

# laSalle

UNIVERSITAT RAMON LLULL

**Escola Tècnica Superior d'Enginyeria  
Electrònica i Informàtica La Salle**

Treball Final de Màster

Màster Universitari en Ingenieria de Telecomunicaciones

**Modelaje de mallas faciales 3D y  
automatización de landmarking  
en pacientes con esquizofrenia y  
bipolaridad**

Nombre Alumno

Nombre Profesor Ponente

David Ureña Gómez

Xavier Sevillano Domínguez



---

# ACTA DEL EXAMEN DEL TRABAJO FINAL DE MÁSTER

---

Reunido el Tribunal calificador en la fecha indicada, el alumno

**David Ureña Gómez**

expuso su Trabajo Final de Máster, titulado:

**Modelaje de mallas faciales 3D y automatización de landmarking en pacientes con esquizofrenia y bipolaridad**

Acabada la exposición y contestadas por parte del alumno las objeciones formuladas por los Sres. miembros del tribunal, éste valoró dicho Trabajo con la calificación de

Barcelona,

VOCAL DEL TRIBUNAL

VOCAL DEL TRIBUNAL

PRESIDENTE DEL TRIBUNAL



# *Abstract*

La esquizofrenia (ESQ) y la bipolaridad (BP) son dos de las enfermedades mentales más complejas de diagnosticar actualmente. Ambas tienen asociadas dismorfologías faciales que permiten realizar estudios sobre la forma facial, ayudando así al sector clínico para su diagnóstico. El análisis de la forma facial se desarrolla mediante *landmarks*, a menudo colocados a mano. Los *landmarks* son puntos anatómicos discretos que se pueden reconocer de la misma manera en todas las muestras, y que sirven para obtener biomarcadores de enfermedades que tengan asociadas dichas dismorfologías faciales. En este proyecto se han estudiado algoritmos para automatizar el proceso de detección de *landmarks* faciales. Para su obtención, se trabajará con resonancias magnéticas craneales, permitiendo así obtener una reconstrucción detallada de los órganos y tejidos del cráneo. Partiendo del modelo craneal 3D y sin ningún tipo de referencia con respecto a la orientación y alineación, se logrará obtener una malla facial 3D. Una vez obtenida dicha malla, mediante descriptores de curvatura y de localización, se desarrollará un modelo de aprendizaje automático para la detección automática de *landmarks* faciales.



# Resumen

La esquizofrenia (ESQ) y la bipolaridad (BP) son dos de las enfermedades mentales más comunes en todo el mundo. Los costos económicos asociados a la enfermedad son el motivo para identificar predictores clínicos fiables que tengan una aplicación directa para la prevención, el diagnóstico, y la terapia de la enfermedad. Desafortunadamente, tales predictores biológicos siguen siendo una necesidad clínica no satisfecha.

Tanto la ESQ como la BP, son causadas principalmente por factores genéticos y ambientales que alteran el desarrollo del cerebro durante la embriogénesis. Estos insultos pueden inducir también anomalías físicas menores (AFM), alteraciones sutiles del desarrollo. Entre las AFM, los biomarcadores faciales emergen como uno de los factores de riesgo biológicos más prometedores para la ESQ y BP. Dado que el cerebro y la cara están íntimamente integrados en el desarrollo, la dismorfología facial se puede considerar como un índice clínicamente accesible de la dismorfología cerebral y del deterioro cognitivo en pacientes con ESQ y BP.

Una estrategia para descubrir predictores biológicamente informados, consiste en construir un biomarcador multisistémico. Para cada individuo, se recopilarn datos morfológicos cerebrales y faciales a partir de imágenes de resonancia magnética (RM), y se evaluará su puntaje de riesgo poligénico para dichas enfermedades. Basándonos en resultados preliminares, nuestra hipótesis es que la presencia y la intensidad de los rasgos faciales se correlacionarán significativamente con las dismorfologías cerebrales, así como con un mayor riesgo genético. Aplicando métodos de aprendizaje automático, se evaluará el potencial diagnóstico de los biomarcadores faciales, cerebrales y genéticos, y se buscaran combinaciones que maximicen el potencial discriminante.

A fin de obtener los biomarcadores, es necesario partir de un conjunto de puntos de referencia llamados *landmarks*. En este proyecto de fin de máster, se trabajará solamente con *landmarks* faciales facilitados por el grupo de investigación de biología con el que trabajamos. El objetivo será desarrollar una herramienta basada en algoritmos de visión computacional y aprendizaje automático, la cual realizará la ubicación automática de *landmarks* faciales a partir de modelos craneales 3D, extraídos de RM. Esta herramienta permitirá que en un futuro se puedan calcular automáticamente los biomarcadores faciales, y así establecer una comprensión de la relación entre las dismorfologías faciales y la ESQ y BP. Además, esto será un complemento clínico de alto valor para el diagnóstico de dichas enfermedades, que actualmente se basa en entrevistas clínicas.



# Agradecimientos

Este apartado va dirigido a todas aquellas personas que han hecho posible que realizara este proyecto, ya sea por su apoyo, motivación, o supervisión.

En primer lugar, me gustaría dirigirme a mi tutor de proyecto Xavi Sevillano. Quería agradecerte primero la confianza que tuviste desde el primer momento al presentarme el proyecto, y en consecuencia, verme capaz de realizarlo. También quería agradecerte el apoyo continuo recibido durante todas las fases del proyecto. Nos hemos topado con obstáculos en el camino, algunos que se recordaran en la historia de la humanidad, el COVID-19. Aún así, hemos seguido trabajando sin que nada ni nadie nos detenga, siempre bajo nuestro lema *Seguim!*. Agradecerte también las "terapias" que hemos tenido en semanas difíciles y de confusión, ya que me han ayudado a calmar mi pensamiento continuo y preocupación en el proyecto. Finalmente, agradecerte también la oportunidad y confianza de poder hacer la tesis doctoral contigo, ojalá sea así.

También me gustaría agradecer a mi familia el apoyo incondicional que me han brindado durante toda la etapa estudiantil. Mis padres Cristina y Luís, mi hermana Vanessa, y mi abuela Amparo. Sin ellos nada de esto hubiera sido posible.

Finalmente agradecer también a aquellos amigos/amigas más cercanos que también me han apoyado y animado en la realización del proyecto.



# Índice general

Abstract	I
Resumen	II
Agradecimientos	III
Lista de tablas	VI
Lista de figuras	VII
Lista de acrónimos	IX
<b>I Parte Teórica</b>	<b>1</b>
<b>1. Morfometría Geométrica</b>	<b>3</b>
1.1. Introducción . . . . .	3
1.2. Landmarks . . . . .	5
1.3. Variables de forma y tamaño: <i>Bookstein shape coordinates</i> . . . . .	6
1.4. Teoría de la forma . . . . .	8
1.5. Métodos de Superposición . . . . .	11
1.6. Métodos de Ordenación . . . . .	13
<b>2. Descripción de Datos</b>	<b>16</b>
2.1. Esquizofrenia . . . . .	16
2.2. Bipolaridad . . . . .	17
<b>II Parte Práctica</b>	<b>18</b>
<b>3. Preprocesamiento de los datos</b>	<b>19</b>
3.1. Origen y descripción de los datos . . . . .	19

3.1.1. Introducción . . . . .	19
3.1.2. Origen . . . . .	19
3.1.3. Formato de Datos . . . . .	20
3.2. Problemáticas . . . . .	21
3.2.1. Orientación . . . . .	23
3.2.2. Alineación . . . . .	29
3.2.3. Ruido . . . . .	31
3.3. Obtención de la malla facial 3D . . . . .	32
3.3.1. $AMF_{3D}$ - Basado en cuantificación angular . . . . .	33
3.3.2. $AMF_{3D}$ - Basado en trazado de rayos . . . . .	38
<b>4. Automatización de Landmarking 3D</b>	<b>42</b>
4.1. Introducción . . . . .	42
4.2. Metodología . . . . .	43
4.2.1. Introducción . . . . .	43
4.2.2. LSM . . . . .	43
4.2.3. LLM . . . . .	49
4.2.4. Combinación entre métodos . . . . .	52
<b>Conclusiones</b>	<b>X</b>
<b>Líneas de futuro</b>	<b>XI</b>
<b>Bibliografía</b>	<b>XIII</b>

# Índice de cuadros

3.1. Diversidad de la base de datos . . . . .	20
3.2. Clasificación de la base de datos según la diversidad de orientación craneal .	23



# Índice de figuras

1.1. Operaciones matemáticas realizadas previamente para la obtención de la descripción de forma. (A) Modelos originales; (B) después de eliminar las diferencias en la ubicación; (C) después de eliminar las diferencias de escala; (D) después de eliminar las diferencias de orientación, dejando solo diferencias de forma. Extraído de [7]. . . . .	4
1.2. Conjunto de 20 <i>landmarks</i> faciales anatómicos. Extraído de [3]. . . . .	6
1.3. Operaciones que no alteran la forma, aplicadas a un triángulo: (A) traslación; (B) rotación; (C) escalado. Extraído de [7]. . . . .	7
1.4. Representación geométrica del cálculo del tamaño del centroide. Calculado como la raíz cuadrada de las longitudes L1, L2, L3 cuadradas y sumadas. Extraído de [7]. . . . .	7
1.5. <i>Landmarks</i> de un tipo de especie acuática visualizada por dos tipos de superposiciones: (A) Coordenadas de forma <i>Bookstein</i> , (B) Procrustes. Extraído de [7]. . . . .	12
3.1. Modelo PLY de un paciente masculino con esquizofrenia . . . . .	21
3.2. Modelo <i>Point Cloud</i> de un paciente masculino con esquizofrenia . . . . .	21
3.3. Casuísticas que presentan los pacientes respecto la orientación . . . . .	22
3.4. Representación craneal 3D con presencia de ruido . . . . .	22
3.5. Diagrama de bloques del método de detección barbilla/nariz . . . . .	24
3.6. Representación del 5% de puntos más lejanos respecto la nube de puntos original . . . . .	24
3.7. Representación del conjunto de vecinos más lejanos y con mayor <i>Peakness score</i> de un punto en concreto . . . . .	25
3.8. Representación del 15% de vecinos más lejanos y con mayor <i>Peakness score</i> . . . . .	26
3.9. <i>Clusters</i> del 15% de vecinos más lejanos y con mayor <i>Peakness score</i> . . . . .	26
3.10. Método <i>Fit Spheres</i> para diferenciar entre los <i>clusters</i> de la coronilla y los de la nariz y barbilla. . . . .	27
3.11. Comparación entre esferas al aplicar el método <i>Fit Spheres</i> . . . . .	27
3.12. Orientación de un modelo 3D originalmente orientado hacia el eje de coordenadas Z positivas . . . . .	28

3.13. Orientación de un modelo 3D originalmente orientado hacia el eje de coordenadas Y positivas con el cráneo invertido. . . . .	29
3.14. Diagrama de bloques del método de detección del punto máximo de la nariz	30
3.15. Rango para hallar el punto con Y máxima . . . . .	30
3.16. Ejemplo de alineación de un modelo <i>PointCloud</i> . . . . .	31
3.17. Ejemplo de aplicación del método de <i>Clustering</i> . . . . .	32
3.18. Modelo <i>PointCloud</i> con presencia de ruido. . . . .	33
3.19. Modelo <i>PointCloud</i> después de haber aplicado <i>clustering</i> para la corrección de ruido. . . . .	33
3.20. Reducción facial definido por el rango de un modelo <i>PointCloud</i> . . . . .	34
3.21. Zona facial de un modelo <i>PointCloud</i> . . . . .	35
3.22. Mallas faciales 3D de un modelo <i>PointCloud</i> con cuantificación espacial 128 y 192 respectivamente. . . . .	35
3.23. Malla facial 3D de un modelo <i>PointCloud</i> . . . . .	36
3.24. Comparación modelo 3D original VS modelo 3D reducido . . . . .	36
3.25. Malla facial 3D de un modelo <i>PointCloud</i> con <i>landmarks</i> . . . . .	37
3.26. Cuadrículas generadas en el plano XZ y YZ para establecer el origen des de donde se lanzarán las proyecciones frontales y laterales respectivamente. . .	38
3.27. Proyección de rayos frontales incidiendo sobre el cráneo. . . . .	39
3.28. Intersecciones entre la cuadrícula y el cráneo. . . . .	39
3.29. Representación de los <i>clusters</i> que presenta el modelo facial 3D. . . . .	40
3.30. Malla facial 3D aplicando el método AMF <sub>3D</sub> basado en trazado de rayos. . .	41
4.1. Malla facial 3D, (A) <i>Point Cloud</i> , (B) malla triangular. . . . .	44
4.2. Representación de los vectores normales $\hat{n}$ de la nariz. . . . .	44
4.3. Escala de <i>Shape Index</i> que muestra los ocho tipos básicos de formas de superficie	46
4.4. Nube de puntos 3D proyectada en una cuadrícula 2D. Extraído de [14]. . .	47
4.5. LSD de los tres <i>landmarks</i> marcados en rojo en el modelo facial 3D. (1) Extremo interno del ojo derecho, (2) punta de la nariz, y (3) extremo derecho de la boca, respectivamente. . . . .	48
4.6. Representación de diferentes espacios involucrados en ALF <sub>3D</sub> . (1) Alineación de los modelos faciales 3D hacia una ubicación de referencia común mediante el método de Procrustes. (2) Representación de los modelos faciales como vectores 60D. Cada uno representado como un punto en el espacio 60D. (3) Definición del espacio PCA dimensionalmente reducido. Extraído de [13]. .	51
4.7. Diagrama de bloques de los métodos LSM y LLM. Extraído de [14]. . . . .	53

# Lista de acrónimos

- **ESQ:** Esquizofrenia
- **BP:** Bipolaridad
- **RM:** Resonancia magnética
- **AFM:** Anomalías físicas menores
- **PCA:** Principal Components Analysis
- **PC:** Principal Components
- **PLY:** Polygon File Format
- **AMF<sub>3D</sub>:** Automatic Meshing of faces in 3D
- **ALF<sub>3D</sub>:** Automatic Landmarking of Faces in 3D
- **LSM:** Landmark Shape Model
- **LLM:** Landmark Location Model
- **ASM:** Active Shape Models
- **PDM:** Point Distribution Model
- **SVD:** Singular Value Decomposition
- **ICP:** Iterative Closest Point



Parte I

Parte Teórica



# Capítulo 1

## Morfometría Geométrica

### 1.1. Introducción

La morfometría geométrica desempeña un papel muy importante en muchos tipos de estudios biológicos. Principalmente, se encarga de analizar la forma de individuos y estudiar las diferencias entre estos.

La variedad de procesos biológicos entre individuos, generan diferencias de forma, produciendo así diferentes tipos de enfermedades o lesiones. Las diferencias en la forma, pueden indicar diferentes roles funcionales, o diferencias en los procesos de crecimiento y morfogénesis. El análisis de forma es un enfoque para comprender las diversas causas de variación y transformación morfológica existentes hoy en día entre los individuos.

Las diferencias de forma se resumen comparando formas observadas con objetos familiares. Aún así, dichas comparaciones pueden ser inexactas o engañosas, especialmente si se trata del análisis de formas complejas. Cuando necesitamos cierta precisión, procedemos a la medición.

La morfometría geométrica es una forma cuantitativa de abordar las comparaciones de formas que siempre han interesado al mundo de la biología. El lenguaje utilizado en este tipo de análisis es muy abstracto y matemático. Es por esto, que la morfometría es una rama del análisis matemático de formas. La información que se extrae de los datos morfométricos, implican operaciones matemáticas en lugar de conceptos biológicos.

Tradicionalmente, los datos morfométricos se han basado en medidas de longitud, profundidad, y ancho. Tal conjunto de datos, contiene poca información. Por ello, se desarrolló la definición de los *landmarks*, puntos anatómicos discretos que se pueden reconocer de la misma manera en todas las muestras.

#### Forma

En morfometría geométrica, la forma se define como "toda la información geométrica que queda cuando la ubicación, la escala y los efectos de rotación de un objeto se filtran" [6]. Es decir, para poder abordar solamente las diferencias de forma, y obtener la información geométrica, previamente se deben eliminar las diferencias en ubicación, escala, y orientación.

Estas operaciones y sus consecuencias se pueden ver en la Figura 1.1. En la Figura 1.1A

vemos las dos estructuras originales (dos modelos de  $N$  *landmarks*). En la Figura 1.1B vemos que ambas configuraciones comparten ubicación, pero estas aún difieren en escala y rotación. En la Figura 1.1C vemos que ambas configuraciones se re-escalan por igual. I finalmente, como la rotación tampoco influye en las diferencias y análisis de forma, en la Figura 1.1D ambas configuraciones se rotan para poder alinearse.

Representar a un objeto únicamente mediante *landmarks*, omite algunos aspectos claves de lo que normalmente podríamos definir como forma. Por ello, la curvatura es un descriptor importante para poder realizar un análisis de forma con precisión.

La curvatura es una característica del objeto que permanece después de filtrar la ubicación, la escala y los efectos de rotación, pero no necesariamente es capturada correctamente por las coordenadas de un conjunto de puntos de referencia, *landmarks*.

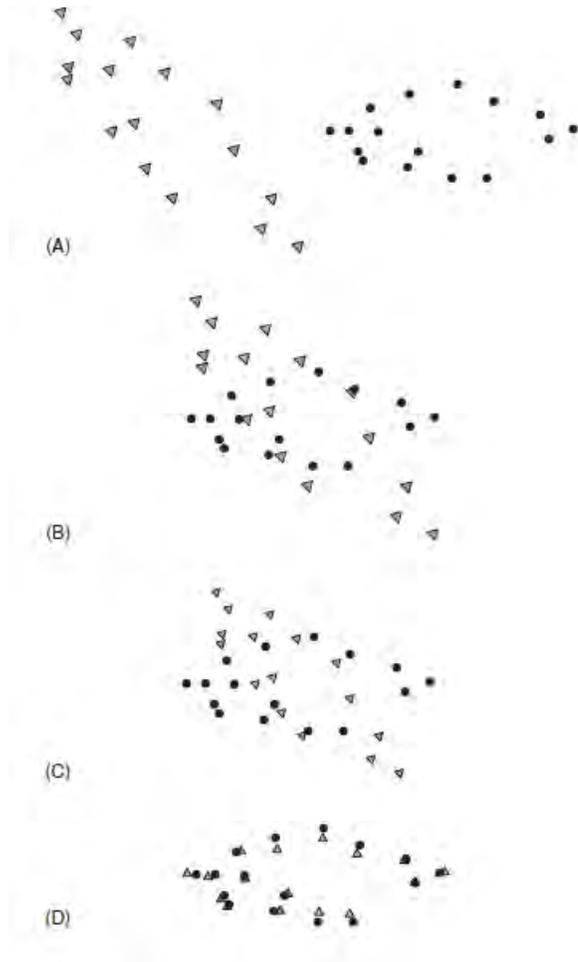


Figura 1.1: Operaciones matemáticas realizadas previamente para la obtención de la descripción de forma. (A) Modelos originales; (B) después de eliminar las diferencias en la ubicación; (C) después de eliminar las diferencias de escala; (D) después de eliminar las diferencias de orientación, dejando solo diferencias de forma. Extraído de [7].

## 1.2. Landmarks

Tal como hemos mencionado anteriormente, los *landmarks* son puntos de referencia anatómicos discretos que se pueden reconocer de la misma forma en todas las muestras de un estudio. El hecho de escoger estos *landmarks* no es tarea fácil, ya que han de cumplir una serie de requisitos.

Idealmente, los puntos de referencia son (1) anatómicamente homólogos que (2) no alteran sus posiciones relativas topológicas a otros puntos de referencia, (3) proporcionan una cobertura adecuada de la morfología, (4) se pueden encontrar repetidamente, y (5) se encuentran dentro del mismo plano.

1. La homología es la relación que existe entre dos partes orgánicas diferentes cuando sus determinantes genéticos tienen el mismo origen evolutivo. Comprender el papel que juega la homología tanto matemática, como biológicamente, requiere una sensación intuitiva tanto para las matemáticas como para los problemas biológicos.
2. Los métodos morfométricos no se pueden aplicar correctamente cuando las formas son demasiado diferentes. Estos métodos desarrollan aproximaciones lineales. Por lo tanto, cuando las formas difieren demasiado, las aproximaciones lineales general problemas y error.
3. Evidentemente, para poder detectar posibles cambios entre diferentes estructuras, estas han de partir de una cobertura uniforme y completa. Una forma de decidir si ha cumplido este criterio, es dibujar una imagen de los puntos de referencia sin rastrear el resto del organismo. Si la forma se puede apreciar con claridad, este punto se cumple. Sino es así, no podremos detectar cambios de forma entre estructuras.
4. Los *landmarks* deben poder ubicarse y reubicarse repetidamente sin error. Si este punto es irreplicable, se ha de descartar.
5. El quinto y último criterio para la selección de *landmarks*, está relacionado con la orientación en el momento de tomar las muestras en los individuos. En casos de organismos tridimensionales, no es necesario reducirlos a dos dimensiones, ya que siempre se puede examinar más de un plano, a pesar de ser más complejo de analizar. En cualquier caso, es importante tener en cuenta que el la orientación siempre ha de ser hacia el mismo plano.

Particularizando en este proyecto, se han obtenido los datos de una selección de 20 *landmarks* faciales anatómicos. Estos *landmarks* corresponden a estructuras faciales homólogas que se pueden registrar de manera confiable para todos los individuos.

En la Figura 1.2 podemos apreciar la distribución de *landmarks* de la cual partiremos.

