

*Análisis comparativo de descriptores de forma 3D para
detección de características faciales*

ALEXANDER CERÓN CORREA
INGENIERO DE SISTEMAS
CÓDIGO: 299736



UNIVERSIDAD NACIONAL DE COLOMBIA
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE INGENIERÍA DE SISTEMAS E INDUSTRIAL
BOGOTÁ, D.C.
2011

*Análisis comparativo de descriptores de forma 3D para
detección de características faciales*

ALEXANDER CERÓN CORREA
INGENIERO DE SISTEMAS
CÓDIGO: 299736

TRABAJO DE TESIS PARA OPTAR AL TÍTULO DE
MAGISTER EN INGENIERÍA - INGENIERÍA DE SISTEMAS Y COMPUTACIÓN

DIRECTOR
FLAVIO PRIETO, PH.D.

CODIRECTOR
AUGUSTO SALAZAR.
MASTER EN INGENIERÍA



UNIVERSIDAD NACIONAL DE COLOMBIA
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE INGENIERÍA DE SISTEMAS E INDUSTRIAL
BOGOTÁ, D.C.
2011

Título en español

Análisis comparativo de descriptores de forma 3D para detección de características faciales

Title in English

Comparative analysis of 3D shape descriptors for facial feature detection

Resumen: El rostro humano presenta una gran cantidad de características que actualmente pueden ser modeladas mediante un simple patrón 2D, un conjunto complejo de vértices 3D que forman una malla poligonal o un conjunto de parámetros para cada grado de libertad o variación. La caracterización del rostro tiene gran cantidad de aplicaciones dentro de las cuales se tienen: identificación de rostros, modelado de la cara, síntesis de voz, identificación de expresiones y cirugía facial. Los modelos tridimensionales del rostro contienen gran cantidad de la información geométrica, la cual puede ser extraída por diversos métodos. Entre ellos se pueden mencionar los basados en modelos de superficies cuádricas que buscan clasificar los puntos de una superficie como planar o no planar, de ahí se estima si puede ser esférica, cilíndrica o cónica; otros permiten obtener el perfil del rostro en imágenes de rango aprovechando la simetría del mismo. Existen diversos métodos para caracterizar superficies 3D, los cuales han sido aplicados al reconocimiento de objetos 3D, pero aun no existe un estudio comparativo representativo que discrimine las ventajas y desventajas de los mismos en las características faciales. El objetivo de este trabajo es realizar un análisis de relevancia de los siguientes descriptores: curvaturas principales, curvatura media, curvatura gaussiana, índice de forma (*shape index*), curvacidad (*curvedness*), las imágenes spin y el índice de forma multi-escala para determinar cuales ofrecen información más relevante en ubicaciones específicas del rostro mediante la utilización del análisis discriminante de Fisher. Esto permite definir cual es el descriptor más adecuado en superficies complejas mostrando una aplicación en caracterización y reconocimiento de rostros en imágenes 3D. Las características consideradas corresponden a 182 puntos y a 28 regiones del rostro obtenidos mediante un modelo compuesto de curvas que se ajustan a la morfología del rostro. Al obtener los descriptores más relevantes en cada una de las regiones del rostro humano se espera contribuir a la solución del problema de reconocimiento de rostros en imágenes 3D y al de la obtención de medidas para antropometría facial. La caracterización mediante descriptores de forma puede ser extendida a diferentes partes del cuerpo y puede ser aplicada en diversos sectores que incluyen la animación, la medicina forense, la cirugía reconstructiva, entre otras.

Abstract: The human face presents a big amount of morphological features that can be modeled by using a 2D basic template, a complex set of vertex that can be structured in a polygonal mesh or a set of parameters for each freedom degree or variation. The face featuring has several applications: face identification, face modeling, voice synthesis, identification of expressions and facial surgery. The 3D face models have a lot of geometric information, which can be extracted by using different methods as quadric-surfaces-based, which classifies the surface points as planar, spherical, conical, or cylindrical. Other models are useful for obtaining the face profile in range images by using face symmetry. There exist a lot of methods for 3D surface featuring, which have been applied to 3D objects recognition, but it does not exist a meaningful comparative study which discriminate the advantages and disadvantages of these methods in facial featuring. The main goal of this work is to develop a relevance analysis of the following descriptors: principal curvatures,

mean curvature, gaussian curvature, shape index, curvedness, spin images, and multi-scale shape index; in order to determine which descriptor offers more information in specific locations of the face by using the Fisher Discriminant Analysis–FDA. In addition, this work try out to define which descriptor is more suitable in a complex surface showing an application in featuring and face recognition in 3D range images. The considered features correspond to 182 points and 28 face regions, obtained by using a model composed of curves that can be adjusted to the face’s morphology. This work demonstrates that the identification of the more relevant descriptors obtained in the face regions may contribute to solve problems related to face recognition in 3D images and facial anthropometry characterization. As a perspective of this study, face featuring by using shape descriptors can be extended to different parts of the body. Likewise, it can be applied to different areas as animation, forensic medicine, reconstructive surgery, etc.

Palabras clave: descriptores de forma 3D, superficie facial 3D, capacidad discriminante, extracción de características.

Keywords: 3D shape descriptors, 3D facial surface, discriminant capacity, feature extraction.

Nota de aceptación

Trabajo de tesis

Jurado

Fabio Augusto González Osorio

Jurado

Jean Pierre Charalambos Hernandez

Director

Flavio Augusto Prieto

Bogotá, D.C., Enero 31 de 2011

Dedicado a

A la memoria de mi padre Enrique Cerón Lopez a quien le guardo toda mi admiración y respeto.

A mi madre Aura Librada Correa Chaparro quien siempre me ha mostrado el valor del esfuerzo y el sacrificio.

A mis hermanos por ser siempre el ejemplo a seguir.

A Alejandra por acompañarme, entenderme y darme tantos momentos felices.

DEDICADO A

Agradecimientos

El autor agradece a Flavio Prieto quien condujo el desarrollo de este trabajo de tesis con la paciencia que se necesitaba y a Augusto Salazar quien por su colaboración en la coodirección ayudo a generar un trabajo de gran calidad.

El autor agradece a los desarrolladores del software Meshlab por proveer a la comunidad de una gran herramienta que permitió ahorrar tiempo valioso.

El autor agradece a la Dirección de Bienestar de la Universidad Nacional de Colombia sede Bogotá, por facilitar apoyo económico en la difusión de resultados obtenidos en desarrollo de este trabajo.

En general el autor expresa sus agradecimientos a todos aquellos que de una u otra forma han colaborado, contribuido o aportado en el desarrollo de este trabajo.

Índice General

Índice General	III
Lista de Tablas	VII
Lista de Figuras	IX
Lista de Algoritmos	XI
Introducción	XIII
1. Descriptores de Forma en 3D	1
1.1. Definición de curvatura	1
1.2. Descriptores de forma basados en curvatura	2
1.3. Imágenes spin	3
1.4. Imágenes spin multi-resolución	5
1.5. Índice de forma multi-escala	7
1.5.1. Variación local de escala	7
1.5.2. Extracción de características multi-escala	8
1.6. Estimación del costo computacional	8
1.6.1. Costo computacional de las curvaturas principales	9
1.6.2. Costo computacional de las imágenes spin	10
1.6.3. Costo computacional del índice de forma multi-escala	10
1.7. Consideraciones finales	11
2. Extracción de descriptores en áreas de interés	13
2.1. Plantilla del rostro	13
2.2. Extracción de descriptores en puntos de interés	14
2.2.1. Registro semi-automático	16

2.3. Extracción de descriptores en contornos	18
2.4. Extracción de descriptores en regiones	18
2.5. Consideraciones finales	21
3. Materiales y Métodos	23
3.1. Análisis discriminante de Fisher	23
3.2. Bases de datos de rostros	24
3.3. Experimentos	26
3.3.1. Análisis Global	26
3.3.2. Análisis local	26
4. Resultados del Análisis de Relevancia	29
4.1. Resultados del Análisis Global	29
4.1.1. Análisis en los contornos y las regiones	30
4.2. Resultados del Análisis Local	30
4.2.1. Prueba del área de los ojos	30
4.2.2. Prueba del área de la nariz	31
4.2.3. Prueba del área de las mejillas	32
4.2.4. Prueba del área de la boca	32
4.2.5. Prueba del área de la quijada	34
4.2.6. Ranking de descriptores por área del rostro	34
4.3. Consideraciones Finales	35
5. Aplicación al reconocimiento de rostros	37
5.1. Sistema de reconocimiento de rostros	37
5.1.1. Adquisición	37
5.1.2. Preprocesamiento	38
5.1.3. Registro	38
5.1.4. Segmentación	39
5.1.5. Extracción de características	39
5.1.6. Entrenamiento	40
5.1.7. Clasificación	40
5.2. Experimentos de reconocimiento	40
5.2.1. Bases de datos	40
5.2.2. Esquema de reconocimiento	40
5.2.3. Prueba identificación del individuo	41

5.2.4. Prueba identificación del género	41
5.2.5. Prueba identificación de la expresión	41
5.3. Resultados	41
5.3.1. Identificación de un individuo	41
5.3.1.1. Identificación a partir de puntos característicos	41
5.3.1.2. Identificación a partir de puntos y características relevantes	42
5.3.1.3. Identificación a partir de regiones	43
5.3.1.4. Identificación con regiones más relevantes y reducción de características	43
5.3.2. Identificación de género	44
5.3.3. Identificación de expresión facial	44
5.4. Consideraciones finales	45
Conclusiones	47
Trabajo futuro	49
Bibliografía	51
A. Estimación del tensor de curvatura de una superficie	55
A.1. Tensor de curvatura	55
A.2. Estimación del tensor de curvatura	56
A.3. Algoritmo para la obtención de curvaturas direccionales	58

Lista de Tablas

1.1. Costo de realizar operaciones involucradas.	9
2.1. Rangos de variación de descriptores de curvatura sobre la superficie del rostro. 15	
2.2. Error obtenido al cambiar la subdivisión en la búsqueda.	21
3.1. Regiones consideradas para cada prueba.	27
4.1. Coeficientes de Fisher para la prueba <i>Global</i> . R indica el ranking.	29
4.2. Coeficientes de Fisher para la prueba <i>Global</i> en la base de datos BU-3DFE. R indica el ranking.	30
4.3. Coeficientes de Fisher para la prueba de los <i>Ojos</i>	31
4.4. Coeficientes de Fisher para la prueba de los <i>Ojos</i> en la BD <i>BU-3DFE</i>	31
4.5. Coeficientes de Fisher para la prueba de la <i>Nariz</i>	32
4.6. Coeficientes de Fisher para la prueba de las <i>Mejillas</i>	33
4.7. Coeficientes de Fisher para la prueba de los <i>Mejillas</i> en la BD <i>BU-3DFE</i>	33
4.8. Coeficientes de Fisher para la prueba de la <i>Boca</i>	33
4.9. Coeficientes de Fisher para la prueba de la <i>Quijada</i>	34
4.10. Descriptores de acuerdo a las regiones del rostro.	34
5.1. Resultados identificación con 168 puntos y 9 descriptores con diferentes clasificadores	42
5.2. Resultados identificación con 64 puntos y 4 descriptores con diferentes cla- sificadores	42
5.3. Resultados identificación para 26 regiones con diferentes clasificadores	43
5.4. Resultados identificación para el área de los ojos y nariz con reducción de características con diferentes clasificadores	43
5.5. Resultados identificación de género con diferentes clasificadores	44
5.6. Resultados identificación de expresiones faciales con diferentes clasificadores	45

Lista de Figuras

1.1. Curvaturas principales.	2
1.2. Generación de imagenes spin.	3
1.3. Ejemplo de imagenes spin.	5
1.4. Dos formas de generar imágenes spin multi resolución.	7
2.1. Regiones del rostro humano que se analizan en este trabajo.	14
2.2. Modelo parametrizado del rostro.	14
2.3. Extracción de 53 puntos [1]	15
2.4. Descriptores de forma basados en curvatura calculados sobre la superficie del rostro.	15
2.5. Ajuste de la plantilla	16
2.6. Etiquetado de las regiones y un contorno obtenido.	18
2.7. Obtención de los puntos interior a la región 9	20
3.1. Valores del índice de forma (base de datos <i>3DImDB</i>)	25
3.2. Imágenes de las bases de datos. <i>FemDB</i> (primera fila), <i>MaleDB</i> (segunda fila), <i>3DImDB</i> (tercer fila) y <i>BU-3DFE</i> (cuarta fila).	25
5.1. Esquema del sistema de reconocimiento.	38
5.2. Ajuste de la plantilla y segmentación obtenida	39
5.3. Imágenes de las bases de datos <i>BU-3DFE</i> con expresiones, primera columna expresión neutral, segunda columna expresión de enojo y tercera columna expresión alegría	45

Lista de Algoritmos

- 1. Índice de Forma Multi-escala 8
- 2. Ajuste y obtención de puntos de interés 17
- 3. Obtención de contornos 19
- 4. Obtención de regiones 20

Introducción

Los modelos tridimensionales del rostro dan gran cantidad de la información sobre la morfología del mismo, la cual puede ser extraída por diversos métodos. Entre ellos se pueden mencionar los basados en modelos de superficies cuádricas que buscan clasificar los puntos de una superficie como planar o no planar, de ahí se estima si puede ser esférica, cilíndrica o cónica [2], otros permiten obtener el perfil del rostro en imágenes de rango aprovechando la simetría del mismo [3]. Por otra parte, existen descriptores basados en las curvaturas que están presentes en cualquier objeto 3D, los cuales han mostrado gran efectividad ya que la curvatura como propiedad de superficie local tiene la capacidad de ser invariante al punto de vista. Los descriptores de curvatura, el cual es un concepto proveniente de la geometría diferencial, permiten identificar ocho formas fundamentales: hiperbólica silla valle, parabólica cóncava hacia arriba, elíptica cóncava hacia abajo, hiperbólica mínima, plana, hiperbólica silla arista, parabólica cóncava hacia arriba y elíptica cóncava hacia arriba [4].

A nivel general se pueden mencionar algunos trabajos relacionados con el análisis y reconocimiento del rostro mediante imágenes 3D. En [5] se presenta un reconocimiento de rostros basado en mapas de profundidad y la medición de la curvatura principal de superficies sobre un conjunto de datos específico. El cálculo de descriptores de superficies generales se basa en la curvatura y el cálculo de características del rostro basados en conocimiento a priori sobre la estructura del mismo. En [6] se busca una correspondencia entre imágenes en regiones convexas, haciendo uso de la curvatura Gaussiana y de un mapeo para obtener imágenes extendidas Gaussianas. En [7] se presenta una evaluación de algoritmos de reconocimiento de rostros mediante imágenes de rango y se propone el uso de la información de la curvatura, la cual puede ser adquirida mediante enfoques geométricos en el proceso de proyección lineal.

Pese a que en varios trabajos se realiza detección de características con este tipo de descriptores, no es muy común encontrar aquellos donde se realizan comparaciones cuantitativas entre ellos. Dentro de los trabajos más relacionados con el reconocimiento de rostros en imágenes de rango utilizando descriptores de forma, se puede contar el de [8], donde se hace uso de una segmentación basada en los signos de las curvaturas media y gaussiana, para encontrar las regiones de interés y así caracterizarlas posteriormente, haciendo uso de un conjunto de descriptores. Adicionalmente se realiza una clasificación de 86 descriptores aplicados en las regiones encontradas, mediante el coeficiente de Fisher [9, 10]. Los descriptores usados fueron áreas de regiones, relaciones entre áreas, media de áreas, distancias entre centros de masas de regiones, media de distancias simétricas, ángulos entre centros de masas de regiones, media de ángulos de regiones, promedio y varianza de la curvatura media y Gaussiana de los puntos de una región y líneas de la cara. Los

resultados obtenidos en el trabajo en mención muestran que los mejores desempeños en las técnicas utilizadas ocurren en imágenes de pose frontal. Sin embargo, también es bueno notar que el autor sólo utilizó una segmentación basada en los signos de las curvaturas media y gaussiana, para encontrar las regiones de interés.

Se pueden mencionar otros métodos para caracterizar superficies 3D, que en algunos casos involucran otras medidas de curvatura como las que se presentan a continuación:

- COSMOS. El enfoque presentado en [11] busca representar y reconocer objetos de forma libre pero sin estar ocluidos, se caracteriza por: ser independiente del punto de vista, manejar objetos que pueden variar tanto en forma como en complejidad y no hacer suposiciones restrictivas acerca de los tipos de superficies presentes en el objeto.

Un objeto es representado concisamente en términos de parches de superficie maximal de *índice de forma* constante. Los parches maximales que representan el objeto son mapeados sobre una esfera unitaria mediante sus orientaciones, son agregados mediante funciones espectrales de forma. Propiedades de curvatura, conectividad y área de superficie son requeridas para obtener información global y local. Un aporte de este trabajo es el concepto de *espectro de forma*.

- Curvatura de conos ponderados (WCC). Este método hace una reducción del modelo del objeto sin pérdida de información [12]. Una medida de similaridad basada en las características WCC ha sido definida para comparar modelos, de tal forma que una matriz de similaridad basada en la correspondencia WCC es la entrada a un algoritmo de agrupamiento difuso, el cual agrupa automáticamente los objetos de acuerdo a la similaridad. El método usado se basa en una teselación, en la cual pueden variar los nodos que representan círculos concéntricos, que se adaptan a la curvatura del objeto en forma de cono.
- Curvaturas en nubes de puntos no organizados. En el enfoque presentado en [4], se presenta un método para estimar las curvaturas y direcciones principales en nubes de puntos no organizados, los cuales han sido muestreados a partir de una superficie 3D. La información de curvaturas se obtiene mediante una estimación de un vecindario geodésico local alrededor de cada punto, para lo cual se propone un método para calcularlo basado en una ponderación. En el trabajo en mención se utilizan las curvaturas Gaussiana y Media para identificar formas fundamentales o tipos de superficies de forma que se mencionaron anteriormente, se utiliza un método basado en análisis de componentes principales PCA, para aproximar el plano tangente en un vecindario de los puntos más cercanos alrededor de un punto p , con la información ponderada de vecindad se hace una matriz de covarianza. Los valores propios de dicha matriz miden la variación de los puntos en el vecindario. Las curvaturas se obtienen de manera simple haciendo una combinación lineal de los valores propios.

Es bueno mencionar que la mayoría de los enfoques para el reconocimiento de imágenes 3D mediante la curvatura, están basados en el análisis de las curvaturas Gaussiana y media [13], por tal razón se incluyeron y validaron otros enfoques para caracterizar las regiones del rostro.

El objetivo de este trabajo es realizar un análisis de relevancia de los siguientes descriptores: curvaturas principales, curvatura media, curvatura gaussiana, índice de Forma

(*shape Index*), curvacidad (*curvedness*), las imágenes spin, el índice de forma multi-escala y las imágenes spin multi-resolución, para determinar cuales ofrecen más información en ubicaciones específicas del rostro. Esto permitirá definir cual es el descriptor más adecuado en superficies complejas, mostrando una aplicación en caracterización y reconocimiento de rostros en imágenes 3D.

Dentro del proceso experimental que se planteó, se utilizó una plantilla compuesta de 28 regiones anatómicas, para segmentar un modelo 3D del rostro y extraer los puntos que hacen parte de las regiones y contornos que las define, en lugar de definir agrupaciones de vértices. En principio, el problema consiste en establecer cuales características son relevantes en cada punto. El proceso de caracterización está basado en las diferentes maneras de representar los puntos específicos, utilizando solamente la información disponible. La plantilla es ajustada a cada modelo 3D, para obtener diferentes grupos de vértices: los vértices de la plantilla, los que hacen parte del contorno (aquellos que caen sobre las curvas que definen la región) y los que se encuentran dentro de la región. Gracias al proceso experimental se desarrollaron algunas estrategias de búsqueda, las cuales pueden ser reconfiguradas de acuerdo a la región del rostro sobre la cual se quiera detectar una característica.

Los resultados de este trabajo son de gran interés en diversos sectores como la cirugía facial, la identificación de personas, la representación paramétrica de características faciales, la antropometría, la animación de personajes, entre otros. Este trabajo es la primera etapa de un sistema para detectar puntos específicos sobre la superficie facial la cual está definida mediante diferentes regiones.

