que se alcanzó el 100% de reconocimiento en 30 individuos considerando variaciones de postura vertical e inclinaciones de la cabeza utilizando la metodología propuesta bajo ambas segmentaciones.

**Tabla 11: Resultados de reconocimiento frente a variaciones de postura utilizando segmentación nasal de la parametrización conforme y curvatura media para 30 individuos.**

Conjunto de entrenamiento	Horizontal		Vertical		Inclinación	
	Izquierda	Derecha	Arriba	Abajo	Izquierda	Derecha
Neutral + Horizontal	-	-	93.33	93.33	93.33	93.33
Neutral + Vertical	13.33	93.33	-	-	100	100
Neutral + Inclinación	20	93.33	100	100	-	-

## 6.6. Resumen

En este capítulo se presentó la evaluación de la metodología propuesta y la comparación de su rendimiento contra el uso tradicional de nubes de puntos tridimensionales. Los experimentos se diseñaron con el fin de valorar la capacidad de reconocimiento del método propuesto variando las condiciones en que fueron adquiridos los modelos 3D de los rostros respecto a iluminación, expresiones y variaciones de postura de la cabeza.

Se sabía de antemano que podría haber errores en el reconocimiento a causa de una segmentación incorrecta de las zonas de interés, ya que la etapa de localización de puntos que sirve de base para la segmentación de los rostros se realizó de manera automática. Se esperaba una mayor influencia negativa para modelos en presencia de expresiones y variaciones de pose, sin embargo, la metodología propuesta mostró ser robusta ante estas últimas.

Acercas de los experimentos que manejaban rostros con expresiones, se pudo observar que la intensidad de las expresiones en la base de datos varía mucho de una expresión a otra, lo cual no permite extraer información con capacidad de discriminación suficiente tanto para el caso de la evaluación de forma 3D original como para la metodología propuesta. Además, el rendimiento decrece significativamente al incrementar el número de identidades a reconocer: tratándose de 10 identidades se obtuvo un máximo de 90 % de reconocimiento, mientras que con 30 identidades el máximo fue de 73.33 %.

Para el caso de la evaluación de la metodología sobre rostros con expresión neutra deben considerarse los resultados del experimento 2 y 4. El mejor rendimiento en el segundo experimento fue de 90 % y 83.33 % de reconocimiento sobre 10 y 30 individuos, respectivamente, utilizando un conjunto de entrenamiento de modelos adquiridos con distinta iluminación (solo uno de ellos coincidía con la iluminación del rostro de prueba). Por otro lado, el cuarto experimento mostró la robustez del algoritmo propuesto al manejar modelos con expresión neutra cuando estos fueron adquiridos con iluminación uniforme; descartando una variante que presentaba errores de localización de la nariz, los resultados pasaron a estar dentro de un rango de 93.33 % a 100 % de reconocimiento sobre 30 identidades.

## Capítulo 7. Conclusiones

---

Para el caso de individuos cooperativos, la detección y el seguimiento de personas bajo buena iluminación es un problema que se considera resuelto asumiendo que el rostro se captura con suficiente resolución. La verificación o autenticación funciona bien para rostros con postura frontal usando el espectro visible, pero en el caso de reconocimiento (también denominado identificación) los métodos aún no alcanzan los requerimientos necesarios para aplicaciones prácticas aún tratándose del caso cooperativo. La principal desventaja a la que se enfrentan los sistemas comerciales es la iluminación, motivo por el cual el uso de infrarrojo está en el mercado desde el 2008 (Jain y Li (2011)).

Así, el reconocimiento facial en condiciones no controladas es un problema abierto. En este trabajo se abordó el problema de reconocimiento de rostros mediante modelos 3D en presencia de expresiones faciales, variaciones de iluminación y cambios de postura. La motivación de esta tesis fue diseñar un método que incluyese distintas técnicas que fueran aptas para combinarse y permitieran el reconocimiento de rostros de manera automática y robusta.

Se utilizó un enfoque híbrido que combina técnicas de métodos estadísticos, métodos basados en regiones y métodos basados en transformaciones. El algoritmo propuesto consistió en transformar la información tridimensional al plano mediante una parametrización espectral conforme después de una etapa de preprocesamiento que incluye localización de puntos característicos, segmentación automática de los rostros y un análisis de curvatura local. Se consideró la segmentación de los rostros seleccionando la parte central del rostro que cubre información de ojos y nariz con el objeto de evitar la influencia de variaciones faciales por expresiones. La correspondencia se hizo a partir de dos técnicas ampliamente utilizadas en el reconocimiento facial: el análisis de componentes principales y la búsqueda de vecinos cercanos.