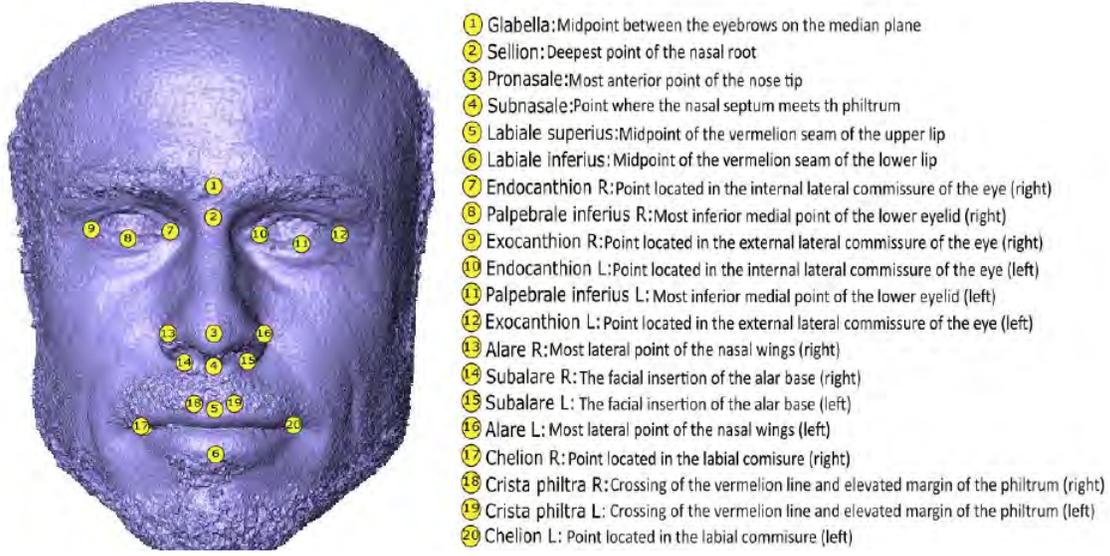


Figura 1.2: Conjunto de 20 *landmarks* faciales anatómicos. Extraído de [3].

### 1.3. Variables de forma y tamaño: *Bookstein shape coordinates*

En el apartado 1.1 discutimos sobre los significados de forma y tamaño, tal como se definen en morfometría geométrica. Definimos la forma en términos de operaciones que no la alteran, específicamente, traslación, rotación y reescalado. Estas operaciones se pueden aplicar a una forma simple como por ejemplo, un triángulo, obteniendo así un sistema de coordenadas.

Para el triángulo que se muestra en la Figura 1.3, podemos trasladarlo para que un punto de referencia esté en el origen (0,0) (Figura 1.3A). A continuación, podemos rotar para que el lado AB esté a lo largo del eje X (Figura 1.3B), y podemos escalarlo de manera que la coordenada del punto de referencia B esté en el punto (1,0) (Figura 1.3C). Finalmente, la coordenada del tercer punto de referencia, C, se puede calcular en el sistema de coordenadas que acabamos de definir.

Todas estas operaciones se pueden aplicar sin problemas, ya que estas no repercutirán ninguna consecuencia en la forma del objeto o individuo en cuestión.

Para hallar las coordenadas del punto C en el nuevo sistema de coordenadas, se pueden aplicar las siguientes ecuaciones:

$$\begin{aligned}
 SC_x &= \frac{(B_x - A_x)(C_x - A_x) + (B_y - A_y)(C_y - A_y)}{(B_x - A_x)^2 + (B_y - A_y)^2} \\
 SC_y &= \frac{(B_x - A_x)(C_y - A_y) - (B_y - A_y)(C_x - A_x)}{(B_x - A_x)^2 + (B_y - A_y)^2}
 \end{aligned} \tag{1.1}$$

donde  $A_x$ ,  $A_y$ ,  $B_x$ ,  $B_y$ ,  $C_x$  y  $C_y$  son las coordenadas del triángulo, y  $SC_x$ , y  $SC_y$  son las coordenadas del punto de referencia C en el nuevo sistema de coordenadas.

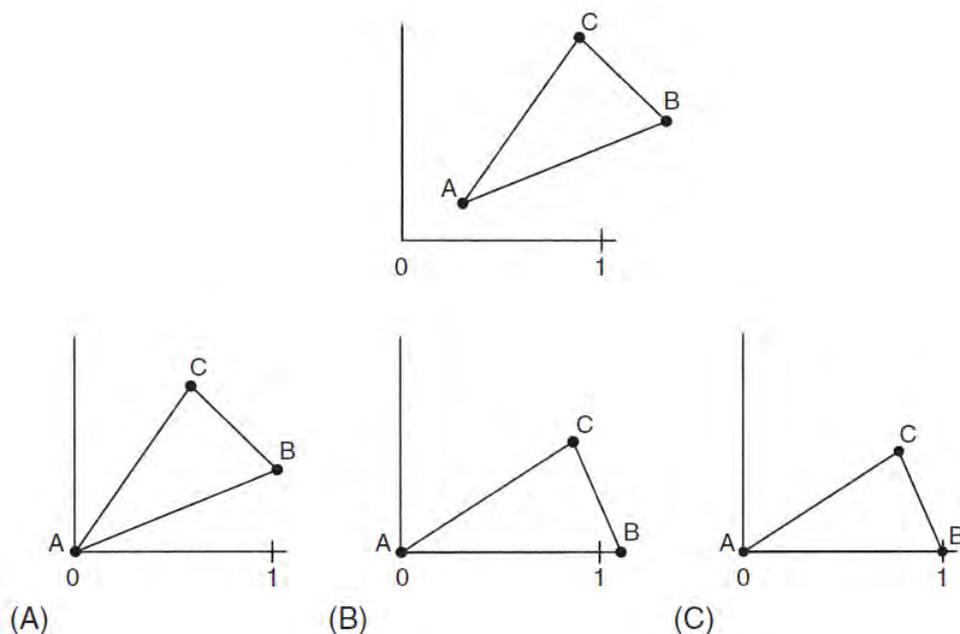


Figura 1.3: Operaciones que no alteran la forma, aplicadas a un triángulo: (A) traslación; (B) rotación; (C) escalado. Extraído de [7].

$SC_x$  y  $SC_y$  son las coordenadas de forma del punto de referencia C, llamado  $C_{xy}$ . Este conjunto de operaciones será importante cuando queramos comparar la forma de dos objetos, en este ejemplo, triángulos.

Es importante tener en cuenta que no perdemos información sobre la forma cuando realizamos las operaciones anteriores, pero si que eliminamos información sobre el tamaño.

Para restaurar la información sobre el tamaño, lo haremos mediante el concepto de escala. Por escala, nos referimos a la propiedad que se produce cuando una imagen se amplía o se reduce. Este concepto de escala es independiente de la forma.

La variable de tamaño que define el concepto de escala, es el tamaño del centroide (ver Figura 1.4).

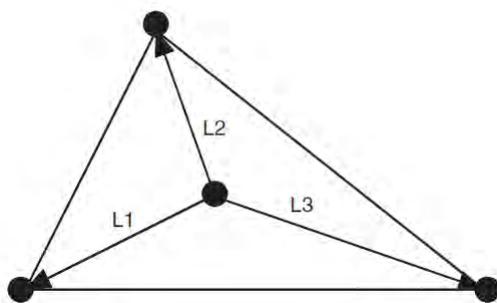


Figura 1.4: Representación geométrica del cálculo del tamaño del centroide. Calculado como la raíz cuadrada de las longitudes  $L_1$ ,  $L_2$ ,  $L_3$  cuadradas y sumadas. Extraído de [7].

Por lo tanto, el tamaño se mide independientemente de la forma, y estadísticamente este no está correlacionado con la forma.

Esta independencia de la forma es una de las principales razones por las cuales el tamaño del centroide se usa como una variable de tamaño.

## 1.4. Teoría de la forma

En esta sección vamos a profundizar sobre la teoría de la forma a nivel matemático, aún así no nos extenderemos, es por ello que si se requiere de una mayor explicación podemos consultar el libro de *Zelditch* [7].

Tal y como hemos definido en el apartado 1.1, la forma se define como "toda la información geométrica que queda cuando la ubicación, la escala y los efectos de rotación se filtran de un objeto" [6]. Toda la teoría de la morfometría geométrica se deriva de esta definición, por lo que en este apartado vamos a desarrollarlo matemáticamente.

Si recordamos la definición de *landmark* que explicamos en el apartado 1.2, estos son puntos de referencia que se hallan por igual entre poblaciones.

El conjunto de objetos a comparar son estructuras formadas por  $K$  *landmarks* (número de *landmarks*), cada uno de los cuales tiene  $M$  coordenadas ( $M = 3$  para formas 3D). Por ejemplo, para la estructura de la Figura 1.2 sería de  $K = 20$  y  $M = 3$ .

En geometría morfométrica, las mediciones individuales no son variables ni aportan información. Es decir, una variable de forma consta del vector completo de *landmarks* de un individuo, pudiendo así medir la covarianza entre otro vector de *landmarks* de otro individuo.

### Espacios de la Morfometría

Dada la definición de forma y de *landmark*, ya podemos desarrollar la idea matemática de los espacios de la morfometría.

#### *Matriz de configuración*

La matriz de configuración se representa mediante los *landmarks* y sus coordenadas. Tal que  $K \times M$  es una matriz de  $K$  coordenadas cartesianas que describe un conjunto particular de *landmarks* en  $M$  dimensiones.

Una matriz simple de  $K = 3$  y  $M = 3$  quedaría como:

$$X = \begin{bmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \\ X_3 & Y_3 & Z_3 \end{bmatrix} \quad (1.2)$$

### ***Espacio de configuración***

El espacio de configuración se describe como el conjunto de todas las matrices  $K \times M$  posibles que describen todos los conjuntos de *landmarks*.

Por lo tanto, el espacio de  $K$  *landmarks* con  $M$  coordenadas por *landmark*, tiene  $K \times M$  dimensiones. En el ejemplo anterior estaríamos hablando de un espacio 9D.

### ***Localización de la matriz de configuración***

La posición de una matriz de configuración se puede definir como la ubicación del centroide de dicha matriz.

La posición del centroide se define tal que así:

$$\begin{aligned} X_C &= \frac{1}{K} \sum_{j=1}^K X_j \\ Y_C &= \frac{1}{K} \sum_{j=1}^K Y_j \\ Z_C &= \frac{1}{K} \sum_{j=1}^K Z_j \end{aligned} \tag{1.3}$$

La matriz de configuración está centrada si el promedio de todas las coordenadas es cero. Para ello centraremos la matriz como:

$$XC = \begin{bmatrix} (X_1 - X_C) & (Y_1 - Y_C) & (Z_1 - Z_C) \\ (X_2 - X_C) & (Y_2 - Y_C) & (Z_2 - Z_C) \\ \vdots & \vdots & \vdots \\ (X_K - X_C) & (Y_K - Y_C) & (Z_K - Z_C) \end{bmatrix} \tag{1.4}$$

### ***Tamaño de la matriz de configuración***

Anteriormente, en el apartado 1.3 hablamos sobre el tamaño y la escala. En este apartado definiremos que entendemos por tamaño a nivel matemático.

La medida de tamaño más utilizada en morfometría geométrica es la del tamaño del centroide definida como:

$$CS(X) = \sqrt{\sum_{i=1}^K \sum_{j=1}^M (X_{ij} - C_j)^2} \tag{1.5}$$

donde  $i$  y  $j$  son las filas y columnas respectivamente de la matriz  $X$ , y  $C$  es la localización del centroide.

Por lo tanto, multiplicando la matriz de configuración  $X$  por un factor constante, el tamaño del centroide aumenta en el mismo factor. En el caso de que dos configuraciones de *landmarks* difieran solo en el tamaño del centroide, estas no lo hacen en forma, sino que solo lo harían en escala.

### ***Espacio de pre-forma***

Cada conjunto de  $K$  *landmarks* con número de coordenadas  $M$ , se puede representar como un punto en un espacio con dimensiones  $K \times M$ . Cuando hablamos de "punto", nos referimos a una forma individual, es decir, a todo un conjunto de *landmarks* y no a uno en particular.

Para poder crear este *pre-espacio* de forma, y que todos los conjuntos se puedan expresar de la misma manera y no difieran en ubicación o tamaño, la matriz de configuración deberá (1) estar centrada y (2) que el tamaño del centroide sea uno.

Podemos determinar el número de dimensiones de este *pre-espacio*, considerando el número de dimensiones que se pierden en las transición hasta la creación de este. Una dimensión se pierde al fijar el tamaño del centroide a uno. Y otras  $M$  dimensiones se pierden al centrar el conjunto de *landmarks* como:

$$KM - (M + 1) = KM - M - 1 \quad (1.6)$$

Para el caso presentado anteriormente, tendríamos un *pre-espacio* de  $3K - 4$  dimensiones. Si los *landmarks* tuvieran 2 coordenadas (espacio 2D), el *pre-espacio* sería de  $2K - 3$  dimensiones.

En este espacio todavía no se han eliminado los efectos de rotación, es por eso, que en el *pre-espacio* los *landmarks* solo difieren en rotación.

### ***Espacio de forma***

Para construir un espacio de forma, seleccionamos un punto en cada *fiber*, eliminando así las diferencias de rotación. La *fiber* es un arco circular que contiene el conjunto de todos los puntos del *pre-espacio*, y que se pueden hallar mediante la rotación de la matriz del *pre-espacio*. En este subapartado no entraremos en riguroso detalle sobre como se definen estas *fiber* ni que conllevan, solamente veremos como proceder a hacer dicha rotación (ver capítulo 4 del libro de *Zelditch* [7]).

El espacio de forma se describirá con menos dimensiones que el *pre-espacio*, ya que en este no se habían contemplado las diferencias de rotación. El cálculo de dimensiones queda tal que así:

$$\frac{M(M - 1)}{2} \quad (1.7)$$

En nuestro ejemplo de  $K$  *landmarks* de tres dimensiones, podemos resolver la Ecuación (1.7) dando como resultado 3. Este hecho refleja que una forma tridimensional, puede girar sobre su centroide en tres ejes ortogonales distintos en un espacio de coordenadas tridimensional.

Restando este resultado a  $3K - 4$  definido por el espacio de *pre-forma* (Ecuación (1.6)), obtenemos  $3K - 7$  dimensiones para espacios de forma tridimensionales.

Las siete dimensiones se describen de las diferencias de tamaño (-1), traslación (-3), vistas anteriormente, y de rotación (-3).

Las matrices de rotación en tres dimensiones se definen como:

$$R = R_z(\alpha)R_y(\beta)R_x(\gamma) = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\gamma) & -\sin(\gamma) \\ 0 & \sin(\gamma) & \cos(\gamma) \end{bmatrix} \quad (1.8)$$

dando lugar a

$$R = \begin{bmatrix} \cos(\alpha)\cos(\beta) & \cos(\alpha)\sin(\beta)\sin(\gamma) - \sin(\alpha)\cos(\gamma) & \cos(\alpha)\sin(\beta)\cos(\gamma) + \sin(\alpha)\sin(\gamma) \\ \sin(\alpha)\cos(\beta) & \sin(\alpha)\sin(\beta)\sin(\gamma) + \cos(\alpha)\cos(\gamma) & \sin(\alpha)\sin(\beta)\cos(\gamma) - \cos(\alpha)\sin(\gamma) \\ -\sin(\beta) & \cos(\beta)\sin(\gamma) & \cos(\beta)\cos(\gamma) \end{bmatrix} \quad (1.9)$$

Para hallar la distancia mínima entre puntos del *pre-espacio*, es suficiente con rotar la primera forma (conjunto de *landmarks* de  $M = 3$ ) para minimizar la distancia de sumas cuadradas entre ella (objetivo), y la referencia (el conjunto de *landmarks*) como:

$$D^2 = \sum_{j=1}^K [(X_{Rj} - (X_{Tj}R(1,1) + Y_{Tj}R(1,2) + Z_{Tj}R(1,3)))^2 + (Y_{Rj} - (X_{Tj}R(2,1) + Y_{Tj}R(2,2) + Z_{Tj}R(2,3)))^2 + (Z_{Rj} - (X_{Tj}R(3,1) + Y_{Tj}R(3,2) + Z_{Tj}R(3,3)))^2] \quad (1.10)$$

donde  $(X_{Rj}, Y_{Rj}, Z_{Rj})$  son las coordenadas del *landmark* de referencia,  $(X_{Tj}, Y_{Tj}, Z_{Tj})$  es la primera forma (objetivo), y  $R$  la matriz de rotación calculada anteriormente. Los ángulos  $\alpha$ ,  $\beta$ , y  $\gamma$  se pueden resolver como:

$$\begin{aligned} \alpha &= \arctan\left(\frac{\sum_{j=1}^K Y_{Rj}X_{Tj} - X_{Rj}Y_{Tj}}{\sum_{j=1}^K X_{Rj}X_{Tj} + Y_{Rj}Y_{Tj}}\right) \\ \beta &= \arctan\left(\frac{\sum_{j=1}^K Z_{Rj}X_{Tj} - X_{Rj}Z_{Tj}}{\sum_{j=1}^K X_{Rj}X_{Tj} + Z_{Rj}Z_{Tj}}\right) \\ \gamma &= \arctan\left(\frac{\sum_{j=1}^K Z_{Rj}Y_{Tj} - Y_{Rj}Z_{Tj}}{\sum_{j=1}^K Y_{Rj}Y_{Tj} + Z_{Rj}Z_{Tj}}\right) \end{aligned} \quad (1.11)$$

## 1.5. Métodos de Superposición

En el apartado 1.3, *Variables de forma y tamaño: Bookstein shape coordinates*, presentamos un método simple para obtener las variables de forma, fijando dos puntos y obteniendo

así las coordenadas de forma de *Bookstein*. Este método se explicó, ya que es fácil de entender, incluso sin saber nada de la teoría desarrollada anteriormente en el apartado 1.4, *Teoría de la forma*.

Una vez presentada la teoría de forma, ya podemos desarrollar un enfoque más general, uniendo formas antes de su comparación. Esta unión de formas es denominada "superposición", ya que todos los conjuntos de formas se posicionan una encima de otras (mediante las operaciones matemáticas que no alteran la forma). En la Figura 1.5 podemos ver la diferencia entre los dos métodos respecto la precisión de los mismos.



Figura 1.5: *Landmarks* de un tipo de especie acuática visualizada por dos tipos de superposiciones: (A) Coordenadas de forma *Bookstein*, (B) *Procrustes*. Extraído de [7].

El objetivo de este apartado es explicar el método más importante y usado en la "superposición", el método de *Procrustes*.

El nombre *Procrustes* proviene de la mitología griega; *Procrustes* estiraba a sus víctimas en una cama para que estos se ajustaran a ella. Al hacerlo, *Procrustes* minimizaba la diferencia entre sus víctimas y la cama. Este es un mito que se explica, pero el sentido matemático que contiene es que *Procrustes* minimiza las diferencias entre las configuraciones de diferentes formas.

La superposición de *Procrustes* se basa en la traslación, escala, y rotación para poder eliminar toda la información que no aporte sobre la descripción de la forma. Tal como mencionamos en apartados anteriores, los pasos para poder hacer estas operaciones son:

1. Centre cada configuración de forma (cada conjunto de *landmarks*) al origen. Para ello, se ha de restar cada coordenada de cada *landmark* con la coordenada del centroide.

Esto provocará que todo el conjunto de *landmarks* se traslade hacia el centroide.

2. Escale cada conjunto de *landmarks* dividiendo cada coordenada de cada *landmark* por el tamaño del centroide.
3. Seleccione un conjunto de *landmarks* como referencia y gire el segundo conjunto para minimizar la distancia parcial de Procrustes entre ambas.

Cuando existe más de un conjunto de formas, estas se rotan minimizando la distancia a partir de la misma referencia; se calcula el promedio de de todas las formas, y este pasa a ser la nueva referencia. Este proceso es iterativo hasta que la nueva referencia y la anterior son iguales. La referencia final es la que minimiza el promedio de las distancias de todas las formas.

## 1.6. Métodos de Ordenación

En este apartado hablaremos de como describir la diversidad de formas mediante el análisis de componentes principales (PCA).

PCA es una herramienta matemática para simplificar descripciones de variación entre un conjunto de individuos. Este análisis produce nuevo conjunto de variables que son combinaciones lineales de las variables originales.

Los patrones de variación y covariación en variables morfométricas son a menudo complejos y difíciles de interpretar. Es por ello que el objetivo de PCA es simplificar estos patrones, y hacerlos más fáciles de interpretar reemplazando las variables originales, con otras nuevas (componentes principales, PC) que son combinaciones lineales de las variables originales e independientes entre sí.

Un dato a favor de este análisis, es que con pocas PCs se puede estimar la mayoría de variación que presenta un conjunto de formas. Otro beneficio que tiene este análisis es la visualización de los datos. Es mucho más fácil reproducir en una gráfica las tres primeras PC's y ver su variación, que no visualizar por separado cada una de las variables que engloban el conjunto de formas.

Para poder explicar la descripción algebraica de PCA, imaginemos que tenemos un conjunto de observaciones de  $P$  variables de forma, y  $N$  individuos (ver Ecuación (1.12)).

Queremos obtener un conjunto de nuevas variables (PC) en el que la suma de sus varianzas sea igual al del conjunto original.

$$X = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1P} \\ X_{21} & X_{22} & \dots & X_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ X_{N1} & X_{N2} & \dots & X_{NP} \end{bmatrix} \quad (1.12)$$

donde  $X_{NP}$  es el valor de la  $P$  coordenada en el  $N$  individuo.

El espacio definido tendrá  $P$  dimensiones con  $N$  puntos proyectados en este espacio.

Nuestro objetivo es substituir nuestras variables originales ( $X_1, X_2, \dots, X_P$ ) representadas como columnas en nuestra matriz  $X$ , en nuevas variables (PC) como ( $Y_1, Y_2, \dots, Y_P$ ).

Tal como hemos mencionado anteriormente, las PC son combinaciones lineales de las variables originales, por eso podemos escribir  $Y_j$  como una combinación lineal de las variables originales como:

$$Y_j = A_{1j}X_1 + A_{2j}X_2 + \dots + A_{Pj}X_P \quad (1.13)$$

o en forma matricial como:

$$Y_j = A_j^T X \quad (1.14)$$

donde  $[A_{1j}, A_{2j}, \dots, A_{Pj}]$  es un vector de constantes y nuestra incógnita actual.