La organización de este trabajo es la siguiente: en primer lugar se enuncian los descriptores de forma 3D utilizados, en el Capítulo 1. Luego se presenta la extracción de los descriptores y los algoritmos involucrados en las búsquedas, en el Capítulo 2. El marco experimental, es presentado en el Capítulo 3. Posteriormente se desarrolla un análisis discriminante usando puntos, contornos y regiones del rostro junto con los resultados correspondientes, en el Capítulo 4. Posteriormente, se muestra una aplicación básica de la utilización de los descriptores evaluados en el reconocimiento de rostros y de características faciales como género y expresión, en el Capítulo 5. Finalmente, se exponen las conclusiones y el trabajo futuro.

Descriptores de Forma en 3D

Los descriptores de forma son un conjunto de herramientas matemáticas y procedimentales que permiten obtener diversas medidas sobre los contornos y las estructuras morfológicas de los modelos o imágenes 3D. Dentro de los descriptores de forma existen diferentes tipos: los que están basados en los modelos de superficies cuádricas que buscan clasificar puntos dentro de las diferentes formas geométricas en 3D (planos, esferas, cilindros, conos, entre otros); los que están basados en curvaturas principales, los que están basados en transformaciones de información 3D a 2D como es el caso de las imágenes spin que pueden ser calculadas a diferentes resoluciones, los multi-escala, entre otros. Este capítulo se enfoca en definir los descriptores utilizados en este trabajo los cuales son: curvaturas principales, curvatura media, curvatura gaussiana, índice de forma (*shape index*), curvacidad (*curvedness*), las imágenes spin, el índice de forma multi-escala y las imágenes spin multi-resolución, los cuales se presentarán a continuación y se mostrarán los métodos, procedimientos para su obtención y el costo computacional que tienen.

1.1. Definición de curvatura

Una curva plana es definida como $\kappa = \frac{d\alpha}{ds}$, donde α , es el ángulo formado por el vector tangente a la curva (en la dirección de desplazamiento) con una dirección fija y s representa la longitud de arco a lo largo de la curva. La curvatura no es únicamente un valor, el signo indica la dirección de desviación. Es bueno mencionar que no es posible darle a una curva un signo en el espacio.

Una curva en el espacio, se puede definir mediante el uso del parámetro s , designando a x como el vector de posición de la curva, obteniendo el vector unitario tangente $T = \frac{dx}{ds}$ y la curvatura $k = \left| \frac{dT}{ds} \right| = \left| \frac{d^2x}{ds^2} \right|$.

En una superficie $S \subset \mathbf{R}^3$ el doblés máximo y el mínimo son medidos por las dos curvaturas principales k_1 y k_2 , las cuales corresponden a los valores máximos y mínimos de curvaturas ubicadas en los planos ortogonales que contienen al vector normal al punto a evaluar (ver Figura 1.1).

Las conexiones entre las curvaturas principales y las dos mediciones clásicas de forma, la curvatura Gaussiana K y la curvatura media H , están dadas por:

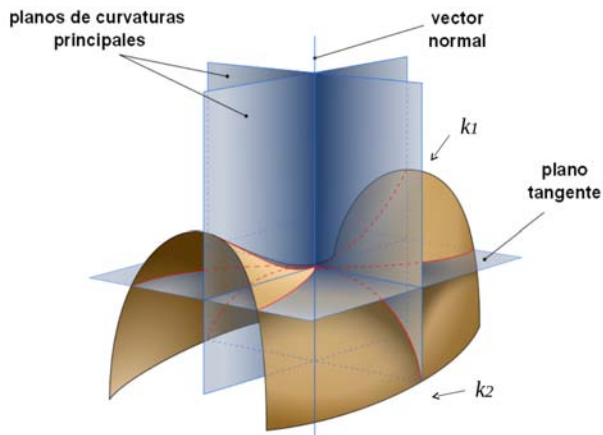


Figura 1.1: Curvaturas principales.

$$\begin{aligned} k_1 &= H + \sqrt{H^2 - K} \\ k_2 &= H - \sqrt{H^2 - K} \end{aligned} \quad (1.1)$$

Mediante el análisis de los valores de curvatura es posible extraer el plano de simetría del rostro [3], detectar las regiones de los ojos y la boca [5, 14], segmentar un modelo facial como entrada de un sistema de reconocimiento [15, 16] o detectar diversas marcas sobre la superficie facial [17, 18]. Otras aproximaciones combinan la información de curvatura con otras características obtenidas de información 2D [19] y/o un conocimiento a priori de la geometría del rostro [20].

El proceso de estimación de las curvaturas principales se puede observar en el Apéndice A.

1.2. Descriptores de forma basados en curvatura

En reconocimiento de rostros y detección de características faciales, la curvatura ha sido usada ampliamente, especialmente cuando la información de geometría 3D del rostro está disponible debido a que es una propiedad de superficie local, que tiene la ventaja de ser invariante a traslaciones y rotaciones.

Se pueden encontrar una amplia gama de descriptores basados en las curvaturas principales utilizados en análisis y reconocimiento del rostro, los cuales se mencionarán a continuación.

Utilizando las ecuaciones 1.1 se puede obtener la curvatura media $H = \frac{k_1+k_2}{2}$ y curvatura Gaussiana $K = k_1k_2$ [21]. Estas curvaturas han sido usadas en [8], para la segmentación y el análisis de características en un sistema de reconocimiento de rostros.

Con base en las curvaturas principales k_1, k_2 , se han propuesto otros descriptores como el índice de forma S_I [22]. Para un punto \mathbf{p} sobre una superficie, el S_I se define como:

$$S_I(p) = \frac{1}{2} - \frac{1}{\pi} \tan^{-1} \frac{k_1(p) + k_2(p)}{k_1(p) - k_2(p)}. \quad (1.2)$$

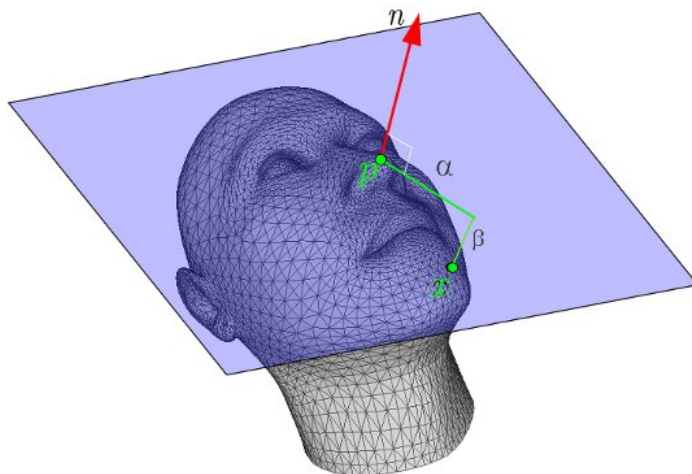


Figura 1.2: Generación de imágenes spin.

El S_I ha sido usado para localizar características faciales en [23, 24, 25, 26]. Otro descriptor es la curvatura (*curvedness*) definida como:

$$R(p) = \sqrt{(k_1^2(p) + k_2^2(p)) / 2}. \quad (1.3)$$

R representa la cantidad de curvatura en una región, lo cual hace posible la percepción de variación en escala de la forma de los objetos. R ha sido utilizado como criterio para segmentación automática de mallas triangulares [27].

1.3. Imágenes spin

La imagen spin [28] es un descriptor local, calculado en un punto orientado (\mathbf{p}, \mathbf{n}) (un punto y su normal respectivamente), que codifica 2 de las 3 coordenadas de un sistema cilíndrico en su vecindad. La imagen spin X para un punto \mathbf{p} en una superficie, es un histograma 2D en el cual cada píxel es un *bin*¹ que almacena el número de vecinos que están a una distancia α desde \mathbf{n} y a una profundidad β desde su plano tangente \mathbf{P} (ver Figura 1.2).

El parámetro α puede ser estimado como la distancia perpendicular desde \mathbf{n} hasta el punto \mathbf{x} . El segundo parámetro β , es la distancia relativa paralela entre \mathbf{n} y \mathbf{x} . La representación matemática de estos parámetros conocidos como *coordenadas spin* es:

$$\begin{aligned} \alpha &= \sqrt{\|\mathbf{x}-\mathbf{p}\|^2 - (\mathbf{n} \cdot (\mathbf{x}-\mathbf{p}))^2}, \\ \beta &= \mathbf{n} \cdot (\mathbf{x}-\mathbf{p}). \end{aligned} \quad (1.4)$$

Para preservar la exactitud en el proceso de generación de imágenes spin, las coordenadas spin (α, β) , son calculadas usualmente desde la nube de puntos [29], dado que éstas son medidas en coordenadas espaciales que no están de acuerdo a una grilla. La reducción de

¹unidad de almacenamiento de información

dimensionalidad asociada desde coordenadas 3D al dominio de la imagen spin es descrita por:

$$[x, y, z] = [r, \theta, z] \rightarrow [r, z] \approx [\alpha, \beta]. \quad (1.5)$$

Asumiendo una estimación normal precisa, (α, β) deberían aproximarse a las coordenadas cilíndricas (r, z) , relativas al punto \mathbf{p} . La coordenada cilíndrica θ es ignorada en el dominio de las imágenes spin para reducir la dimensión de los datos. Las imágenes spin son generadas de acuerdo al cumplimiento de tres criterios C_s de generación que se describen a continuación.

El primer criterio W , es usado para especificar el rango relativo a \mathbf{p} , sobre el cual los puntos pueden contribuir a la imagen spin. El segundo criterio es el tamaño del *bin* b_s . En cualquier histograma, b_s especifica el rango sobre el cual las coordenadas spin se mapean en un *bin* sencillo.

El primer paso en la construcción de la imagen spin es determinar cuales puntos pueden contribuir a la imagen spin $S\mathbf{p}$. Esto se logra ajustando los datos 3D en una grilla en el espacio de parámetros spin, y aplicando el criterio de generación de imágenes spin a la lista de puntos $L\{\mathbf{x}\}_{\mathbf{p}}$. Con base en las coordenadas spin (α, β) asociadas con todos los puntos en $L\{\mathbf{x}\}_{\mathbf{p}}$, los datos son ajustados a la grilla, de acuerdo con:

$$i = \left\lceil \frac{\alpha}{b_s} \right\rceil, \quad j = \left\lceil \frac{\frac{W}{2} + \beta}{b_s} \right\rceil. \quad (1.6)$$

En la práctica, el tamaño del *bin* es típicamente cuatro veces la media de la resolución de muestreo asociada con la nube de puntos. Esto es realizado con el fin de asegurar que múltiples puntos puedan ser mapeados al mismo *bin*, creando características de contraste. Estas características son importantes cuando se realiza un emparejamiento de imágenes spin.

Sólo los puntos $L\{\mathbf{x}\}_{\mathbf{p}}$ que cumplan las siguientes condiciones:

$$i < \left\lceil \frac{W}{b_s} \right\rceil, \quad j < \left\lceil \frac{W}{b_s} \right\rceil, \quad (1.7)$$

pueden contribuir a la $S\mathbf{p}$. El tercer criterio de generación de imágenes spin es el ángulo spin Λ_x , el cual está definido como:

$$\Lambda_x = \cos^{-1}(\mathbf{n} \cdot \mathbf{n}_{\mathbf{x}}), \quad (1.8)$$

Λ_x es la medición del ángulo entre \mathbf{n} y el vector normal $\mathbf{n}_{\mathbf{x}}$ asociado al punto \mathbf{x} . Λ_x es calculado en cada punto de $L\{\mathbf{x}\}_{\mathbf{p}}$ de acuerdo a las condiciones especificadas en la Ecuación 1.7. Por lo tanto, únicamente los puntos con ángulos spin menores que un umbral especificado por el usuario Λ_s , son guardados para la lista final de puntos disponibles $L\{\mathbf{x} \cap C_s\}_{\mathbf{p}}$ (únicamente puntos que cumplen los tres criterios).

Ya que $L\{\mathbf{x} \cap C_s\}_{\mathbf{p}}$ ha sido establecido, se utiliza la interpolación bilineal para reducir los errores generados en el ajuste de la grilla, dado que el error de ajuste introducido por el operador de ajuste en la Ecuación 1.6, es directamente proporcional a b_s . Para

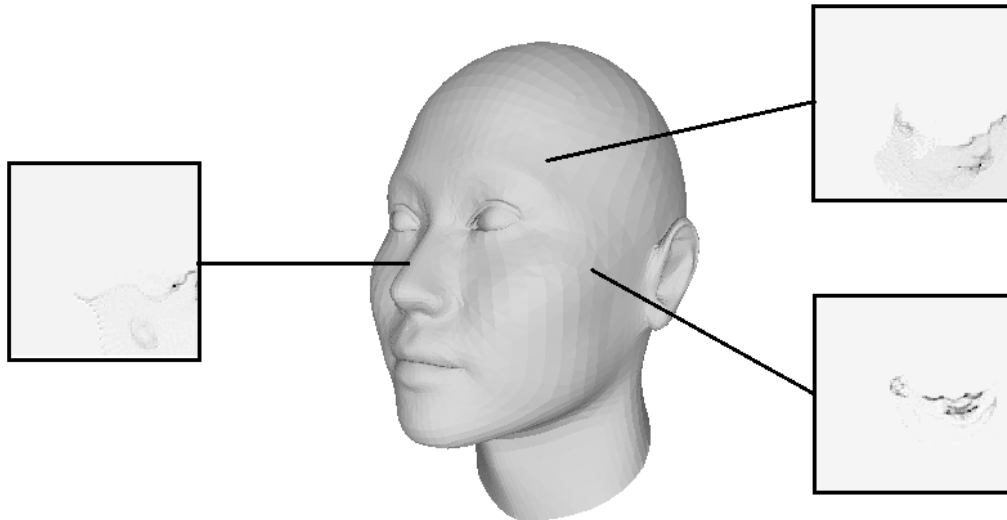


Figura 1.3: Ejemplo de imagenes spin.

mitigar este error, las coordenadas spin (α, β) y las coordenadas de ajuste de grilla son interpoladas bilinealmente para calcular la contribución de un punto con su ubicación de grilla adyacente. Los pesos de interpolación son calculados de acuerdo a:

$$a = \alpha - ib_s, \quad b = \beta + \frac{W}{2} - jb_s. \quad (1.9)$$

Una contribución a $S\mathbf{p}$ resultante desde la interpolación bilineal es calculada mediante:

$S\mathbf{p}(i, j + 1) = S\mathbf{p}(i, j + 1) + a(b_s - b)$	$S\mathbf{p}(i + 1, j + 1) = S\mathbf{p}(i + 1, j + 1) + (ab)$
$S\mathbf{p}(i, j) = S\mathbf{p}(i, j) + (b_s - a)(b_s - b)$	$S\mathbf{p}(i + 1, j) = S\mathbf{p}(i + 1, j) + b(b_s - a)$

(1.10)

Finalmente, todos los puntos 3D en $L\{\mathbf{x} \cap C_s\}_{\mathbf{p}}$ son ajustados en el *bin* de acuerdo con la Ecuación 1.10. Es importante reconocer que una imagen spin puede ser generada para cada punto en una nube de puntos 3D. La Figura 1.3 ilustra tres ejemplos de imágenes spin generadas desde una malla 3D, en tres puntos representativos de un rostro sintético. En ciertos casos, una representación de un objeto 3D mediante imágenes spin puede ser reconocida fácilmente, dado que las características de la imagen spin son directamente relacionadas a las simetrías sobre \mathbf{p} (ver Figura 1.3). Debido a que las coordenadas (α, β) son medidas relativas a \mathbf{p} y \mathbf{n} , las imágenes spin son invariantes a la orientación del objeto [28].

1.4. Imágenes spin multi-resolución

El muestreo uniforme es requerido para coincidir puntos en dos diferentes mallas mediante imágenes spin. En la implementación original de imágenes spin, únicamente vértices en la malla son agregadas en imágenes spin [28]. En [30] se aplican algoritmos de remues-

treo para asegurar la uniformidad de las superficies. En [31, 32] se agregan en la superficie, puntos uniformemente distribuidos entre los vértices de las mallas, en las imágenes spin.

En [33] se muestra que imágenes spin obtenidas mediante diferentes resoluciones, pueden ser utilizadas para emparejar superficies con densidades de puntos no uniformes, mediante un proceso denominado imágenes spin multi-resolución.

Para encontrar cuales puntos de un objeto corresponden a un punto de prueba, sea en el mismo o en otro objeto, se obtiene la imagen spin en dicho punto. Los coeficientes de correlación de la imagen spin en un punto de prueba, para todas las imágenes spin de los objetos, proveen una medida de que tan similar al punto de prueba son los puntos en la superficie. El coeficiente de correlación c es obtenido mediante la siguiente relación, donde s_1 y s_2 son las dos imágenes spin comparadas y n es el número de píxeles en una imagen spin:

$$c = \frac{n \sum s_1 s_2 - \sum s_1 \sum s_2}{((n \sum s_1^2) - (\sum s_1)^2)(n \sum s_2^2 - (\sum s_2)^2)}^{\frac{1}{2}}. \quad (1.11)$$

Como se mencionó anteriormente, existen 3 criterios para la generación de las imágenes spin, de los cuales 2 parámetros son de gran importancia en la generación de las imágenes spin multi-resolución: la longitud de soporte W y el tamaño del *bin* de una imagen. La longitud soporte determina la localidad de la imagen spin.

En [33] la longitud soporte es determinada como una fracción del tamaño del objeto, el cual se aproxima como la diagonal del *bounding box*². Una longitud de soporte pequeña conlleva a que solamente los puntos cercanos al punto central sean tenidos en cuenta, resultando en un descriptor local. De esta forma, las imágenes spin pueden variar desde un descriptor global hasta uno local (ver Figura 1.4(a)).

El tamaño del *bin* determina si los puntos vecinos están agrupados. Con un tamaño pequeño de *bin*, los puntos vecinos en la superficie caerán en *bins* separados. Si el *bin* es grande, es más probable que los puntos vecinos caigan en el mismo *bin*.

Cuando las imágenes spin están correlacionadas, diferencias pequeñas entre las imágenes son más aparentes cuando el tamaño del *bin* es pequeño. Por el contrario, con un tamaño de *bin* grande, pequeñas diferencias en los valores de *bin* no contribuyen significativamente a la correlación (ver Figura 1.4(b)).

El tamaño del *bin*, así como la longitud de soporte, pueden ser variadas para cambiar la imagen spin desde un descriptor con un mayor o menor grado de discriminancia. Variar la longitud de soporte y el tamaño del *bin* es similar a un submuestreo o sobremuestreo en imágenes corrientes [33]. En [33] se explota este aspecto, creando una pirámide de imágenes spin con la mayor resolución en el fondo y la menor resolución en la torre. Es posible generar dos pirámides, una mediante la variación del tamaño del *bin* b , y otra mediante la variación de la longitud de soporte l como una fracción del tamaño del objeto m .

El tamaño de la imagen spin resultante (donde el ancho es el número α de bins, y la altura es el número β de bins), está dado por las ecuaciones que se muestran a continuación.

$$ancho = \frac{ml}{b} + 1, \quad altura = \frac{2ml}{b} + 1. \quad (1.12)$$

²paralelepipedo que encierra al objeto

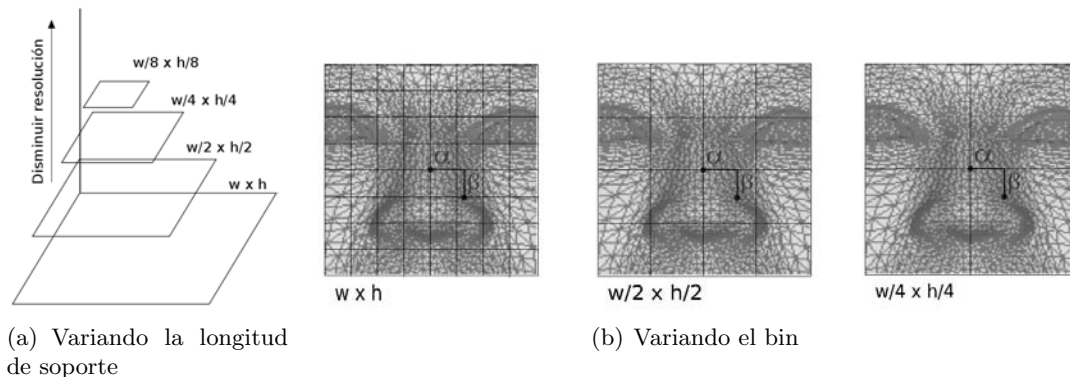


Figura 1.4: Dos formas de generar imágenes spin multi resolución.

El número β de bins es aproximadamente el doble del número α , ya que la distancia al punto central del plano tangente puede ser positiva o negativa y el radio es siempre positivo.