Con base en resultados preliminares de experimentos con pocos individuos, se optó por descartar la etapa de corrección de postura previa al cálculo de la parametrización conforme. La idea detrás de esta etapa era facilitar el reconocimiento cuando el individuo cambiaba la orientación de su cabeza, sin embargo, la inclusión del procedimiento de co-

rección de postura convertía al rostro con variación de pose en un modelo con información faltante.

La evaluación de la metodología destaca la importancia del diseño del conjunto de entrenamiento utilizado en las pruebas dependiendo de las variaciones a las que se hace frente. Los resultados resaltan también la dependencia del método propuesto hacia la etapa previa de segmentación; por ello, se ve la importancia de refinar el proceso de selección de puntos característicos sobre ojos y nariz.

La comparación de los resultados se hizo contra información de forma 3D original resultando en que el rendimiento del método propuesto fue semejante y en algunos casos superior (tratándose de modelos con ciertos gestos). Para el caso del manejo de rostros con expresiones neutrales puede concluirse que la metodología propuesta alcanza un rendimiento comparable al uso de coordenadas 3D al aplicar una segmentación circular sobre la parametrización. Para el caso del manejo de rostros con expresiones no se puede hacer una generalización, ya que los resultados obtenidos fueron muy variantes de una expresión a otra.

El algoritmo propuesto no obtuvo resultados espectaculares, sin embargo, obtuvo mejores resultados en comparación a la parametrización conforme en la cual estuvo basado cuando se usó en combinación con segmentaciones de regiones de interés y curvatura media. En particular, se mostró que es posible reconocer hasta 30 individuos aún en presencia de variaciones moderadas en la postura de la cabeza cuando se manejan modelos adquiridos bajo buena iluminación y con un cambio de expresión moderado. Así, la contribución de este trabajo fue una implementación que rivaliza con el uso tradicional de nubes de puntos 3D para realizar reconocimiento de rostros en presencia de variaciones de postura y expresiones.

Por último, estos resultados son difíciles de comparar directamente con los que se encuentran en la literatura, pues el estudio realizado en esta tesis contempla conjuntos de entrenamiento y prueba pequeños.

## 7.1. Publicaciones en congresos

- Echeagaray-Patrón, B. A. y Kober, V. 3D face recognition based on matching of facial surfaces. 2015. En *Proceedings of SPIE, Optics and Photonics for Information Processing IX*. SPIE. 95980V-95980V-8 (10.1117/12.2186695).
- Echeagaray-Patrón, B. A., Miramontes-Jaramillo, D. y Kober, V. Conformal parameterization and curvature analysis for 3D facial recognition. *2015 International Symposium on Signal and Image Processing, Computer Vision and Pattern Recognition*. CSCI (10.1109/CSCI.2015.133).

## 7.2. Trabajo futuro

A partir del trabajo realizado y resultados obtenidos, se presentan a continuación algunas ideas que se habrían de considerar en el futuro, así como oportunidades de trabajo, siguiendo la misma línea de investigación. Dado que en la literatura se han presentado resultados considerando conjuntos de entrenamiento más grandes, se recomienda incluir más rostros de entrenamiento y prueba en los experimentos. Podría continuarse por construir una base de datos propia para tener control sobre las variaciones que presenten los rostros o extender la técnica propuesta a imágenes infrarrojas. Para adecuar el método a condiciones naturales, se podría incluir una etapa de detección de rostros robusta aún cuando la base de datos presuma condiciones ideales.

Algunas otras maneras de perfeccionar la metodología podrían ser mejorar el procedimiento de localización de puntos (nariz y ojos) con el fin de volver más robusta la segmentación del rostro verificando también que los puntos localizados correspondan a la ubicación verdadera encontrándolos manualmente; agregar una etapa de normalización en la cual las expresiones (en el conjunto de prueba) sean atenuadas; extender la metodología con otros algoritmos de clasificación y la implementación de las técnicas en GPU con el fin de reducir el tiempo de ejecución.