La varianza total no ha cambiado y esta se puede expresar mediante la matriz  $S$ :

$$S = \begin{bmatrix} S_{11} & S_{12} & \dots & S_{1P} \\ S_{21} & S_{22} & \dots & S_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ S_{P1} & S_{P2} & \dots & S_{PP} \end{bmatrix} \quad (1.15)$$

donde  $s_{ii}$  es la varianza observada en la variable  $X_i$  y se puede calcular mediante:

$$s_{ii} = \frac{1}{N} \sum_{k=1}^N (X_{ki} - \text{mean}(X_i))^2 \quad (1.16)$$

A partir de la Ecuación (1.14) y (1.15) podemos hallar la varianza de  $Y_j$  como:

$$\text{Var}(Y_j) = \text{Var}(A_j^T X) = A_j^T S A_j \quad (1.17)$$

Una vez llegado este punto, se pondrán en práctica dos restricciones. Una de las restricciones a cumplir es que la varianza no cambie, por eso el producto de cada vector ha de ser igual a 1:

$$\sum_{k=1}^p A_{kj}^2 = 1 \quad (1.18)$$

La segunda restricción es que los ejes de PC tienen covarianza = 0, es decir, que los ejes son ortogonales. Esta restricción se expresa como:

$$A_i^T A_j = 0 \quad (1.19)$$

Si aplicamos la primera restricción (Ecuación (1.18)) a la Ecuación (1.17) nos queda que la varianza de  $Y_j$  es  $S$ . Aplicando la segunda restricción (Ecuación (1.19)) a la matriz restante  $S$ , esta queda como:

$$\lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_p \end{bmatrix} \quad (1.20)$$

donde cada  $\lambda$  es el *eigenvalue*, y da solución a la ecuación:

$$S - \lambda_i I = 0 \quad (1.21)$$

representando I la matriz identidad de  $P \times P$ :

$$I = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \quad (1.22)$$

Para cada *eigenvalue* hay un vector  $A_i$  asociado llamado *eigenvector*:

$$(S - \lambda_i I)A_i = 0 \quad (1.23)$$

reorganizado:

$$SA_i = \lambda_i A_i \quad (1.24)$$

Por lo tanto, los *eigenvectors* son un nuevo grupo de variables con las misma varianza que sus *eigenvalues* y con covarianza igual a 0.

Si mostramos cual seria la varianza de PC1, juntando las Ecuaciones (1.17) y (1.24) obtenemos:

$$Var(Y_1) = A_1^T \lambda_1 A_1 \quad (1.25)$$

## Capítulo 2

# Descripción de Datos

En este capítulo haremos una breve descripción sobre los datos que trataremos en este proyecto, así como un breve análisis de sus características.

En este proyecto principalmente trataremos con datos (en concreto, resonancias magnéticas craneales) de pacientes con enfermedades mentales; esquizofrenia (ESQ) y bipolaridad (BP).

El estudio *Global Burden of Disease* realizado en 2010 destacó que los trastornos mentales representan el 7,5% de las enfermedades a nivel mundial. Entre todos los trastornos mentales, la ESQ y el trastorno BP, ocupan la quinta y sexta posición en términos de años vividos con dicha discapacidad [2].

Los trastornos mentales en general, y los trastornos de ESQ y BP en particular, son factores de riesgo para el comportamiento suicida. La ESQ aguda se considera el trastorno con mayor discapacidad individual entre todas las enfermedades debido a la gravedad de sus síntomas. En cuanto a la BP, la discapacidad individual asociada a un episodio maníaco también es alta, siendo comparable a las de afecciones más predominantes como el asma o la enfermedad de Alzheimer.

A continuación profundizaremos brevemente sobre las principales características de cada una de las enfermedades mentales mencionadas.

### 2.1. Esquizofrenia

La esquizofrenia (ESQ) es un trastorno grave, complejo, heterogéneo, crónico y de discapacidad causado por una alteración genética y / o ambiental del desarrollo cerebral [3].

Comúnmente, los primeros episodios psicóticos aparecen en la adolescencia tardía o en la edad adulta joven (entre 16 y 30 años), y principalmente afecta al género masculino. Más del 50% de las personas diagnosticadas con ESQ, sufren problemas psiquiátricos a largo plazo, y alrededor del 20% de los pacientes muestran cronicidad y discapacidad.

Además, la tasa de desempleo entre las personas afectadas aumenta hasta 80-90%, y la esperanza de vida también se reduce significativamente en 10-20 años. Debido a esta enorme carga social, la investigación médica se ha centrado en comprender la etiología y la patogénesis de ESQ para mejorar su prevención, diagnóstico y terapia.

La psicopatología en ESQ es diversa y heterogénea; sin embargo, los síntomas más comunes se clasifican en 1) Síntomas positivos: alucinaciones y delirios, que implican una pérdida de la percepción de la realidad; y 2) Síntomas negativos: deterioro del pensamiento y motor, capacidad de hablar reducida, funcionamiento ejecutivo deficiente y memoria de trabajo reducida [3].

## 2.2. Bipolaridad

La bipolaridad (BP) es una enfermedad crónica y compleja que se clasifica dentro de los trastornos afectivos. Esta enfermedad es causada por factores genéticos y ambientales, y presenta una tasa de heredabilidad relativamente alta que oscila entre 60-85% [2].

La BP se caracteriza por fluctuaciones en el estado de ánimo y episodios pronunciados de elevación del estado de ánimo, seguidos por episodios de depresión.

Según el *Diagnostic and Statistical Manual of Mental Disorders* DSM-V, el trastorno bipolar se puede dividir en cuatro subgrupos: trastorno bipolar I (BP-I), trastorno bipolar II (BP-II), trastorno bipolar no especificado (BP-NOS) y trastorno ciclotímico. Cada subgrupo varía en severidad y duración de los síntomas maníacos y depresivos.

Los estudios epidemiológicos mostraron que la BP tiene una incidencia del 1% en la población general, independientemente de su origen étnico, nacionalidad, o nivel socio-económico. La edad media de aparición en BP es alrededor de los 20 años. Sin embargo, algunos autores sugieren dos momentos vitales para su inicio: entre 15-24 años y entre 45-54 años [2]. Debido a que el diagnóstico es complejo, hace que sea difícil establecer una edad de inicio precisa.

Con respecto a la relación sexual de la BP, a pesar de que es difícil evaluar las diferencias entre sexos en el diagnóstico de BP a causa de los criterios de diagnóstico utilizados, se ha observado que la BP-II es más común en mujeres que en hombres, mientras que BP-I tiene la misma prevalencia en ambos géneros.

**Parte II**

**Parte Práctica**



## Capítulo 3

# Preprocesamiento de los datos

### 3.1. Origen y descripción de los datos

#### 3.1.1. Introducción

El segundo capítulo del proyecto tiene como objetivo presentar la base de datos con la que trataremos. Se explicará al detalle el preprocesado que es necesario hacer antes de desarrollar los algoritmos automáticos de detección de *landmarks*.

La base de datos con la que trabajaremos está formada por un conjunto de resonancias magnéticas (RM) craneales, y un conjunto de *landmarks*, 20 en concreto, asociados a estas. Nuestro principal objetivo será extraer una malla facial 3D, a partir de los datos craneales representados por las RM.

Este preprocesado de datos requiere de una gran importancia, ya que el tipo de datos con el que trataremos necesita de una normalización previa. Este hecho es debido principalmente a dos factores:

1. El proceso de adquisición no está en nuestro dominio.
2. Necesario para poder desarrollar *landmarking*, y el posterior análisis de la forma facial en base a los *landmarks*.

Nuestro objetivo principal será lograr la automatización de todos estos procesos. La finalidad de que estos procesos sean automáticos, consiste en minimizar las tareas manuales repetitivas que no aporten valor añadido, y que puedan provocar lugar a error. De esta manera, conseguimos que la escalabilidad de la base de datos sea viable y factible.

A lo largo del capítulo se describirá con extensión la base de datos mencionada, así cómo las principales problemáticas que se han debido afrontar para lograr su normalización de forma automática.

#### 3.1.2. Origen

La base de datos con la que trabajaremos, consta de datos reales y está formada por un total de 185 resonancias magnéticas craneales. Las muestras fueron recolectadas íntegramente, y pertenecen a la Fundación de Investigación FIDMAG Hermanas Hospitalarias en

Barcelona (España). Los datos utilizados en esta colaboración se obtuvieron con el consentimiento del Director de la Fundación de Investigación FIDMAG, Dra. Edith Pomarol-Clotet, y de los investigadores Dr. Raymond Salvador y Dr. Mar Fatjó-Vilas [3].

La diversidad que presenta nuestra base de datos está compuesta por pacientes con patologías con bipolaridad, esquizofrenia y pacientes control, tanto para el género masculino como femenino (ver Cuadro 3.1).

	<b>Bipolaridad</b>	<b>Control</b>	<b>Esquizofrenia</b>
<b>Género Femenino</b>	24	29	39
<b>Género Masculino</b>	21	46	26
<b>TOTAL (185)</b>	45	75	65

Cuadro 3.1: Diversidad de la base de datos

Una resonancia magnética (RM) es una técnica de imágenes médicas que utiliza un campo magnético y ondas de radio generadas por computadora para crear imágenes detalladas de los órganos y tejidos del cuerpo [9].

La RM no es la técnica más común para evaluar la morfología facial, debido a su naturaleza y los costos asociados. Sin embargo, ya se disponía de las imágenes de RM, y por lo tanto se pudo proceder a hacer las reconstrucciones faciales en 3D a partir de las RM craneales. El protocolo de neuroimagen se realizó en Hospital Sant Joan de Déu en un escáner 1.5T GE Sigma [2]. Y la reconstrucción facial se hizo a partir un software de análisis 3D para científicos, Amira 5.2 [1].

Además de las RM, la base de datos consta de 20 *landmarks* faciales por paciente (expresados en coordenadas XYZ), anotados manualmente sobre las RM. Estos *landmarks* han sido definidos por las propias biólogas que nos han facilitado la base de datos. Los datos de los *landmarks*, nos permitirán en un futuro conseguir su detección automática en mallas faciales 3D.

#### 3.1.3. Formato de Datos

En este apartado hablaremos del tipo de formato en el que trataremos las RM. Cada RM está representada por un archivo con formato PLY (ver Figura 3.1). PLY o *Polygon File Format* es una extensión utilizada para almacenar datos recopilados con escáneres 3D. También se conoce como formato del triángulo de Stanford porque fue creado en la universidad de Stanford en los mediados de los años 90, con un diseño inspirado por el formato del vector de Wavefronts OBJ [17].

Para poder preprocesar estos datos, hemos trabajado con *Point Clouds* a partir de los archivos PLY (ver Figura 3.2). Un archivo *Point Cloud* es una colección de puntos en el espacio 3D resultado de lo que captura un escáner al digitalizar un objeto físico. Los puntos representan la lectura de las coordenadas X, Y, Z de la superficie del objeto. Por lo tanto, nuestros datos estarán representados por un conjunto de puntos, alrededor de 1 millón, que nos proporcionaran la información necesaria para conocer la estructura craneal del paciente y así poder extraer la malla facial 3D.

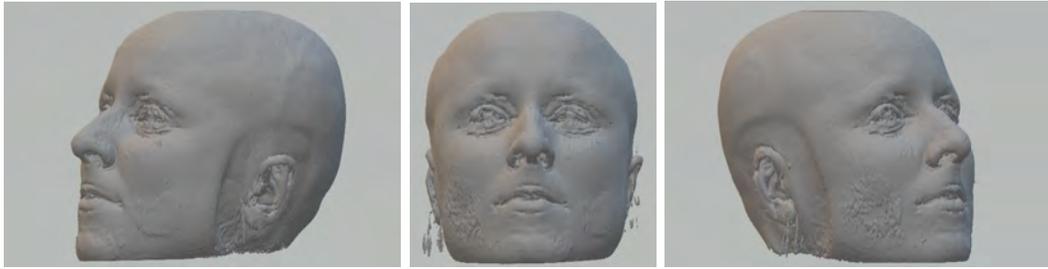


Figura 3.1: Modelo PLY de un paciente masculino con esquizofrenia.

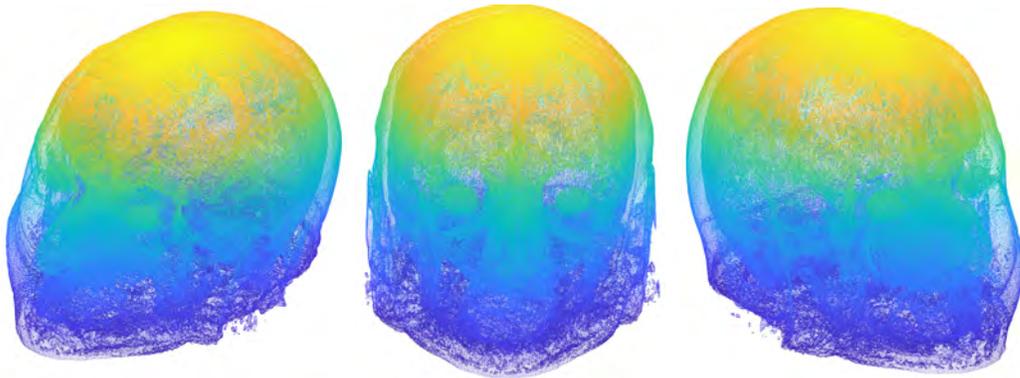


Figura 3.2: Modelo *Point Cloud* de un paciente masculino con esquizofrenia.

### 3.2. Problemáticas

Una vez descrita la base de datos, veremos a nivel general las diferentes problemáticas que nos han surgido al tratar con los modelos 3D facilitados. A continuación las particularizaremos para ver cómo solventarlas.

Antes de presentar las problemáticas, citar que los datos tratados son de gran complejidad, ya que la nube de puntos que presenta cada paciente en su representación craneal, es totalmente aleatoria, es decir, cada paciente presenta una morfología craneal diferente, hecho que dificulta notablemente la normalización de los modelos 3D.

Las problemáticas que nos encontramos son tres: la existencia de distintas orientaciones y alineaciones de las nubes de puntos en el espacio 3D (lo que impide su tratamiento uniforme), la presencia de ruido (lo que dificulta el procesamiento de los datos, pudiendo inducir a error), y el volumen de datos (dificulta el tiempo de cálculo computacional). En los siguientes párrafos se detallan en más profundidad dichas problemáticas.

Analizando los datos, la primera problemática que hallamos es la orientación 3D con respecto a los ejes de coordenadas XYZ, ya que existe cierta variabilidad entre modelos. Cada modelo craneal 3D representado por una nube de puntos, contiene una orientación diferente del cráneo. Este hecho hace que halla que implementar un algoritmo para orientar a todos los pacientes hacia un mismo eje (ver Figura 3.3).

Otra problemática vinculada a la anterior es la alineación. La nube de puntos se presenta alineada de diferente manera según la posición que tuvo el paciente cuando procedió a hacerse la RM. Esta problemática se relaciona con la anterior, ya que sin saber la orientación

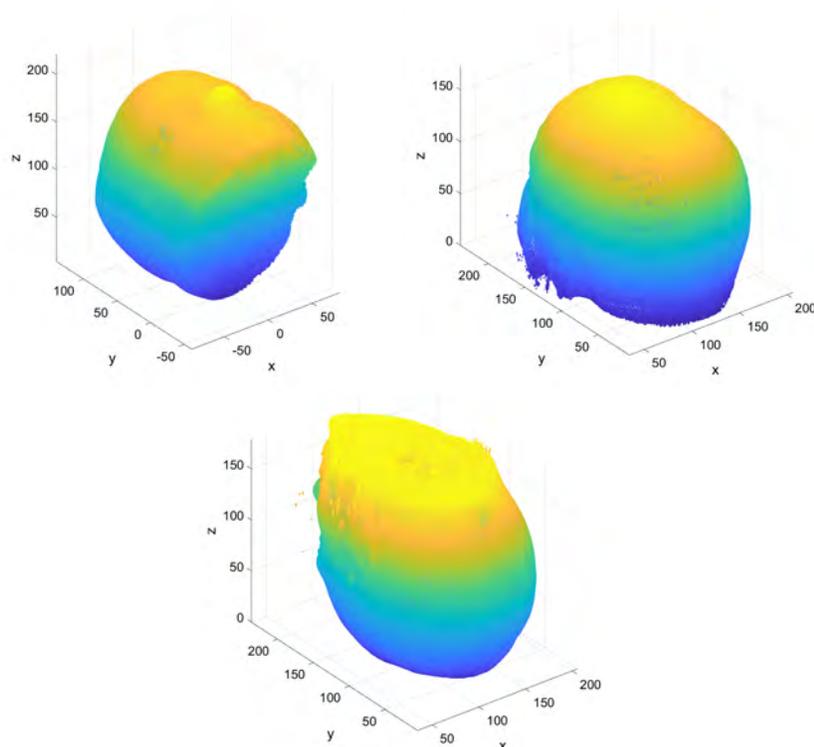


Figura 3.3: Casuísticas que presentan los pacientes respecto la orientación

no podemos proceder a hacer la corrección de la alineación.

La tercera problemática se debe al ruido que presentan algunos pacientes en las RM. Este puede ser causado por diferentes elementos de protección que se facilitaron a los pacientes en el momento de hacerse la RM o simplemente a presencia de artefactos en el momento de hacer la reconstrucción facial.

Este factor puede resultar perjudicial cuando implementemos nuestro sistema de aprendizaje para la detección automática de *landmarks* faciales (ver Figura 3.4).

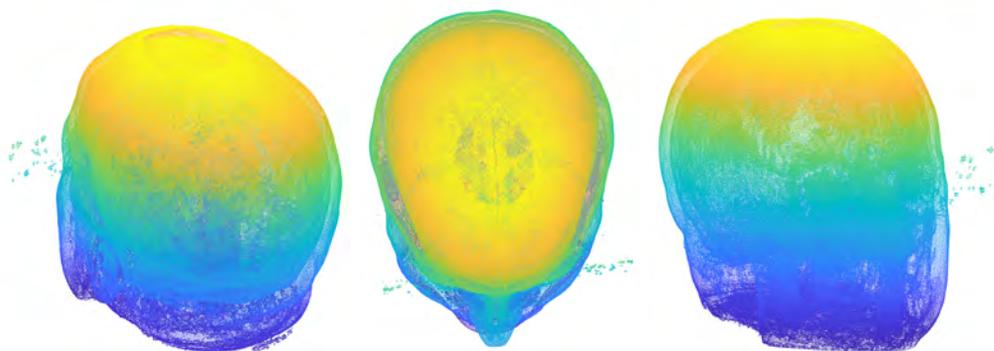


Figura 3.4: Representación craneal 3D con presencia de ruido

Los datos también presentan una complejidad con respecto a su volumen. Cada RM está formada por una nube de puntos que forman todo el cráneo. Esto produce que tengamos

una sobre representación de cada paciente, ya que en nuestro caso solamente nos interesa la información que pertenece a la superficie del rostro. Además, el hecho de tener todo el cráneo, hace que tengamos aproximadamente un millón de puntos por cada RM, y por lo tanto, que el coste computacional sea elevado y dificulte el procesado. Para ello, eliminaremos información de forma selectiva para acabar obteniendo una malla facial, y al mismo tiempo reducir el número de puntos.

A continuación, en los próximos apartados se describen detenidamente las soluciones ideadas e implementadas para hacer frente a las problemáticas mencionadas anteriormente.

### 3.2.1. Orientación

Tal como vemos en la Figura 3.3, existe variabilidad en la orientación de los modelos 3D, es decir, cada nube de puntos (cada representación craneal) está orientada hacia un eje de coordenadas diferente.

La casuística que presentan nuestros modelos 3D, son objetos orientados hacia el eje de coordenadas Y o Z, tanto positivas como negativas (ver Cuadro 3.2). En nuestro caso, se ha determinado orientar a todos los pacientes hacia el eje de coordenadas Y positivas. La finalidad de esta determinación es que la mayoría de pacientes tengan que aplicar por el método implementado, y que este tenga la fiabilidad de ser robusto.

Diversidad	Bipolaridad		Control		Esquizofrenia	
	Mujeres	Hombres	Mujeres	Hombres	Mujeres	Hombres
<b>Y+</b>	8	7	7	0	13	0
<b>Y+ (cráneo invertido)</b>	16	14	22	0	13	0
<b>Z+</b>	0	0	0	46	0	39

Cuadro 3.2: Clasificación de la base de datos según la diversidad de orientación craneal

Para poder realizar la orientación pertinente, el primer paso es centrar el cráneo en el punto (0,0,0) basándonos en la media del eje de coordenadas XYZ. Una vez centrada la nube de puntos, el próximo paso es detectar, de forma no supervisada, la parte facial de nuestro objeto para poder orientarla hacia el eje de coordenadas Y positivas. Este paso es complejo, ya que los datos no presentan ninguna referencia que ayude a detectar dónde se sitúa el rostro del individuo. Para ello, se ha implementado un método que detecta donde se encuentra la nariz o la barbilla definido en el diagrama de bloques de la Figura 3.5.

Si pensamos en la morfología del cráneo, podemos visualizar como la nariz y barbilla son dos de las regiones más alejadas del cráneo respecto el centroide. Es por ello que el primer paso de este método consistirá en hallar el 5% de puntos más lejanos respecto al centroide de nuestra nube de puntos. Para realizarlo, calcularemos la distancia Euclidiana de todos los puntos respecto al punto (0,0,0) mediante la Ecuación (3.1), y haremos una selección acotada. Este conjunto de puntos lo etiquetaremos como *Conjunto del 5%* (ver Figura 3.6).

$$D(P) = \sqrt{x_p^2 + y_p^2 + z_p^2} \quad (3.1)$$

Donde  $P = (x_p, y_p, z_p)$  representan las coordenadas d'un punto genérico de nuestra nube de puntos.