1.5. Índice de forma multi-escala

La representación espacio-escala de una superficie puede ser estimada mediante el uso de un kernel de suavizado gaussiano. Aunque de esa forma se pueden generar artefactos no deseados tales como cambios en la estructura geométrica de los modelos.

Un enfoque de obtención de características multi-escala que resuelve el inconveniente anterior, fue propuesto en [34], el cual se basa en la obtención del índice de forma a múltiples escalas, ajustando la superficie a vecindades de diferente tamaño, para aplicaciones de registro. Este método busca medir la variación local de superficie calculando la variación del índice de forma de cada punto en la vecindad. Los puntos correspondientes a la variación máxima de la superficie son seleccionados como características adecuadas, de acuerdo a los valores de la desviación estandar.

1.5.1. Variación local de escala

Dado que la topología de superficie local en un punto \mathbf{p} puede ser representada por su índice de forma S_I , la variación de forma local alrededor de \mathbf{p} puede ser estimada por la desviación estándar del índice de forma de los vecinos de \mathbf{p} , basados en la suposición de que la distribución es Gaussiana:

$$\sigma_{N_{\mathbf{p}}} = \sqrt{\frac{\sum_{\mathbf{p}_j \in N_r} (S_I(\mathbf{p}_j) - \mu_{N_r})^2}{n}}, \quad (1.13)$$

donde N_p es el conjunto de todos los n vértices en la vecindad de \mathbf{p} y μ_{N_p} es la media del índice de forma N_p :

$$\mu_{N_{\mathbf{p}}} = \frac{\sum_{\mathbf{p}_j \in N_{\mathbf{p}}} S_I(\mathbf{p}_j)}{n}. \quad (1.14)$$

1.5.2. Extracción de características multi-escala

En este marco de trabajo, el tamaño de la vecindad de un punto usado para obtener los puntos como la entrada al algoritmo de obtención de las curvaturas es definido como la escala. La escala r podría ser la distancia euclidiana o la distancia geodésica para una nube de puntos no estructurada. Los puntos en la superficie que son cercanos al punto a evaluar serán guardados. En el caso de mallas 3D, la escala r es escogida como el número de anillos alrededor de el vértice v .

El proceso de obtención del índice de forma multi-escala se muestra en el Algoritmo 1, el cual puede ser usado no sólo con el índice de forma multi-escala, sino con otros descriptores como la curvatura.

Algoritmo 1 Índice de Forma Multi-escala

Datos:

Superficie $S = \{\mathbf{V}, \mathbf{F}\}$ donde $\mathbf{V} = \{\mathbf{v}_i\} \in \mathbb{R}^3$: es un conjunto de vértices y $\mathbf{F} = \{\mathbf{f}_i\} \in \mathbb{N}^3$: es un conjunto de caras. $\mathbf{R} = \{r_k\}$: es un conjunto de escalas.

Algoritmo:

- 1: **for** $r \in \{r_k\}$ **do**
 - 2: **for** $\mathbf{v} \in \{\mathbf{v}_i\}$ **do**
 - 3: Encontrar la vecindad N_r en la escala r
 - 4: Ajuste un flujo a N_r
 - 5: Calcule las curvaturas principales k_1 y k_2
 - 6: Calcule el índice de forma $S_I(\mathbf{v})$
 - 7: $S_I(\mathbf{v}) = \frac{1}{2} - \frac{1}{\pi} \tan^{-1} \frac{k_1(p) + k_2(p)}{k_1(p) - k_2(p)}$.
 - 8: Calcule la media μ_{N_r} del índice de forma
 - 9: $\mu_{N_r} = \frac{\sum_{\mathbf{v}_j \in N_r} S_I(\mathbf{v}_j)}{n_r}$
 - 10: donde n_r es el número de vértices en N_r
 - 11: Calcular la variación local σ_{N_r}
 - 12: $\sigma_{N_r} = \sqrt{\frac{\sum_{\mathbf{v}_j \in N_r} (S_I(\mathbf{v}_j) - \mu_{N_r})^2}{n_r}}$
 - 13: **end for**
 - 14: Declarar marcas de escala r como vértices \mathbf{v} teniendo valores máximos σ_{N_r} , en la vecindad de radio r .
 - 15: **end for**
-

1.6. Estimación del costo computacional

La estimación de los descriptores presentados en este capítulo requieren de un conjunto de procedimientos, dentro de los cuales se realiza una gran cantidad de operaciones matemáticas, y de estructuras de datos que pueden variar de acuerdo a el sistema operativo, los tipos de datos usados y los algoritmos definidos para funciones matemáticas. En esta sección se hará un cálculo aproximado del número de operaciones que se requieren, para obtener cada uno de los descriptores planteados en los procedimientos implementados. El costo computacional de cada procedimiento estará dado en términos del tiempo de cómputo necesario para realizar las siguientes operaciones: suma (ts), resta (tr), multiplicación (tm), división (td) de números flotantes, independientemente del sistema operativo. No

se consideraron los tiempos de asignaciones de variables ni de corrimientos de bits para generar una expresión de costo computacional más sencilla.

Dentro de los procedimientos para obtener los descriptores es necesario realizar operaciones compuestas como: producto punto, producto cruz, normalización, norma, raíz cuadrada, multiplicación de matrices de 2 por 2, de 3 por 3, multiplicación de una matriz por un escalar y recorridos por la malla.

Como el código de algunas funciones no es accesible para ver que operaciones se realizan, ya que se encuentran en librerías, se estimaron algunos procedimientos. Esto se realizó de acuerdo a un análisis de las operaciones que se necesitan para obtener funciones simples; por ejemplo un producto punto en \mathbf{R}^3 requiere 3 multiplicaciones y 2 sumas. Como el calculo de la norma y normalización de un vector requieren calcular la raíz cuadrada de un valor flotante, se estimo la cantidad de operaciones de acuerdo al procedimiento para la obtención de la raíz inversa presentado en [35].

De esta forma se definió el conjunto de conjunto de valores de la Tabla 1.1.

Tabla 1.1: Costo de realizar operaciones involucradas.

Función	Operaciones realizadas
$Raiz^{-1}$	$4 \times tm + 3 \times tr$
Normalizar	$6 \times tm + 2 \times ts + Raiz^{-1}$
Raíz	$Raiz^{-1} + td$
Norma	$3 \times tm + Raiz$
t_{acos}	$10 \times tm$
t_{asen}	$10 \times tm$
t_{atan}	$10 \times tm$
t_{abs}	tm
M3x3	$27 \times tm + 18 \times ts$
M2X2	$8 \times tm + 4 \times ts$
Pcruz	$6 \times tm + 3 \times tr$
Ppunto	$3 \times tm + 2 \times ts$

1.6.1. Costo computacional de las curvaturas principales

De acuerdo al procedimiento para la estimación del tensor de curvatura presentado en el Apéndice A, que se encuentra implementado en el software de código abierto MehsLab [36], se realizó una estimación del costo computacional de la función implementada en C++.

En este procedimiento se tiene en cuenta que una malla está compuesta de N vértices y las caras tienen M vértices (típicamente 3 o 4). En primer lugar se realiza un ciclo sobre toda la malla para calcular el vector normal en cada vértice obteniendo que el costo del cálculo de los vectores normales en la malla es $N \times ((3,0) \times tr + (2,0) \times ts + (10,0) \times tm)$ de acuerdo a la Tabla 1.1.

Posteriormente, se realiza el procedimiento de estimación de curvaturas, el cual requiere recorrer toda la malla de nuevo y realizar una gran cantidad de ciclos anidados, normas, productos punto, productos cruz y multiplicaciones de matrices del cual se obtuvo la siguiente expresión:

$$cc_{k_1k_2} = N \times ((2,0) \times ts + (3,0) \times tr + (10,0) \times tm) + N \times (((7,0) \times tm + (3,0) \times tr + td + ts) \times M + (252,0) \times tm + (8,0) \times tm \times M + (47,0) \times tr + ((31,0) \times tm + (6,0) \times tr + td + (13,0) \times ts) \times M + (17,0) \times td + (tm + td) \times M + (51,0) \times ts + (16,0) \times tm^2).$$

En el caso de tener una malla triangular de 6292 vértices, el cual es el número de vértices de las imágenes sintéticas (ver Figura 2.5(a)), se obtiene el siguiente costo computacional:

$$cc_{k_1k_2} = (597740,0) \times ts + (100672,0) \times tm^2 + (2535676,0) \times tm + (484484,0) \times tr + (163592,0) \times td. \quad (1.15)$$

Para una imagen real (ver Figura 2.5(b)), compuesta de una malla triangular de 76490 vértices se obtiene el siguiente costo computacional:

$$cc_{k_1k_2} = (7266550,0) \times ts + (3,082547E7) \times tm + (1223840,0) \times tm^2 + (5889730,0) \times tr + (1988740,0) \times td.$$

Con base en lo anterior, se puede decir que descriptores basados en curvaturas principales como H , K , R y S_I tienen un costo computacional derivado del de calcular las curvaturas principales (ver Ecuación 1.15), siendo necesario realizar algunas operaciones adicionales. Por ejemplo el costo de la media es $cc_H = cc_{k_1k_2} + ts + tm$ el de la gaussiana, $cc_K = cc_{k_1k_2} + tm$, el de la curvatura $cc_R = cc_{k_1k_2} + ts + 2 \times tm + t_{raiz}$ y el del índice de forma $cc_{S_I} = cc_{k_1k_2} + ts + tr + 2 \times td + t_{atan}$.

1.6.2. Costo computacional de las imágenes spin

El proceso de obtención de las imágenes spin tiene una menor cantidad de instrucciones que el de curvaturas principales, pero depende de un factor adicional el cual es el tamaño de la imagen spin a generar (tam).

El procedimiento comienza por realizar un recorrido por la malla o por el conjunto de los puntos a los cuales se les va a calcular la imagen spin. En este recorrido, es necesario realizar los cálculos presentados en las ecuaciones 1.4, 1.6, 1.7, 1.8, 1.9 y 1.10, obteniendo una matriz de $tam \times tam$. De acuerdo al análisis de las operaciones involucradas en el algoritmo implementado, se obtiene que el costo computacional es el siguiente: $cc_{I_s} = (5 \times td + (15,0) \times ts + ts \times tam^2 + (34,0) \times tm + (13,0) \times tr) \times N$.

Para el caso de la obtención de la media de las imágenes spin (con $tam = 50$), en una malla sintética de 6292 vértices se obtiene un costo computacional de: $cc_{I_s} = (1,582438E7) \times ts + (213928,0) \times tm + (81796,0) \times tr + (31460,0) \times td$.

Para una imagen real de 76490 vértices se obtiene un costo computacional de:

$$cc_{I_s} = (2600660,0) \times tm + (994370,0) \times tr + (382450,0) \times td + (1,9237235E8) \times ts.$$

1.6.3. Costo computacional del índice de forma multi-escala

Para obtener el costo computacional de este procedimiento es necesario analizar la implementación del Algoritmo 1, en la cual se tiene en cuenta el costo computacional de la obtención del S_I (cc_{S_I}) y considerar que se deben ordenar N_r datos correspondientes a los valores del descriptor S_I en la vecindad de cada punto que hace parte de la lista a evaluar, la cual generalmente es una región.

$$cc_{S_{Ims}} = rk \times (to + N \times (cc_{S_I} + (tm + tr) \times Nr + ts \times Nr + 2 \times td)).$$

Como se puede observar este es el algoritmo más costoso dado que requiere calcular al descriptor S_I varias veces dependiendo el tamaño de vecindad elegido N_r y a que se requiere tener en cuenta un tiempo de ordenamiento para N_r datos (to).

1.7. Consideraciones finales

Se puede observar que algunos descriptores generan una información que corresponde a un número real como es el caso de las curvaturas principales y los derivados de ellas, otros generan una imagen 2D para cada punto como es el caso de las imágenes spin y otros generan una lista de puntos como es el caso del índice de forma multi-escala. Por tal razón, es necesario definir como y para que calcularlos, dado que algunos son más adecuados para representar información global y otros para local. Debido al tipo de datos que procesa el índice de forma multi-escala resulta adecuado para el análisis de regiones dado que requiere una lista de puntos para generar una lista con los de mayor varianza.

Se observa que el costo computacional de las imágenes spin puede aumentar de acuerdo al tamaño de la matriz de la imagen. Para los valores utilizados de tamaño y cantidad de vértices, las imágenes spin pueden requerir de una cantidad inferior de multiplicaciones que el procedimiento de curvaturas principales pero una cantidad superior de sumas, por tal razón su costo es variable y depende de la implementación de estas operaciones a bajo nivel. Se puede decir que el índice de forma multi-escala es el descriptor con mayor costo computacional, dado que implica la obtención de las curvaturas principales en vecindades de diferente tamaño y un ordenamiento de valores. Se debe evaluar no solo el aspecto de costo computacional, también se debe evaluar si un descriptor llega a ser lo suficientemente discriminante para una aplicación específica.

Dado que cada descriptor tiene diferentes características, se hace necesario en primer lugar, encontrar mecanismos que permitan la extracción de los mismos en diferentes puntos, contornos y regiones específicas del rostro, esto se presentara en el Capítulo 2. En segundo lugar es necesario encontrar mecanismos que permitan analizar las ventajas de cada uno lo cual se presenta en el Capítulo 4.

Extracción de descriptores en áreas de interés

El rostro humano está compuesto de diversas estructuras morfológicas cuyo análisis, segmentación, representación y visualización tiene utilidad en diversos campos, dentro de los cuales se tienen: cirugía del rostro, medicina forense, antropometría, reconocimiento de rostros, síntesis de voz, el modelado y animación facial. En este trabajo se realiza un análisis de un conjunto de áreas de interés que contienen un grupo significativo de regiones representativas del rostro. Este análisis fue realizado en conjuntos de rostros sintéticos y reales. La utilidad de trabajar con rostros sintéticos es que permiten realizar una validación inicial de los procedimientos, asegurando la obtención de resultados congruentes en términos de confiabilidad y tiempo de ejecución, dado que sus geometrías tridimensionales son de baja complejidad (menor cantidad de vértices que un rostro real), y no presentan discontinuidades, huecos, o asimetría. Por otra parte, la utilización de rostros reales permite que el análisis tenga mayor validez y sea susceptible de ser utilizado e incorporado en una aplicación práctica, como un sistema de reconocimiento de rostros. En este capítulo se define el modelo del rostro, base para la definición de las regiones sobre la cual se realizara la extracción de puntos, contornos y regiones.

2.1. Plantilla del rostro

Cada modelo del rostro es segmentado en 28 regiones (ver Figura 2.1), las cuales corresponden a regiones anatómicas del tejido blando del rostro, usadas para describir una lesión en medicina forense y/o para planear una cirugía en muchos otros contextos médicos. El contorno de una región es definido mediante curvas de Bézier, cada una con dos puntos de quiebre y dos puntos de control. La implementación de la curvas de Bézier se realizó en C++, tomando como base la presentada en [37] y se adicionó al software MeshLab [36]. La plantilla completa está compuesta por 68 contornos de regiones, 46 puntos de quiebre, 136 puntos medios y 338 triángulos; tanto los puntos de quiebre como los puntos medios son vértices de la plantilla del rostro 3D (ver Figura 2.2) [38].

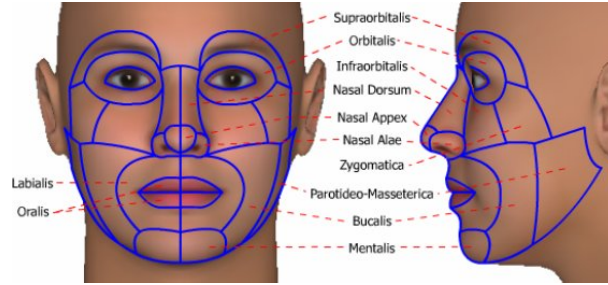


Figura 2.1: Regiones del rostro humano que se analizan en este trabajo.

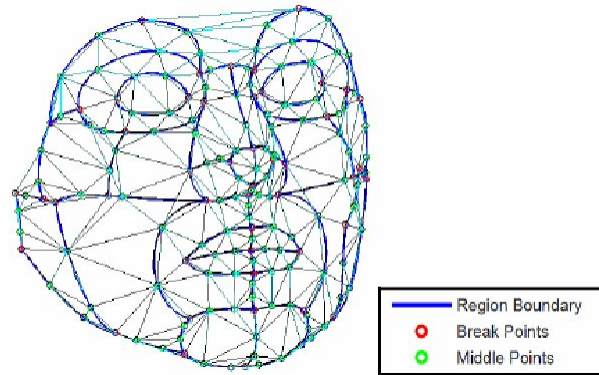


Figura 2.2: Modelo parametrizado del rostro.

2.2. Extracción de descriptores en puntos de interés

El proceso de extracción de descriptores de curvaturas k_1 , k_2 , H , K , R y S_I en un punto seleccionado de la malla, se realiza haciendo uso de las librerías que hacen parte del MeshLab [36], en las cuales se encuentra un algoritmo de obtención de curvaturas principales k_1 y k_2 , basado en [39].

Las imágenes spin se calcularon para 3 resoluciones, de acuerdo a un esquema multi-resolución (ver Sección 1.4), el cual puede ser utilizado como descriptor local al realizar un promedio de la imagen, para lo cual se realizó una implementación propia y se incluyó a un complemento del *MeshLab* [36] llamado *editpickpoints*. El índice de forma multi-escala no se utilizó como descriptor local, dado que su algoritmo necesita un conjunto de vértices, por lo tanto se extraerá y se analizará en regiones.

En un trabajo anterior enfocado al análisis de descriptores de curvatura, se realizó una obtención manual de descriptores en 24 puntos de interés (color azul), en mallas sintéticas y reales, con los cuales se podía obtener otros 29 puntos (color rojo), gracias a la simetría y las proporciones del rostro, llegando a 53 puntos (ver Figura 2.3) [1].

La Figura 2.4, muestra el resultado del cálculo de los histogramas de los diferentes descriptores de forma basados en curvaturas. Los valores de variación de cada paleta dependen de el mínimo y el máximo del descriptor en cada modelo. En la Tabla 2.1, se presentan los rangos de variación del valor medio de cada descriptor, aplicado a imágenes reales y sintéticas del rostro, que se consideraron en el marco experimental para la extracción de los descriptores (ver Sección 3.2 del Capítulo 2).

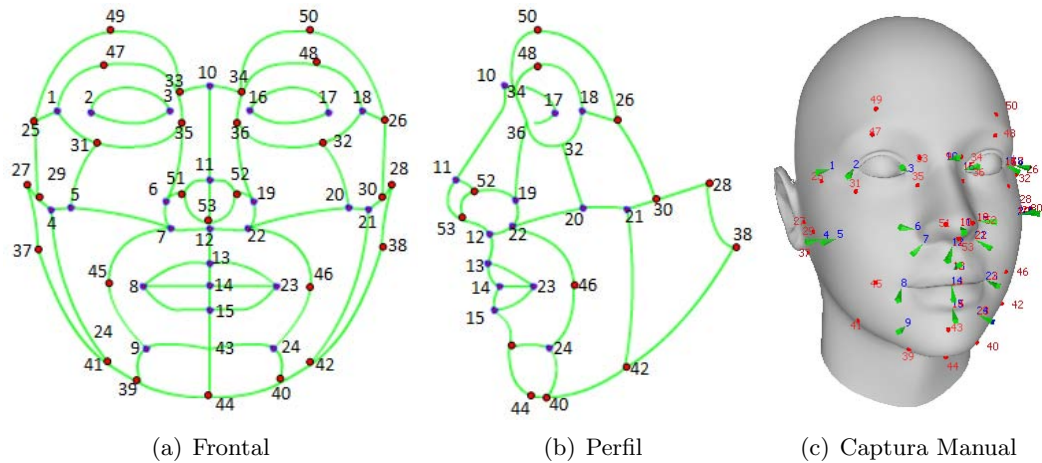


Figura 2.3: Extracción de 53 puntos [1]

La paleta del histograma es dividida en cuatro rangos, desde rojo (mínimo) hasta azul (máximo) como se puede observar en la Figura 2.4(g).