## Lista de referencias bibliográficas

- Al-Osaimi, F., Bennamoun, M., y Mian, A. (2009). An expression deformation approach to non-rigid 3D face recognition. *International Journal of Computer Vision*, **81**(3): 302–316.
- Alyuz, N., Gokberk, B., y Akarun, L. (2008). A 3D face recognition system for expression and occlusion invariance. En: *2nd IEEE International Conference on Biometrics: Theory, Applications and Systems, 2008. BTAS 2008*. IEEE, pp. 1–7.
- Berretti, S., Bimbo, A. D., y Pala, P. (2011). Facial curves between keypoints for recognition of 3D faces with missing parts. En: *2011 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. CVPRW 2011*. IEEE, pp. 46–51.
- Bowyer, K. W., Chang, K., y Flynn, P. (2006). A survey of approaches and challenges in 3D and multi-modal 3D+ 2D face recognition. *Computer Vision and Image Understanding*, **101**(1): 1–15.
- Bronstein, A. M., Bronstein, M. M., y Kimmel, R. (2003). Expression-invariant 3D face recognition. En: *Audio-and Video-Based Biometric Person Authentication*. Springer, pp. 62–70.
- Chang, K. I., Bowyer, K. W., y Flynn, P. J. (2006). Multiple nose region matching for 3D face recognition under varying facial expression. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **28**(10): 1695–1700.
- Colombo, A., Cusano, C., y Schettini, R. (2006). 3d face detection using curvature analysis. *Pattern recognition*, **39**(3): 444–455.
- Devadoss, S. L. y O'Rourke, J. (2011). *Discrete and computational geometry*. Princeton University Press.
- Do Carmo, M. P. (1976). *Differential geometry of curves and surfaces*. Prentice Hall.
- Drira, H., Amor, B. B., Daoudi, M., y Srivastava, A. (2010). Pose and expression-invariant 3D face recognition using elastic radial curves. En: *British Machine Vision Conference*. pp. 1–11.
- Du, S., Zheng, N., Xiong, L., Ying, S., y Xue, J. (2010). Scaling iterative closest point algorithm for registration of m-D point sets. *Journal of Visual Communication and Image Representation*, **21**(5-6): 442–452.
- Faltemier, T. C., Bowyer, K. W., y Flynn, P. J. (2008). A region ensemble for 3-D face recognition. *IEEE Transactions on Information Forensics and Security*, **3**(1): 62–73.
- Gu, D. X., Luo, F., y Yau, S.-T. (2010). Fundamentals of computational conformal geometry. *Mathematics in Computer Science*, **4**(4): 389–429.
- Jain, A. K. y Li, S. Z. (2011). *Handbook of face recognition*. Springer.
- Kakadiaris, I., Passalis, G., Toderici, G., Murtuza, M. N., Lu, Y., Karampatziakis, N., Theoharis, T., *et al.* (2007). Three-dimensional face recognition in the presence of facial expressions: An annotated deformable model approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **29**(4): 640–649.

- Kjer, H. M. y Wilm, J. (2010). Evaluation of surface registration algorithms for PET motion correction. Tesis de doctorado. Technical University of Denmark. Lyngby, Denmark.
- Lei, Y., Bennamoun, M., Hayat, M., y Guo, Y. (2014). An efficient 3D face recognition approach using local geometrical signatures. *Pattern Recognition*, **47**(2): 509–524.
- Lévy, B. y Zhang, H. R. (2010). Spectral mesh processing. En: *ACM SIGGRAPH 2010 Courses*. ACM, p. 8.
- Lévy, B., Petitjean, S., Ray, N., y Maillot, J. (2002). Least squares conformal maps for automatic texture atlas generation. En: *ACM Transactions on Graphics. TOG 2002*. ACM, Vol. 21, pp. 362–371.
- Li, X., Jia, T., y Zhang, H. (2009). Expression-insensitive 3D face recognition using sparse representation. En: *IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2009*. IEEE, pp. 2575–2582.
- Lu, X. y Jain, A. K. (2008). Deformation modeling for robust 3D face matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **30**(8): 1346–1357.
- Mahmood, S. A., Ghani, R. F., y Kerim, A. A. (2013). Nose tip detection using shape index and energy effective for 3D face recognition. *International Journal of Modern Engineering Research. IJMER 2013*, **3**(5): 3086–3090.
- Mian, A. S., Bennamoun, M., y Owens, R. (2007). An efficient multimodal 2D-3D hybrid approach to automatic face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **29**(11): 1927–1943.
- Mullen, P., Tong, Y., Alliez, P., y Desbrun, M. (2008). Spectral conformal parameterization. *Eurographics Symposium on Geometry Processing*, **27**(5): 1487–1494.
- Queirolo, C. C., Silva, L., Bellon, O. R., y Segundo, M. P. (2010). 3D face recognition using simulated annealing and the surface interpenetration measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **32**(2): 206–219.
- Rusinkiewicz, S. y Levoy, M. (2001). Efficient variants of the ICP algorithm. En: *3rd International Conference on 3-D Digital Imaging and Modeling, Proceedings 2001*. IEEE, pp. 145–152.
- Sheffer, A., Praun, E., y Rose, K. (2006). Mesh parameterization methods and their applications. *Foundations and Trends® in Computer Graphics and Vision*, **2**(2): 105–171.
- Smeets, D., Claes, P., Hermans, J., Vandermeulen, D., y Suetens, P. (2012). A comparative study of 3-D face recognition under expression variations. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, **42**(5): 710–727.
- Szeptycki, P., Ardabilian, M., y Chen, L. (2009). A coarse-to-fine curvature analysis-based rotation invariant 3D face landmarking. En: *3rd International Conference on Biometrics: Theory, Applications, and Systems. BTAS 2009*. IEEE, pp. 1–6.