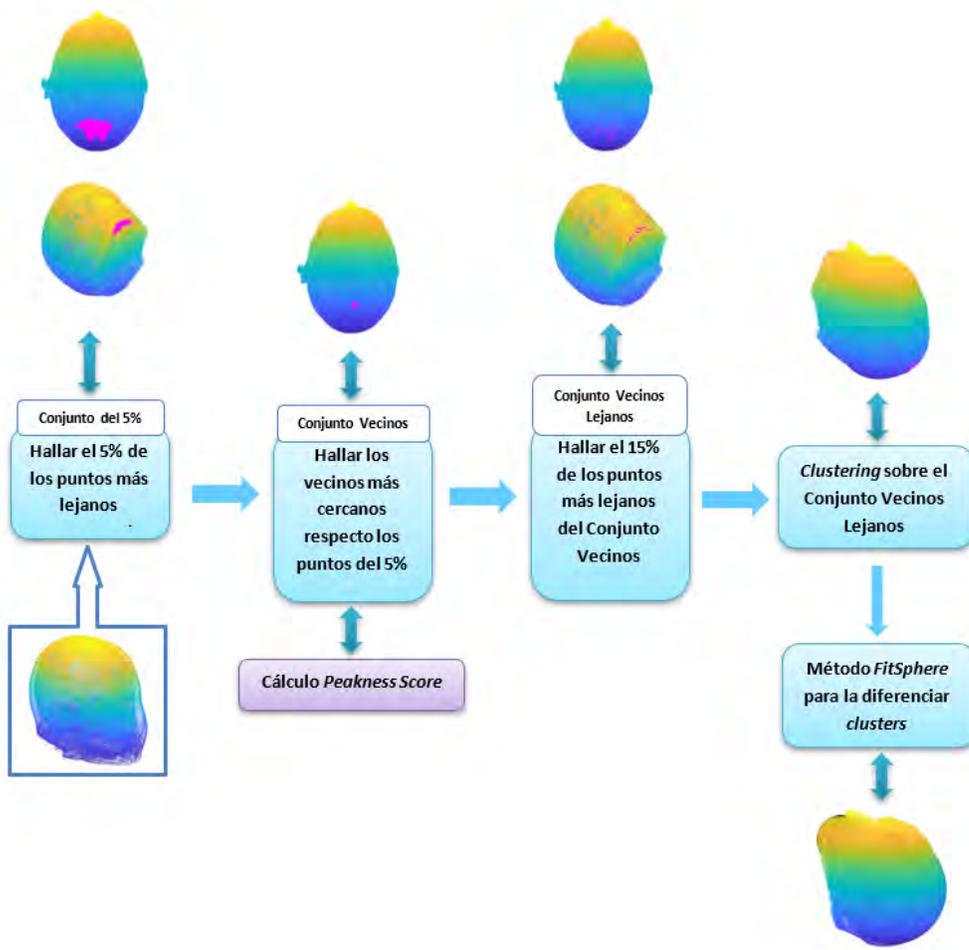


Figura 3.5: Diagrama de bloques del método de detección barbilla/nariz.

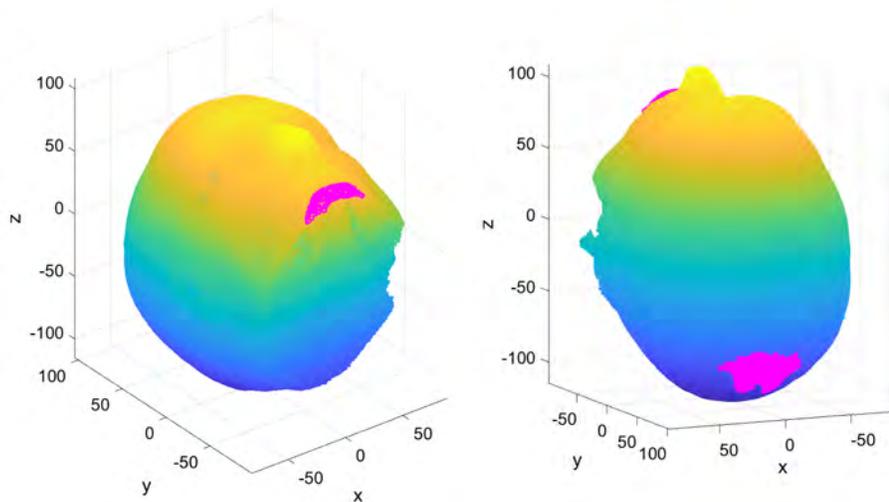


Figura 3.6: Representación del 5 % de puntos más lejanos respecto la nube de puntos original.

Una vez obtenido este conjunto, el segundo paso consiste en hallar los vecinos más cercanos de cada punto que engloba este *Conjunto del 5 %* [10]. Para cada punto, calcularemos la distancia Euclidiana (Ecuación (3.1)) de los puntos vecinos y su *Peakness score*.

El *Peakness score* es un estimador de la curvatura que tiene como objetivo buscar las zonas puntiagudas más lejanas del cráneo, con tal de hallar la zona de la nariz/barbilla y así poder detectar la zona facial del cráneo. Para ello, se ha hecho una selección de los vecinos que se encuentran más lejos del centroide y con mayor *Peakness score*, es decir, con mayor estimación de curvatura entre puntos. Este nuevo grupo de puntos lo etiquetaremos como *Conjunto Vecinos*.

En la Figura 3.7 podemos apreciar los vecinos que presenta un punto del *Conjunto del 5 %*.

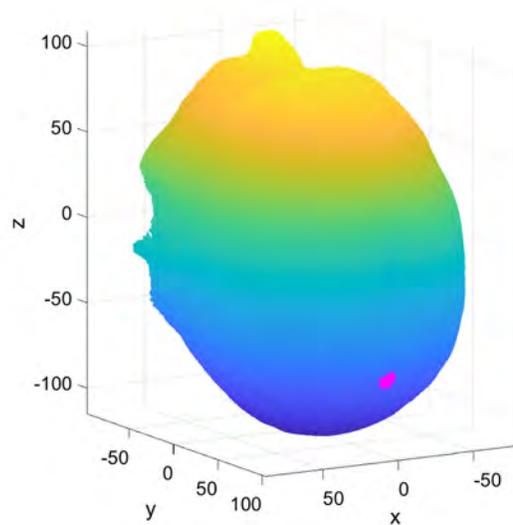


Figura 3.7: Representación del conjunto de vecinos más lejanos y con mayor *Peakness score* de un punto en concreto.

El tercer paso consistirá en hacer una búsqueda por el *Conjunto Vecinos* y quedarnos con el 15 % de puntos más lejanos. Este nuevo conjunto lo etiquetaremos como *Conjunto Vecinos lejanos* (ver Figura 3.8).

Tal como hemos avanzado anteriormente, según la morfología del cráneo, los puntos más lejanos respecto al centroide se sitúan en la coronilla, nariz o barbilla. Nuestro objetivo en el cuarto paso, será diferenciar la coronilla vs nariz y barbilla. Para ello, aplicaremos técnicas de *Clustering* en el *Conjunto Vecinos lejanos* para segregar los datos en diferentes *clusters*.

La técnica de *Clustering* consiste en agrupar un conjunto de objetos, de tal manera que los miembros del mismo grupo (llamado *cluster*) sean más similares, en algún sentido u otro. De tal manera que en nuestro caso, acabemos obteniendo un *cluster* en la coronilla, otro en la nariz, y otro en la barbilla (ver Figura 3.9).

En la Figura 3.9 podemos apreciar los diferentes *clusters* hallados después de aplicar el método de *Clustering*.

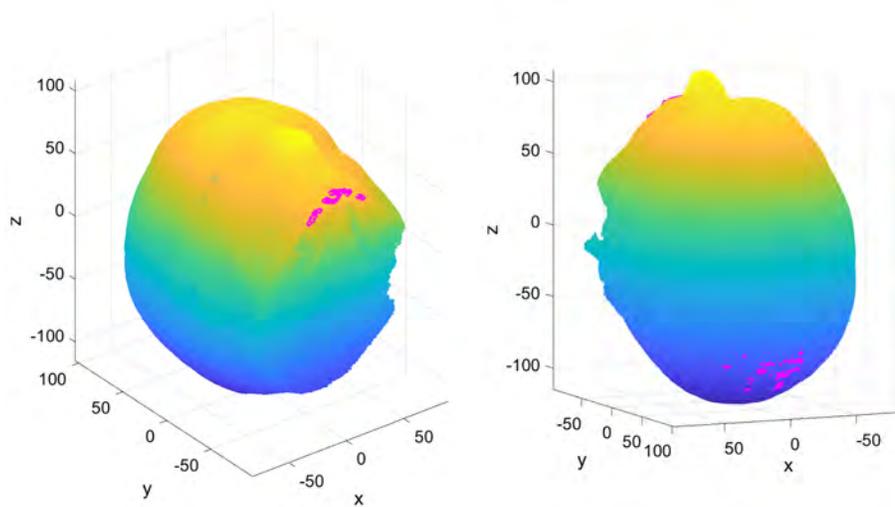


Figura 3.8: Representación del 15% de vecinos más lejanos y con mayor *Peakness score*.

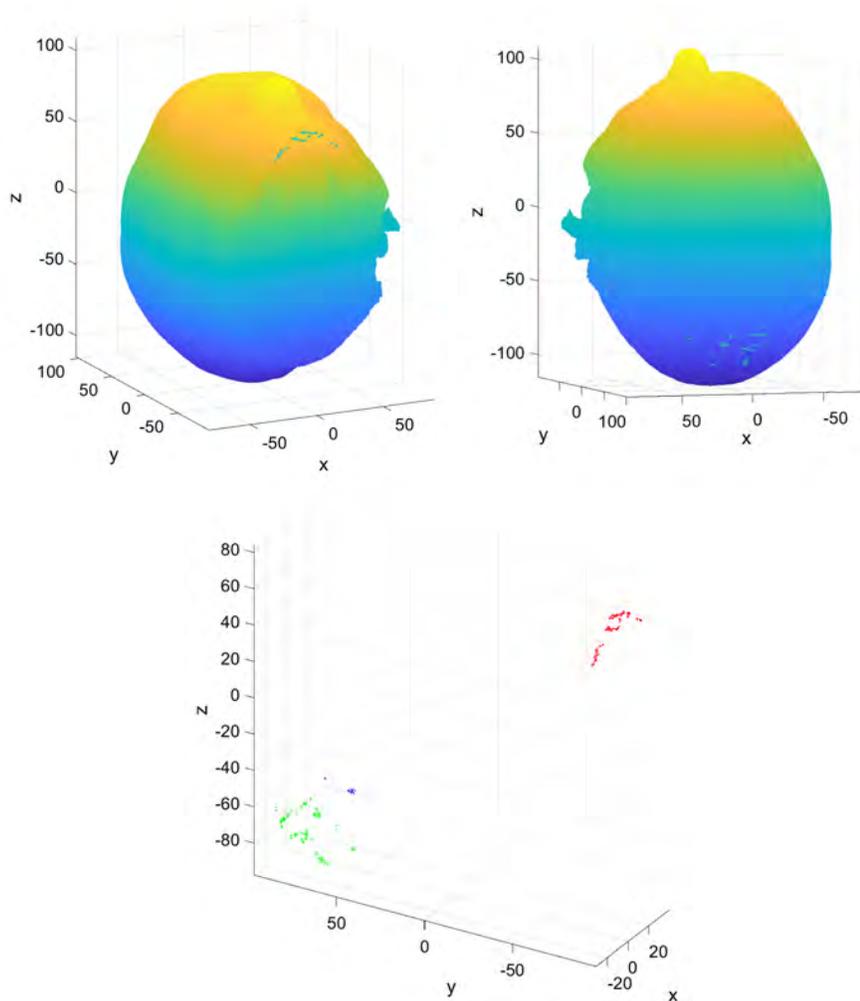


Figura 3.9: *Clusters* del 15% de vecinos más lejanos y con mayor *Peakness score*.

Una vez obtenidos los diferentes *clusters*, en el quinto y último paso, aplicaremos el método *Fit Spheres* [16] para proceder a la diferenciación entre *clusters*. Nuestro objetivo ahora es diferenciar la coronilla, de nariz y barbilla. Este método se basa en aplicar un algoritmo que adapta una esfera a una nube de puntos (ver Figura 3.10).

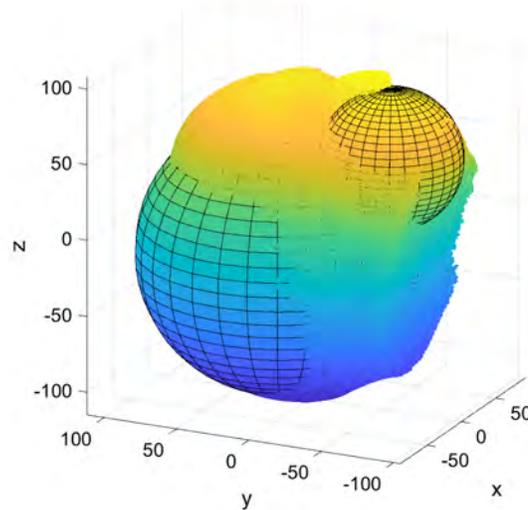


Figura 3.10: Método *Fit Spheres* para diferenciar entre los *clusters* de la coronilla y los de la nariz y barbilla.

Creadas las esferas para cada *cluster*, a fin de hacer la diferenciación entre ambas, nos basaremos en el radio que estas presentan. La morfología del cráneo nos define que la región de la nariz y barbilla es mucho más pequeña que la región de la coronilla (ver Figura 3.11). Esto produce que el radio de la esfera en la coronilla sea mucho más grande que el radio de la nariz y barbilla. Comparando ambos radios podremos diferenciar ambas partes.

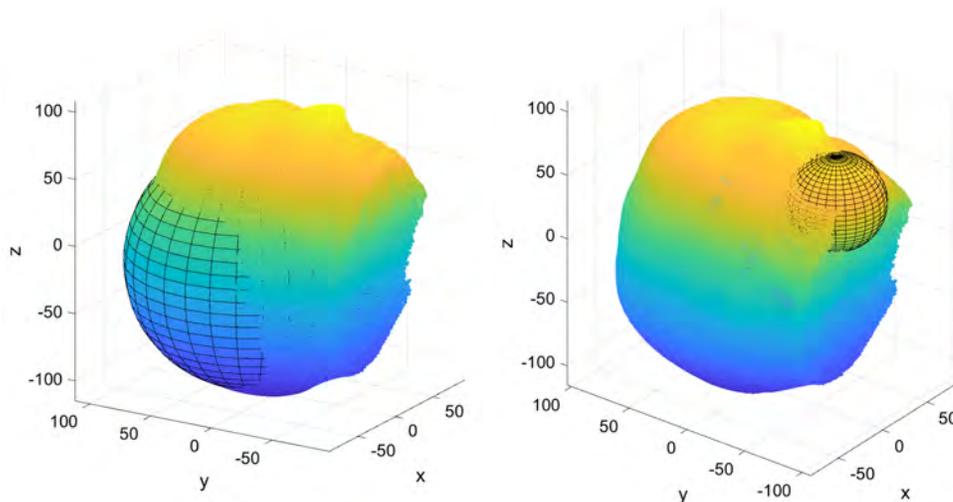


Figura 3.11: Comparación entre esferas al aplicar el método *Fit Spheres*.

Detectado y eliminado el *cluster* de la coronilla, solamente nos quedaran las nubes de puntos que forman el *cluster* de la barbilla y de la nariz. Las coordenadas de estas nubes de puntos, nos servirán para orientar y alinear el objeto.

Una vez aplicado este método de forma totalmente no supervisada, conseguimos detectar nariz o barbilla correctamente en un 98% de los pacientes. Esto permite que podamos orientar a todos los pacientes hacia el mismo eje de coordenadas (en nuestro caso hacia el eje de coordenadas Y positivo). Para ello, buscando el máximo entre coordenadas X, Y, Z del conjunto de puntos que forman el *cluster* de la nariz o barbilla, sabremos hacia dónde está orientado el cráneo del paciente.

Una vez detectada la orientación del cráneo del paciente, procedemos a orientarlo hacia el eje de coordenadas Y positivas en caso que esté orientado en otra dirección. Esta rotación se aplicará a todo el *PointCloud* mediante una transformación geométrica 3D formada por las matrices de rotación definidas en la Ecuación (3.2). Los *landmarks* asociados a cada RM también se rotaran mediante estas matrices.

$$M_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\alpha) & \sin(\alpha) & 0 \\ 0 & -\sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} M_y = \begin{bmatrix} \cos(\alpha) & 0 & -\sin(\alpha) & 0 \\ 0 & 1 & 0 & 0 \\ \sin(\alpha) & 0 & \cos(\alpha) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} M_z = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) & 0 & 0 \\ -\sin(\alpha) & \cos(\alpha) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.2)$$

Donde cada matriz pertenece a la rotación sobre la coordenada X, Y, Z respectivamente, y  $\alpha$  representa el ángulo de rotación.

Tal como podemos apreciar en la Figura 3.12, la nube de puntos del *PointCloud* original estaba orientada hacia el eje de coordenadas Z positivas, una vez aplicado el método implementado, la nube de puntos se ha orientado hacia el eje de coordenadas que habíamos determinado previamente (eje de coordenadas Y positivas).

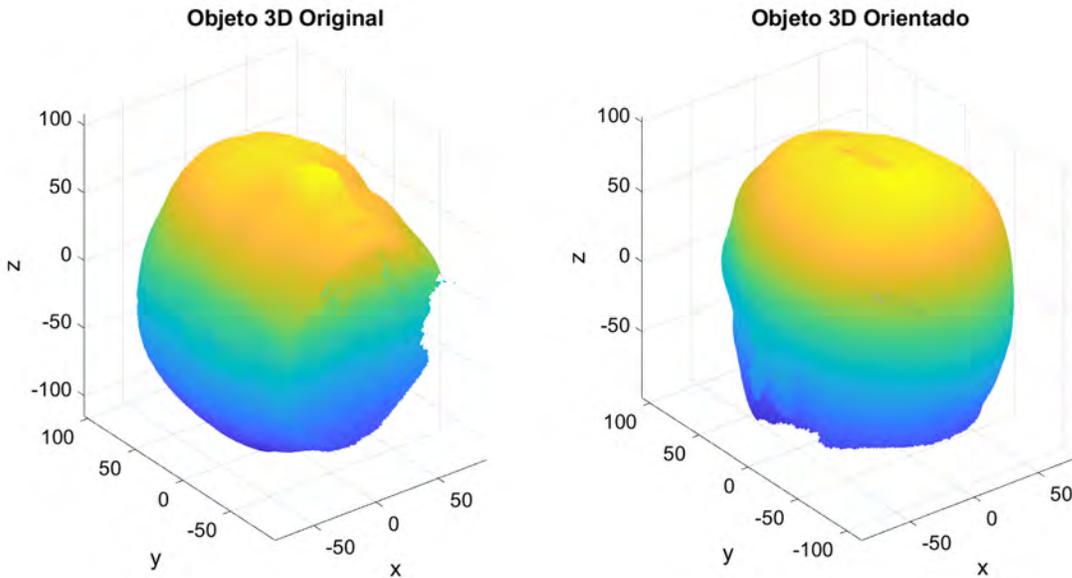


Figura 3.12: Orientación de un modelo 3D originalmente orientado hacia el eje de coordenadas Z positivas.

En la Figura 3.13 podemos ver la otra casuística que presentaba la base de datos (nube de puntos orientada hacia el eje de coordenadas Y positivas y con el cráneo invertido).

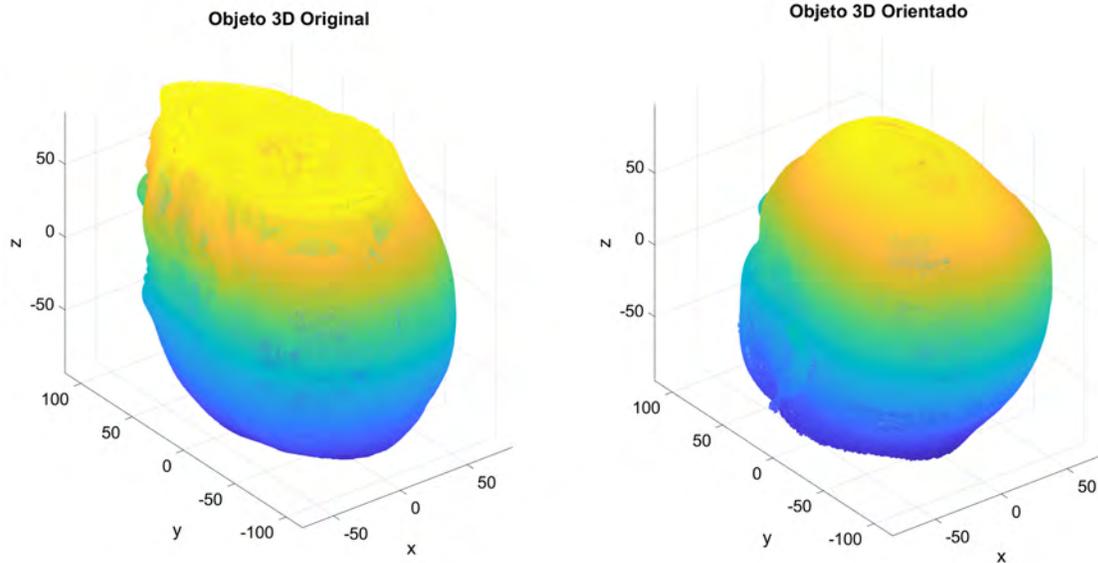


Figura 3.13: Orientación de un modelo 3D originalmente orientado hacia el eje de coordenadas Y positivas con el cráneo invertido.

### 3.2.2. Alineación

Una vez realizado el método para detectar la ubicación de la barbilla o nariz, y haber orientado a todos los pacientes hacia el eje de coordenadas Y positivo, el siguiente paso es alinearlos.

El problema que existe con respecto a la alineación es que el cráneo de los pacientes puede estar ligeramente rotado en cualquier eje. Esto provoca que cada nube de puntos esté alineada de manera diferente. Es por esto que en este apartado vamos a tratar de alinear a todos los pacientes respecto el mismo punto de referencia. En nuestro caso, este punto será el que pertenece a la punta de la nariz.

Llegados a este punto, y habiendo aplicado el método de orientación explicado en el apartado 3.2.1, tenemos la seguridad de afirmar que todos los pacientes están orientados hacia el eje de coordenadas Y positivas. Este hecho nos da cierta ventaja para afrontar la implementación del siguiente método, detección del punto máximo en el eje de coordenadas Y positiva.

Tal como hemos visto en el apartado 3.2.1, los puntos más lejanos se encuentran tanto en la nariz como en la barbilla. Es por ello que habrá que excluir los puntos de la barbilla en el eje de coordenadas Z, y hacer una selección entre los puntos de la nariz para hallar el punto máximo. El método se basará en los siguientes pasos (ver Figura 3.14):

- 1 Ordene la nube de puntos respecto el eje de coordenadas Z.