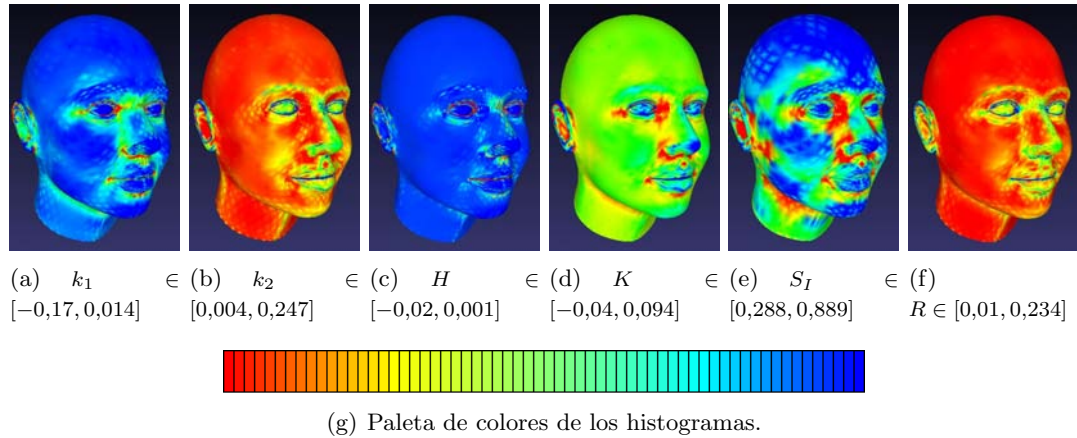
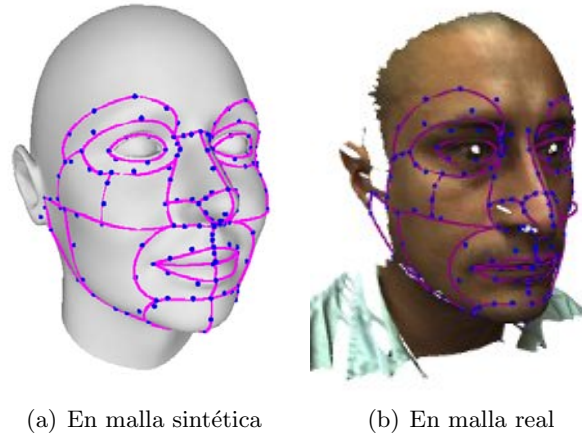


Figura 2.4: Descriptores de forma basados en curvatura calculados sobre la superficie del rostro.

Tabla 2.1: Rangos de variación de descriptores de curvatura sobre la superficie del rostro.

D	<i>FemDB</i>		<i>MaleDB</i>		<i>3DImDB</i>	
	min	max	min	max	min	max
k_1	-0,489	0.0879	-0,376	0.0850	-0,653	0.0611
k_2	-0,018	0.2327	-0,012	0.2278	-0,047	0.2554
H	-0,222	0.1225	-0,172	0.1154	-0,292	0.1326
K	-0,045	0.0125	-0,036	0.0116	-0,078	0.0140
S_I	0,146	0.9599	0,185	0.9666	0,148	0.9012
R	0,004	0.3525	0,004	0.2800	0,009	0.4681



(a) En malla sintética (b) En malla real

Figura 2.5: Ajuste de la plantilla

2.2.1. Registro semi-automático

Se desarrolló un proceso de registro semi automático para ajustar la plantilla a cada uno de los rostros, usando una cantidad inferior de puntos extraídos manualmente p, q, r y s , que permiten tener una proporción del ancho y el alto del rostro, y un cálculo de su orientación en el espacio. Gracias a esto, se logró obtener un muestreo mayor, compuesto de 182 puntos por modelo.

El proceso comienza con la obtención del centro geométrico del objeto (x_c, y_c, z_c) , mediante los puntos obtenidos manualmente.

$$x_c = \frac{1}{N} \sum_{i=0}^{N-1} x_i, \quad y_c = \frac{1}{N} \sum_{i=0}^{N-1} y_i, \quad z_c = \frac{1}{N} \sum_{i=0}^{N-1} z_i. \quad (2.1)$$

Utilizando los datos capturados del rostro se genera un vector horizontal $\vec{V}_1 = q - p$, un vector vertical $\vec{V}_2 = q - p$ y se calculan sus magnitudes. Luego se obtiene un tercer vector \vec{V}_3 , como el producto cruz de los anteriores. Se realiza una traslación con parámetros $(-x_c, -y_c, -z_c)$ en la plantilla. Se normalizan los vectores anteriores denominándolos \vec{v}_1, \vec{v}_2 y \vec{v}_3 , para agruparlos como columnas de una nueva matriz de transformación M , la cual es una rotación 3D, que permite alinear la plantilla con el rostro mediante la matriz inversa. Es necesario obtener los mismos vectores y magnitudes en la plantilla, así como la matriz de transformación, denominada M' , la cual define la alineación con los ejes x, y y z .

Posteriormente, se realiza un escalamiento basado en las magnitudes obtenidas, donde los factores de escalamiento se obtienen como un cociente entre las distancias del modelo real y el de la plantilla, de tal forma que la plantilla tenga dimensiones cercanas al modelo. Este procedimiento puede ser observado en el Algoritmo 2.

Finalmente, la plantilla es ajustada manualmente sobre cada modelo (malla del rostro), para asegurar la integridad de los resultados (ver Figura 2.5).

La obtención de los descriptores en los puntos de interés después del registro se puede realizar ya sea en forma manual [1] o mediante el uso del Algoritmo 2. Los problemas de encontrar cuales puntos de la imagen 3D (rostro) pasan por cualquiera de los contor-

Algoritmo 2 Ajuste y obtención de puntos de interés

Datos:

Superficie $S = \{\mathbf{V}, \mathbf{F}\}$ donde $\mathbf{V} = \{\mathbf{v}_i\} \in \mathbb{R}^3$: es un conjunto de vértices.

$\mathbf{V}' = \{\mathbf{v}'_j\} \in \mathbb{R}^3$: es un conjunto de puntos de la plantilla.

p, q, s, r : puntos obtenidos manualmente.

Algoritmo:

```

1:  $\vec{U} = q - p.$ 
2:  $\vec{V} = s - r.$ 
3:  $\vec{N} = \vec{U} \times \vec{V}.$ 
4:  $ancho = |\vec{U}|.$ 
5:  $alto = |\vec{V}|.$ 
6:  $\vec{u} = \frac{\vec{U}}{\|\vec{U}\|}, \vec{v} = \frac{\vec{V}}{\|\vec{V}\|}, \vec{n} = \frac{\vec{N}}{\|\vec{N}\|}$ 
7:  $M = [\vec{u}, \vec{v}, \vec{n}]$ 
8: for all  $v'$  do
9:    $v' = T(-x_c, -y_c, -z_c)v'$ 
10:   $v' = Sc(s_x, s_y, s_z)v'$ 
11:   $M_2 = M'M^{-1}$ 
12:   $v' = M_2v'$ 
13:   $v' = T(x_c, y_c, z_c)v'$ 
14:  for  $v \in \{S\}$  do
15:    if  $dist(v, v') < \epsilon$  then
16:      Almacenar  $v'$  en una lista
17:    end if
18:  end for
19: end for

```

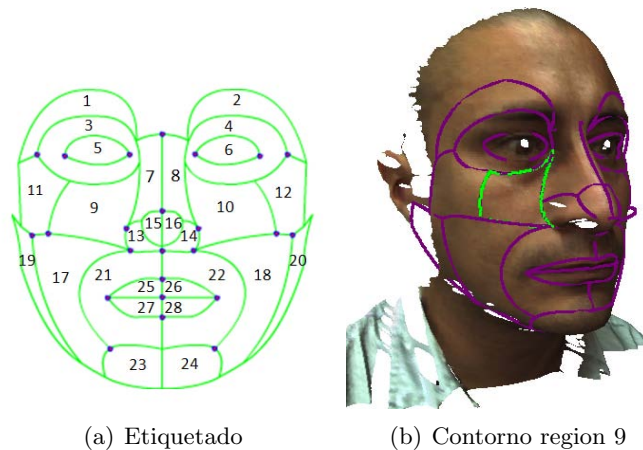


Figura 2.6: Etiquetado de las regiones y un contorno obtenido.

nos (plantilla ajustada), y de encontrar cuales pertenecen a las regiones, requirieron el desarrollo de procedimientos específicos que se explican a continuación.

2.3. Extracción de descriptores en contornos

Se desarrolló un procedimiento para obtener los contornos de cada región, obteniendo los puntos más cercanos a las curvas para todas las 28 regiones, las cuales fueron etiquetadas, como se puede observar en la Figura 2.6(a).

Para el cálculo de las curvas de Bézier se define un incremento en su parámetro independiente, inferior a la mitad de la distancia promedio de los puntos de la malla, con el objeto de realizar un muestreo que capture suficiente cantidad de puntos, evitando saltar puntos que pertenezcan a un contorno y realizar muestreos innecesarios. Con estos criterios se obtienen los puntos a evaluar en las mallas, tanto sintéticas como reales, que se consideraron para los experimentos. También se pueden variar los parámetros de los procedimientos, para obtener un muestreo de los puntos en la región. En el Algoritmo 3 se puede observar el procedimiento empleado para la extracción de los puntos en los contornos.

Se realiza una lista de vértices para cada región, denominada contorno, en la cual se almacenan los vértices pertenecientes a la malla que estén más cercanos a cada una de las curvas que pertenecen al modelo. Este procedimiento se desarrolla siguiendo un orden consecutivo hasta volver al punto inicial, formando una figura cerrada, lo cual puede ser considerado una polilínea en 3D.

En la Figura 2.6(b) se puede observar la obtención del contorno para la región infra orbital, la cual fue etiquetada como 9.

2.4. Extracción de descriptores en regiones

Se desarrolló un procedimiento para obtener los puntos que están dentro de cada una de las 28 regiones acotadas, basada en una búsqueda incremental de puntos al interior de

Algoritmo 3 Obtención de contornos**Datos:**

Superficie $S = \{\mathbf{V}, \mathbf{F}\}$, donde $\mathbf{V} = \{\mathbf{v}_i\} \in \mathbb{R}^3$: es un conjunto de vértices.

Estructura $S' = \{\mathbf{V}', \mathbf{CTR}, \mathbf{CUR}, \mathbf{REG}\}$, donde $\mathbf{V}' = \{\mathbf{v}'_i\} \in \mathbb{R}^3$: es un conjunto de puntos de la plantilla, $\mathbf{CTR} = \{\mathbf{ctr}_j\} \in \mathbb{R}^3$: es un conjunto de puntos de control de la plantilla, $\mathbf{CUR} = \{\mathbf{cur}_k\} \in \mathbb{N}$: es un conjunto índices de curvas en la plantilla y $\mathbf{REG} = \{\mathbf{reg}_l\} \in \mathbb{N}$: es un conjunto índices de regiones del rostro.

Algoritmo:

```

1: Calcular la resolución de la malla res.
2: Definir incremento inc  $< \frac{res}{2}$ .
3: for reg  $\in \{\mathbf{reg}_l\}$  do
4:   for cur  $\in \{\mathbf{cur}_k\}$  do
5:     for v  $\in \{\mathbf{v}_i\}$  do
6:       if distancia(v, v')  $< \epsilon$  then
7:         Almacenar v' en la listal
8:       end if
9:     end for
10:   end for
11: end for

```

cada región definida en la plantilla, para luego extraer los descriptores que se mencionaron en el Capítulo 1. Para obtener los puntos que están dentro de una región, se desarrolló un procedimiento que realiza una búsqueda en dirección, que apunta hacia el punto central de la región. Este punto es definido como el punto perteneciente a la malla más cercano a la distancia media entre dos puntos opuestos de la región en la plantilla. Cuando la región es convexa, el centro de la región *cr* es cercano al centro geométrico. Cuando la región no es convexa, se puede separar en secciones que sean convexas, y aplicar el mismo procedimiento encontrando puntos centrales en las mismas.

El recorrido se realiza en contornos internos cada vez más pequeños, usando una subdivisión de la distancia entre cada punto del contorno de la región y el centro de la región. La subdivisión depende de la resolución de la malla. El valor de la subdivisión indica la cantidad de incrementos que se deben hacer para llegar al centro de la región. La subdivisión en general se obtiene al dividir la distancia de cada punto del contorno hasta el punto central. Se encontró, empíricamente, que la subdivisión puede ser 15 para imágenes reales y puede reducirse hasta 7 en imágenes sintéticas.

La ventaja de este procedimiento basado en subdivisiones de la distancia, con respecto a uno basado en vecindades, es que no se queda estancado en regiones que presenten discontinuidades o huecos, lo cual es un caso común en imágenes 3D. Otra ventaja es que se puede disminuir la subdivisión, y realizar un muestreo de puntos en un tiempo mucho menor.

En el Algoritmo 4 se observa el procedimiento empleado para la extracción de los puntos interiores a los contornos, el cual debe aplicarse a cada uno de los contornos obtenidos con el Algoritmo 3.

En la Figura 2.7 se muestra la obtención de los puntos dentro de la región infra orbital.

Como validación del proceso de extracción de regiones, se obtuvieron los descriptores en las regiones para diferentes valores de subdivisión, partiendo desde el valor 30 hasta 3.

Algoritmo 4 Obtención de regiones**Datos:**

Superficie $S = \{\mathbf{V}, \mathbf{F}\}$, dónde $\mathbf{V} = \{\mathbf{v}_i\} \in \mathbb{R}^3$: es un conjunto de vértices.

Estructura $S' = \{\mathbf{V}', \mathbf{CTR}, \mathbf{CUR}, \mathbf{REG}\}$, dónde $\mathbf{V}' = \{\mathbf{v}_i\} \in \mathbb{R}^3$: es un conjunto de puntos de la plantilla, $\mathbf{CTR} = \{\mathbf{ctr}_j\} \in \mathbb{R}^3$: es un conjunto de puntos de control de la plantilla, $\mathbf{CUR} = \{\mathbf{cur}_k\} \in \mathbb{N}$: es un conjunto índices de curvas en la plantilla y $\mathbf{REG} = \{\mathbf{reg}_l\} \in \mathbb{N}$: es un conjunto índices de regiones del rostro.

contorno_l es el conjunto de puntos $p_m \in \mathbb{R}^3$: que forman el contorno de la región l .

lista_1 , lista_2 y listacompleta son listas para almacenar de puntos de \mathbb{R}^3 .

Algoritmo:

- 1: Calcular la resolución de la malla res .
- 2: $inc = \frac{res}{2}$.
- 3: $\text{lista}_1 \leftarrow \text{contorno}_l$
- 4: $\text{listacompleta} \leftarrow \text{lista}_1$
- 5: $cr = \text{CentroRegion}(\text{contorno}_l)$
- 6: **repeat**
- 7: **for** $\mathbf{p} \in \text{lista}_1$ **do**
- 8: $\vec{v}_{dir} = cr - p$
- 9: $\vec{v}_{dir} \leftarrow \frac{\vec{v}_{dir}}{\|\vec{v}_{dir}\|}$
- 10: $p_{new} = p + \vec{v}_{dir} \cdot inc$
- 11: **for** $\mathbf{v} \in \{\mathbf{v}_i\}$ **do**
- 12: **if** $\text{dist}(p_{new}, \mathbf{p}) < \epsilon$ AND $p_{new} \notin \text{listacompleta}$ **then**
- 13: Almacenar p_{new} en la lista lista_2
- 14: **end if**
- 15: **end for**
- 16: $\text{lista}_1 \leftarrow \text{lista}_2$
- 17: **end for**
- 18: $\text{listacompleta} \leftarrow \text{listacompleta} + \text{lista}_1$
- 19: **until** $\text{distancia}(cr, \mathbf{p}) < \epsilon$

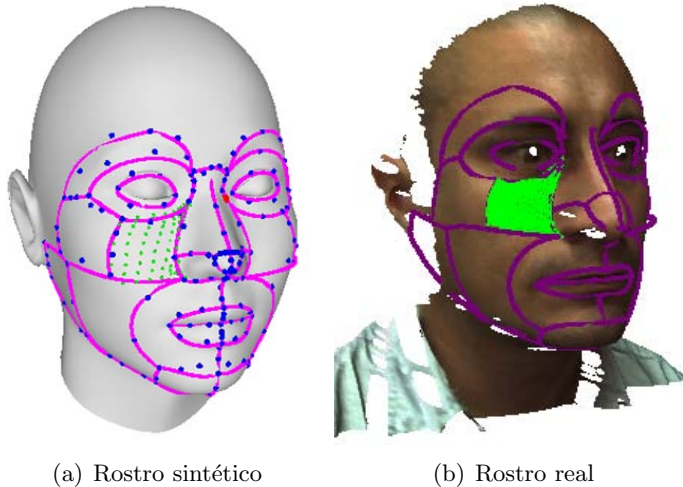


Figura 2.7: Obtención de los puntos interior a la región 9

El valor 30 corresponde a una búsqueda en intervalos muy pequeños para asegurar que se capturan todos los descriptores. En algunos casos puede existir sobre muestreo pero se tiene cuidado de no repetir los puntos encontrados. Cuando se llega a 3, se presenta un sub-muestreo en el cual se empiezan a perder puntos. Nótese que el error cuadrático medio (EMS), entre el descriptor calculado con una subdivisión 30 y los demás, aumentan a medida que disminuye el valor de subdivisión (ver Tabla 2.2).

Tabla 2.2: Error obtenido al cambiar la subdivisión en la búsqueda.

Subdivisión	EMS
15	0.0011418
7	0.0034175
5	0.0072800
3	0.0100000

2.5. Consideraciones finales

Los métodos desarrollados para la obtención de puntos, contornos y regiones permiten tener una base experimental sólida para un análisis de discriminancia de los descriptores, lo cual tiene gran aplicación en diferentes áreas.

Se decidió desarrollar un conjunto de procedimientos que respondan a las necesidades del problema de la segmentación de regiones, dadas las limitaciones que presenta el software comercial para adaptarse a procedimientos específicos que nacen en temas de investigación, y la intención de desarrollar sistemas independientes para futuras fases.

Aunque los métodos de registro planteados facilitan el trabajo de un ajuste preliminar de la plantilla a la imagen 3D, es necesario trabajar en el desarrollo un método de alta precisión que permita realizar un registro de la imagen 3D, automatizando así las primeras etapas de un sistema de análisis y reconocimiento de rostros.

Se espera, en un trabajo posterior, incorporar estas rutinas en un sistema para el análisis y reconocimiento de rostros, el cual puede extenderse a otras regiones anatómicas del cuerpo, y a otro tipo de aplicaciones en las cuales se requiera.

Materiales y Métodos

En éste capítulo se introducen los conjuntos de datos utilizados, los experimentos diseñados y el método para evaluar la relevancia de los descriptores de forma 3D que se estimaron en las regiones del rostro, mediante los procedimientos expuestos en el Capítulo 2. Esto se realiza, con el objeto de determinar cuales ofrecen información más relevante en ubicaciones específicas del rostro, mediante la utilización del análisis discriminante de Fisher. Esto permitirá definir cual es el descriptor más adecuado en superficies complejas, mostrando una aplicación en caracterización y reconocimiento de rostros en imágenes 3D.

3.1. Análisis discriminante de Fisher

Algoritmos como el de Análisis de Componentes Principales encuentran un conjunto de características que son útiles para la representación de datos, estos no permiten encontrar las características que tienen mayor relevancia, lo cual es un paso importante antes de realizar un proceso de clasificación o reconocimiento. Por otra parte, el análisis discriminante de Fisher encuentra las características que contienen la mayor información relevante, proyectando los datos en un espacio con el menor solapamiento entre clases.

Considerando un conjunto de datos con d dimensiones y n medidas $\mathbf{x}_1, \dots, \mathbf{x}_n$, el cual esta compuesto de l clases C_i . El análisis discriminante de Fisher es realizado con dos (subconjuntos o) clases con N_1 y N_2 elementos, obteniendo un criterio de separación.

Cada clase tiene media:

$$m_i = \frac{1}{N_i} \sum_{x \in C_i} \mathbf{x}. \quad (3.1)$$

Los datos son proyectados en un nuevo espacio.

$$y = \mathbf{w}^T \mathbf{x}. \quad (3.2)$$

Esas dos clases proyectadas tienen medias μ_1 y μ_2 , respectivamente mediante el uso de:

$$\mu_i = \frac{1}{N_i} \sum_{y \in C_i} \mathbf{y}. \quad (3.3)$$

La separación entre clases es obtenida encontrando un \mathbf{w} que maximice la función de acuerdo a [10]

$$m_2 - m_1 = \mathbf{w}^T (\mu_2 - \mu_1), \quad (3.4)$$

donde $m_i = \mathbf{w}^T \mu_i$. La varianza intra-clases de los datos transformados de la clase C_i es obtenida con la Ecuación 3.5.

$$\sigma_k = \sum_{n \in C_i} (y_i - m_k)^2. \quad (3.5)$$

El total de la varianza intra-clases para todo el conjunto de datos es definida como: $\sigma_i^2 + \sigma_j^2$.