- Szeptycki, P., Ardabilian, M., Chen, L., Zeng, W., Gu, D., y Samaras, D. (2010). Conformal mapping-based 3D face recognition. En: *3D Data Processing, Visualization and Transmission Symposium. 3DPVT 2010. Paris, France*. Vol. 226.
- Wang, S., Wang, Y., Jin, M., Gu, X. D., y Samaras, D. (2007). Conformal geometry and its applications on 3d shape matching, recognition, and stitching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **29**(7): 1209–1220.
- Xu, C., Wang, Y., Tan, T., y Quan, L. (2004). Automatic 3D face recognition combining global geometric features with local shape variation information. En: *6th IEEE international conference on automatic face and gesture recognition, 2004.*. IEEE, pp. 308–313.
- Zeng, W. y Gu, X. D. (2009). Surface matching and registration using symmetric conformal mapping. En: *11th IEEE International Conference on Computer-Aided Design and Computer Graphics. CAD/Graphics 2009*. IEEE, pp. 249–254.

## Apéndice A. Comparación de mapeos conformes

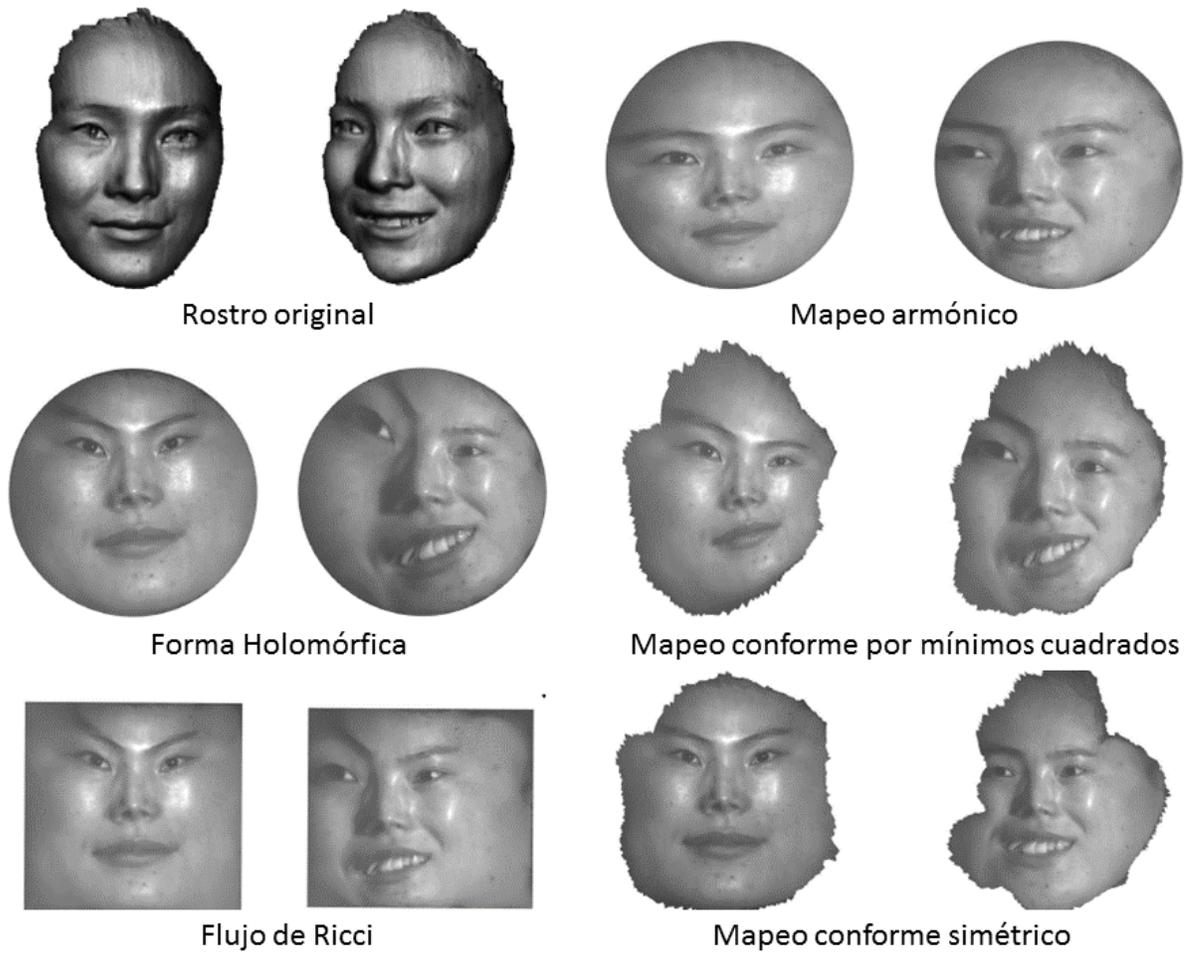
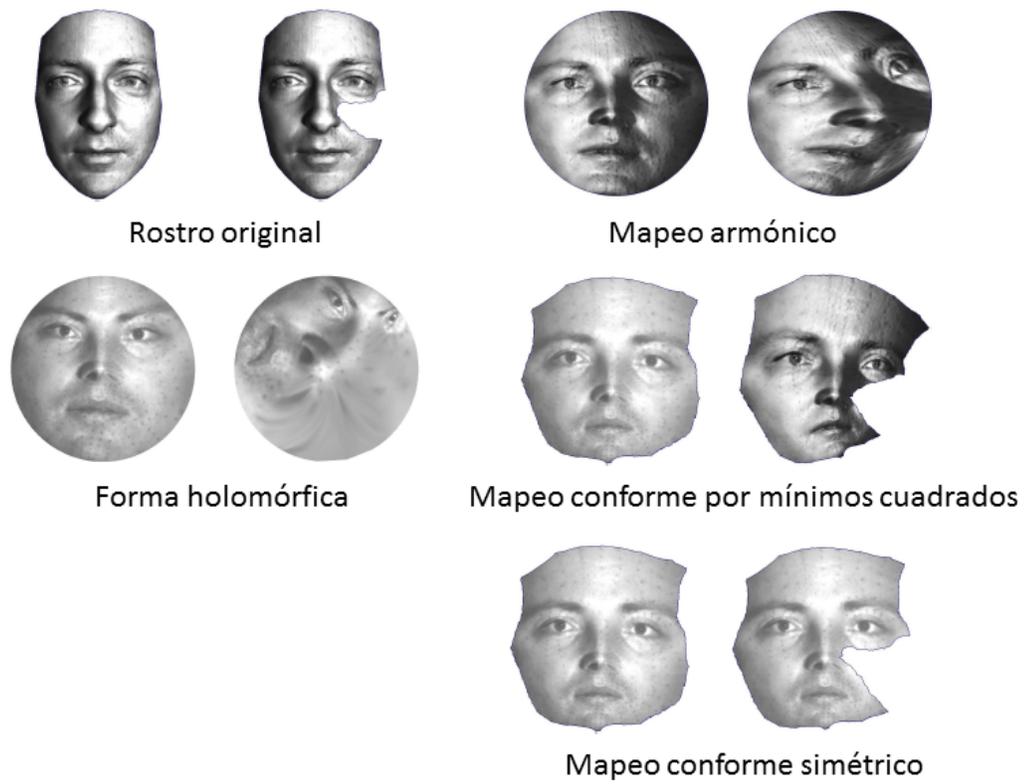


Figura 26: Comparación de distintos métodos para mapeos conformes en un rostro con variación de pose. (Extraído de Zeng y Gu (2009)).



**Figura 27: Comparación de distintos métodos para mapeos conformes en presencia de oclusiones. (Extraído de Zeng y Gu (2009)).**