### 3.2. Problemáticas

- II Halle la distribución de puntos mediante el valor de  $Z$  máximo, mínimo y el número de puntos totales.
- III Calcule el 15 % respecto los puntos totales.
- IV Calcule el número de puntos a eliminar respecto la distribución y el 15 % estipulado.
- v Halle el punto con valor de  $Y$  máxima respecto los puntos restantes (ver Figura 3.15).
- VI Alinee el objeto 3D a partir del punto hallado anteriormente

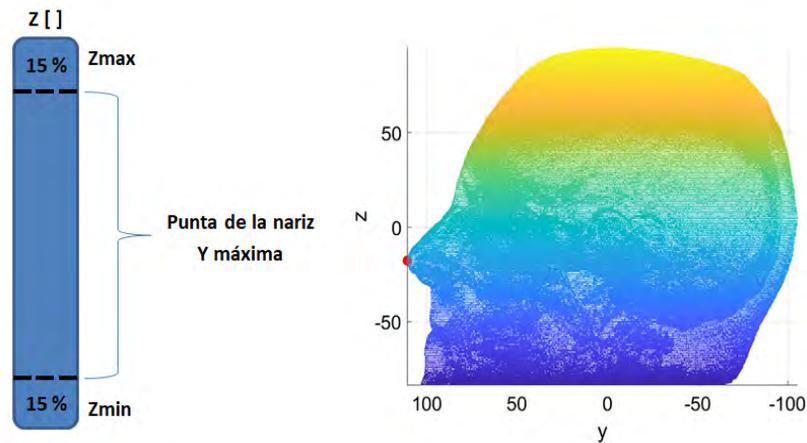


Figura 3.14: Diagrama de bloques del método de detección del punto máximo de la nariz.

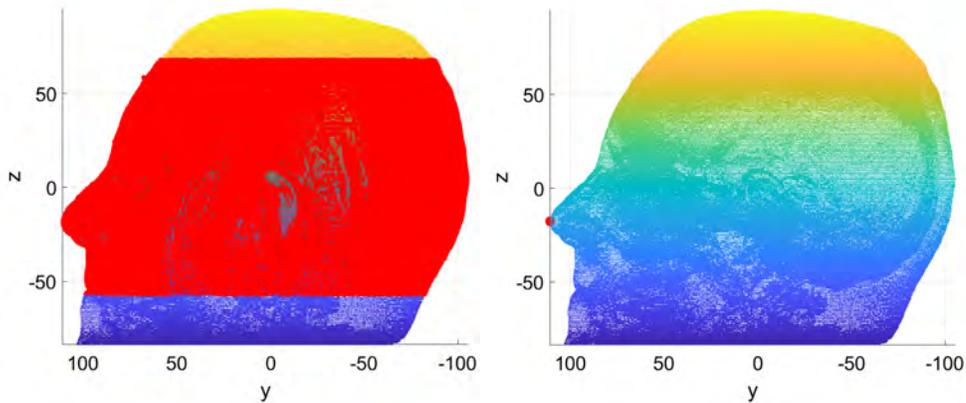


Figura 3.15: Rango para hallar el punto con  $Y$  máxima.

Una vez aplicado este método, conseguimos hallar el punto que hace referencia a la punta de la nariz del paciente en un 99 % de los casos. Esto permite que podamos alinear el cráneo a partir de la información de este punto. Para ello calcularemos los ángulos respecto los planos  $YZ$  y  $XZ$ , y rotaremos el conjunto de puntos original para alinear la cabeza (ver Figura 3.16).

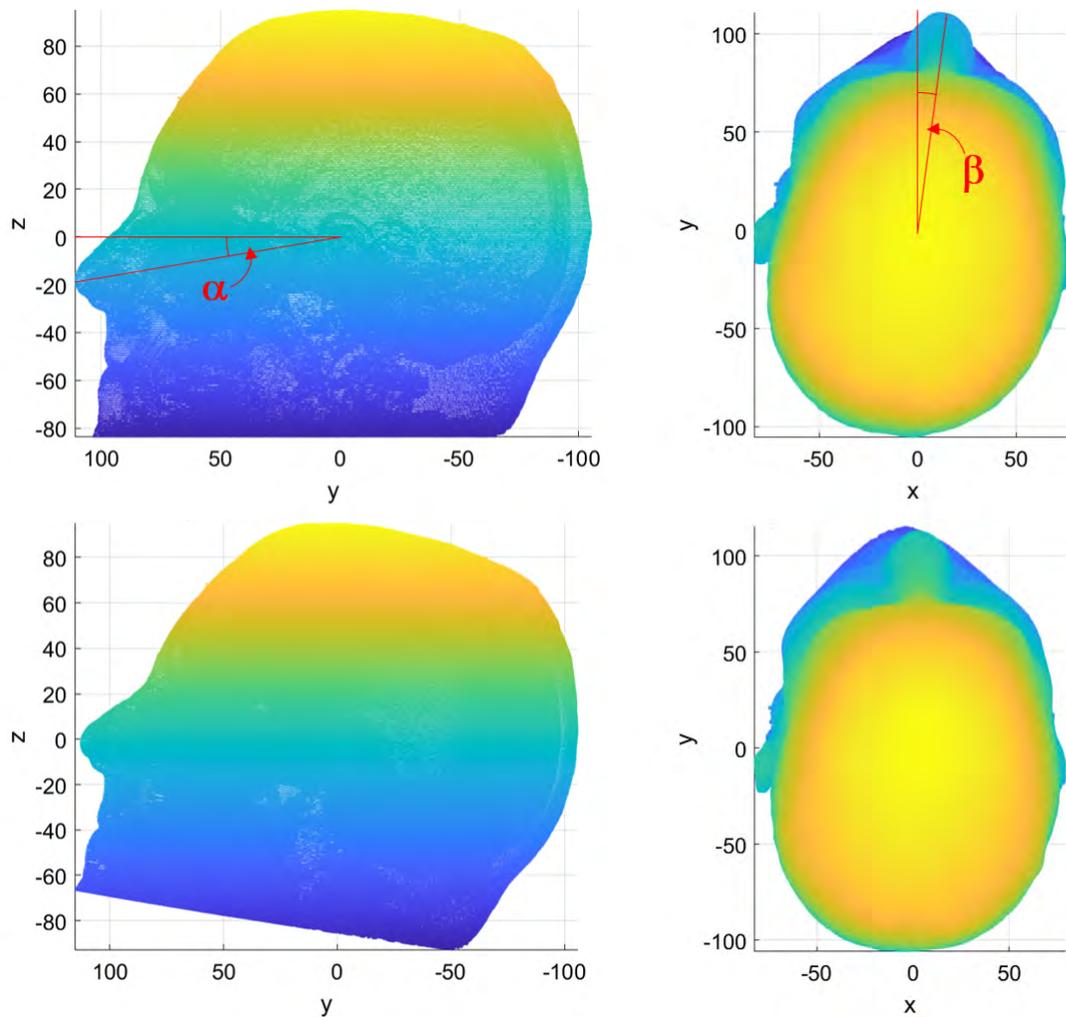


Figura 3.16: Ejemplo de alineación de un modelo *PointCloud*.

La rotación sobre la nube de puntos y *landmarks*, se basará en la misma metodología que se empleó en el apartado 3.2.1 mediante las matrices de rotación representadas en la Ecuación (3.2).

### 3.2.3. Ruido

Como hemos podido apreciar en la Figura 3.4, muchos pacientes presentan ruido en sus RM. Este factor es un inconveniente al tratar con los datos, ya que estos puntos añadidos por el ruido, no pertenecen al cráneo, generan dificultades durante el análisis, y pueden conllevar a resultados erróneos.

En nuestro caso, uno de los inconvenientes que nos produce el ruido, es la construcción de la malla facial 3D, ya que tal y cómo veremos más adelante, a fin de obtener dicha malla, realizaremos proyecciones hacia el cráneo, y des del centro del cráneo.

Para eliminar estos puntos no deseados hemos aplicado de nuevo técnicas de *Clustering*. La finalidad es generar un *cluster* para cada artefacto o residuo que presente el modelo 3D.

Tal como podemos observar en la Figura 3.17, cada *cluster* se muestra de diferente color. Existe un *cluster* de color rojo que engloba a todo el cráneo y varios *clusters* de diferente color que representan los residuos varios. El objetivo será quedarnos con el *cluster* más grande (*cluster* que pertenece al cráneo), y eliminar aquellos que estén muy separados del cráneo (*clusters* que pertenecen a residuos). Este proceso se basa en definir una distancia máxima entre *clusters* y eliminar aquellos que estén fuera del rango deseado.

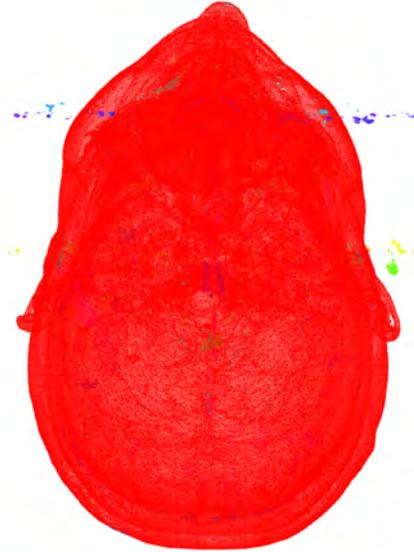


Figura 3.17: Ejemplo de aplicación del método de *Clustering*.

Tal como podemos apreciar en la Figura 3.18, vemos como la nube de puntos original presenta ruido en forma de residuos cerca de los ojos, a diferencia de la Figura 3.19 que ya no presenta ningún residuo.

### 3.3. Obtención de la malla facial 3D

Una vez resueltas las problemáticas de orientación, alineación, y ruido, podemos asegurar que nuestras muestras se orientan y alinean en el mismo eje de referencia, y sin artefactos o puntos no deseados.

El objetivo en este apartado será construir una malla facial 3D a partir del modelo craneal 3D obtenido posteriormente a las problemáticas existentes en su origen.

Como hemos mencionado anteriormente, las RM presentan más información de la necesaria para construir la malla facial 3D, ya que están formadas por una nube de puntos que engloba todo el cráneo. Además, este hecho produce que tengamos alrededor de 1 millón de puntos por RM, y que a nivel computacional dificulte el procesamiento de datos.

Para ello, se ha implementado el método llamado *Automatic Meshing of faces in 3D*,  $AMF_{3D}$ . Este método es capaz de obtener una malla facial, obteniendo así los puntos que pertenecen únicamente a la superficie del rostro.

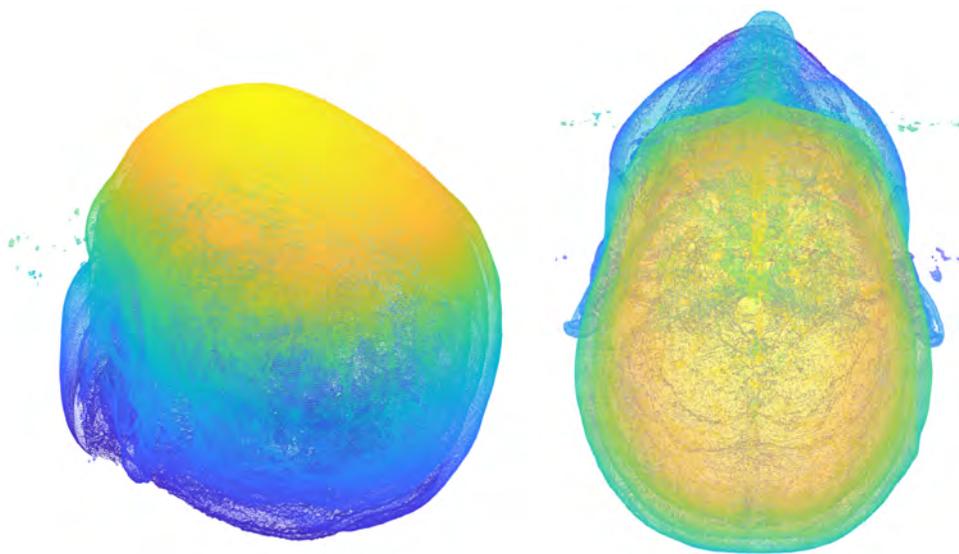


Figura 3.18: Modelo *PointCloud* con presencia de ruido.

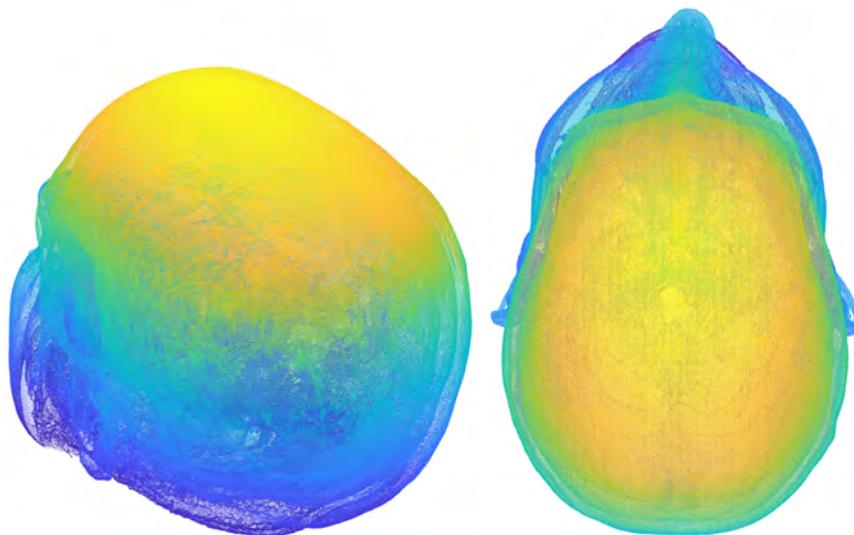


Figura 3.19: Modelo *PointCloud* después de haber aplicado *clustering* para la corrección de ruido.

A continuación, se presentan dos variantes del método  $AMF_{3D}$ . Ambas obtienen una malla facial 3D, pero la primera variante presenta una malla sin tener en cuenta la uniformidad de la nube puntos, a diferencia de la segunda que si que tiene en cuenta este factor.

### 3.3.1. $AMF_{3D}$ - Basado en cuantificación angular

El primer el paso para obtener la malla facial 3D, será quedarnos con aquellos puntos que solamente pertenezcan al rostro del paciente. El método para lograrlo se basará en hacer

filtrado angular de todo el cráneo y quedarnos solamente con la parte facial.

A causa de la geometría del problema, haremos una transformación de datos y trabajaremos con coordenadas esféricas en vez de con cartesianas. Para ello utilizaremos las Ecuaciones (3.3) y (3.4).

El rango del filtrado angular que definiremos será de  $[0 \pi]$  respecto el plano XY (parte facial del cráneo) y de  $[-\pi/4 \pi/4]$  respecto el plano YZ (elevación). Aplicando este filtrado, eliminamos todos aquellos puntos que se encuentran fuera de la zona facial (ver Figura 3.20).

$$x = R \sin \theta \cos \phi \quad y = R \sin \theta \sin \phi \quad z = R \cos \theta \quad (3.3)$$

$$R = \sqrt{x^2 + y^2 + z^2} \quad \phi = \tan^{-1}\left(\frac{y}{x}\right) \quad \theta = \cos^{-1}\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) \quad (3.4)$$

Donde  $x, y, z$  representan las coordenadas cartesianas,  $\theta$  y  $\phi$  los ángulos de elevación y azimut respectivamente, y  $R$  el radio.

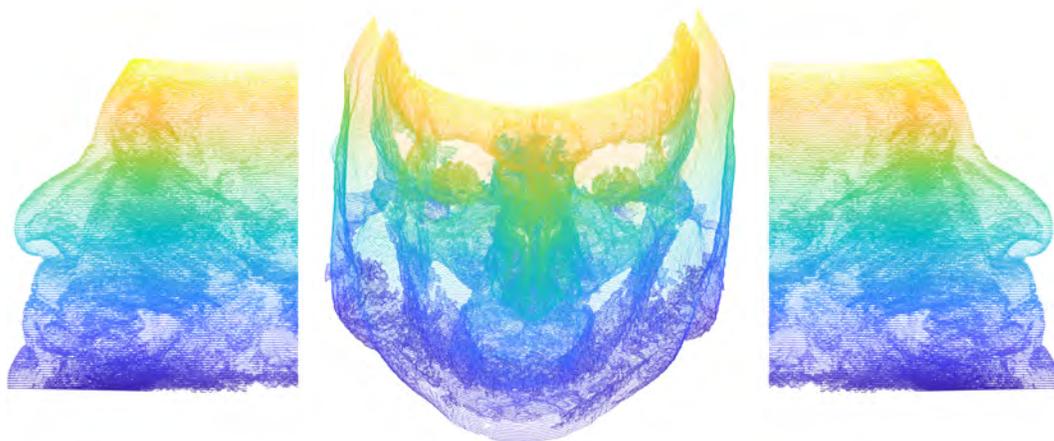


Figura 3.20: Zona facial de un modelo *PointCloud*.

Haciendo esta reducción, conseguimos reducir la nube de puntos hasta la cifra de 300.000 puntos. Tal como podemos apreciar en la Figura 3.21, nuestra nube de puntos está formada solamente por la parte frontal del cráneo del paciente.

Aún haciendo una primera reducción, y habiendo reducido notablemente el número de puntos, todavía tenemos puntos no relevantes para nuestro análisis que necesitamos eliminar.

En este punto, tenemos la zona facial del cráneo, pero incluyendo los puntos internos que pertenecen al cerebro, mandíbula, cavidades oculares... Nuestro objetivo será extraer una malla facial, es decir, quedarnos solamente con el contorno y eliminar los puntos internos.

Para ello, se ha implementado un método que solamente se quede con aquellos puntos que tengan distancia máxima respecto el centro  $(0,0,0)$ , y por lo tanto, elimine los puntos que pertenezcan a las cavidades internas del cráneo.

Este método consiste en hacer una cuantificación espacial con una resolución angular predeterminada y ajustable por el usuario. En la Figura 3.22 podemos apreciar dos mallas

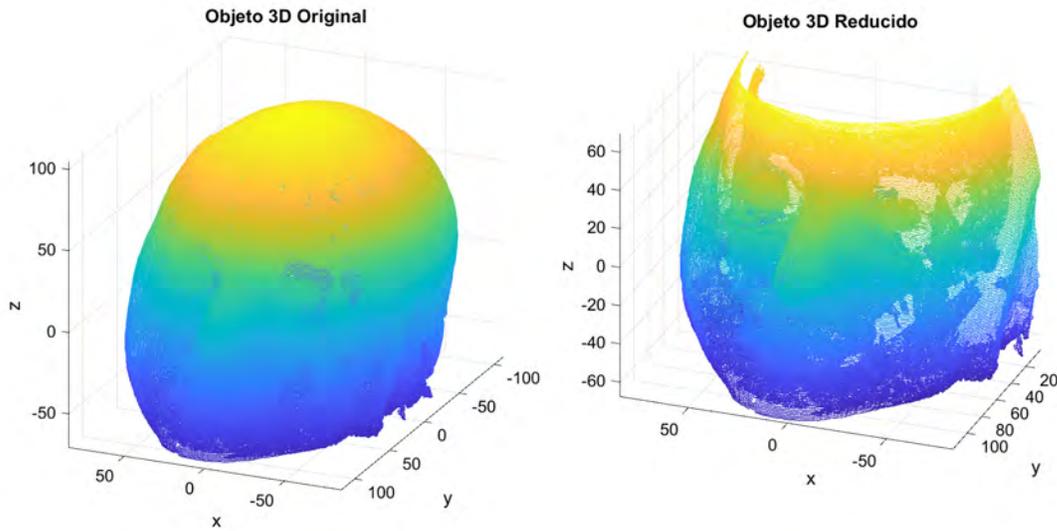


Figura 3.21: Zona facial de un modelo *PointCloud*

faciales con diferente resolución.

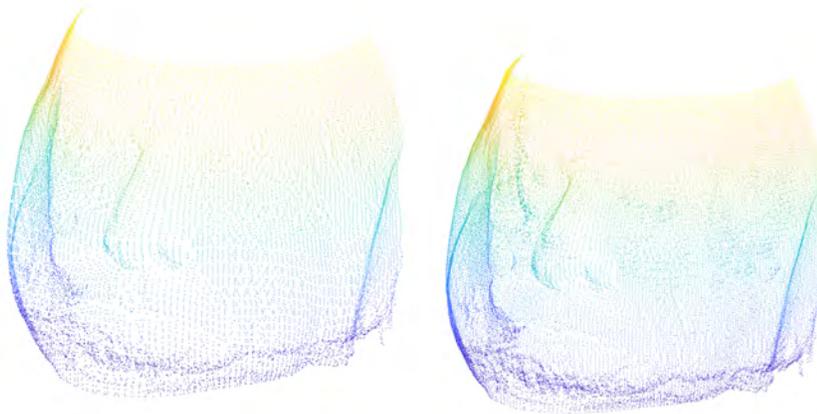


Figura 3.22: Mallas faciales 3D de un modelo *PointCloud* con cuantificación espacial 128 y 192 respectivamente.

Una vez hecha esta cuantificación, la idea es hacer una proyección desde el centro del cráneo, hasta el punto máximo en todos los puntos. La proyección con distancia máxima en ese ángulo determinado, será la que pertenecerá a la superficie del rostro del paciente.

Por lo tanto, el algoritmo que desarrollaremos y aplicaremos, hará un barrido por todas las cuantificaciones, y se quedará con el punto que se encuentre más lejos del centroide en cada proyección. De esta manera, conseguimos tener una malla facial y quedarnos con el contorno del objeto 3D (Figura 3.23).

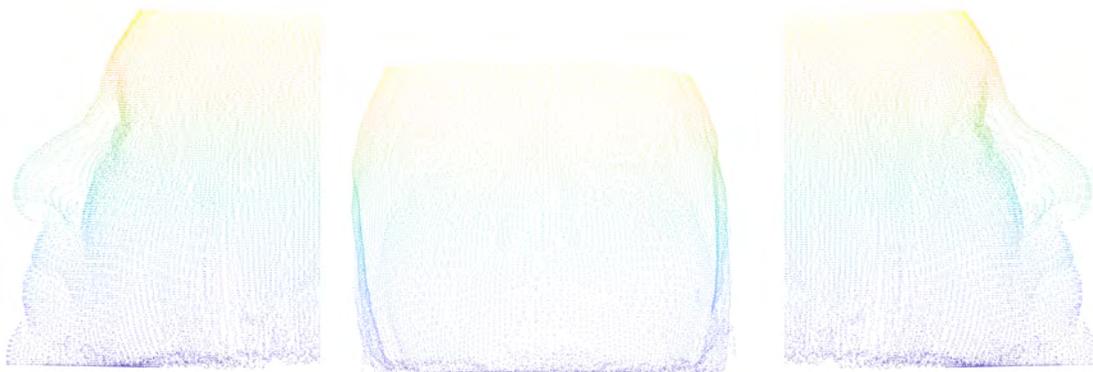


Figura 3.23: Malla facial 3D de un modelo *PointCloud*

Aplicado este método, conseguimos obtener una malla facial 3D solo con puntos de interés (superficie del rostro). También conseguimos reducir aún más la nube de puntos original hasta quedarnos con alrededor de unos 10.000 puntos por paciente. En la Figura 3.24 podemos apreciar la diferencia entre el modelo 3D original, y el modelo 3D reducido después de aplicar el método anterior.