El criterio de Fisher es obtenido como el cociente de la varianza entre-clases y la varianza intra-clases usando la Ecuación 3.6.

$$F_{ij} = \frac{(\mu_i - \mu_j)^2}{\sigma_i^2 + \sigma_j^2}. \quad (3.6)$$

Finalmente, el coeficiente de Fisher es calculado como la media de todas las combinaciones de la Ecuación 3.6, evaluada para cada característica.

3.2. Bases de datos de rostros

En este trabajo se utilizaron tres bases de datos: una compuesta de 10 modelos sintéticos llamada *DBs*, con diferentes características raciales (Africano, Asiático, Europeo e Indio), generados con un software comercial; otra compuesta de 5 imágenes de rango llamada *3DImDB*, obtenidas mediante un escaner laser 3D Minolta Vivid 9i en una pose estándar, sin expresión (ver Figura 3.2) y una última compuesta de 10 individuos proveniente de la base de datos *BU-3DFE* [40].

En una prueba preliminar para la obtención del descriptor S_I sobre 24 puntos del rostro en 10 individuos, los valores obtenidos se mostraban con valores dispersos (ver Figura 3.1(a)) [1]. Para mejorar la calidad de la malla se utilizó un filtro Laplaciano de suavizado, obteniendo mayor regularidad en los valores de los descriptores (ver Figura 3.1(b)).

En este trabajo, se utilizó el software Meshlab [36] para procesar la información 3D. La Figura 3.2 muestra algunos de los modelos de las bases de datos.

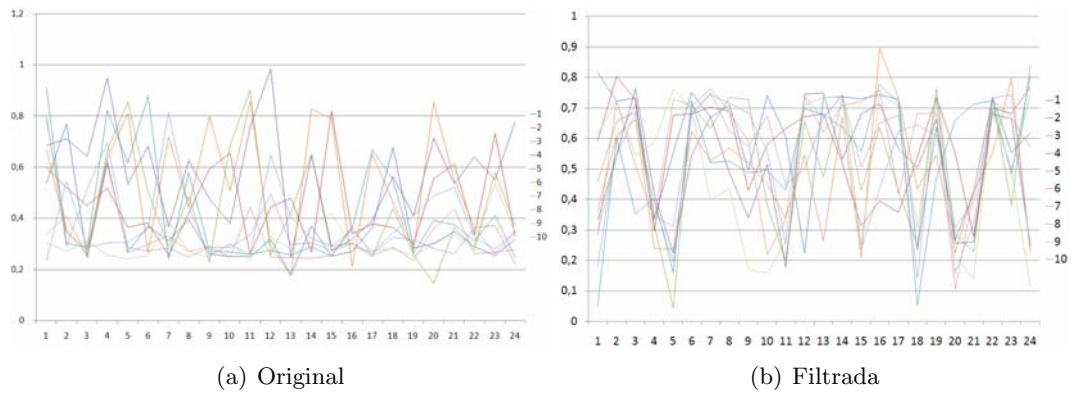


Figura 3.1: Valores del índice de forma (base de datos *3DImDB*)



Figura 3.2: Imágenes de las bases de datos. *FemDB* (primera fila), *MaleDB* (segunda fila), *3DImDB* (tercer fila) y *BU-3DFE* (cuarta fila).

3.3. Experimentos

Se realizaron dos tipos de experimentos de relevancia. El primero (relevancia global) para establecer cual descriptor es el más representativo sobre todo el conjunto de puntos y regiones que están presentes en el modelo facial. El segundo experimento es compuesto de un conjunto de pruebas (relevancia local) realizadas sobre regiones faciales.

3.3.1. Análisis Global

El análisis global se realiza obteniendo los coeficientes de Fisher de los 182 puntos del rostro, de las medias de los contornos de todas las regiones y sobre las medias de todas las regiones. En este experimento intervienen los descriptores de curvatura presentados anteriormente (k_1 , k_2 , H , K , R y S_I) en [41, 42] y las medias de las imágenes spin obtenidas con 3 diferentes subdivisiones (5, 10, 20) μI_{S_5} , $\mu I_{S_{10}}$, $\mu I_{S_{20}}$. Estas subdivisiones hacen que el tamaño del *bin* disminuya. Dado que los descriptores de curvatura están compuestos de un solo valor en cada punto, y la imagen spin es una matriz de valores, se evaluará la media de la imagen spin, para poder usar el esquema del análisis discriminante de Fisher en una comparación de 9 descriptores, siendo los 3 últimos las medias de las imágenes spin.

3.3.2. Análisis local

En este experimento se realiza un análisis de discriminancia de contornos y regiones sobre áreas representativas del rostro (ojos, mejillas, boca, nariz y barbilla), mediante el uso del coeficiente de Fisher. La Tabla 3.1 muestra las regiones consideradas en cada prueba de acuerdo a la Figura 2.6(a). En este caso, se realizó un análisis discriminante del promedio de los contornos y las regiones de cada una de las áreas. Se adicionó la media del descriptor S_I de la lista de puntos obtenidos mediante el procedimiento del índice de forma multi escala S_{Ims} . Dado que tanto la entrada como la salida del algoritmo de índice de forma multi-escala es un conjunto de puntos, se puede pensar que es un descriptor adecuado para regiones, por lo que se calcula la media de los descriptores en la lista de salida.

Se utilizó el Algoritmo 1 con un conjunto de 4 escalas $\mathbf{R} = \{r_k\}$, llegando a una vecindad de radio máximo igual a la cuarta parte de la distancia entre los ojos. Como el algoritmo requiere los puntos con mayor varianza, se tomó como criterio utilizar la mitad de los puntos en cada iteración, obteniendo una cantidad de puntos cercana al 10%.

Tabla 3.1: Regiones consideradas para cada prueba.

Prueba	Lado	Regiones
<i>Ojos</i>	Izquierdo	1, 3, 5
	Derecho	2, 4, 6
<i>Nariz</i>	Frontal	7,8,13,14,15,16
<i>Boca</i>	Frontal	25,26,27,28
<i>Barbilla</i>	Frontal	23,24
<i>Mejillas</i>	Izquierdo	9,11,17,19
	Derecho	10,12,18,20

Resultados del Análisis de Relevancia

En este capítulo se presenta la discusión de los resultados obtenidos de las pruebas planteadas en el análisis de relevancia, para establecer cuales características son relevantes a nivel global y local. El proceso planteado involucra la evaluación del ranking de los coeficientes de Fisher de los descriptores en: puntos, contornos y regiones representativas del rostro.

4.1. Resultados del Análisis Global

En esta prueba se obtiene el análisis de la relevancia de los descriptores en 182 puntos de la cara. En primer lugar se compara el valor de cada descriptor de cada punto con el otro, por pares. En este caso es necesario realizar 16471 comparaciones que corresponden a hacer las combinaciones de 182 elementos. Se encontró que para las imágenes reales el mejor descriptor es el S_I y que en ultimo lugar se encuentran las imágenes spin multi-resolución (ver Tabla 4.1). Sin embargo, para las imágenes sintéticas se observa que las imágenes spin multi-resolución son las de mejor ranking para los puntos y el descriptor $\mu I s_5$ es el más discriminante, lo cual indica que las resoluciones analizadas, son adecuadas para representar las variaciones en los puntos calculados en mallas sintéticas.

Tabla 4.1: Coeficientes de Fisher para la prueba *Global*. **R** indica el ranking.

D	Puntos			DBs Contornos			3DImDB Contornos			Regiones							
	Valor	R		Valor	R		Valor	R		Valor	R						
k_1	2,6044	6		11,2320	7		16,3860	5		0,9326	3		1,4084	4		4,6065	2
k_2	2,3436	7		22,1453	2		68,7680	1		0,7918	4		1,8121	3		2,7684	4
H	4,9909	5		9,8122	8		23,3560	3		0,9911	2		1,3659	5		2,3763	5
K	0,8933	9		3,8367	9		14,0190	6		0,3542	6		0,7285	6		0,9236	6
S_I	5,6349	4		13,7236	3		21,7780	4		1,7530	1		2,4251	1		4,3807	3
R	2,3147	8		29,0561	1		50,3680	2		0,7868	5		1,9145	2		4,6826	1
$\mu I s_5$	15,0607	1		12,4209	4		14,3740	7		0,1720	7		0,1926	7		0,2098	7
$\mu I s_{10}$	15,0578	2		12,2397	5		13,8200	8		0,1625	8		0,1875	8		0,2037	8
$\mu I s_{20}$	14,8695	3		12,1969	6		13,6190	9		0,1581	9		0,1847	9		0,2004	9

Tabla 4.2: Coeficientes de Fisher para la prueba *Global* en la base de datos BU-3DFE. **R** indica el ranking.

D	<i>Puntos</i>		<i>BU-3DFE</i> <i>Contornos</i>		<i>Regiones</i>	
	Valor	R	Valor	R	Valor	R
k_1	2,2282	3	4,6174	5	6,9139	3
k_2	2,1037	4	7,4371	2	8,4633	2
H	2,6959	2	8,4489	1	10,8471	1
K	0,8045	9	1,4859	7	2,8033	6
S_I	5,2520	1	6,3468	3	6,5048	5
R	1,6823	5	5,9997	4	6,8636	4
$\mu I s_5$	0,9640	6	1,5304	6	1,5270	7
$\mu I s_{10}$	0,8755	7	1,4406	8	1,4308	8
$\mu I s_{20}$	0,8362	8	1,3986	9	1,3869	9

4.1.1. Análisis en los contornos y las regiones

También se obtuvo la relevancia de los descriptores para las 28 regiones. Para esto se usaron las medias de los descriptores en los contornos y en las regiones. Para el caso de la obtención de los coeficientes de Fisher en los contornos y regiones se obtuvieron grandes diferencias el ranking en comparación con la prueba anterior. Se estima que, está diferencia se debe a que las mallas de la base de datos *DBs* están generadas con un espaciamiento diferente al de las mallas reales *3DimDB*, que consiste en una concentración de polígonos en algunas partes del rostro, lo cual conlleva a que la media de los descriptores en un contorno o región en una imagen sintética sea diferente a la media en una malla real. Se encontró que las imágenes obtenidas con cámaras 3D tienen mallas en forma de cuadrícula, mientras que las sintéticas presentan estructuras generadas por curvas paramétricas.

De esta prueba se destaca que el ranking de los coeficientes de Fisher es muy similar en las pruebas realizadas a los 182 puntos en las bases de datos *3DimDB* y *BU-3DFE* (ver columna 8 de la tabla 4.1 y columna 2 de la tabla 4.2) que corresponden a imágenes reales.

Sin embargo, como se pudo evidenciar, existen variaciones entre los niveles de algunos descriptores, y ciertas diferencias entre los rankings obtenidos, por lo tanto es necesario obtener la relevancia en áreas del rostro y así establecer cual descriptor, o combinación de los mismos, puede ser más adecuado para cada área o región.

4.2. Resultados del Análisis Local

4.2.1. Prueba del área de los ojos

En esta prueba la variable de lado fue incluida. Los resultados son mostrados en la Tabla 4.3 (los subíndices *I* y *D* corresponden a los lados izquierdo y derecho respectivamente). Se observó una gran diferencia entre el ranking de los contornos de ambos lados en la base de datos *3DimDB*. Esto se debe a que los contornos del área de los ojos presentan una gran variación en forma, pueden ser asimétricos, pueden no ser suficientes y estar localizados en diferentes lugares de la cara. También, el ranking de los contornos y las regiones fue muy diferente. Esto implica que los contornos y las regiones deben ser analizados en forma separada.

En las regiones de la base de datos $3DImDB$ se obtuvo que los lados tienen un ranking con ciertas similitudes, entre las que se destaca que en los contornos del lado izquierdo y derecho, el k_1 es el de mayor relevancia. También se destaca un conjunto de descriptores con valores significativos en las bases de datos de imágenes reales compuesto de H , k_1 , S_I , R y k_2 . Sin embargo, el ranking obtenido no presenta gran regularidad en las bases de datos.

Observando la Tabla 4.3 se puede inferir que la región de los ojos presenta gran cantidad de información de contornos, dado que existen valores altos de los coeficientes de Fisher para varios descriptores. Esto muestra que en regiones de alto detalle, es necesario tener una gran cantidad de puntos y descriptores, para caracterizar de forma adecuada los detalles presentes en áreas asimétricas del rostro.

Tabla 4.3: Coeficientes de Fisher para la prueba de los *Ojos*.

D	DB_s		$3DImDB_I$				$3DImDB_D$					
	Contorno Valor	R	Región Valor	R	Contorno Valor	R	Región Valor	R	Contorno Valor	R	Región Valor	R
k_1	6,6076	3	12,7501	6	1,5383	1	1,8563	3	0,1796	1	1,6345	2
k_2	14,1181	1	107,7870	1	0,4311	6	1,1240	5	0,7042	3	1,1722	4
H	3,1297	5	35,8472	4	0,4832	5	0,2410	7	0,4581	5	0,5111	7
K	1,4958	9	35,9046	3	0,5602	4	0,4559	6	0,0324	7	0,6447	6
S_I	3,1723	4	18,6456	5	0,9422	2	1,6957	4	0,5476	4	1,2111	3
R	11,9966	2	54,9540	2	0,8784	3	2,2541	1	0,7528	2	1,8417	1
μIs_5	1,9780	6	1,7662	7	0,0026	10	0,0036	10	0,0012	10	0,0016	10
μIs_{10}	1,6553	7	1,4359	8	0,0032	9	0,0046	9	0,0013	9	0,0020	9
μIs_{20}	1,5561	8	1,3342	9	0,0036	8	0,0051	8	0,0014	8	0,0022	8
S_{Ims}	0,2910	10	0,8152	10	0,3005	7	1,8744	2	0,3310	6	0,9621	5

Tabla 4.4: Coeficientes de Fisher para la prueba de los *Ojos* en la BD $BU-3DFE$.

D	$BU-3DFE_I$		$BU-3DFE_D$					
	Contorno Valor	R	Región Valor	R	Contorno Valor	R	Región Valor	R
k_1	2,8327	1	3,2271	3	3,4373	1	3,3930	1
k_2	0,5948	5	2,9371	4	0,2977	6	2,0835	4
H	2,6950	2	7,8714	1	2,3873	3	3,2648	2
K	0,5807	6	0,3117	7	0,3810	5	2,1298	5
S_I	2,1016	3	4,1230	2	2,4071	2	3,0082	3
R	1,5614	4	1,4274	5	2,0842	4	1,8721	6
μIs_5	0,0225	9	0,0092	10	0,0204	10	0,0069	10
μIs_{10}	0,0299	8	0,0202	9	0,0250	9	0,0153	9
μIs_{20}	0,0335	7	0,0267	8	0,0279	8	0,0203	8
S_{Ims}	0,1930	10	0,5360	6	0,0295	7	0,2778	7

4.2.2. Prueba del área de la nariz

Los resultados muestran que los mejores descriptores para esta área son los S_I y k_1 para los contornos y regiones de todos los *data sets* (ver Tabla 4.5). Los altos valores obtenidos indican que esta región tiene una gran cantidad de información. Adicionalmente, se puede observar que el valor del coeficiente discriminante de Fisher del S_I es muy alto, comparado con el resto de descriptores en casi todos los *data sets*. El descriptor k_1 también presenta gran relevancia para los contornos en todos los *data sets*. A pesar de que la forma de la

nariz varia drásticamente de un individuo a otro, la variación de forma ocurre en una baja escala, entonces las diferencias en resolución de las bases de datos evita que variaciones sutiles puedan ser comparadas entre sí.

Tabla 4.5: Coeficientes de Fisher para la prueba de la *Nariz*.

D	<i>DBs</i>				<i>3DIMDB</i>				<i>BU-3DFE</i>			
	Contorno Valor	R	Región Valor	R	Contorno Valor	R	Región Valor	R	Contorno Valor	R	Región Valor	R
k_1	18,6193	2	28,7061	3	2,0561	2	4,1290	2	3,4373	1	3,3930	1
k_2	4,0887	7	19,7612	4	1,0669	4	1,6593	5	0,2977	6	2,0835	5
H	11,7729	3	29,1349	2	1,8785	3	2,3465	4	2,3873	3	3,2648	2
K	7,0173	6	11,2663	6	0,6492	5	0,5534	7	0,3810	5	2,1298	4
S_I	25,1290	1	42,7492	1	7,2177	1	9,5492	1	2,4071	2	3,0082	3
R	8,9752	4	11,8045	5	0,4356	6	2,4090	3	2,0842	4	1,8721	6
μIs_5	0,2948	8	0,2221	8	0,0020	8	0,0021	10	0,0204	10	0,0069	10
μIs_{10}	0,2139	9	0,1876	9	0,0012	9	0,0023	9	0,0250	9	0,0153	9
μIs_{20}	0,1865	10	0,1819	10	0,0011	10	0,0025	8	0,0279	8	0,0203	8
S_{Ims}	7,8116	5	4,6214	7	0,2315	7	0,8583	6	0,0295	7	0,2778	7

4.2.3. Prueba del área de las mejillas

En esta prueba se notó que a pesar de que la superficie de las mejillas es suave, los valores de los coeficientes de Fisher fueron altos (ver Tabla 4.6). De manera contraria a la prueba de los ojos, los contornos y regiones tuvieron un ranking muy similar para ambos lados del rostro en imágenes reales. Para el caso de las regiones en imágenes reales, el ranking fue bastante similar dado que los primeros 4 descriptores tuvieron un ranking muy similar. Sin embargo, existen diferencias en los demás descriptores, lo cual sigue evidenciando que la asimetría afecta la capacidad de representación de los descriptores. En este caso, las variaciones de forma ocurren a una gran escala. Por lo tanto, para datos reales, los descriptores fueron capaces de caracterizar cambios de morfología.

Se destaca que los descriptores k_1 , S_I y H tuvieron valores sobresalientes en el nivel de discriminancia para las imágenes reales, y que los descriptores de imágenes spin multi resolución tuvieron valores altos para los contornos de las imágenes sintéticas, lo cual lleva a pensar que la resolución elegida permite obtener un buen nivel de discriminancia. Sin embargo, es necesario realizar todo un estudio con base en una variación más amplia de los parámetros de este tipo de descriptores.

4.2.4. Prueba del área de la boca

Los valores de los coeficientes de Fisher obtenidos en esta prueba fueron los más bajos (ver Tabla 4.8). Desde el punto de vista cuantitativo se puede observar que los descriptores S_I y R tienen gran importancia en contornos y regiones en las imágenes sintéticas. Por otra parte, se observa que el k_1 y el S_{Ims} son los de mayor valor de coeficiente de Fisher para las regiones de las imágenes reales. Sin embargo, sus índices de relevancia no son lo suficientemente altos en comparación con el de los otros descriptores. Por esta razón se estima que se requieren más de 3 descriptores para una buena representación de la región, los cuales pueden incluir a k_1 , S_{Ims} , K y R . Se observa que para este tipo de regiones y contornos, el S_{Ims} tuvo una alta relevancia para las imágenes reales, lo cual indica que el

Tabla 4.6: Coeficientes de Fisher para la prueba de las *Mejillas*.

D	DB_s				$3DIMDB_I$				$3DIMDB_D$			
	Contorno Valor	R	Región Valor	R	Contorno Valor	R	Región Valor	R	Contorno Valor	R	Región Valor	R
k_1	13,6058	6	18,2945	2	1,6820	2	3,2554	1	3,4547	2	13,5725	1
k_2	0,7193	10	0,9150	9	0,2184	6	0,2076	6	0,0121	10	0,0118	9
H	17,5711	2	16,8992	4	1,4026	4	1,4732	3	3,0056	3	2,9628	3
K	0,8127	9	0,9334	8	0,9146	5	1,0225	10	1,4751	4	0,7778	5
S_I	37,3392	1	29,0382	1	3,5102	1	2,5686	2	4,5442	1	7,1488	2
R	10,5720	7	17,5287	3	1,4029	3	1,1773	4	1,4433	5	1,0570	4
μIs_5	14,5944	5	14,7440	7	0,1071	9	0,1041	7	0,1057	8	0,0968	7
μIs_{10}	15,4960	4	15,6221	6	0,1092	8	0,1037	8	0,1089	7	0,0971	5
μIs_{20}	15,7088	3	15,7792	5	0,1099	7	0,1034	9	0,1099	6	0,0970	6
S_{Ims}	2,4700	8	0,2890	10	0,0551	10	0,2260	5	0,0898	9	0,0533	8

Tabla 4.7: Coeficientes de Fisher para la prueba de los *Mejillas* en la BD *BU-3DFE*.

D	$BU-3DFE_I$				$BU-3DFE_D$			
	Contorno Valor	R	Región Valor	R	Contorno Valor	R	Región Valor	R
k_1	6,2115	3	9,1906	1	5,7646	2	9,7526	1
k_2	0,1201	10	0,0920	10	0,1628	10	0,0907	10
H	6,9612	2	6,8259	2	5,1957	3	8,3876	2
K	0,3919	9	2,0244	5	0,1703	9	0,4222	8
S_I	9,2900	1	3,7831	3	6,2808	1	3,9229	3
R	3,1425	4	2,9512	4	2,8666	4	2,4320	4
μIs_5	0,9338	7	0,4646	9	0,8111	8	0,3835	9
μIs_{10}	0,9442	6	0,5131	8	0,8173	7	0,4280	7
μIs_{20}	0,9474	5	0,5406	7	0,8181	6	0,4496	6
S_{Ims}	0,6127	8	0,6954	6	0,9452	5	0,7538	5

tamaño de la región a analizar puede influir al calcular descriptores multi-escala, ya que las regiones anteriores presentaban un tamaño superior al de la boca.

Se presentan grandes diferencias entre los ranking de contornos y regiones de las bases de datos utilizadas. Las razones para este comportamiento tienen que ver con el tamaño de la región, dado que es pequeña comparada con otras regiones del rostro, otro factor es que la superficie presenta cambios suaves, excepto en las esquinas de la boca, haciendo que la tarea de localizar un punto sea complicada.

Tabla 4.8: Coeficientes de Fisher para la prueba de la *Boca*.

D	DB_s				$3DIMDB$				$BU-3DFE$			
	Contorno Valor	R	Región Valor	R	Contorno Valor	R	Región Valor	R	Contorno Valor	R	Región Valor	R
k_1	0,1872	4	0,0766	6	0,0137	6	0,0725	1	0,0222	6	0,0553	4
k_2	0,2248	3	0,1433	3	0,0270	4	0,0186	6	0,0566	5	0,0432	6
H	0,1413	6	0,1404	4	0,0096	7	0,0271	5	0,2040	1	0,3103	1
K	0,1452	5	0,0714	7	0,0558	2	0,0509	3	0,0601	4	0,0559	3
S_I	0,3287	1	1,3310	1	0,0189	5	0,0048	7	0,0757	3	0,2625	2
R	0,2761	2	0,1510	2	0,0430	3	0,0406	4	0,0020	7	0,0015	10
μIs_5	0,0349	10	0,0431	10	0,0005	8	0,0009	8	0,0011	8	0,0045	9
μIs_{10}	0,0401	9	0,0461	9	0,0005	9	0,0009	9	0,0011	9	0,0040	8
μIs_{20}	0,0417	8	0,0465	8	0,0005	10	0,0009	10	0,0010	10	0,0035	7
S_{Ims}	0,0526	7	0,1269	5	0,0820	1	0,0769	2	0,0799	2	0,0383	5

4.2.5. Prueba del área de la quijada

Los resultados se muestran en la Tabla 4.9. Las proporciones obtenidas indican similitud entre las formas presentes en estas regiones como se mencionó en los casos anteriores. Para los contornos y regiones de las imágenes sintéticas se obtuvieron similitudes en los descriptores k_2 y R . El descriptor H , es el mejor para contornos en ambos conjuntos de datos. Sin embargo, se notó que el S_{Ims} fue el de mayor coeficiente de Fisher para las regiones en imágenes sintéticas y el segundo mejor en imágenes reales, lo cual deja a este descriptor como opcionado para representar este tipo de regiones.

Por otra parte, los resultados en las regiones de los conjuntos de datos son muy diferentes, no se evidencia la prevalencia de algún descriptor para las regiones. Por esta razón, esta es la zona donde la identificación de puntos es más difícil, especialmente en la $3DImDB$. En este caso, se sugiere utilizar una combinación de al menos los descriptores K , S_{Ims} , k_2 y k_1 . Sin embargo, al igual que en el área de la boca, puede ser necesario realizar un estudio que incluya otros descriptores que sean más adecuados en esta área.

Tabla 4.9: Coeficientes de Fisher para la prueba de la *Quijada*.

D	DBs				3DImDB				BU-3DFE			
	Contorno Valor	R	Región Valor	R	Contorno Valor	R	Región Valor	R	Contorno Valor	R	Región Valor	R
k_1	0,0260	4	0,0048	5	0,0004	5	0,3457	4	0,0109	5	0,0447	2
k_2	0,0660	2	0,0466	2	0,2021	2	0,3484	3	0,0277	2	0,0157	4
H	0,0766	1	0,0273	4	0,2757	1	0,0466	7	0,0647	1	0,0417	3
K	0,0097	9	0,0047	6	0,0236	4	0,9831	1	0,0090	6	0,0003	8
S_I	0,0210	5	0,0004	10	0,0000	10	0,1072	5	0,0199	3	0,0579	1
R	0,0307	3	0,0392	3	0,0494	3	0,0998	6	0,0089	7	0,0080	5
μ_{Is_5}	0,0020	10	0,0023	9	0,0001	9	0,0001	8	0,0012	10	0,0001	9
$\mu_{Is_{10}}$	0,0035	8	0,0031	8	0,0002	6	0,0001	9	0,0015	9	0,0001	10
$\mu_{Is_{20}}$	0,0043	7	0,0038	7	0,0002	7	0,0000	10	0,0020	8	0,0009	7
S_{Ims}	0,0118	6	0,0859	1	0,0101	8	0,4048	2	0,0140	4	0,0035	6

4.2.6. Ranking de descriptores por área del rostro

Gracias a las pruebas anteriores, se definieron cuales son los descriptores más adecuados en los contornos y regiones del rostro de las 5 áreas propuestas para el análisis local, estos se pueden observar en la Tabla 4.10. Estas características pueden ser evaluadas en un sistema de reconocimiento de rostros, mediante imágenes 3D para cada una de las regiones, así como en mediciones de antropometría facial.

Tabla 4.10: Descriptores de acuerdo a las regiones del rostro.

Área	Contornos	Regiones
<i>Ojos</i>	k_1, S_I, R, H, k_2	H, k_1, R, S_I, S_{Ims}
<i>Nariz</i>	k_1, S_I, H, R	k_1, S_I, H, R
<i>Mejillas</i>	k_1, S_I, H, R	k_1, S_I, H, R
<i>Boca</i>	S_{Ims}, K, R, k_2	k_1, S_{Ims}, K, R
<i>Quijada</i>	H, k_2, R, K	R, S_{Ims}, k_2, k_1

4.3. Consideraciones Finales

En este capítulo se realizó un análisis de relevancia en diferentes regiones del rostro, las cuales están definidas con la intención de representar una gran cantidad de estructuras morfológicas, lo cual es de interés en diversos campos como la cirugía facial, la antropometría y los sistemas de reconocimiento de rostros mediante imágenes 3D. Para este análisis se utilizó el ranking de los coeficientes de Fisher.

En las pruebas donde se incluyó la lateralidad como variable, se encontró que los descriptores evaluados son aptos para capturar variaciones a causa de la asimetría natural del rostro. Sin embargo, es muy temprano para caracterizar dichas variaciones, para esto se debe realizar un análisis sobre una base de datos amplia, donde también se tenga en cuenta que dependiendo de la zona evaluada se requiere de un análisis a mayor o menor resolución (e.g., en las regiones de la nariz y la boca).

Aunque en las regiones que conforman la boca y el mentón, los descriptores estudiados presentaron diferencias en sus niveles de discriminancia, es necesario realizar un estudio que incluya otro tipo de descriptores que capturen de mejor forma la información local de superficie, o tener en cuenta que para poder identificar dichas regiones, la estrategia de búsqueda debe considerar la poca variabilidad en las superficies.

Es de resaltar, que en este trabajo se realizó un análisis discriminante sobre descriptores de forma para encontrar los de mayor relevancia. Se considero cada una de las características independientemente y no como una combinación. Razón por la cual se diferencia de otro tipo de análisis en los que se realiza reducción de dimensionalidad como PCA.

Aplicación al reconocimiento de rostros

El reconocimiento facial tridimensional es un tema de investigación reciente y busca subsanar las falencias que presentan los métodos basados en imágenes bidimensionales. Este tipo de sistemas tienen la ventaja de no ser afectados por cambios en la iluminación, dado que la representación viene dada como una nube de puntos o una malla 3D. En este capítulo se presenta un sistema de reconocimiento de rostros, basado en el modelo antropométrico del rostro mencionado en la Sección 2.1 (ver Figuras 2.1 y 2.2), los métodos de registro y extracción de características planteados en el Capítulo 2, los descriptores analizados en el Capítulo 1, y en los resultados del análisis de relevancia realizado en el Capítulo 4.

5.1. Sistema de reconocimiento de rostros

Un sistema de reconocimiento de rostros está compuesto de un conjunto de procesos, que permiten alimentar una base de datos con una o más imágenes de un conjunto de individuos etiquetados, y realizar un entrenamiento mediante técnicas de aprendizaje de máquina, que permitan identificar a los individuos con imágenes diferentes a las que fue entrenado. Un sistema más sofisticado podrá determinar características propias de los individuos, como género, edad, expresión facial, entre otros.

Un sistema de reconocimiento de rostros consta de las siguientes etapas: adquisición, preprocesamiento, registro, segmentación, extracción de características y reconocimiento (clasificación o interpretación). Se puede decir que estas etapas permiten generar una base de conocimiento propia del sistema, en la Figura 5.1 se muestran las principales etapas del sistema de reconocimiento.

5.1.1. Adquisición

El objetivo del sensor es crear una nube de puntos a partir de muestras geométricas tomadas en la superficie del objeto. A partir de estos puntos se puede generar una malla 3D. Un escáner 3D al igual que una cámara tradicional tiene un campo de visión en forma de cono, pero puede obtener información acerca de superficies sin iluminación. La etapa de adquisición de la imagen depende de las características propias del sensor utilizado (tipo

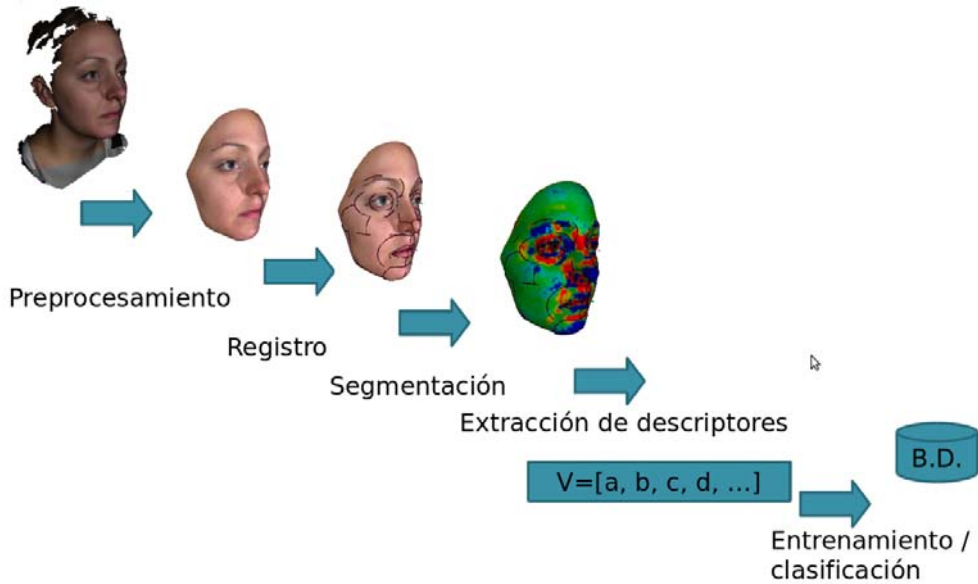


Figura 5.1: Esquema del sistema de reconocimiento.

de radiación que emite y recibe), distancia de captura, capacidad de apertura, posición. También depende del punto de referencia y distancia al objeto.

5.1.2. Preprocesamiento

En esta etapa se deben considerar los siguientes aspectos: filtrado, mejora y refinamiento de la malla, los cuales facilitan la manipulación de las imágenes y la obtención de descriptores con mayor confiabilidad. Dado que las imágenes 3D obtenidas mediante sensores de rango contienen ruido, se hace necesario el uso de técnicas de filtrado para las mallas, que mejoran los resultados obtenidos, como se observó en la Figura 3.2 del Capítulo 2. Se han logrado buenos resultados al filtrar las imágenes con los filtros Laplaciano y Taubin. Adicionalmente, existen mallas en las que se presentan discontinuidades o huecos debido al proceso de captura, estas se generan debido a limitaciones del sensor para capturar todas las estructuras morfológicas presentes en el rostro humano, como es el caso de la región bajo la nariz o las cejas, donde por lo general se generan huecos en las mallas 3D y se pierde información real, que puede ser completada mediante diversas técnicas que incluyen interpolación. Por otra parte, es posible que tanto en la captura como en los pasos anteriores se generen vértices repetidos, aislados, caras repetidas, mal conformadas, o con área cero, las cuales deben ser eliminadas para evitar tener mediciones incongruentes en los descriptores.

5.1.3. Registro

El proceso de registro que implica el ajuste de dos superficies 3D, es un problema que tiene básicamente tres enfoques. (1) Registrar la superficie de búsqueda a la superficie de referencia. (2) Extraer y ajustar representaciones centradas en el observador. (3) Extraer y ajustar representaciones centradas en el objeto de las superficies.

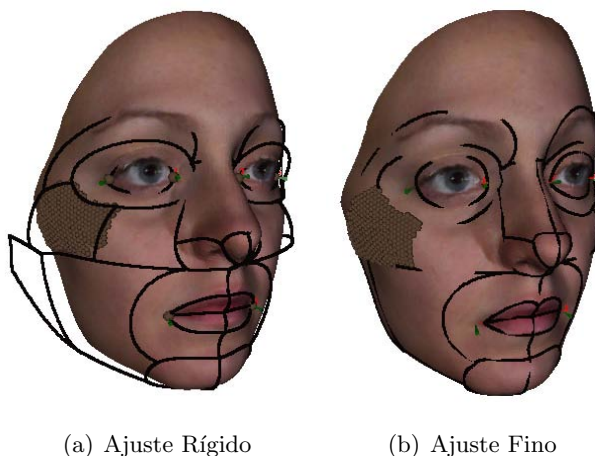


Figura 5.2: Ajuste de la plantilla y segmentación obtenida

En algunos trabajos relacionados como en [8], se realiza un ajuste de cada rostro en lo que se conoce como: normalización del rostro en cuanto a la pose. Lo cual tiene como objeto que todas las caras se encuentren en la misma posición y con la misma orientación en el espacio 3D. Luego del proceso anterior, es necesario ajustar una grilla 3D, o un modelo 3D, que permita la segmentación e identificación de áreas y regiones de interés en las imágenes 3D.

En este trabajo se realizó un ajuste de un modelo a cada rostro, lo cual evita modificar cada rostro. En este caso lo que se modifica es el modelo anatómico del rostro planteado, el cual está compuesto de una menor cantidad de vértices (182 puntos) que un rostro real, que puede estar compuesto de 4000 vértices y llegar a más de 54000 vértices, cuando el modelo es de alta resolución. El proceso de registro se realizó, con el proceso presentado en la Sección 2.2.1 del Capítulo 2.

5.1.4. Segmentación

En esta etapa se realiza una segmentación de las regiones de cada rostro, con el procedimiento planteado en la sección 2.4, de acuerdo al ajuste del modelo planteado en la Sección 2.1 del Capítulo 2. En principio, se realiza la segmentación buscando capturar todos los puntos que se encuentran dentro de las regiones, pero se evalúa que factibilidad tiene el realizar una segmentación basada en un ajuste rígido (ver Figura 5.2(a)), con el objeto de reducir el tiempo de cómputo que implica el realizar el ajuste fino (ver Figura 5.2(b)), de cada una de las curvas de Bézier que componen al modelo. Como se puede observar en la Figura 5.2 el resultado de segmentación de la región 11 (ver Figura 2.6(a)) tanto con el ajuste rígido como con el fino presentan gran similitud.

5.1.5. Extracción de características

Después de tener los puntos pertenecientes a cada región, es posible realizar la extracción de los descriptores planteados, y almacenarlos en un vector de características, de acuerdo al ranking obtenido en el análisis discriminante de Fisher, con el objeto de extraer información relevante para la clasificación a partir de las entidades cuantificables.

5.1.6. Entrenamiento

Haciendo uso de técnicas de aprendizaje de máquina como: redes bayesianas, redes neuronales, máquinas de vectores de soporte (SVM), se realiza un entrenamiento del sistema con base en los datos de entrada etiquetados. En una fase inicial se utiliza el vector de características para generar una colección (posiblemente ordenada y estructurada) de descriptores, denominada patrón, el cual define a cada individuo de una base de datos BD . La BD es compuesta de una cantidad de N individuos con M imágenes por individuo.

5.1.7. Clasificación

En esta etapa se evalúan patrones correspondientes a la información de algún individuo en particular con el objeto de determinar: 1) si el individuo hace parte de la base de datos o no, 2) identificar al individuo y 3) si el individuo posee alguna características de interés.

5.2. Experimentos de reconocimiento

Para evaluar la utilidad del procedimiento de extracción de características y el análisis de relevancia se plantea un conjunto de experimentos enfocados al reconocimiento de individuos, genero y expresiones.

5.2.1. Bases de datos

Se utilizaron 10 individuos de la base de datos $BU-3DFE$ de los cuales 5 son hombres y 5 son mujeres, y se tienen tres muestras de cada individuo que corresponden a tres expresiones faciales básicas: neutral, enojo y alegría.

5.2.2. Esquema de reconocimiento

En este caso se realiza el ajuste de la plantilla compuesta de 182 puntos y 26 regiones sobre cada uno de los modelos, se realiza la segmentación para obtener las características de cada región. Es bueno mencionar que aunque anteriormente se definieron 28 regiones (ver 2.1), existen 2 regiones laterales que no pueden obtenerse con precisión (estas corresponden a las regiones 19 y 20). Para los individuos en expresión neutral se realizó un ajuste de cada una de las curvas del modelo para que pasen por los contornos de las imágenes 3D. En las imágenes con expresiones de enojo y alegría se utilizó un ajuste rígido para evaluar su funcionamiento en la segmentación, dado que implica un menor tiempo de cómputo que hacer que las curvas pasen exactamente por el rostro. El ajuste rígido planteado consiste en escalar los puntos que componen la plantilla al tamaño de un rostro particular y la segmentación de las regiones se realiza buscando los puntos más cercanos a las regiones definidas dentro de las superficie que hacen parte de cada región.

Se evaluaron tres clasificadores: una red bayesiana con un estimador simple y el algoritmo de búsqueda $K2$, una red neuronal y una máquina de vectores de soporte utilizando el algoritmo de entrenamiento SMO (Sequential Minimal Optimization) [43].

Para el caso de la red neuronal se utilizó una capa de entrada con un número de neuronas correspondiente a la cantidad de descriptores multiplicada por la cantidad de

puntos o regiones a evaluar, una capa de salida que indica la cantidad de clases (número de individuos (10), género (2), número de expresiones (3)), para la capa intermedia se utilizaron dos alternativas: la primera con la misma cantidad de neuronas de salida y la segunda con el promedio entre el número de neuronas de entrada y salida.

Existen varios métodos para evaluar la capacidad del clasificación, en este caso se realizó una validación cruzada con un pliegue (*folding*) de 10, lo cual permite tener un indicio del error de generalización.

5.2.3. Prueba identificación del individuo

Esta prueba se divide en 4 experimentos: 1) identificación a partir de todos puntos característicos y todos los descriptores locales, 2) identificación a partir de puntos y características relevantes, 3) identificación con regiones y 4) identificación con regiones más relevantes y reducción de características.

5.2.4. Prueba identificación del género

En ésta prueba se etiquetaron todos los individuos como hombre ó mujer, según el caso, con el objeto de analizar la factibilidad de reconocer el género a partir de los descriptores más relevantes extraídos, también, en las regiones más relevantes.

5.2.5. Prueba identificación de la expresión

Dado que el dataset analizado estaba compuesto de individuos con 3 expresiones: neutral, enojo y alegría. Se etiquetaron las 3 instancias de los 10 individuos según el caso, con el objeto de encontrar la factibilidad de reconocer una expresión, a partir de los descriptores más relevantes extraídos en las regiones más relevantes.