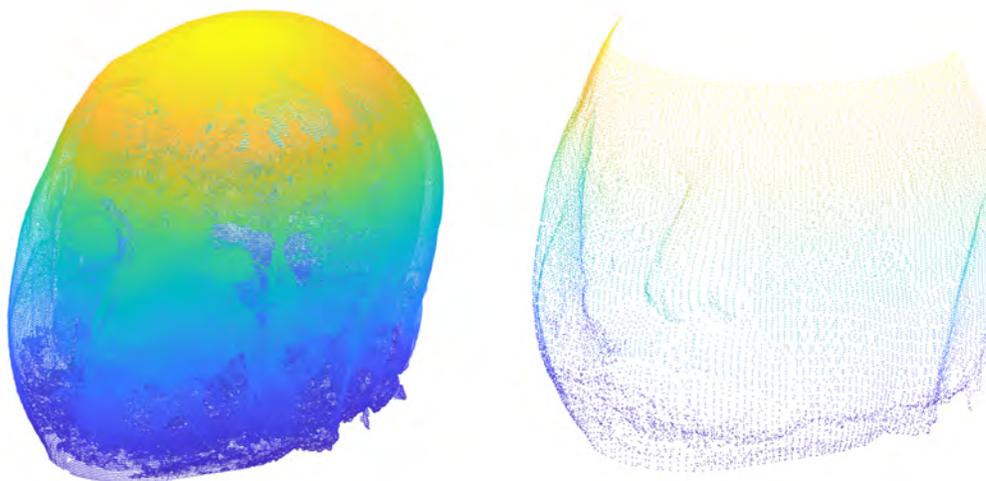


Figura 3.24: Comparación modelo 3D original VS modelo 3D reducido

Finalmente, hemos conseguido obtener una malla facial 3D sin puntos internos, y una cantidad de puntos 100 veces más pequeña que la del modelo original. Esto producirá que podamos tratar con fluidez los datos de cara al proceso de automatización de *landmarks*, tratando únicamente con los puntos de interés.

En la Figura 3.25 podemos apreciar la malla facial 3D con *landmarks*.

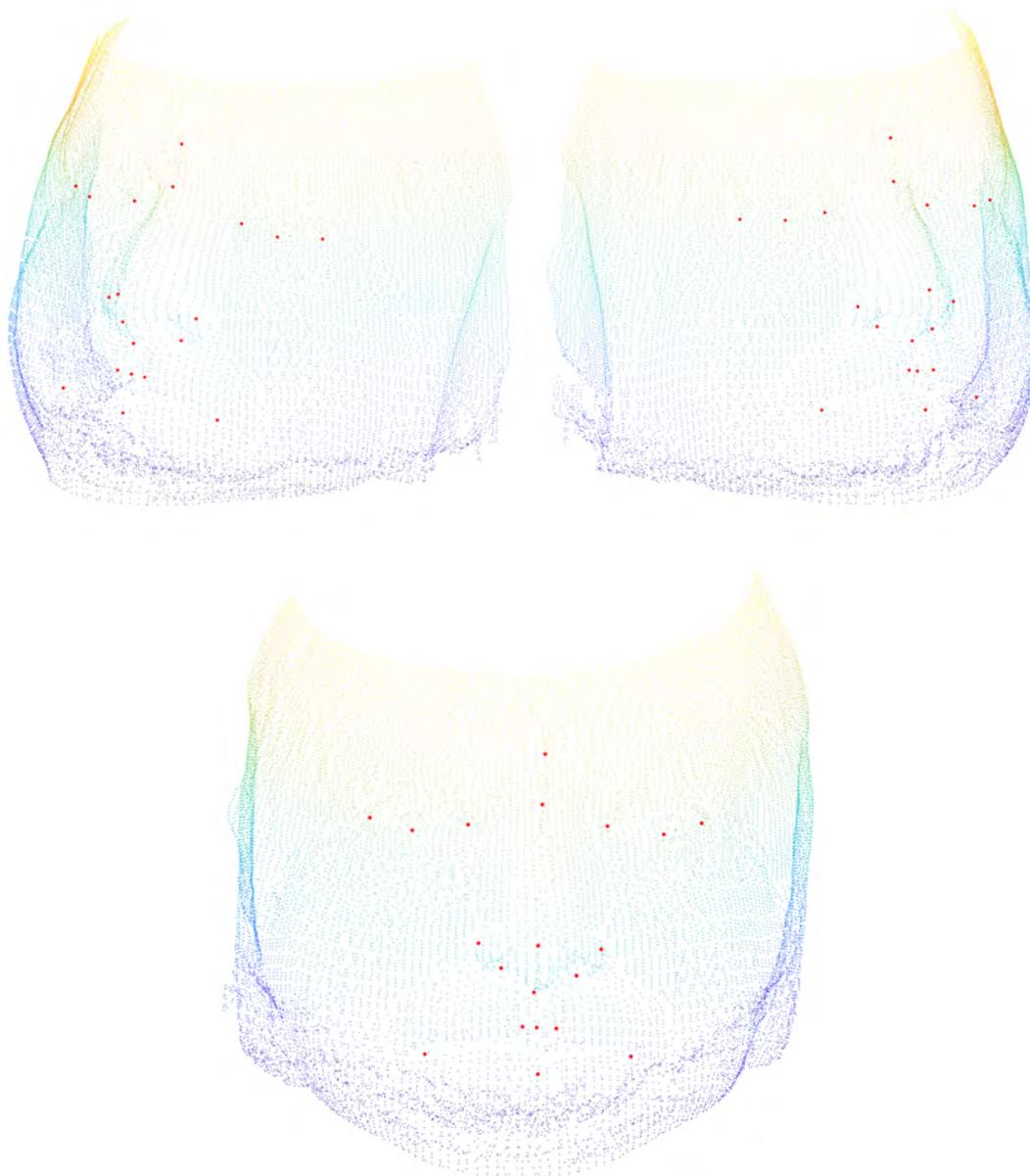


Figura 3.25: Malla facial 3D de un modelo *PointCloud* con *landmarks*

Un inconveniente que presentan las mallas faciales obtenidas a partir del algoritmo explicado anteriormente, es que no existe uniformidad con respecto a la nube de puntos, es decir, existen zonas donde la malla facial contiene más puntos que en otras.

Teniendo en cuenta que más adelante trataremos con la automatización de los *landmarks*, nos interesa tener uniformidad en cada región perteneciente a cada *landmark*. Por ello, a continuación presentamos una variante al método  $AMF_{3D}$  donde conseguiremos obtener una malla facial uniforme.

### 3.3.2. $AMF_{3D}$ - Basado en trazado de rayos

Tal como hemos explicado, la primera variante del método  $AMF_{3D}$  consistía en hacer proyecciones desde el centro del cráneo, hacia el exterior, quedándonos así con los puntos más lejanos. La variante que explicaremos a continuación, consiste en hacer lo contrario, lanzando así proyecciones del exterior, hacia el interior del cráneo.

El primer paso a realizar, es compartido con respecto a la anterior variante del método  $AMF_{3D}$ . Tal como se explicó en el subapartado 3.3.1, este consiste en realizar un filtrado angular de todo el cráneo, quedándonos así con la zona facial. Para ello utilizaremos las Ecuaciones (3.3) y (3.4), obteniendo así un modelo como el que se presenta en la Figura 3.20 y 3.21.

Una vez hecho el filtrado angular, procederemos a hacer las proyecciones desde el exterior, hacia el interior del cráneo. Estas proyecciones se pueden hacer en tres direcciones, proyecciones frontales, laterales, y verticales. En nuestro caso, realizaremos proyecciones frontales y laterales, ya que la parte superior e inferior del cráneo no nos supone ninguna información útil para tratar con la zona facial.

Para poder realizar proyecciones desde el exterior del cráneo, necesitamos un punto de referencia que establezca el origen de la proyección. Para ello se generarán dos cuadrículas en el plano XZ (para las proyecciones frontales) y en el plano YZ (para las proyecciones laterales).

En la Figura 3.26 se pueden apreciar las cuadrículas que se generan en el plano XZ para las proyecciones frontales, y en el plano YZ para las proyecciones laterales.

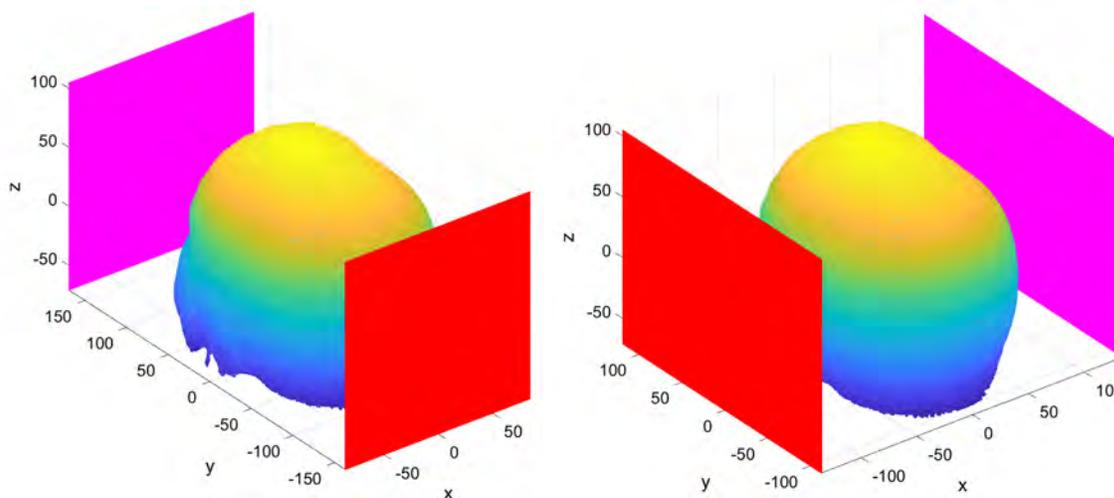


Figura 3.26: Cuadrículas generadas en el plano XZ y YZ para establecer el origen de donde se lanzarán las proyecciones frontales y laterales respectivamente.

El tamaño de las cuadrículas será proporcional al del cráneo, es decir, se definirá mediante el límite de las coordenadas XYZ de la nube de puntos que presenta el modelo 3D. En el caso de la cuadrícula para las proyecciones frontales, el tamaño de esta se calculará mediante el límite de las coordenadas XZ. A diferencia de la cuadrícula para las proyecciones laterales que definirá su tamaño mediante el límite de las coordenadas YZ.

Una vez construidas dichas cuadrículas, crearemos un conjunto de rayos a proyectar

que incidirán sobre el cráneo. En la Figura 3.27 podemos ver los rayos que se proyectan frontalmente des de la cuadrícula generada, hacia el cráneo.

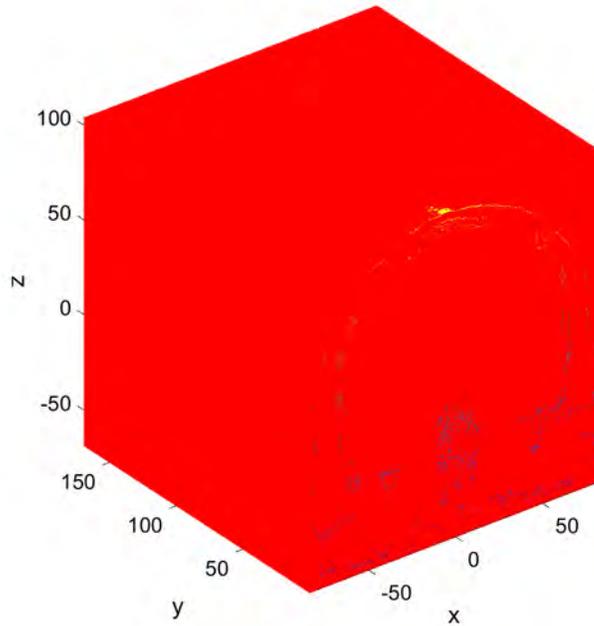


Figura 3.27: Proyección de rayos frontales incidiendo sobre el cráneo.

Tal como podemos apreciar en la Figura 3.27, no todas las proyecciones inciden en el cráneo. Por eso, el siguiente paso, será hallar solamente las proyecciones que inciden en el cráneo, y su primera intersección con este (ver Figura 3.28). De esta manera, lo que hacemos es quedarnos solamente con aquellos puntos que pertenecen al rostro del modelo 3D.

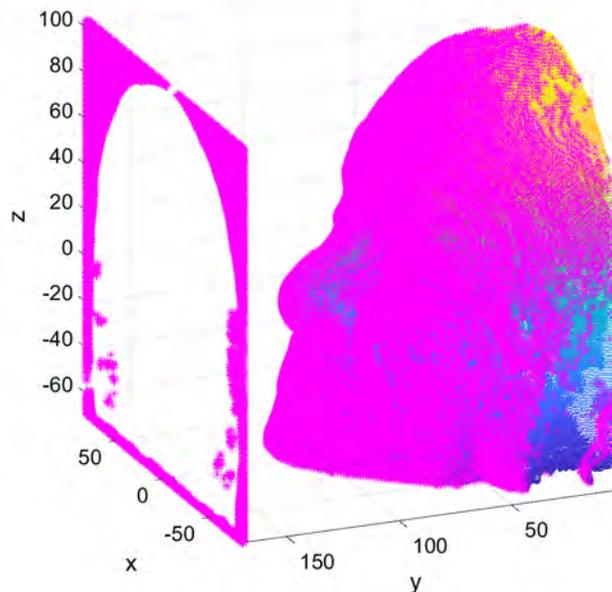


Figura 3.28: Intersecciones entre la cuadrícula y el cráneo.

Finalmente, aplicaremos técnicas de *clustering* para quedarnos con el *cluster* más grande, cual corresponde a la superficie facial. En la Figura 3.29 podemos apreciar como existen diferentes *clusters* representados en colores diferentes. Todo aquel *cluster* que no pertenezca al *cluster* de color rojo, será eliminado.



Figura 3.29: Representación de los *clusters* que presenta el modelo facial 3D.

En la Figura 3.30 podemos apreciar la malla facial finalmente obtenida. Se puede apreciar que esta se caracteriza por su uniformidad en referencia a la nube de puntos. Comparándola pues con la malla obtenida en el subapartado 3.3.1, vemos como esta no presenta uniformidad en su nube de puntos (ver Figura 3.25). Es por ello, que a fin de obtener la malla facial de todos nuestros modelos 3D, utilizaremos la segunda variante del método  $AMF_{3D}$ .

Si recordamos en el subapartado 3.2.1, se orientaron todos los modelos 3D hacia el eje referencia Y. Esta decisión, tal como se explicó en dicho subapartado se hizo para ver que el algoritmo implementado era robusto. Aun así, una vez obtenidas las mallas faciales de todos los pacientes, estas serán orientadas hacia el eje de referencia Z. El motivo de ello es básicamente porque la literatura existente sobre formas tridimensionales, habitualmente esta orientada hacia el eje de coordenadas Z. Por lo tanto, el último paso para obtener nuestra malla facial será orientarla hacia el eje de coordenadas Z mediante las matrices de rotación de la Ecuación (3.2).

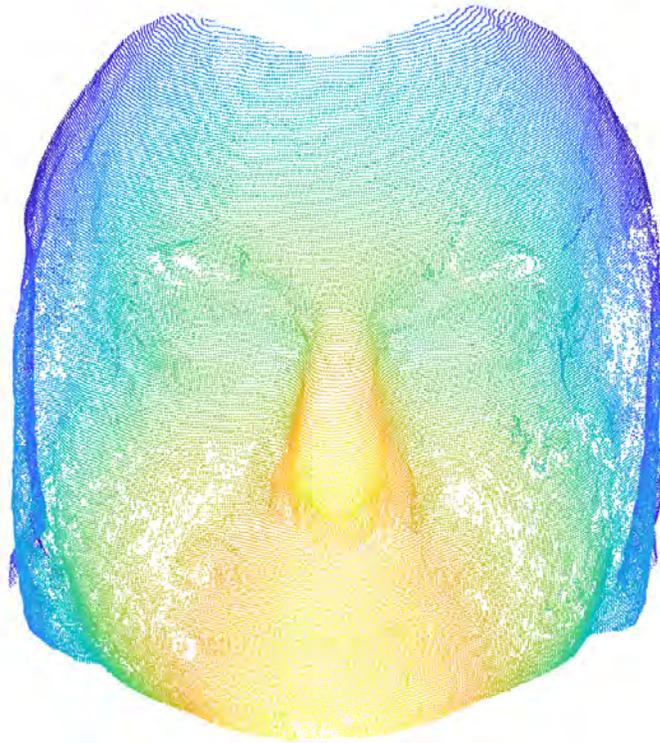


Figura 3.30: Malla facial 3D aplicando el método  $AMF_{3D}$  basado en trazado de rayos.

## Capítulo 4

# Automatización de Landmarking 3D

### 4.1. Introducción

Los *landmarks* o puntos de referencia en su traducción rigurosa, son características de identificación provenientes de un objeto. En el caso de las caras, son elementos espaciales como la punta de la nariz, las comisuras de los ojos, los bordes de la boca, la frente, entre otras.

La localización de *landmarks* es un primer paso importante en muchos algoritmos de reconocimiento facial. En nuestro caso, nos proporcionaran una serie de datos que permitirán posteriormente desarrollar un estudio sobre enfermedades mentales mediante biomarcadores.

Sin embargo, la detección automatizada y la ubicación de estos puntos de referencia son una tarea compleja y difícil de llevar a cabo, ya que la apariencia de estos puede tener un alto grado de variabilidad. Este hecho se puede atribuir a diferencias intrínsecas entre individuos, es decir, una persona puede tener la forma de nariz redonda, mientras que la de otra puede ser más puntiaguda. Otra motivo se debe a factores extrínsecos, la pose y las condiciones de iluminación cuando se captura el modelo pueden perjudicar a dicha detección. Este factor ya se trató en el capítulo 3, por lo tanto no nos proporcionará dificultades añadidas.

Al fin y al cabo, la cara no es un objeto rígido y los cambios en la expresión facial pueden causar grandes cambios en la ubicación de los *landmarks* y la forma 3D.

Tal y como presentamos en el capítulo 1, apartado 1.2, tenemos una selección de 20 *landmarks* que corresponden a estructuras faciales homólogas, y que se pueden registrar en todos los individuos por igual (ver Figura 1.2). El objetivo será conseguir la detección automática de los *landmarks* de un paciente a partir de un modelo entrenado previamente.

Existen múltiples métodos para desarrollar dicha automatización. Trabajos recientes de *landmarking* 3D han hecho uso de la información de curvatura del rostro sobre regiones locales, utilizados posteriormente para clasificar los puntos de interés. Es por ello, que la segunda parte del proyecto consistirá en implementar una metodología basada en álgebra lineal para hallar la curvatura y realizar así la automatización de *landmarking* 3D. La metodología que implementaremos se basará en el estudio científico presentado por *M.C.Ruiz* [14].

## 4.2. Metodología

### 4.2.1. Introducción

En esta sección se presenta el algoritmo  $ALF_{3D}$ , *Automatic Landmarking of Faces in 3D* implementado por Ruiz [14]. Este sistema está basado en estadística y álgebra lineal combinando dos métodos: *Landmark Shape Model* (LSM) y *Landmark Location Model* (LLM).

LSM se basa en un modelo de forma el cual aprende cada región por separado alrededor de una referencia determinada. Posteriormente, este modelo de forma local de cada uno de los puntos de referencia, se modelará individualmente usando un modelo PCA. LLM en cambio, aprende de las ubicaciones en 3D de una colección *landmarks*. Para ello utilizaremos *Active shape models* (ASM) [15].

Cada uno de estos modelos se detallará a lo largo de este capítulo.

### 4.2.2. LSM

En este apartado detallaremos el método LSM basado en la obtención de un descriptor de superficie de forma local en un punto de interés. La forma de los *landmarks* es lo que hace que un punto sea distintivo de otro. Es por eso, que este modelo está diseñado para capturar las similitudes y variaciones de forma de estas regiones en la población.

Para poder modelar la forma de las regiones de cada *landmark*, trabajaremos con métodos de ajuste cuadrático. Estos métodos, permiten estimar propiedades de la superficie como la curvatura. Un estudio sobre estos métodos es presentado por McIvor y Valkenburg [8] y más recientemente por Petitjean [11].

Típicamente, la curvatura se estima de dos maneras. En primer lugar, por geometría diferencial. Sin embargo, esta no es la opción preferida para datos discretos, ya que implica calcular segundas derivadas las cuales son muy sensibles al ruido. Es por esto que utilizaremos la segunda manera. Esta se estima ajustando una superficie cuadrática en el punto de interés, englobando así un conjunto de puntos vecinos.

Para llevar a cabo esto, se requiere de una malla triangular y al ser un ajuste local, es importante que se trabaje con una superficie suave desde la cual se pueda estimar la curvatura con precisión.

Los métodos de ajuste cuadrático se basan en la idea de que la geometría de la superficie, pueda ser localmente aproximada a una superficie polinómica cuadrática. Para realizar este método y acabar obteniendo así la superficie polinómica cuadrática, nos basaremos en la implementación realizada por Tena [12] y extendida por Garimella y Swartz [5].

### Descriptor de superficie

Con tal de proceder a desarrollar la implementación que presentan Tena [12] y Garimella y Swartz [5], se requiere partir de una malla triangular como hemos mencionado anteriormente. Por lo tanto, el primer paso será convertir la nube de puntos procesada en el capítulo 3, a una malla de triangulación. En la Figura 4.1 podemos apreciar el *Point Cloud* y su transformación a una malla triangular de alta resolución.

Antes de presentar el algoritmo y ver como podemos calcular el descriptor de superficie,

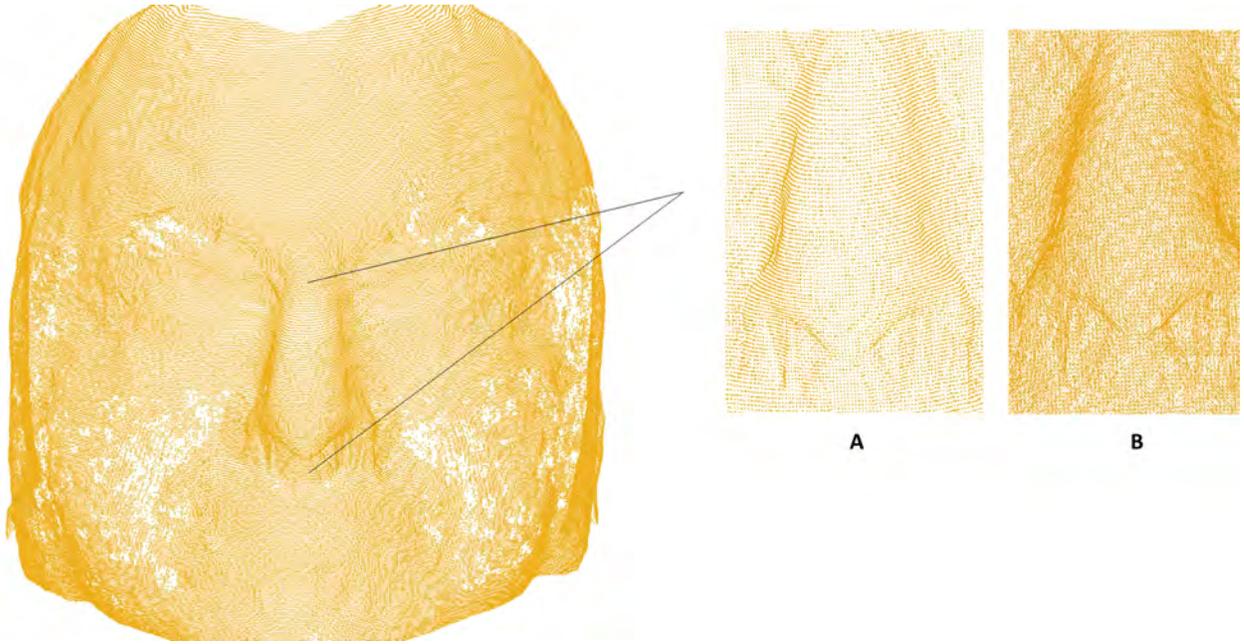


Figura 4.1: Malla facial 3D, (A) *Point Cloud*, (B) malla triangular.

mencionar que este no se aplicará en toda la malla, sino que solo se hará en la región del punto de interés. En nuestro caso definiremos un descriptor de forma local, *Local Shape Descriptor* (LDS), para cada *landmark*. Este descriptor se explicará más adelante.

A continuación presentamos el algoritmo implementado:

- I Estime la superficie normal  $\hat{n}$  en el punto de interés  $\mathbf{p}$ . Esta estimación se puede hacer mediante el promedio simple o ponderado de las normales del triángulo vecino. El vector normal de una superficie da la orientación de la superficie. En nuestro caso, la superficie es un triángulo formado por tres puntos. Pudiéndose calcular como el producto vectorial de la Ecuación (4.1), teniendo en cuenta que el orden de los vectores en el producto vectorial, hará que la normal  $\hat{n}$  apunte hacia afuera o hacia adentro (ver Figura 4.2).

$$\hat{n} = (p_2 - p_0) \times (p_1 - p_0) \quad (4.1)$$

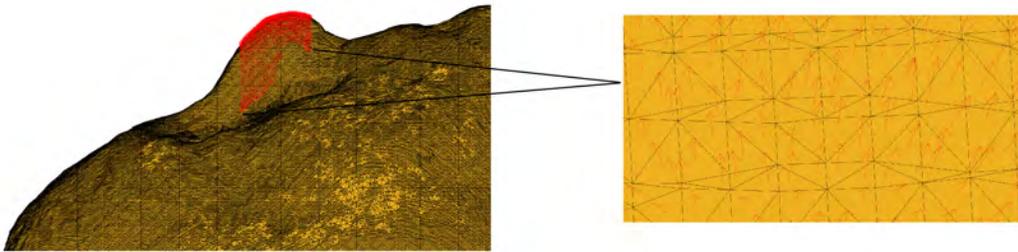


Figura 4.2: Representación de los vectores normales  $\hat{n}$  de la nariz.

- II Alinee la normal respecto el eje de coordenadas  $Z$ . Para poder hacer esto, McIvor y Valkenburg [8] sugieren utilizar la matriz de rotación,  $R = [r_1, r_2, r_3]^T$  como:

$$r_3 = \hat{n} \quad r_1 = \frac{(I - \hat{n}\hat{n}^T)\hat{i}}{\|(I - \hat{n}\hat{n}^T)\hat{i}\|} \quad r_2 = r_3 \times r_1 \quad (4.2)$$

en el que  $I$  es la matriz identidad, y  $\hat{i}$  el vector unitario de  $Y$   $[0,1,0]^T$ . En caso que el vector  $r_1$  sea nulo, el vector  $\hat{i}$  pasa a ser el vector unitario de  $X$   $[1,0,0]^T$ .

III Seleccione el conjunto de puntos que se utilizarán para hacer el ajuste cuadrático. La opción más simple es seleccionar el conjunto de vecinos conectados al punto de interés.

IV Mapee las coordenadas de los puntos seleccionados y aplique la matriz de rotación calculada en el punto II tal que así

$$Q' = R(Q - p) \quad (4.3)$$

donde  $R$  es la matriz de rotación,  $Q$  es el conjunto de puntos seleccionado, y  $p$  es el punto de interés.

v Calcule los coeficientes cuadráticos a partir del polinomio cuadrado. La forma simple del polinomio cuadrado que presenta Tena [12] se puede expresar como:

$$Z = aX^2 + cXY^2 + cY^2 \quad (4.4)$$

Mientras que la forma extendida presentada por Garimella y Swartz [5] se expresa como:

$$Z = aX^2 + bXY + cY^2 + dX + eY \quad (4.5)$$

Para resolver el sistema de Ecuaciones (4.6), se puede expresar como  $\mathbf{Ax} = \mathbf{b}$  y resolver mediante el método de los mínimos cuadrados. Teniendo en cuenta, que para poder resolver el sistema, hemos de seleccionar como mínimo seis vecinos respecto el punto de interés  $\mathbf{p}$ .

$$\begin{bmatrix} x_1^2 & x_1y_1 & y_1^2 & x_1 & y_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^2 & x_ny_n & y_n^2 & x_n & y_n \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} = \begin{bmatrix} z_1 \\ \vdots \\ \vdots \\ \vdots \\ z_n \end{bmatrix} \quad (4.6)$$

Nota: la solución de mínimos cuadrados para un sistema sobredeterminado  $\mathbf{Ax} = \mathbf{b}$  viene dada por  $\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$ .

VI Recalcule la normal como:

$$\hat{n} = \frac{1}{(d^2 + e^2 + 1)}[-d, -e, 1]^T \quad (4.7)$$

VII Repita el punto II hasta que la normal converja.

VIII Estime la curvatura Gaussiana  $K$ , la curvatura media  $H$ , y las principales curvaturas  $k1$  y  $k2$ , todo en referencia al punto de interés  $\mathbf{p}$  mediante:

$$K = \frac{4ac - b^2}{(1 + d^2 + e^2)^2} \quad (4.8)$$

$$H = \frac{a + c + ae^2 + cd^2 - bde}{(1 + d^2 + e^2)^{\frac{3}{2}}} \quad (4.9)$$

$$k_{1,2} = H \pm \sqrt{H^2 - K} \quad (4.10)$$

### Shape Index

Una descripción más compacta de la forma de superficie calculada anteriormente, es mediante el cálculo del índice de forma, *Shape Index*. Este se calcula a partir de las curvaturas principales  $k_1$  y  $k_2$  tal que así:

$$S(p) = \frac{1}{2} - \frac{1}{\Pi} \arctan \frac{k_1(p) + k_2(p)}{k_1(p) - k_2(p)} \quad (4.11)$$

Alguna de las principales características de este descriptor son:

- Las formas de superficie o valor de *Shape Index* se describen con un solo valor en el rango de [0,1].
- El valor de *Shape Index* describen los diferentes tipos de superficie para cada rango como presenta la Figura 4.3.

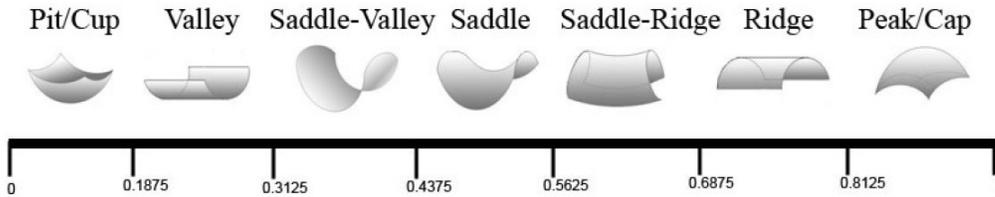


Figura 4.3: Escala de *Shape Index* que muestra los ocho tipos básicos de formas de superficie

### LSD

Si recordamos el objetivo que tiene el método LSM, este se basa en un modelo de forma el cual aprende cada región por separado, alrededor de una referencia determinada. Esta región se definirá como *Local Shape Descriptor* (LSD), un descriptor de forma local.

El LSD que definiremos, consiste en una imagen representada en escala de grises, donde cada píxel adoptará el valor de *Shape Index* que le pertenezca. Para ello, la imagen se generará a partir de la proyección de la nube de puntos 3D, a una cuadrícula 2D (ver Figura 4.4).

Para cada *landmark* se construye una imagen definida por la región a la que pertenece, proyectando la forma 3D, en una cuadrícula 2D perpendicular al eje Z.

Una vez definida la cuadrícula, esta se dividirá en diferentes celdas, en nuestro caso, el tamaño será de  $31 \times 31$  celdas. Para cada celda se asignará un valor de *Shape Index*, y este corresponderá al que pertenezca al punto que se halla proyectado en esa celda.

En caso de que una celda contenga más de un punto, se le asignará el valor de *Shape Index* del punto que se encuentre más cerca del centro de dicha celda. También se puede

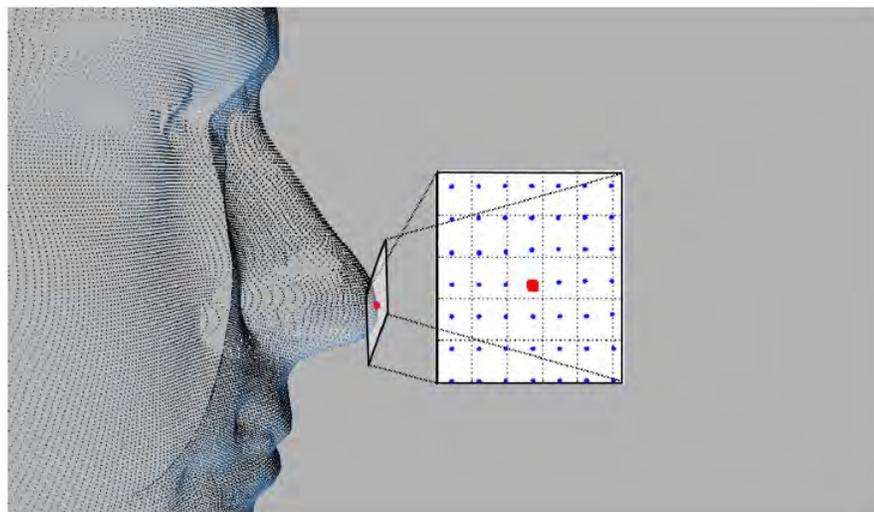


Figura 4.4: Nube de puntos 3D proyectada en una cuadrícula 2D. Extraído de [14].

dar el caso que alguna celda no contenga ningún punto, en tal caso se hará una media de los 4-vecinos más cercanos y se asignará el valor medio.

Haciendo este proceso para cada *landmark*, acabamos obteniendo 20 cuadrículas por paciente, donde cada cuadrícula representará el LSD de cada *landmark*.

En la Figura 4.5 se puede apreciar el LSD que presentan tres *landmarks* en concreto, representando en escala de grises las zonas más oscuras como *cups*, y las zonas más blancas como *cap* según los tipos de forma definidos en la Figura 4.3.

Tal como podemos observar, este hecho se cumple en los resultados presentados. El LSD de la punta de la nariz, principalmente tiene valores de *Shape Index* que en la escala de grises pertenecerían al blanco, y por lo tanto, que presentan curvatura hacia fuera (*cap shape*). A diferencia del LSD del extremo del ojo y de la boca, donde existen curvaturas hacia dentro (*cup shape*), y por eso sus valores de *Shape Index* tienden a negro en la escala de grises.

Una vez obtenidos los 20 LSD por cada paciente, aplicaremos métodos de ordenación para describir este conjunto de datos en términos de nuevas variables (componentes) no correlacionadas mediante *Principal component analysis* (PCA).

## PCA

El objetivo actual es representar todos los datos obtenidos anteriormente, 20 LSD por paciente, en un espacio dimensional global. Para ello, utilizaremos PCA con el objetivo de representar cada *landmark* como un punto en el espacio.

Cuando hablamos de representar cada *landmark* como un punto en el espacio, nos referimos a representar el *landmark* de la nariz de todos los pacientes como un punto en el espacio, y así para los 20 *landmarks* que forman la cara.

La forma de cada *landmark* se modelará independientemente, obteniendo así localizadores de *landmarks* individualizados. Tal como hemos visto en el apartado anterior, cada

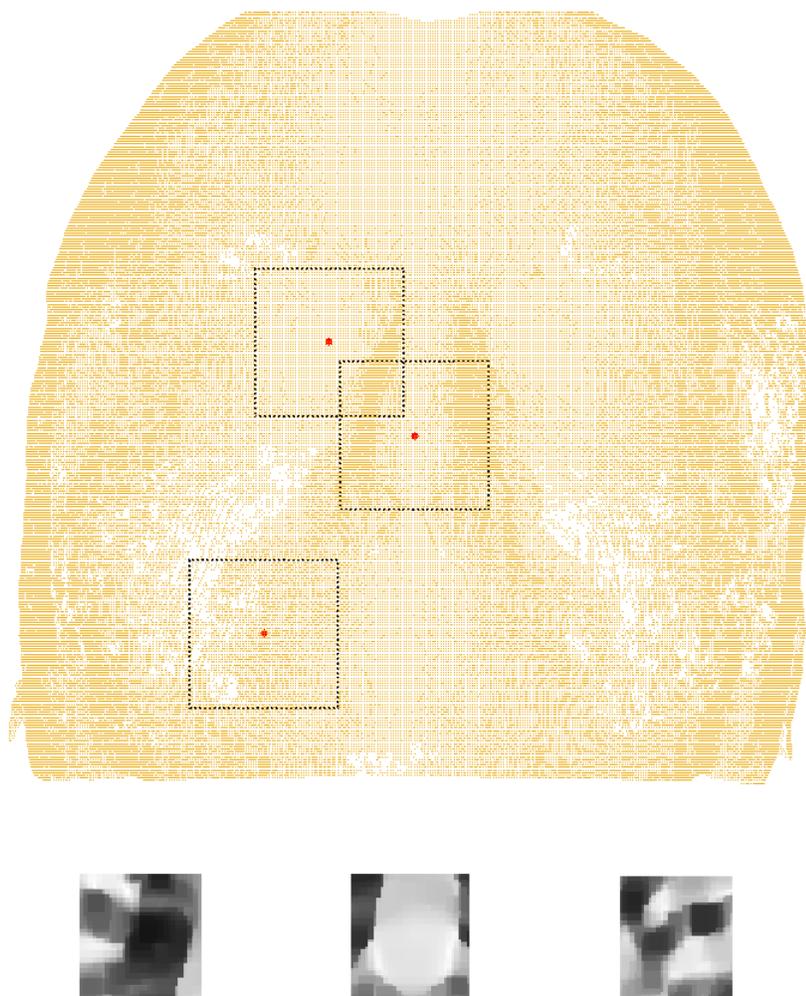


Figura 4.5: LSD de los tres *landmarks* marcados en rojo en el modelo facial 3D. (1) Extremo interno del ojo derecho, (2) punta de la nariz, y (3) extremo derecho de la boca, respectivamente.

*landmark* obtiene su LSD definido por una cuadrícula. Esta se puede expresar como un vector columna, tal que  $31 \times 31 = 961$  vectores dimensionales. De esta manera, podemos realizar PCA en un conjunto de  $N$  vectores para cada *landmark*.

El objetivo de PCA es encontrar un conjunto de ejes ortogonales donde los datos muestran la mayor variabilidad. Los ejes son independientes y cualquier *landmark* se puede escribir como una combinación lineal de proyecciones sobre los ejes. De esta manera, expresando cada LSD como un vector componente, obtenemos la representación de cada LSD como un punto en el espacio de 961D.

Si realizamos el análisis de PCA en estos vectores, obtenemos un descriptor medio de forma, un conjunto de *eigenvectors*, y los valores propios correspondientes a cada *landmark*. La formulación matemática se encuentra en detalle en el capítulo 1, apartado 1.6.

Una vez obtenido el modelo PCA, podemos generar un detector de *landmarks*. Para ello, tendríamos que proyectar el LSD nuevo del *landmark*  $j$  que queremos hallar, en el espacio

definido por los vectores propios tal que así:

$$h_j = P_j - (s_j - \mu_j) \quad (4.12)$$

donde  $\mu_j$  representa el descriptor medio de forma,  $P_j$  es la matriz que contiene los *eigenvectors*,  $s_j$  es la cuadrícula LSD del nuevo paciente que queremos hallar, y  $h_j$  son las componentes del nuevo paciente en la base de *eigenvectors*.

Obtenidas las nuevas proyecciones, estas son comparadas con el modelo PCA entrenado mediante la distancia *Mahalanobis* tal que así:

$$D = \sum_{i=1}^N \frac{h_{ji}}{\lambda_i} \quad (4.13)$$

donde  $N$  es el numero de *eigenvectors* que contiene el modelo,  $h_{ji}$  son las proyecciones a comparar, y  $\lambda_i$  representan los *eigenvalues* del modelo.

Cuanto más pequeño sea el valor de  $D$ , mejor será el ajuste de los datos observados al modelo.

### 4.2.3. LLM

El segundo método que se emplea en el algoritmo ALF<sub>3D</sub> que presenta Ruiz [14] es *Landmark Location Model*. Su objetivo es aprender las ubicaciones en 3D de una colección de *landmarks*.

Este método está basado en el modelo *Active Shape Model* (ASM) presentado por Cootes [15], pero en este caso se extenderá a modelos 3D. Los ASM son modelos paramétricos deformables capaces de generar un modelo estadístico de la variación global de la forma de un objeto.

Este modelo es generado a partir de un conjunto de entrenamiento formado por imágenes y tiene como objetivo, hallar la mejor alineación entre el modelo y los datos de una una imagen.

La forma del objeto en cuestión, será representada a partir de un conjunto de *landmarks*, formando así un *Point Distribution Model* (PDM). En nuestro caso, adaptaremos el modelo a nuestros datos faciales 3D.

Tal como se muestra en la Figura 1.2, trabajaremos con la localización de 20 *landmarks* que construyen un modelo facial 3D. Cada *landmark* se describe mediante las coordenadas 3D  $l = (x, y, z)$ .

Estos *landmarks* nos van a permitir generar el PDM para generar un conjunto de entrenamiento inicial. Este conjunto lo expresaremos como:

$$F_i = l_1, l_2, \dots, l_M \quad (4.14)$$

representado  $i$  el modelo facial 3D del conjunto de entrenamiento,  $l$  las coordenadas 3D de los *landmarks* del modelo  $i$ , y  $M$  el total de *landmarks*.

Este conjunto de muestras que generan el entrenamiento, se han de situar en un marco de referencia común. Para ello, emplearemos Procrustes, método presentado en el capítulo 1, apartado 1.5.

Procrustes se basa en emplear un análisis estadístico de forma que analiza la distribución de un conjunto de formas. El objetivo es eliminar los efectos de traslación, rotación y escala de las muestras de entrenamiento para que solo modele la variabilidad de la forma.

El algoritmo se presenta a continuación:

- I Alinee todas las muestras de entrenamiento partiendo de una muestra de referencia. Para ello se calculará la rotación, traslación y escala que mejor se adapte a la muestra de referencia. La descomposición de valores singulares (SVD) es un método útil para poder realizar este apartado.
- II Calcule la media de todas las muestras de entrenamiento transformadas, obteniendo así un vector de  $1 \times M$ .
- III Use como muestra de referencia la media calculada anteriormente.
- IV Reitere el algoritmo hasta que la muestra de referencia nueva y antigua no varíen.

Concluido el proceso iterativo, podemos reconstruir los modelos 3D originales como:

$$F_i = T_i^{-1}q \quad (4.15)$$

donde  $T_i$  es la acumulación de rotaciones, traslaciones y escalas, y  $q_i$  es el modelo alineado representado como un vector de  $1 \times M$ .

$$q = \begin{bmatrix} x_1^i \\ y_1^i \\ z_1^i \\ \vdots \\ x_M^i \\ y_M^i \\ z_M^i \end{bmatrix} \quad (4.16)$$

Obtenidos los PDM, el objetivo actual será representar todas estas formas faciales 3D en una región de un mismo espacio. Para eso se utilizará PCA, un método que ya se utilizó para LSM, y que explicamos en el capítulo 1, apartado 1.6. En la Figura 4.6 podemos ver una representación gráfica de dichos métodos.

El objetivo de PCA es hallar un conjunto de ejes ortogonales para ver la variabilidad que presentan los datos, en nuestro caso, modelos faciales 3D. Estos modelos se pueden escribir como una combinación lineal de las proyecciones sobre los ejes. En nuestro problema, cada modelo facial se puede escribir como un vector de  $3 \times M = 60$  el cual corresponde a un punto en el espacio de 60D.