5.3. Resultados

5.3.1. Identificación de un individuo

5.3.1.1. Identificación a partir de puntos característicos

En este experimento se evaluaron los puntos que componen a la plantilla, evitando puntos laterales, los cuales no se encuentran presentes en las imágenes procesadas. Los descriptores empleados fueron k_1 , k_2 , H , K , R y S_I y las medias de las imágenes spin obtenidas con 3 diferentes subdivisiones (5, 10, 20) μI_{s_5} , $\mu I_{s_{10}}$, $\mu I_{s_{20}}$. En este caso se cuenta con 168 puntos característicos y 9 descriptores por cada uno llegando a 1512 características por instancia, contando con 10 individuos y 3 instancias de cada uno.

Los resultados de utilizar un clasificador tipo red bayesiana, una red neuronal con 10 neuronas en la capa oculta (igual al de la capa de salida), una red neuronal con 761 neuronas en la capa oculta (el promedio entre neuronas en la capa de entrada y de salida) y una máquina con vectores de soporte usando el algoritmo SMO, se pueden observar en la Tabla 5.1. En este caso se puede observar que el uso de los 168 puntos para representar

el rostro humano, no presenta un rendimiento adecuado para identificar rostros con los clasificadores evaluados, dado que los resultados obtenidos no superan el 60% y pueden tomar un tiempo considerable debido a la cantidad de puntos y de descriptores. Esto también indica que una red neuronal puede tomar un tiempo importante, lo cual la hace inviable para un sistema que tenga que entrenar los descriptores en los puntos de una cantidad considerable de imágenes.

Tabla 5.1: Resultados identificación con 168 puntos y 9 descriptores con diferentes clasificadores

Resultado	Red Bayesiana	RN (10)	RN (761)	SMO
Instancias clasificadas correctamente	53.33 % (16)	60 % (22)	53.33 % (16)	46.66 % (14)
Instancias clasificadas incorrectamente	46.66 % (14)	40 % (8)	46.66 % (14)	53.33 % (16)
Error medio absoluto	0.0984	0.1147	0.1175	0.1695
Tiempo tomado en construir el modelo	0.17 seg	6.85 seg	487.99 seg	0.14 seg

5.3.1.2. Identificación a partir de puntos y características relevantes

En este caso se tomaron solo 62 puntos de los 182 de la plantilla, los cuales corresponden a las regiones de los ojos y la nariz, que son las que tienen coeficientes de Fisher más altos, como se expuso en la Sección 4.2.6. Teniendo en cuenta lo obtenido en la Sección 4.2.6 se seleccionaron solo 4 descriptores en dichas regiones, correspondientes a k_1 , H , R y S_I . Debido a lo anterior se cuenta con 248 características por individuo, 10 individuos y 3 instancias por individuo. Los resultados de este experimento se pueden observar en la Tabla 5.2, en la columna dos se presentan los resultados de clasificación haciendo uso de una red bayesiana, en la columna tres con una red neuronal de 10 neuronas en la capa oculta (número de individuos) y en la cuarta columna se presentan los resultados de clasificación haciendo uso de una red neuronal con 133 neuronas en la capa oculta (el promedio entre neuronas en la capa de entrada y de salida).

Como se puede observar, los tiempos de entrenamiento fueron mucho menores que en la prueba anterior y se llegó a una mejor clasificación.

Adicionalmente, se puede llegar a una mejor clasificación, al aumentar el número de neuronas en la capa oculta, a expensas de gastar más tiempo en el entrenamiento (12 segundos con 133 neuronas en la capa oculta y 0.92 segundos utilizando solo 10 neuronas en la capa oculta) lo cual hace que se tenga que evaluar más a fondo el uso de redes neuronales para grandes conjuntos de datos.

Tabla 5.2: Resultados identificación con 64 puntos y 4 descriptores con diferentes clasificadores

Resultado	Red Bayesiana	RN (10)	RN (133)	SMO
Instancias clasificadas correctamente	56.66 % (17)	73.33 % (22)	76.66 % (23)	70 % (21)
Instancias clasificadas incorrectamente	43.33 % (13)	26.66 % (8)	23.33 % (7)	30 % (9)
Error medio absoluto	0.086	0.0916	0.0818	0.1652
Tiempo tomado en construir el modelo	0.01 seg	0.92 seg	12 seg	0.12 seg

5.3.1.3. Identificación a partir de regiones

Para realizar este experimento, se utilizaron 26 regiones y 10 descriptores por región. En primer lugar se experimentó con una red bayesiana, posteriormente se utilizó una red neuronal con 10 neuronas en la capa oculta, luego con una red neuronal con 135 neuronas en la capa oculta (la cantidad de neuronas en la capa intermedia se calculo de la misma forma que en las pruebas anteriores) y finalmente se utilizó una maquina con vectores de soporte obteniendo los resultados presentados en la Tabla 5.3.

En este caso no se encuentran resultados superiores de clasificación al método anterior. Sin embargo, se pueden obtener resultados en menor tiempo. Esto indica que al promediar la información de descriptores en una región pueden existir pérdidas de información local, la cual puede ser necesaria para la identificación. En un estudio futuro se espera analizar otras características de las regiones faciales que guarden mayor información sobre el comportamiento de los descriptores en superficie de cada región.

Tabla 5.3: Resultados identificación para 26 regiones con diferentes clasificadores

Resultado	Red Bayesiana	RN (10)	RN (135)	SMO
Instancias clasificadas correctamente	50 % (15)	63.33 % (19)	63.33 % (19)	60 % (18)
Instancias clasificadas incorrectamente	50 % (15)	36.66 % (11)	36.66 % (11)	40 % (12)
Error medio absoluto	0.0982	0.1019	0.0933	0.167
Tiempo tomado en construir el modelo	0.01 seg	0.9 seg	11.66 seg	0.32 seg

5.3.1.4. Identificación con regiones más relevantes y reducción de características

Gracias a los resultados presentados en la Sección del Capítulo 4, para realizar el entrenamiento, se tomaron solamente las regiones que tenían los coeficientes de Fisher más altos, las cuales hacen parte de las áreas de los ojos y la nariz. Teniendo en cuenta lo obtenido en la Sección 4.2.6 se seleccionaron solo 4 descriptores en dichas regiones, correspondientes a k_1 , H , R y S_I . En este caso se entrenó una red bayesiana, una red neuronal con 10 neuronas en la capa oculta, una red neuronal con 57 neuronas en la capa oculta (la cantidad de neuronas en la capa intermedia se calculo de la misma forma que en las pruebas anteriores), y un SMO obteniendo los resultados de la Tabla 5.4. Aunque en las pruebas anteriores el algoritmo SMO no mostró un buen desempeño, para esta prueba se alcanzó un acierto de 80 %, y un tiempo de construcción del modelo inferior al de las redes neuronales, lo cual indica que puede ser una herramienta a considerar cuando se hace una reducción de características utilizando una análisis de relevancia.

Tabla 5.4: Resultados identificación para el área de los ojos y nariz con reducción de características con diferentes clasificadores

Resultado	Red Bayesiana	RN (10)	RN (57)	SMO
Instancias clasificadas correctamente	60 % (18)	63.33 % (19)	76.66 % (23)	80 % (24)
Instancias clasificadas incorrectamente	40 % (12)	36.66 % (11)	23.33 % (7)	20 % (6)
Error medio absoluto	0.0846	0.0933	0.079	0.162
Tiempo tomado en construir el modelo	0.001 seg	11.66 seg	0.81 seg	0.21 seg

5.3.2. Identificación de género

En éste caso se desarrolló un experimento enfocado a ver la capacidad de discriminancia de los descriptores más relevantes en las regiones más relevantes, para la identificación del genero de cada individuo. Teniendo en cuenta lo obtenido en la Sección 4.2.6 se seleccionaron solo 4 descriptores en dichas regiones, correspondientes a k_1 , H , R y S_I . Se utilizaron 4 clasificadores para evaluar la capacidad representativa de los descriptores en términos de permitir la identificación del género. Los clasificadores utilizados fueron: una red bayesiana, una red neuronal con 2 neuronas en la capa oculta, una red neuronal con 53 neuronas en la capa oculta (la cantidad de neuronas en la capa intermedia se calculo de la misma forma que en las pruebas anteriores), y una máquina con vectores de soporte utilizando el algoritmo SMO. Los resultados obtenidos se pueden observar en la Tabla 5.5. Al igual que en la prueba anterior el algoritmo SMO mostró un buen rendimiento como clasificador y en términos de tiempo, lo cual indica que puede ser tenido en cuenta en un sistema de reconocimiento de características faciales en el cual exista un requerimiento de tiempo de ejecución corto, el cual se vuelve difícil de alcanzar cuando se tiene una gran cantidad de individuos.

Tabla 5.5: Resultados identificación de género con diferentes clasificadores

Resultado	Red Bayesiana	RN (2)	RN (53)	SMO
Instancias clasificadas correctamente	80 % (24)	86.66 % (26)	86.66 % (26)	86.66 % (26)
Instancias clasificadas incorrectamente	20 % (6)	13.33 % (4)	23.33 % (4)	23.33 % (4)
Error medio absoluto	0.1962	0.1698	0.1794	0.1333
Tiempo tomado en construir el modelo	0.01 seg	0.27 seg	11.38 seg	0.01 seg

5.3.3. Identificación de expresión facial

La última prueba consiste en la identificación de la expresión a partir de los descriptores más relevantes en las regiones más relevantes (ojos y nariz), para lo cual se evaluaron 4 descriptores (k_1 , H , R y S_I). La base de datos *BU-3DFE* contiene rostros humanos con diferentes expresiones (neutral, enojo, tristeza, alegría, sorpresa, entre otras) en diferentes niveles de cada una de ellas excepto la neutral. Para esta prueba se utilizaron las expresiones: neutral, enojo y alegría (ver Figura 5.3). Las dos ultimas en el nivel más bajo de la expresión, es decir ligeramente, enojado o alegre. Los resultados obtenidos con los clasificadores utilizados pueden observarse en la Tabla 5.6, donde se puede observar que existe gran factibilidad de reconocer una expresión facial teniendo en cuenta solamente las regiones consideradas. Se destacan estos resultados dado que los rostros no fueron exactamente ajustados a la plantilla, es decir que se uso un ajuste rígido y se utilizaron individuos con expresiones faciales en el nivel menos acentuado que proporcionaba la base de datos *BU-3DFE*. Sin embargo, es bueno considerar el desarrollar un estudio enfocado a encontrar cuales son los descriptores y regiones más relevantes para caracterizar expresiones faciales.

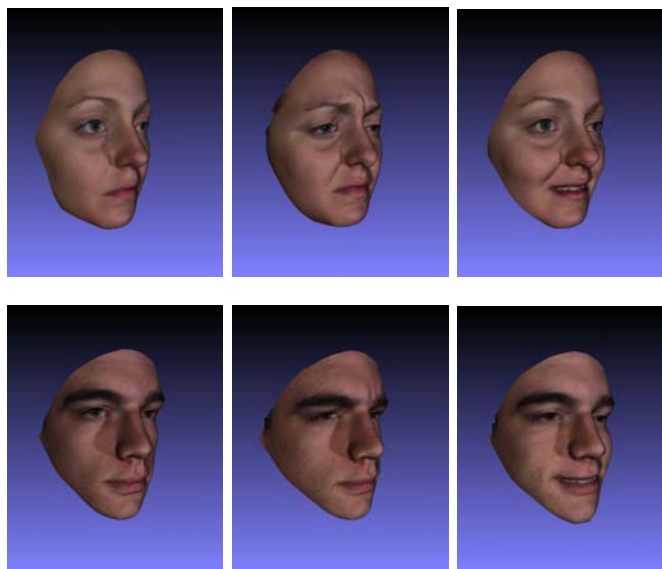


Figura 5.3: Imágenes de las bases de datos *BU-3DFE* con expresiones, primera columna expresión neutral, segunda columna expresión de enojo y tercera columna expresión alegría

Tabla 5.6: Resultados identificación de expresiones faciales con diferentes clasificadores

Resultado	Red Bayesiana	RN (3)	RN (54)	SMO
Instancias clasificadas correctamente	83.33 % (25)	86.66 % (26)	86.66 % (26)	86.66 % (26)
Instancias clasificadas incorrectamente	16.66 % (5)	13.33 % (4)	23.33 % (4)	23.33 % (4)
Error medio absoluto	0.1113	0.1494	0.1458	0.2519
Tiempo tomado en construir el modelo	0.06 seg	0.34 seg	11.35 seg	0.07 seg

5.4. Consideraciones finales

Es interesante observar que con las imágenes empleadas se obtiene un desempeño alto en el reconocimiento de género y expresión facial con todos los clasificadores, a pesar de que se hace un ajuste rígido del modelo utilizado a las imágenes 3D, el cual no es del todo preciso. Por lo tanto se espera mejorar estos resultados al usar un registro más preciso y/o desarrollando plantilla que involucre expresiones faciales.

En los experimentos de reconocimiento realizados se observó que al seleccionar los descriptores más discriminantes se puede lograr aumentar el rendimiento de los clasificadores utilizados y por ende la cantidad de individuos reconocidos correctamente. Se observó que el algoritmo SMO presenta un resultado superior en términos de tiempo y reconocimiento.

Las pruebas de identificación de género y expresión indican que al utilizar un clasificador tipo red neuronal, puede ser suficiente utilizar una cantidad de neuronas en la capa oculta cercana al número de neuronas de salida. Es decir, para identificar la expresión se obtuvo un buen resultado con 3 neuronas en la capa intermedia, las cuales corresponden al número de expresiones a identificar con el número de individuos evaluados.

Estos resultados plantean como trabajo futuro, la obtención de los parámetros óptimos para los clasificadores evaluados en grandes volúmenes de datos. Por ejemplo, la arquitec-

tura de una red neuronal más adecuada para identificar individuos o características del rostro, así como los parámetros que presentan mejor desempeño en la clasificación cuando se usa una maquina de soporte de vectores.

El análisis de relevancia planteado en este trabajo, puede ser extendido al análisis de descriptores en puntos y regiones más relevantes involucradas en la representación de: expresiones faciales, género, raza, edad, entre otros.

Conclusiones

En este trabajo se realizó un análisis de relevancia de 10 descriptores de forma, seis basados en la curvatura, 3 basados en imágenes spin multi resolución y uno basado en un procedimiento multi-escala. Este análisis se realizó sobre puntos, contornos, regiones, áreas y lados del rostro. Con base en el análisis de Fisher, se mostró como la morfología de la superficie del rostro influye en la capacidad de representación de los descriptores. A partir de esto, se determinó de forma específica cuales descriptores son más adecuados para caracterizar cada una de las principales regiones que definen el rostro.

Gracias a este trabajo se analizaron aspectos en la morfología del rostro tales como: asimetría (presente en las regiones de los ojos y mejillas), cantidad de puntos, suavidad e invariancia que pueden ser de importancia en un sistema de detección de características. Se destacan estos resultados pues en la mayoría de los trabajos donde se discute acerca de la asimetría, los análisis son cualitativos, por el contrario, en este trabajo se mostró de una forma cuantitativa, como un descriptor de forma pierde o gana capacidad de representación de una región o contorno de una misma zona del rostro pero de diferente lado.

Los métodos desarrollados para la obtención de puntos, contornos y regiones permiten tener una base experimental para un análisis de relevancia de otro tipo de descriptores. Se espera extender dichos procedimientos para que hagan parte de un sistema de análisis y reconocimiento de rostros.

El aspecto del costo computacional juega un papel crucial en cualquier sistema de extracción de características, especialmente en biometría dado que se tiene una gran cantidad de información que debe ser procesada en corto tiempo como es el caso de los sistemas para identificación de personas. Por tal razón no sólo se debe analizar la capacidad de representación de un descriptor, sino también el costo computacional que conlleva.

Se puede realizar un estudio de otro tipo de descriptores que capturen de mejor forma la información local de superficie, teniendo en cuenta que para poder identificar dichas regiones, la estrategia de búsqueda debe considerar la poca variabilidad en las superficies.

Al encontrar los descriptores en las regiones más relevantes y seleccionar las regiones con mayor relevancia mediante el análisis discriminante de Fisher, se evidenció de una forma cuantitativa la utilidad del mismo en reconocimiento de rostros, llegando a un incremento en la capacidad de los clasificadores y en la reducción del tiempo de entrenamiento de los mismos.

Aunque utilizar los promedios de regiones permitió hacer un reconocimiento de individuos e identificación de diversas características faciales se espera utilizar otros procedimientos utilizando un muestreo uniforme de cada región, que permita mejorar el ren-

dimiento computacional para llegar a sistemas que puedan funcionar eficientemente con menor cantidad de información reduciendo tiempos de computo.

Los descriptores analizados pueden ser parte de un sistema de detección de características en el cual se consideren aspectos tales como: asimetría, cantidad de puntos, suavidad e invarianza que presentan las regiones faciales, entre otros. De igual forma, este trabajo puede ser extendido a otras partes del cuerpo y sus resultados aplicados en estudios antropométricos.

Trabajo futuro

Aunque los procedimientos desarrollados para el registro y extracción de características en puntos de interés permiten obtener un muestreo adecuado de los puntos, contornos y regiones, el tiempo de computo los hace inviables para un sistema en tiempo real por tal razón se espera buscar formas para incrementar la velocidad, tanto desde el punto de vista algorítmico como el de la implementación.

A pesar de que, en las regiones que conforman la boca y el mentón, los descriptores estudiados presentaron diferencias en sus niveles de discriminancia, es necesario realizar un estudio que incluya una variación a los parámetros involucrados en las imágenes spin multi-resolución y el índice de forma multi-escala. Este último se puede extender para obtener los 6 descriptores de curvatura en versión multi-escala.

Cuando se realiza un promedio sobre un conjunto de datos, como el de los valores de un descriptor en particular en una región, es posible que dos regiones diferentes tengan un promedio similar, por tal razón se espera en un trabajo futuro analizar no solo el promedio sino la varianza y los histogramas de las regiones.

Dado que existen otros descriptores de forma 3D que han mostrado gran utilidad para representar superficies 3D se puede realizar un estudio de otro tipo de descriptores para evaluar la capacidad para capturar la información local de superficie del rostro. Algunos de estos descriptores pueden ser los siguientes: folded spin image [44], cono curvatura ponderada [12] y DESIRE [45].

Se espera incorporar los descriptores analizados en un sistema robusto para el análisis y reconocimiento de rostros. En el cual se efectúen todas las tareas realizadas en este trabajo tales como: filtrado, mejora, registro, segmentación, caracterización e identificación.

Bibliografía

- [1] Alexander Ceron, Augusto Salazar, and Flavio Prieto. Relevance analysis of 3d curvature-based shape descriptors on interest points of the face. In *Proceedings of International Conference on Image Processing Theory, Tools and Applications IPTA*, 2010.
- [2] P.J. Flynn and Jain A.K. Hypothesis testing and parameter estimation. In *Computer Society Conference on Computer Vision and Pattern Recognition*, pages 261 – 267, 1988.
- [3] J. Cartoux, J. Lapreste, and M. Richetin. Face authentication or recognition by profile extraction from range images. In *Workshop on Digital Object Identifier*, 1989.
- [4] Esmeide A. Leal, John William Branch, and Oscar Otega Lobo. Estimación de curvaturas y direcciones principales en nube de puntos no organizados. *Dyna*, 153:351–362, 2007.
- [5] Gaile G. Gordon. Face recognition based on depth maps and surface curvature. In *Proceedings of SPIE Geometric methods in Computer Vision*, pages 234–247, 1991.
- [6] John C. Lee and Evangelos Milios. Matching range images of human faces. In *Third International Conference on Computer Vision*, 1990.
- [7] Joonsoo Lee. 3d face recognition using range images. Technical report, Department of Electrical and Computer Engineering, The University of Texas, Austin, TX, 2005.
- [8] Ana Belén Moreno Díaz. *Reconocimiento Facial Automático mediante Técnicas de Visión Tridimensional*. PhD thesis, Universidad Politécnica de Madrid, Facultad de Informática, 2004.
- [9] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification and Scene Analysis*, New York. John Wiley & Sons, 1998.
- [10] Christopher Bishop. *Pattern Recognition and Machine Learning*. Springer Science Business + Media, LLC, 2006.
- [11] Chitra Dorai and Anil K. Jain. Cosmos - a representation scheme for 3d free-form objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:1115–1130, 1997.
- [12] Miguel Adán, Antonio Adán, and Andrés S. Vázquez. *LNCS Structural, Syntactic, and Statistical Pattern Recognition*, volume 3138/2004, chapter Clustering Using WCC Models, pages 921 – 929. Springer, 2004.

-
- [13] Shu Zhan, Toru Kurihara, and Shigeru Ando. Facial orientation and eye-gaze detection based on real-time 3d facial imaging by using correlation image sensor. In *Sixth International Conference on Intelligent Systems Design and Applications (ISDA'06)*, 2006.
- [14] Alessandro Colombo, Claudio Cusano, and Raimondo Schettini. 3d face detection using curvature analysis. *Journal of Pattern Recognition*, 39:444–455, 2006.
- [15] H. T. Tanaka, M. Ikeda, and H. Chiaki. Curvature-based face surface recognition using spherical correlation. principal directions for curved object recognition. pages 372–377, 1998.
- [16] Peter W. Hallinan, Gaile G. Gordon, A. L. Yuille, Peter Giblin, and David Mumford. Two- and three-dimensional patterns of the face. 1999.
- [17] M. O. Irfanoglu, B. Gokberk, and L. Akarun. 3d shape-based face recognition using automatically registered facial surfaces. In *Proceedings of 17th International Conference on Pattern Recognition (ICPR 2004)*, volume 4, pages 183–186, 2004.
- [18] Dhanannjay Deo and Dibakar Sen. Automatic recognition of facial features and landmarking of digital human head. In *Proceedings of 6th International Conference on Computer Aided Industrial Design and Conceptual Design*, pages 506–602, 2005.
- [19] Feng Xue and Xiaoqing Ding. 3d+2d face localization using boosting in multi-modal feature space. In *Proceedings of 18th International Conference on Pattern Recognition (ICPR'06)*, 2006.
- [20] Yi Sun and Lijun Yin. Automatic pose estimation of 3d facial models. In *Proceedings of 19th International Conference on Pattern Recognition. ICPR 2008*, pages 1–4, 2008.
- [21] Robert Osserman. Curvature in the eighties. *The American Mathematical Monthly, Special Geometry Issue*, 97(8):731–756, October 1990.
- [22] Jan J. Koenderink and Andrea J van Doorn. Surface shape and curvature scales. *Image and Vision Computing*, 10:557 – 56, 1992.
- [23] Xiaoguang Lu, D. Colbry, and Anil K. Jain. Three-dimensional model based face recognition. In *Proceedings of 17th International Conference on Pattern Recognition*, volume 1, pages 362–366, 2004.
- [24] D. Colbry, G. Stockman, and Anil K. Jain. Detection of anchor points for 3d face verification. In *Proceedings of IEEE Workshop on Advanced 3D Imaging for Safety and Security*, 2005.
- [25] Xiaoguang Lu, D. Colbry, and Anil K. Jain. Matching 2.5d scans to 3d models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(1):31–43, 2006.
- [26] Zhang Guangpeng and Wang Yunhong. A 3d facial feature point localization method based on statistical shape model. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing. ICASSP 2007*, volume 2, pages II–249 – II–252, 2007.

-
- [27] Anupama Jagannathan and Eric. L. Miller. Three-dimensional surface mesh segmentation using curvedness-based region growing approach. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 29(12):2195–2204, 2007.
- [28] Andrew Johnson. *Spin-Images: A Representation for 3-D Surface Matching*. PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, August 1997.
- [29] Michael Foster. Spin-image target detection algorithm applied to low density 3d point clouds. *Journal of Applied Remote Sensing*, 2:1–15, 2008.
- [30] A. Johson and M. Hebert. Using spin images for efficient object recognition in cluttered 3d scenes. In *IEEE TPAMI*, 1999.
- [31] A. Johnson. Surface landmark selection and matching in natural terrain. In *Computer Vision and Pattern Recognition (CVPR)*, 1998.
- [32] O. Carmichael, D. Huber, and M. Hebert. Large data sets and confusing scenes in 3-d surface matching and recognition. In *International conference on 3-D Digital Imaging and Modeling (3-DIM)*, 1999.
- [33] H. Quynh Dinh and Steven Kropac. Multi-resolution spin-images. In *Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, 2006.
- [34] Huy Tho Ho and Danny Gibbins. Multi-scale feature extraction for 3d surface registration using local shape variation. In *Image and Vision Computing New Zealand, IVCNZ . 23rd International Conference*, 2008.
- [35] Chris Lomont. Fast inverse square root. Technical report, Purdue University, 2003.
- [36] MeshLab. <http://meshlab.sourceforge.net>.
- [37] Donald Hearn and M. Pauline Baker. *Computer Graphics with OpenGL*. Prentice Hall, 2003.
- [38] A.E. Salazar and F.A. Prieto. 3d bsm for face segmentation and landmarks detection. In *Three-Dimensional Image Processing (3DIP) and Applications*, 2010.
- [39] Gabriel Taubin. Estimating the tensor of curvature of a surface from a polyhedral approximation. In *Fifth International Conference on Computer Vision (ICCV)*, 1995.
- [40] Lijun Yin, Xiaozhou Wei, Yi Sun, Jun Wang, and Matthew J. Rosato. A 3d facial expression database for facial behavior research. In *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, 2006.
- [41] Alexander Cerón, Augusto Salazar, and Flavio Prieto. Análisis de relevancia de descriptores de forma 3d sobre la superficie del rostro. In *Memorias del Encuentro Nacional de Investigación y Desarrollo - Enid*, 2010.
- [42] Augusto Salazar, Alexander Ceron, and Flavio Prieto. 3d curvature-based shape descriptors for face segmentation: An anatomical-based analysis. In *6th International Symposium on Visual Computing (ISVC)*, 2010.
- [43] Jian xiong Dong, Adam Krzyzak, and Ching Y. Suen. A practical smo algorithm. In *Proceedings International Conference Pattern Recognition*, 2002.

- [44] Stefanie Wuhrer, Zouhour Azouz, and Chang Shu. Posture invariant surface description and feature extraction. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.
- [45] Dejan V. Vranic. Desire: a composite 3d-shape descriptor. In *IEEE International Conference on Multimedia and Expo ICME*, 2005.
- [46] J.D. Foley, A. van Dam, S.K. Feiner, and J.F. Hughes. *Computer Graphics, Principles and Practice*. Addison Wesley, 1992.
- [47] G. Golub and C.F. Van Loan. *Matrix Computations*. John Hopkins University Press, 1983.

Estimación del tensor de curvatura de una superficie

A.1. Tensor de curvatura

El tensor de curvatura de una superficie S , es el mapeo $p \mapsto k_p$ que asigna cada punto p de S a la función que mide la curvatura direccional $k_p(T)$ de S en p en la dirección de un vector unitario T , tangente a S en p . La curvatura direccional $k_p(T)$ de una superficie S en un punto p en la dirección de un vector unitario tangente T , es definida por la identidad $x''(0) = k_p(T)N$, donde N es el vector unitario normal a S en p , y $x(s)$ es una *sección normal* a S en p y $x'(0) = T$. La función de curvatura direccional $k_p(\cdot)$ es una forma cuadrática que satisface la siguiente identidad:

$$k_p(T) = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}^t \begin{pmatrix} k_p^{11} & k_p^{12} \\ k_p^{21} & k_p^{22} \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}, \quad (\text{A.1})$$

donde, $T = t_1T_1 + t_2T_2$ es un vector tangente a S en p , $\{T_1, T_2\}$ es una base ortonormal del espacio tangente a S en p , $k_p^{11} = k_p(T_1)$, $k_p^{22} = k_p(T_2)$, y $k_p^{12} = k_p^{21}$. Los vectores $\{T_1, T_2\}$ son llamados *direcciones principales* de S en p cuando $k_p^{12} = k_p^{21} = 0$. Las curvaturas direccionales correspondientes son las *curvaturas principales*, las cuales se denotan como k_p^1 y k_p^2 en lugar de k_p^{11} y k_p^{22} de acuerdo a [39].

Si se adiciona el vector normal N a la base $\{T_1, T_2\}$ de direcciones principales, se obtiene una base ortonormal $\{N, T_1, T_2\}$ de espacio tridimensional con cambios de punto a punto. La siguiente ecuación extiende la definición de la curvatura direccional a las direcciones que no son tangentes:

$$k_p(T) = \begin{pmatrix} n \\ t_1 \\ t_2 \end{pmatrix}^t \begin{pmatrix} 0 & 0 & 0 \\ 0 & k_p^1 & 0 \\ 0 & 0 & k_p^2 \end{pmatrix} \begin{pmatrix} n \\ t_1 \\ t_2 \end{pmatrix}, \quad (\text{A.2})$$

donde $T = nN + t_1T_1 + t_2T_2$ es un vector arbitrario. Los vectores tangentes a S en p son aquellos para los cuales $n = 0$. Al escribir el vector T como una combinación lineal $T = u_1U_1 + u_2U_2 + u_3U_3$ de otro sistema de vectores ortonormales, la curvatura direccional

tendrá ahora la expresión:

$$k_p(T) = u^t K_p u, \quad (\text{A.3})$$

donde $u = (u_1, u_2, u_3)^t$, y K_p es una matriz simétrica 3×3 la cual tiene 0 y las dos curvaturas principales k_p^1 y k_p^2 como valores propios. Se utilizara como base $\{U_1, U_2, U_3\}$ el mismo sistema de coordenadas cartesiano usado para especificar las coordenadas de los puntos sobre la superficie, independientemente del punto p en S .

Las curvaturas principales y direcciones principales de S en p pueden ser recuperadas, primero restringiendo la matriz K_p al plano tangente a S en p , y luego, calculando los valores y vectores propios de la matriz resultante. Este cálculo puede ser realizado en forma cerrada, usando operaciones aritméticas de acuerdo a [39].

A.2. Estimación del tensor de curvatura

En [39] se define procedimiento para la obtención de las curvaturas direccionales, las cuales son aproximadas mediante series de Laurent. En primer lugar se define una matriz M_p mediante una integral que tiene los mismos vectores propios que K_p , y sus valores propios están relacionados por una transformación lineal homogénea fija. La estimación de curvaturas principales y direcciones principales de S en p , reduce la diagonalización de la matriz M_p , lo cual puede realizarse en forma cerrada.

La obtención de M_p se realiza teniendo en cuenta que T_θ es el vector unitario tangente $T_\theta = \cos(\theta)T_1 + \sin(\theta)T_2$ para $-\pi \leq \theta \leq \pi$, donde $\{T_1, T_2\}$ son las direcciones principales ortogonales de S en p . Por lo tanto, de acuerdo a la ecuación A.2 se tiene lo siguiente:

$$k_p(T_\theta) = k_p^1 \cos^2(\theta) + k_p^2 \sin^2(\theta), \quad (\text{A.4})$$

se define la matriz simétrica:

$$M_p = \frac{1}{2\pi} \int_{-\pi}^{+\pi} k_p(T_\theta) T_\theta T_\theta^t, d\theta. \quad (\text{A.5})$$

El vector normal N es el vector propio de la matriz asociada con el valor propio 0, ya que $T_\theta T_\theta^t$ es una matriz de rango 1 para todo θ , y T_θ es tangente a S en p . M_p se puede factorizar de la siguiente forma:

$$M_p = T_{12}^t \begin{pmatrix} m_p^{11} & m_p^{12} \\ m_p^{21} & m_p^{22} \end{pmatrix} T_{12}, \quad (\text{A.6})$$

donde $T_{12} = \{T_1, T_2\}$ es la matriz 3×2 construida mediante la concatenación de los vectores columna T_1 y T_2 , y $m_p^{12} = m_p^{21}$ debido a la simetría. Revisando los parámetros m_p^{ij} , se observa que los elementos fuera de la diagonal son cero:

$$m_p^{12} = T_1^t M_p T_2 = \frac{k_p^1}{2\pi} \int_{-\pi}^{+\pi} \cos^3(\theta) \sin(\theta) d\theta + \frac{k_p^2}{2\pi} \int_{-\pi}^{+\pi} \cos(\theta) \sin^3(\theta) d\theta = 0, \quad (\text{A.7})$$

debido a que los dos integrandos son funciones impares de θ . Esto quiere decir que los dos vectores propios remanentes de M_p (diferentes a N) son las direcciones principales T_1 y T_2 . Los valores propios correspondientes no son las curvaturas principales:

$$m_p^{11} = T_1^t M_p T_2 = \frac{k_p^1}{2\pi} \int_{-\pi}^{+\pi} \cos^4(\theta) d\theta + \frac{k_p^2}{2\pi} \int_{-\pi}^{+\pi} \cos^2(\theta) \sin^2(\theta) d\theta = \frac{3}{8} k_p^1 + \frac{1}{8} k_p^2, \quad (\text{A.8})$$

de forma similar se obtiene:

$$m_p^{22} = T_2^t M_p T_2 = \frac{1}{8} k_p^1 + \frac{3}{8} k_p^2, \quad (\text{A.9})$$

Mediante las ecuaciones A.8 y A.9 se obtienen las curvaturas principales, como funciones de los valores propios diferentes de cero de M_p :

$$\begin{aligned} k_p^1 &= 3m_p^{11} - m_p^{22} \\ k_p^2 &= 3m_p^{22} - m_p^{11}. \end{aligned} \quad (\text{A.10})$$

Para estimar la curvatura direccional $k_p(T)$ para un vector unitario T , tangente a S en p , se considerara de nuevo una curva suave $x(s)$ parametrizada por longitud de arco, contenida en S , y tal que $x'(0) = T$. En tal caso se tiene que $x''(0) = k_p(T)N$. Realizando una expansión en series de Laurent de segundo orden:

$$x(s) = x(0) + x'(0)s + \frac{1}{2}x''(0)s^2 + O(s^3) = p + Ts + \frac{1}{2}k_p(T)Ns^2 + O(s^3), \quad (\text{A.11})$$

y observando que

$$2N^t(x(s) - p) = k_p(T)s^2 + O(s^3). \quad (\text{A.12})$$

De las dos ecuaciones anteriores se obtiene:

$$\frac{2N^t(x(s) - p)}{\|x(s) - p\|^2} = k_p(T) + O(s). \quad (\text{A.13})$$

De acuerdo a lo anterior la derivada direccional $k_p(T)$ es igual al límite:

$$k_p(T) = \lim_{s \rightarrow 0} \frac{2N^t(x(s) - p)}{\|x(s) - p\|^2}. \quad (\text{A.14})$$

Si q es otro punto sobre la superficie, cercano pero diferente de p , y T es el vector unitario de la proyección del vector $q - p$ sobre el plano tangente $\langle N \rangle^\perp$, la curvatura direccional puede ser aproximada de la siguiente forma:

$$k_p(T) \approx \frac{2N^t(q - p)}{\|q - p\|^2}. \quad (\text{A.15})$$

A.3. Algoritmo para la obtención de curvaturas direccionales

En esta sección se presenta el procedimiento para la estimación de curvaturas principales para mallas triangulares desarrollada por Taubin [39], el cual se encuentra implementado en el software para la manipulación de mallas de código abierto Meshlab [36] y es utilizado en este trabajo para la obtención de las curvaturas principales.

Para comenzar se define una superficie triangular como un par de listas $S = \{V, F\}$, una lista de vértices $V = \{v_i : 1 \leq i \leq n_v\}$, y una lista de caras $F = \{f_k : 1 \leq k \leq n_F\}$. Cada cara $f_k = (i_1^k, i_2^k, i_3^k)$ es una terna de índices no repetidos, que representa un triángulo tridimensional. Se consideran tanto superficies triangulares cerradas como superficies triangulares con frontera, pero se asume que las superficies son orientadas y consistentes [46]. El conjunto de vértices que comparten una cara con v_i se nota como V^i . Si el vértice v_j pertenece a V^i , se denomina un *vecino* de v_i .

El número de elementos del conjunto V^i es denominado $|V^i|$. El conjunto de caras que contienen el vértice v_i es denominado F_i . Si la cara f_k pertenece a F^i , se denomina *incidente* a v_i . El número de elementos de el conjunto F^i es denominado $|F^i|$.

La primera tarea es estimar los vectores normales a los vértices de la superficie. Dado que las caras de la superficie son planares, cada cara f_k tiene un vector normal unitario bien definido N_{f_k} . Como la superficie es orientada, todos estos vectores normales apuntan al mismo lado de la superficie. Se define el vector normal a un vértice v_i como la suma ponderada normalizada de las normales de las caras incidentes, con pesos proporcionales a las áreas superficiales de las caras:

$$N_{v_i} = \frac{\sum_{f_k \in F^i} |f_k| N_{f_k}}{\|\sum_{f_k \in F^i} |f_k| N_{f_k}\|}. \quad (\text{A.16})$$

La segunda tarea es estimar las matrices M_{v_i} . Como se mencionó anteriormente, se aproxima la matriz M_v , como una suma ponderada sobre la vecindad V_i :

$$\tilde{M}_{v_i} = \sum_{v_j \in V^i} w_{ij} k_{ij} T_{ij} T_{ij}^t. \quad (\text{A.17})$$

Para cada vecino v_j de v_i , se define T_{ij} como la proyección de longitud normalizada del vector $v_j - v_i$ sobre el plano tangente $\langle N \rangle^\perp$:

$$T_{ij} = \frac{(I - N_{v_i} N_{v_i}^t)(v_j - v_i)}{\|(I - N_{v_i} N_{v_i}^t)(v_j - v_i)\|}. \quad (\text{A.18})$$

Se aproxima la curvatura direccional $k_{v_i}(T_{ij})$, usando la Ecuación A.15

$$k_{ij} = \frac{2N^t(v_j - v_i)}{\|v_j - v_i\|^2}. \quad (\text{A.19})$$

Escogiendo un peso w_{ij} proporcional a la suma de las áreas de las superficies de todos los triángulos que son incidentes a ambos vértices v_i y v_j (dos si la superficie es cerrada y uno si la ambos vértices pertenecen al vecindario de la superficie), se ajusta la constante

de proporcionalidad para hacer que la suma de todos los pesos en la vecindad del vértice v_i sea igual a uno:

$$\sum_{v_j \in V_i} w_{ij} = 1. \quad (\text{A.20})$$

Mediante la construcción, el vector normal N_{v_i} es un vector propio de la matriz \tilde{M}_{v_i} asociada con el valor propio 0. Para calcular los dos pares remanentes se restringe la matriz \tilde{M}_{v_i} al plano tangente $\langle N \rangle^\perp$ usando una transformación Householder [47] y luego se diagonaliza la matriz 2×2 resultante en forma cerrada con una rotación Givens [47]. De esta forma el cálculo de las direcciones principales es restringido a ser ortogonal al vector normal N_{v_i} , incluso si uno de los valores propios es cero o cercano a cero. De todas formas, si los dos vectores remanentes de \tilde{M}_{v_i} son iguales, las direcciones principales no tendrán unicidad. Pero este es un problema que tiene todo algoritmo de estimación de curvaturas principales.

Siendo $E_1 = (1, 0, 0)^t$ el primer vector de coordenadas y:

$$W_{v_i} = \frac{E_1 \pm N_{v_i}}{\|E_1 \pm N_{v_i}\|}, \quad (\text{A.21})$$

con un signo menos si $\|E_1 - N_{v_i}\| > \|E_1 + N_{v_i}\|$ y signo más en otro caso. La matriz Householder:

$$Q_{v_i} = I - 2W_{v_i}W_{v_i}^t, \quad (\text{A.22})$$

es ortogonal y tiene su primera columna igual a N_{v_i} o $-N_{v_i}$, dependiendo de la elección del signo mencionada anteriormente. Las otras dos columnas definen una base ortonormal del espacio tangente, pero no necesariamente las direcciones principales. Denotando los dos vectores \tilde{T}_1 y \tilde{T}_2 , N_{v_i} como vector propio de \tilde{M}_{v_i} con valor propio asociado 0 se obtiene:

$$Q_{v_i}^t \tilde{M}_{v_i} Q_{v_i} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \tilde{m}_{v_i}^{11} & \tilde{m}_{v_i}^{12} \\ 0 & \tilde{m}_{v_i}^{21} & \tilde{m}_{v_i}^{22} \end{pmatrix}, \quad (\text{A.23})$$

donde $\tilde{m}_{v_i}^{21} = \tilde{m}_{v_i}^{12}$. Ahora los menores 2×2 diferentes de cero pueden ser diagonalizados en la forma cerrada con una rotación Givens, obteniendo un ángulo θ tal que los vectores

$$\begin{aligned} T_1 &= \cos^2(\theta)\tilde{T}_1 - \sin^2(\theta)\tilde{T}_2 \\ T_2 &= \cos^2(\theta)\tilde{T}_1 + \sin^2(\theta)\tilde{T}_2, \end{aligned} \quad (\text{A.24})$$

son los vectores propios remanentes de \tilde{M}_{v_i} es decir las direcciones principales de la superficie en v_i . Las curvaturas principales son obtenidas de los valores propios correspondientes de \tilde{M}_{v_i} mediante la Ecuación A.10.