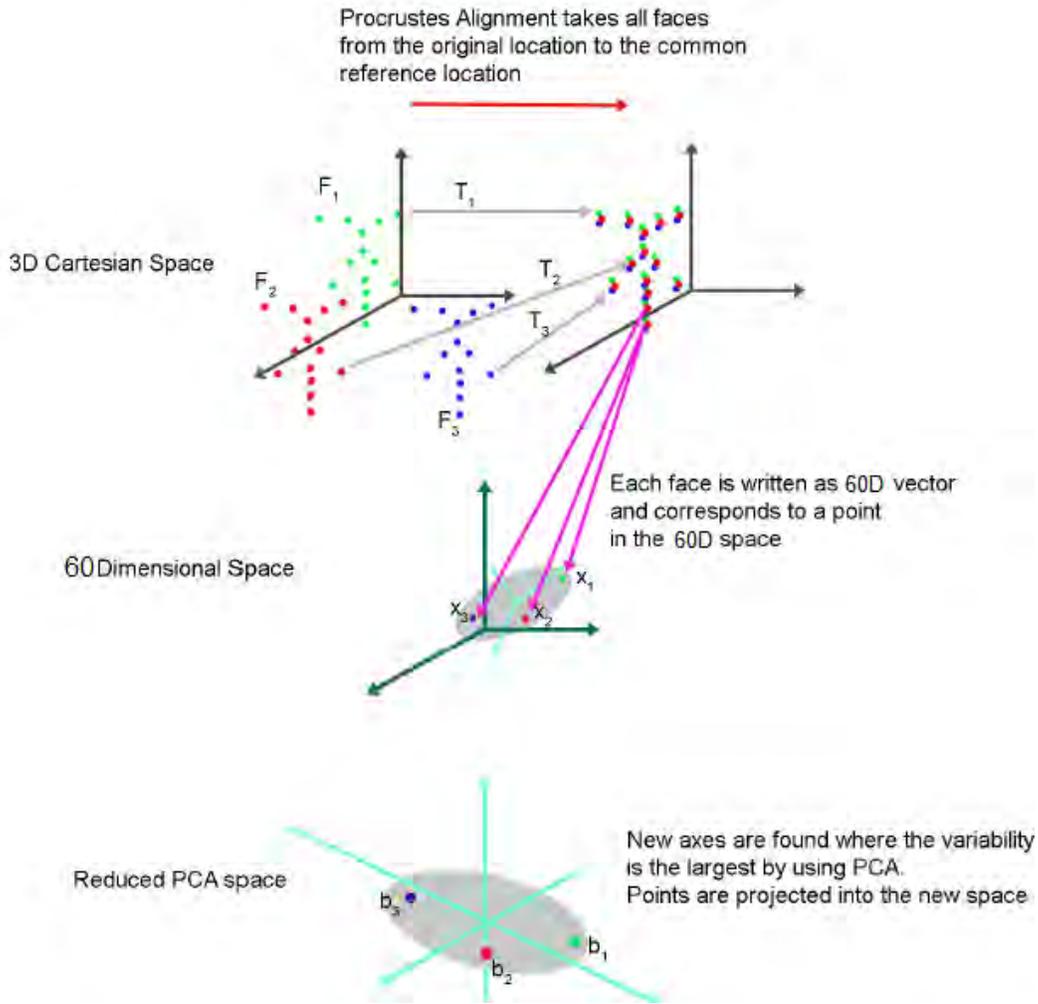


Figura 4.6: Representación de diferentes espacios involucrados en ALF<sub>3D</sub>. (1) Alineación de los modelos faciales 3D hacia una ubicación de referencia común mediante el método de Procrustes. (2) Representación de los modelos faciales como vectores 60D. Cada uno representado como un punto en el espacio 60D. (3) Definición del espacio PCA dimensionalmente reducido. Extraído de [13].

Para poder hallar la variabilidad, el problema se puede formular como la Ecuación (1.24) que presentamos en el capítulo 1, apartado 1.6. Si formulamos de nuevo esta ecuación con nuevas variables, quedaría como:

$$QP = P\lambda \quad (4.17)$$

donde  $P$  es la matriz de *eigenvectors*,  $\lambda$  es la matriz diagonal de *eigenvalues*, y  $Q$  es la matriz de covarianza de los modelos entrenados:

$$Q = \frac{1}{N} \sum_{i=1}^N (q_i - \mu)^T (q_i - \mu) \quad (4.18)$$

donde  $\mu$  es el vector medio de los modelos entrenados:

$$\mu = \frac{1}{N} \sum_{i=1}^N q_i \quad (4.19)$$

Se aplica PCA para poder reducir la dimensionalidad del problema, y poder expresar así los datos como combinaciones lineales de los principales modos de variación representados por los *eigenvectors*.

Una vez aplicado PCA, podemos expresar cada nuevo modelo facial como:

$$b_i = P - (q_i - \mu) \quad (4.20)$$

donde  $\mu$  representa el descriptor medio de forma,  $P$  es la matriz que contiene los *eigenvectors*,  $q_i$  es el modelo facial 3D que queremos describir, y  $b_i$  son las componentes del nuevo paciente en la base de *eigenvectors*. Para obtener el modelo facial original, se puede reconstruir como:

$$q_i = \mu + P b_i \quad (4.21)$$

donde  $\mu$  representa el descriptor medio de forma,  $P$  es la matriz que contiene los *eigenvectors*, y  $b_i$  son las componentes del nuevo paciente calculadas anteriormente.

Una vez se contienen los PDM expresados como *eigenvectors* y *eigenvalues* en el espacio de PCA, se procede a hallar los *landmarks* en los nuevos modelos faciales. Para ello, se empleará ASM.

Llegados a este punto tenemos la implementación tanto de LSM como de LLM (representado como diagrama en la Figura 4.7). El objetivo actual es hallar los *landmarks* de un nuevo paciente combinando ambos métodos.

#### 4.2.4. Combinación entre métodos

En este apartado veremos como hallar los *landmarks* de un nuevo paciente a partir de la combinación de los dos métodos vistos anteriormente.

El objetivo es mover estos nuevos *landmarks* de manera que sean consistentes respecto al modelo PCA presentado en LLM, subapartado 4.2.3. Recordemos que de LLM obtenemos un descriptor medio de forma  $\mu$ , un conjunto de *eigenvectors*  $P$ , y un conjunto de *eigenvalues*  $\lambda$ .

La primera referencia que se fijará en este nuevo paciente se calculará a partir del descriptor medio de forma obtenido en la Ecuación (4.19). Esta es la misma que la última referencia que se utilizó en el paso III del Algoritmo Procrustes presentado en LLM, subapartado 4.2.3. Esta primera referencia se expresará como:

$$F^0 = T^0 q^0 \quad (4.22)$$

donde  $F^0$  representa la referencia de *landmarks* que se le asigna al nuevo paciente,  $T^0$  es la última matriz de rotaciones, traslaciones y escalas obtenida en el método Procrustes, y  $q_0$  es el vector 60D alineado en el espacio Procrustes donde se formó el modelo PCA.

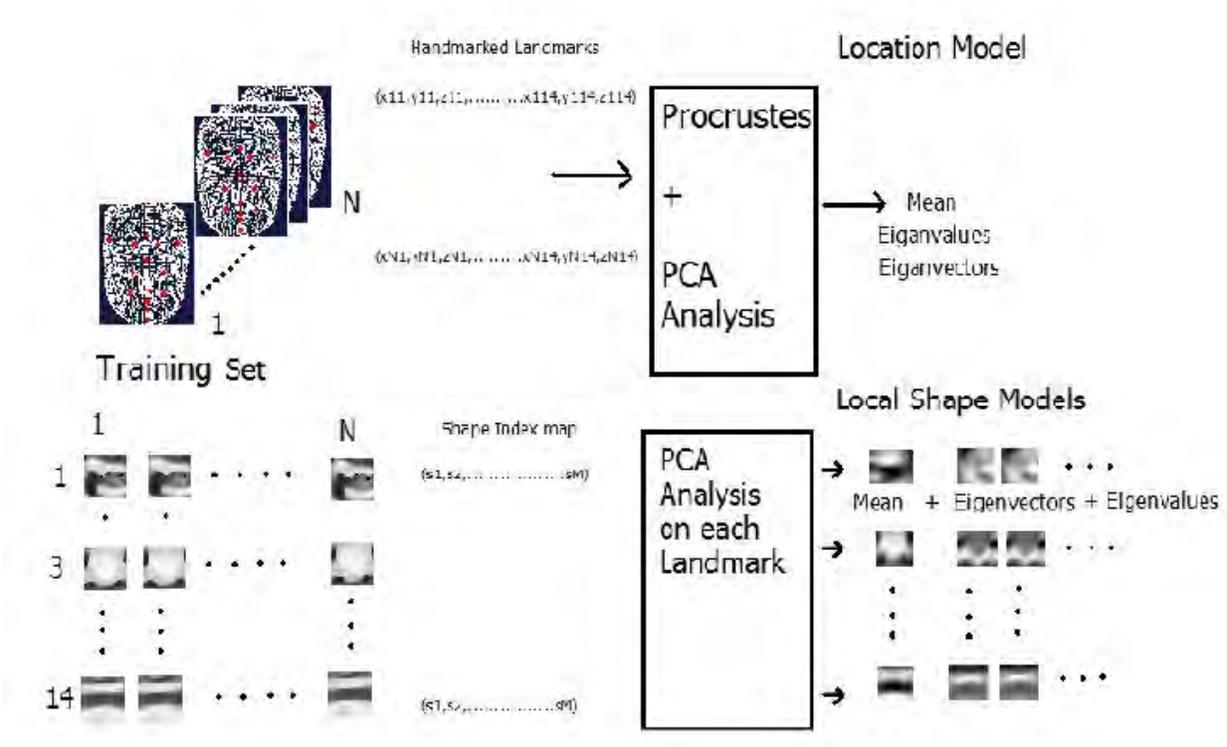


Figura 4.7: Diagrama de bloques de los métodos LSM y LLM. Extraído de [14].

Una vez obtenida la primera referencia, se realizará una búsqueda guiada para hallar la mejor localización de los *landmarks* llamada  $F'$ :

$$F' = F^0 + DF \quad (4.23)$$

donde  $F^0$  es la referencia inicial y  $DF$  es el desplazamiento pertinente con tal de hallar  $F'$ .

### Búsqueda guiada

El objetivo ahora es hallar el valor más cercano de la malla facial 3D del nuevo paciente, respecto  $F^0$ . Para ello se utilizará el algoritmo de *Iterative Closest Point*, ICP, diseñado para realizar alineaciones rígidas (ICP 1).

Podemos expresar  $F'$  como:

$$F' = T^1 T^0 (q^0 + dx) \quad (4.24)$$

donde  $T^1$  pertenece a la transformación rígida mediante ICP, y  $dx$  son las modificaciones que se aplicaran a  $q^0$ .

A continuación presentamos el algoritmo ICP:

Teniendo el nuevo modelo  $M$  en el espacio 3D y la referencia  $D$  ( $F^0$ ) como:

$$M = m_1, m_2, \dots, m_{Ni} \quad D = d_1, d_2, \dots, d_{Nj} \quad (4.25)$$

donde  $M_i$  y  $D_j$  son puntos 3D con coordenadas (x, y, z).

- I Para cada punto  $d_j$  en el conjunto  $D$  encuentre el punto más cercano respecto  $m_i$  del conjunto  $M$ .
- II Ejecute una transformación 3D  $t = [r, t]$ , donde  $r$  es la rotación en 3D, y  $t$  la traslación que mejor alinea  $d_j$  con respecto a  $m_i$ . Esto se puede desarrollar con SVD.
- III Calcule el error medio [4].
- IV Itere este proceso hasta que el error sea menor al de un límite predefinido.

A condición de que estos puntos de referencias obtenidos con ICP se puedan mapear con los pacientes de entrenamiento, estos también aplicarán al método ICP. Por lo tanto, se volverá a aplicar ICP entre  $F^0$  y los *landmarks* (anotados manualmente) de todos los modelos del conjunto de entrenamiento (ICP 2).

Una vez hecho ambos ICP, se calculará la distancia entre las coordenadas que nos proporciona ICP después de cada *landmark* anotado manualmente (ICP 2), y las coordenadas proporcionadas por ICP respecto el nuevo paciente (ICP 1).

Esta distancia se expresará como una función de probabilidad. Si dicha distancia resulta pequeña, indicará que los puntos alrededor del valor extraído de ICP, tienen una probabilidad alta de ser la posición real del *landmark*.

Una vez hallados los puntos con mayor probabilidad a ser la posición real del *landmark*, se calculará el LSD para cada uno de ellos. Obtenidos todos los LSD para cada punto, se mapearán con los LSD obtenidos a partir de los *landmarks* anotados manualmente en LSM. El LSD más parecido al original será nuestra  $F'$ .

Hallado  $F'$  podemos aislar  $dx$  de la Ecuación (4.24) utilizando las Ecuaciones (4.22) y (4.23) como:

$$dx = (T^1 T^0)^{-1} (T^0 q^0 + DF) - q^0 \quad (4.26)$$

Para un nuevo modelo facial  $q^1 = q^0 + dx$ , calculando  $b^1$  con la Ecuación (4.20):

$$b^1 = P(q^1 - \mu) \quad (4.27)$$

y su distancia *Mahalanobis*, podemos comprobar como de robusto es dicho modelo facial respecto al modelo PCA

$$D = \sum_{i=1}^N \frac{b_{ji}}{\lambda_i} \quad (4.28)$$

donde  $N$  es el numero de *eigenvectors* que contiene el modelo,  $b_{ji}$  son las proyecciones a comparar, y  $\lambda_i$  representan los *eigenvalues* del modelo PCA.

A continuación se presenta un resumen del algoritmo iterativo explicado anteriormente:

- I Halle las nuevas localicaciones  $F'$  partiendo de  $F^0$  mediante la búsqueda guiada.

II Calcule  $dx$  mediante la Ecuación (4.26).

III Calcule  $b^1$  mediante la Ecuación (4.27).

IV Calcule la distancia de Mahalanobis mediante la Ecuación (4.28)

V Actualice las siguientes variables:

$$T^0 = T^1 T^0 \quad (4.29)$$

$$q^0 = q^1 \quad (4.30)$$

$$F^0 = T^0 q^0 \quad (4.31)$$

VI El proceso finaliza cuando  $b^1$  converge. En caso contrario, el proceso se itera des del paso I



# Conclusiones

Antes de dar por finalizado el proyecto se añadirán las conclusiones finales y las líneas de futuro a seguir.

El objetivo principal del proyecto consistía en obtener una malla facial 3D a partir de resonancias magnéticas craneales, y realizar la ubicación automática de *landmarks* faciales. En primer lugar, se han preprocesado los datos consiguiendo así uniformidad entre todos los modelos 3D. Para ello se han tenido que implementar diferentes algoritmos a fin de extraer la superficie facial del cráneo. En segundo lugar, se han desarrollado varios métodos basados en álgebra lineal y descriptores de forma con el objetivo de automatizar la ubicación de los *landmarks* faciales.

En la primera parte del bloque práctico se ha tratado la base de datos con el objetivo de que fuera uniforme, es decir, que se pudieran tratar todos los datos por igual. La base consta de un total de 185 resonancias magnéticas craneales junto a 20 *landmarks* cada una, anotados manualmente. La primera apreciación que se hizo fue ver que cada modelo craneal estaba orientado hacia un eje de coordenadas diferente. Este hecho produce que sea imposible realizar cualquier tipo de trato con los datos, ya que no se encuentran en el mismo eje de referencia. Para ello, se diseñó un algoritmo que permite orientar a todos los pacientes hacia el mismo eje de coordenadas. Este consiste en hallar de manera automática la ubicación de la nariz sin tener ningún tipo de referencia, y orientar así el modelo mediante matrices de rotación en el espacio tridimensional. Otra problemática que se observó fue que cada modelo estaba alineado de diferente manera. Para solventar este problema se desarrolló otro mecanismo de automatización que consistía en calcular el ángulo de desviación respecto un punto de referencia. Una vez calculado dicho ángulo, se procedía a alinear el modelo mediante matrices de rotación 3D. El punto de referencia que se asignó fue el de la punta de la nariz del cráneo (previamente orientado). La última problemática que se halló fue que algunas resonancias magnéticas presentaban ruido en sus modelos 3D. Este problema se resolvió aplicando técnicas de *clustering*, aislando y eliminando las zonas con presencia de ruido.

Una vez solventadas las problemáticas citadas anteriormente, se implementó un método que partiendo del modelo craneal construye una malla facial. En primera estancia se desarrolló un método que consistía en realizar una cuantificación angular del modelo y construir así la superficie del rostro a partir de los puntos más lejanos respecto el centro del cráneo. Este método tenía el inconveniente de no mostrar uniformidad en la nube de puntos que formaba la malla facial. Es por ello que se desarrolló una variante del método. Esta consistía en realizar un trazado de rayos hacia el cráneo, es decir, hacer proyecciones hacia el cráneo y construir la superficie del rostro a partir de las primeras intersecciones con este. De esta manera se consiguió obtener una malla facial 3D uniforme, y concluir así la primera parte del bloque práctico.

En la segunda parte el objetivo era implementar un algoritmo que permitiese la ubicación automática de *landmarks* faciales. Este algoritmo está basado en álgebra lineal y descriptores de forma que ayudarán a predecir la ubicación de dichos *landmarks* en nuevos pacientes. La implementación se divide en dos métodos: *Landmark Shape Model* (LSM) y *Landmark Location Model* (LLM). LSM se basa en realizar un estudio de la curvatura que presenta cada punto de la malla facial. Implementando el método, se apreció que el coste computacional era muy alto, por ello, se decidió calcular la curvatura de los puntos que perteneciesen a una región predefinida llamada *Local Shape Descriptor* (LDS). Esta región pertenece a una cuadrícula alrededor de los diferentes *landmarks*, obteniendo así 20 cuadrículas por paciente. Una vez estimado el descriptor de curvatura de dichas regiones, se desarrolló LLM. Este consiste en expresar todos los *landmarks* como un *Point Distribution Model* (PDM). Para ello se alineó todos los *landmarks* de cada paciente en un mismo espacio, el espacio de Procrustes. Una vez implementado Procrustes se implementó *Principal Components Analysis* (PCA) para poder expresar cada PDM como un punto en el espacio dimensional 60D (20 *landmarks* x 3 coordenadas).

Implementados ambos métodos, se combinaron a fin de hacer una búsqueda guiada y hallar la mejor ubicación de cada *landmark* en un nuevo modelo facial (que no contenga *landmarks* anotados manualmente). Para ello se partió de una primera referencia obtenida al hacer PCA la cual se ajustó mediante *Iterative Closest Point* (ICP). A fin de obtener la posición correcta del *landmark*, se trabajó con una función de probabilidad mapeada en el nuevo modelo facial, cual informaba de aquellos puntos con mayor probabilidad de pertenecer a la posición correcta del *landmark*. Una vez hallados los puntos con mayor probabilidad, se calculó los LSD para cada punto y se comparó con los LSD obtenidos en las posiciones de *landmarks* anotados manualmente. El punto con el LSD más parecido sería el candidato a ser la posición del *landmark* del nuevo modelo facial.

# Líneas de futuro

Presentadas las conclusiones, partiremos hacia las líneas de futuro a seguir donde veremos cuales podrían ser los futuros estudios o mejoras relacionadas con el proyecto.

En primer lugar, sería interesante implementar el método  $RALF_{3D}$  que presenta Ruiz en su tesis [13].  $RALF_{3D}$  incluye una variante del algoritmo  $ALF_{3D}$  presentado en este proyecto, que permite mejorar el resultado en la ubicación automática de *landmarks*. Es por ello que sería interesante empezar por desarrollar esta variante y poder realizar así una comparativa entre ambos métodos.

El siguiente paso a seguir sería la implementación de un algoritmo más novedoso tecnológicamente hablando. En este proyecto hemos trabajado con algoritmos muy utilizados en el mundo de procesado y análisis de imagen, basados principalmente en álgebra lineal. Sería interesante pues, poder realizar el proceso de ubicación automática de *landmarks* mediante métodos actuales, basados en *Deep Learning*. Estos métodos son realmente nuevos, pero actualmente ya existen estudios que han trabajado con redes neuronales para desarrollar *landmarking* 3D en mallas faciales.

Ambas líneas de futuro permitirán realizar un estudio completo sobre automatización de *landmarks* 3D, y poder así hacer una comparativa entre métodos y ver así cual presenta mejores resultados con el menor coste computacional posible.



# Bibliografía

- [1] <http://www.fei.com/software/amira-3d-for-life-sciences/>.
- [2] Noemí Hostalet Chavarría. “Facial shape and its neuroanatomical and genetic correlates: analysis of their suitability as a biomarker in schizophrenia and bipolar disorder.” En: 2019.
- [3] Rubèn González Colom. “Evaluation of Artificial Neural Networks and facial biomarkers for diagnostic predictions in schizophrenia.” En: 2018.
- [4] A. Fitzgibbon. “Robust Registration of 2D and 3D Point Sets”. En: 21.1145-1153 (2003).
- [5] Rao Garimella y Blair Swartz. *Curvature Estimation for Unstructured Triangulations of Surfaces*. 2015.
- [6] D. G. Kendall. “The diffusion of shape”. En: *Advances in Applied Probability* 9.3 (1977).
- [7] H.D.Sheets M.L. Zelditch D.L.Swidorski y W.L.Fink. *Geometric Morphometrics for Biologists*. 2004.
- [8] A. M. McIvor y R. J. Valkenburg. *A comparison of local surface geometry estimation methods*. 1997.
- [9] MedlinePlus. <https://medlineplus.gov/spanish/ency/article/003335.htm>.
- [10] M. Muja y David G. Lowe. “Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration”. En: *In VISAPP International Conference on Computer Vision Theory and Applications*. 2009, págs. 331-340.
- [11] Sylvain Petitjean. “A Survey of Methods for Recovering Quadrics in Triangle Meshes”. En: *ACM Comput. Surv.* 34 (2002).
- [12] J. R. Tena Rodríguez. *3D Face Modelling for 2D+3D Face Recognition*. Inf. téc. 2007.
- [13] M.C. Ruiz. “Automatic Face Landmarking in 3D”. Tesis doct. University of Surrey, 2011.
- [14] M.C. Ruiz y J. Illingworth. “Automatic Landmarking of Faces in 3D - ALF3D”. En: *IET Conference Publications* (2008).
- [15] D.H. Cooper T.F. Cootes C.J. Taylor y J. Graham. “Active shape models - Their Training and Application”. En: *Computer Vision and Image Understanding* 61.1 (1995).
- [16] P. H. S. Torr y A. Zisserman. “MLESAC: A New Robust Estimator with Application to Estimating Image Geometry”. En: *Computer Vision and Image Understanding*. 78 (2000), págs. 138-156.
- [17] Wikipedia. [https://en.wikipedia.org/wiki/PLY\\_\(file\\_format\)](https://en.wikipedia.org/wiki/PLY_(file_format)).